

Université de Montréal

**Caractérisation des recombinaisons XerC et XerD  
de *Proteus mirabilis***

par  
Manuela Villion

Département de Microbiologie et Immunologie  
Faculté de Médecine

Thèse présentée à la Faculté des études supérieures  
en vue de l'obtention du grade de docteur  
en microbiologie

Janvier 2005

© Manuela Villion, 2005





## AVIS

L'auteur a autorisé l'Université de Montréal à reproduire et diffuser, en totalité ou en partie, par quelque moyen que ce soit et sur quelque support que ce soit, et exclusivement à des fins non lucratives d'enseignement et de recherche, des copies de ce mémoire ou de cette thèse.

L'auteur et les coauteurs le cas échéant conservent la propriété du droit d'auteur et des droits moraux qui protègent ce document. Ni la thèse ou le mémoire, ni des extraits substantiels de ce document, ne doivent être imprimés ou autrement reproduits sans l'autorisation de l'auteur.

Afin de se conformer à la Loi canadienne sur la protection des renseignements personnels, quelques formulaires secondaires, coordonnées ou signatures intégrées au texte ont pu être enlevés de ce document. Bien que cela ait pu affecter la pagination, il n'y a aucun contenu manquant.

## NOTICE

The author of this thesis or dissertation has granted a nonexclusive license allowing Université de Montréal to reproduce and publish the document, in part or in whole, and in any format, solely for noncommercial educational and research purposes.

The author and co-authors if applicable retain copyright ownership and moral rights in this document. Neither the whole thesis or dissertation, nor substantial extracts from it, may be printed or otherwise reproduced without the author's permission.

In compliance with the Canadian Privacy Act some supporting forms, contact information or signatures may have been removed from the document. While this may affect the document page count, it does not represent any loss of content from the document.

Université de Montréal  
Faculté des études supérieures

Cette thèse intitulée :

Caractérisation des recombinaisons XerC et XerD  
de *Proteus mirabilis*

présentée par :  
Manuela Villion

a été évaluée par un jury composé des personnes suivantes :

Marc Drolet, président-rapporteur et représentant du doyen de la FES  
George Szatmari, directeur de recherche  
Josée Harel, membre du jury  
Paul Roy, examinateur externe

## Résumé

Les recombinaisons spécifiques de site XerC et XerD convertissent les réplicons dimériques en monomères chez *Escherichia coli*. Ces deux recombinaisons font partie de la famille des tyrosine-recombinaisons et agissent de façon coopérative sur des sites particuliers retrouvés sur le chromosome (site *dif*) ou sur des plasmides multicopies (ex : *cer* de ColE1, *psi* de pSC101, etc.). Ainsi ce système, non seulement contribue à la stabilité de ces plasmides, mais assure également la ségrégation de chromosomes monomériques avant que la division cellulaire ne soit enclenchée.

Tout comme c'est le cas pour l'ensemble des espèces bactériennes possédant un chromosome circulaire, les gènes *xer* sont retrouvés chez *Proteus mirabilis*. Ce pathogène opportuniste est reconnu pour les infections urinaires qu'il cause, ainsi que pour son profil de migration caractéristique sur gélose nommé essaimage (« swarming »).

L'ensemble de cette thèse a donc cherché à caractériser davantage le système Xer de cette espèce. Tout d'abord, des réactions *in vitro* ont permis de confirmer les activités de liaison et de clivage de ces protéines aux sites *dif* d'*E. coli* et *cer* du plasmide ColE1. Les recombinaisons de *P. mirabilis* ont également pu reconnaître ces sites en collaboration avec les recombinaisons d'*E. coli*. Cependant, malgré le haut niveau de similarité entre les recombinaisons des deux espèces, la protéine XerC de *P. mirabilis* n'a pas pu agir en collaboration avec la protéine XerD d'*E. coli* pour résoudre et maintenir un vecteur contenant deux sites *cer*, *in vivo*, dans une souche d'*E. coli xer*<sup>-</sup>. Afin de vérifier l'état de cette collaboration au sein de *P. mirabilis*, des tests de complémentation ont également été tentés dans une souche de *P. mirabilis xer*<sup>-</sup>, mais sans succès ; le vecteur rapporteur s'est alors montré instable.

Par la suite, nous nous sommes intéressés à l'expression des protéines Xer chez *P. mirabilis*. Contrairement à ce qui fut supposé chez *E. coli*, la taille de l'ARNm de *xerDPm* ne correspondrait pas à la taille d'un opéron. De plus, grâce à des fusions de transcription à l'opéron *lacZ*, nous avons constaté que le niveau d'activité du promoteur potentiel de ce gène était environ sept fois plus élevé que celui de *xerCPm*.

Dans la poursuite de l'étude du système Xer chez *P. mirabilis*, nous avons voulu vérifier le phénotype engendré par une mutation dans les gènes *xer* de cette espèce. Comme mentionné plus haut, *P. mirabilis* a la capacité d'envahir un milieu gélifié par un phénomène de migration nommé essaimage. Bien que la capacité de migration n'ait pas été affectée par ces mutations, nous avons constaté que le profil obtenu en avait été légèrement modifié. D'un autre côté, le phénotype cellulaire s'est révélé sensiblement le même que celui auparavant observé chez *E. coli* : un plus grand nombre de cellules extraites de géloses inhibant l'essaimage ont formé des filaments au lieu de cellules individuelles.

Finalement, compte tenu du rôle joué par le site *dif* situé dans la région terminale de réplication chez *E. coli*, nous avons recherché ce site chez *P. mirabilis*. Ainsi, nous avons pu cloner et séquencer ce site sur lequel les protéines Xer ont pu ensuite se lier. Une brève analyse des régions flanquantes de ce site a permis de constater que non seulement elles étaient divergentes entre les deux entérobactéries (*E. coli* et *P. mirabilis*), mais également entre deux souches de la même espèce.

**Mots-clés :** *Proteus mirabilis*, recombinaison spécifique de site, tyrosine recombinase, protéines Xer, *cer*, *dif*, essaimage.

## Abstract

XerC and XerD are site-specific recombinases from *Escherichia coli* that convert replicon dimers in monomers. Both recombinases belong to the tyrosine recombinase family and act cooperatively on specific sites which can be found in the chromosome (the *dif* site) or in multicopy plasmids (like the ColE1 *cer* site or the pSC101 *psi* site). This system not only helps to maintain plasmids in the cell but also contributes to the formation of monomeric chromosomes, enabling them to segregate before cell division occurs.

Most bacteria possessing a circular chromosome are shown to have *xer* genes. Not surprisingly, these genes were also found in *Proteus mirabilis*. This opportunistic pathogen is well known for causing infections in the human urinary tract and is well known for its swarming migration which allows the bacteria to colonize certain environments.

In this work, the *P. mirabilis* Xer system was characterized. First of all, it was shown by *in vitro* tests that the two *P. mirabilis* recombinases displayed binding and cleavage activity on the *E. coli dif* site and the ColE1 *cer* site, together or in collaboration with *E. coli* recombinases. *In vivo*, *P. mirabilis* XerC and XerD were able to resolve and monomerize a plasmid containing two *cer* sites, increasing its stability. However, *P. mirabilis* XerC, in combination with *E. coli* XerD, was unable to perform these functions. Complementation tests done using *P. mirabilis xer* mutant strains indicated that, despite the fact that resolution between the two *cer* sites occurred, the plasmid was lost rapidly no matter which combination of Xer proteins were used.

Studies on the expression of the *P. mirabilis xer* genes, using Northern analysis and *lacZ* fusions, demonstrated that the *xerD* promoter activity was seven fold higher than *xerC*. Furthermore, the length of *P. mirabilis xerD* mRNA indicated that *xerD* is monocistronic and is not part of an operon.

The phenotype of *P. mirabilis xer* mutants was also analyzed. Although the ability to swarm was not inhibited, it was observed that the *P. mirabilis xer* mutants displayed slightly different swarming profiles when compared to wild type. On another

hand, these mutants presented a similar cellular phenotype to what was observed in *E. coli xer* mutants: when taken from non swarming agar medium; 10-20% of the cells formed filaments instead of normal short rods.

Finally, the *dif* site, an important region to which both proteins of the Xer system bind and act, was characterized in *P. mirabilis*. Comparisons of the *dif* sites of *E. coli* and *P. mirabilis* showed strong similarities, however significant differences were observed in their flanking sequences. Differences between the flanking sequences were also observed between two strains of *P. mirabilis*.

**Keywords :** *Proteus mirabilis*, site-specific recombination, tyrosine recombinase, Xer proteins, *cer*, *dif*, swarming.

## Table des matières

Résumé.....	iii
Abstract.....	v
Table des matières.....	vii
Liste des tableaux.....	xiii
Liste des figures.....	xiv
Liste des abréviations.....	xvi
Remerciements.....	xxi
Chapitre I : Introduction.....	2
1 Les événements de recombinaison indépendants de RecA.....	2
1.1 La recombinaison spécifique de site.....	2
1.2 Transposases et intégrases rétrovirales.....	2
2 Deux familles de recombinases spécifiques de site.....	3
2.1 Résolvases-Invertases.....	3
2.2 Tyrosine recombinases.....	3
3 Les tyrosine recombinases.....	4
3.1 Quelques exemples :.....	4
3.1.1 Les intégrons.....	4
3.1.2 Les cryptons.....	5
3.1.3 Les protéines Xer.....	5
3.2 Résidus catalytiques des tyrosine recombinases.....	5
3.3 Réaction de recombinaison.....	8
3.3.1 « Strand-Swapping Isomerization model ».....	8
3.3.2 Réaction de clivage.....	10
3.3.3 Arrangement structural des résidus conservés.....	11
3.3.4 Précisions sur le clivage.....	13
3.3.5 Interactions entre les sous-unités.....	14
4 Les recombinases de type Xer.....	16
4.1 Découverte.....	16
4.2 Caractéristiques.....	17
4.3 Sites d'action des recombinases de type Xer.....	18

4.3.1	Liaison aux sites d'action .....	18
4.3.2	Orientation de la réaction.....	19
4.4	La réaction de recombinaison de type Xer .....	21
4.4.1	Petit historique sur le modèle.....	21
4.4.2	Les développements .....	22
4.4.3	Modèle final pour l'action des protéines Xer .....	27
5	Les sites de recombinaison plasmidiques .....	28
5.1	Nécessité de protéines accessoires pour la recombinaison aux sites plasmidiques.....	28
5.1.1	La protéine ArgR.....	28
5.1.2	PepA .....	30
5.1.3	La protéine ArcA.....	31
5.2	Implication topologique des facteurs accessoires.....	31
5.3	Modèles du complexe synaptique au site <i>cer</i> .....	32
5.4	Rôle des protéines accessoires sur le prophage P1 .....	36
5.5	Stabilité plasmidique de ColE1 et le transcrit Rcd .....	36
6	Un autre site d'action pour le système Xer : le site <i>dif</i> chromosomique .....	38
6.1	Généralités .....	38
6.2	La protéine FtsK .....	40
6.2.1	Caractéristiques .....	40
6.2.2	Domaine C-terminal de FtsK .....	41
6.2.3	Similarités de FtsK avec SpoIIIE.....	42
6.3	Études de la recombinaison au site <i>dif</i> d' <i>E. coli</i> avec la protéine FtsK ..	42
6.4	Lien avec XerC et XerD .....	43
6.5	Zone d'activité du site <i>dif</i> .....	44
6.6	Phénomène d'hyper-recombinaison.....	45
6.7	Des éléments polaires dans la DAZ .....	47
6.8	Fragilité du terminus .....	48
6.9	Modèle de co-localisation.....	49
6.10	Les éléments polarisés au terminus : les éléments RAG.....	51
6.11	Orientation de l'activité de FtsK.....	53

6.12	Lien entre la recombinaison et la ségrégation chromosomique.....	54
6.12.1	Modèle d'extrusion-capture .....	57
7	<i>Proteus mirabilis</i> .....	59
7.1	Ubiquité des recombinases de type Xer .....	59
7.1.1	D'autres fonctions pour Xer ?.....	60
7.2	<i>Proteus mirabilis</i> .....	61
7.2.1	La mobilité bactérienne de surface.....	62
7.2.2	L'essaimage chez <i>P. mirabilis</i> .....	62
7.2.3	L'essaimage et la virulence de <i>P. mirabilis</i> .....	65
8	Projet.....	66
Chapitre II : Matériel et méthodes.....		69
1	Souches bactériennes et plasmides employés .....	69
1.1	Milieus de croissance.....	69
1.2	Enzymes employées : .....	74
1.3	Réactions de PCR :.....	74
2	Extractions d'ADN et purifications .....	74
2.1	Méthode de « craquelage (cracking) » .....	74
2.2	Minipreps :.....	75
2.3	Purification de fragments d'ADN extraits de gels :.....	75
2.3.1	Gels d'agarose.....	75
2.3.2	Gels de polyacrylamide .....	76
3	Préparation de cellules compétentes et transformations: .....	76
3.1	Chimiocompétence :.....	76
3.2	Électrocompétence :.....	77
3.2.1	Pour <i>E. coli</i> .....	77
3.2.2	Pour <i>P. mirabilis</i> .....	77
4	Sondes d'ADN (pour hybridation ou pour retards sur gel) :.....	78
4.1	Marquage.....	78
4.1.1	Marquage aléatoire par PCR.....	78
4.1.2	Marquage aléatoire par polymérisation par le fragment Klenow...78	
4.2	Fragments utilisés .....	78

4.2.1	Dans les tests de complémentation dans <i>E. coli xerCD<sup>-</sup></i> .....	78
4.2.2	Dans les réactions <i>in vitro</i> .....	79
5	Recherche et séquençage de <i>xerCPm</i> .....	79
6	Construction de certains plasmides.....	80
6.1	pMVC <sub>mbp</sub> .....	80
6.2	pCS210Km <sup>R</sup> .....	81
6.3	pKΔSacKm <sup>R</sup> .....	81
7	Transfert de type Southern .....	81
7.1	Préhybridation et hybridation d'une membrane récupérée à la suite d'un transfert Southern : .....	82
7.2	Détection et révélation de la digoxygénine .....	82
8	Purifications protéiques.....	83
8.1	MBP-XerCPm et MBP-XerCEc : .....	83
8.2	His-XerDPm et XerDEc : .....	84
8.3	His-XerCPm .....	84
8.4	ArgR.....	84
8.5	PepA .....	85
9	Gels SDS-PAGE .....	85
9.1	Coloration des gels.....	86
10	Test de Bradford .....	86
11	Retards sur gel .....	87
12	Réactions <i>in vitro</i> .....	87
12.1	Avec le site <i>cer</i> .....	87
12.2	Avec le site <i>difEc</i> .....	88
13	Tests de complémentation dans <i>E. coli xer<sup>-</sup></i> .....	88
13.1	Tests de stabilité plasmidique .....	89
13.1.1	Dilutions.....	89
14	Inactivation des gènes <i>xer</i> dans <i>P. mirabilis</i> .....	91
14.1	Inactivation du gène <i>xerDPm</i> : .....	91
14.1.1	Conjugaison bactérienne pour l'inactivation du gène <i>xerDPm</i> : ....	91
14.2	Fusions d'opéron avec <i>lacZ</i> : .....	93

14.2.1	<i>xerCPm</i> : .....	94
14.2.2	<i>xerDPm</i> : .....	95
14.2.3	<i>xerCPm</i> et <i>xerDPm</i> : .....	95
15	Études de l'essaimage (swarming).....	96
16	Microscopie .....	96
17	Dosages $\beta$ -galactosidase.....	97
18	Extraction d'ARN : .....	98
18.1	Électrophorèse d'ARN sur gel d'agarose avec formaldéhyde .....	99
19	Transferts de type Northern.....	100
19.1	Hybridation de membrane récupérée d'un transfert d'ARN .....	101
20	Délétion du site <i>dif</i> dans <i>E. coli</i> :.....	102
21	Détection d'un site <i>dif</i> potentiel chez <i>P. mirabilis</i> .....	103
Chapitre III : Résultats.....		106
1	Première partie.....	106
1.1	Obtention et analyse de la séquence <i>XerCPm</i> .....	106
1.1.1	Organisation des gènes <i>xerC</i> et <i>xerD</i> chez <i>P. mirabilis</i> .....	108
1.2	Expériences de retard sur gel.....	108
1.3	Réactions <i>in vitro</i> .....	110
1.3.1	Au site <i>cer</i> de ColE1.....	110
1.3.2	Au site <i>dif</i> d' <i>E. coli</i> .....	113
1.4	Tests de complémentation <i>in vivo</i> (dans <i>E. coli xerCD<sup>-</sup></i> ).....	115
1.5	Études de stabilité plasmidique.....	116
2	Deuxième partie.....	118
2.1	Tests de stabilité plasmidique dans <i>P. mirabilis</i> .....	119
2.2	Phénotype des souches de <i>P. mirabilis xerC<sup>-</sup></i> et/ou <i>xerD<sup>-</sup></i> .....	120
2.2.1	Essaimage (swarming).....	120
2.2.2	Études microscopiques .....	124
2.3	Expression des gènes <i>xer</i> .....	127
2.3.1	Dosages B-gal .....	127
2.3.2	Hybridations Northern.....	132
2.4	Recherche d'un site <i>dif</i> potentiel chez <i>Proteus mirabilis</i> .....	134

2.4.1	Plusieurs approches essayées .....	134
2.4.2	Le site <i>dif</i> de <i>P. mirabilis</i> .....	141
2.5	Liaison des recombinaisons Xer de <i>P. mirabilis</i> au site <i>difPm</i> :.....	142
Chapitre IV : Discussion.....		146
1	Organisation génomique des recombinaisons de <i>P. mirabilis</i> .....	146
2	Études de complémentation.....	147
3	Études d'expression .....	150
4	Promoteurs potentiels ?.....	153
5	Phénotype des souches de <i>P. mirabilis xerC<sup>-</sup></i> et/ou <i>xerD<sup>-</sup></i> .....	154
6	Le site <i>dif</i> de <i>P. mirabilis</i> .....	157
Conclusion .....		160
Bibliographie.....		162
ANNEXES .....		xxii
Annexe 1 .....		xxiii
Annexe 2 .....		xxiv
Annexe 3 .....		xxx
Annexe 4 : .....		xxxv

## Liste des tableaux

Tableau I: Sites de recombinaison reconnus pas les protéines Xer.....	19
Tableau II : Quelques exemples d'espèces bactériennes possédant les gènes <i>xer</i> .....	60
Tableau III: Souches bactériennes.....	70
Tableau IV: Plasmides.....	71
Tableau V: Amorces pour les diverses réactions de PCR .....	72
Tableau VI: Comparaison de l'activité des promoteurs <i>xerC</i> et <i>xerD</i> à partir des dosages $\beta$ -gal. ....	130
Tableau VII: Moyenne des temps de générations des diverses souches de <i>P. mirabilis</i> . .....	131

## Liste des figures

Figure 1 Alignement des domaines C-terminaux de quelques protéines de type Xer. 8	8
Figure 2: Modèle de la réaction menée par les tyrosine recombinases : « Strand-Swapping Isomerisation model ». ....9	9
Figure 3: Modèle de la réaction de catalyse par Cre sur <i>loxP</i> . ....10	10
Figure 4 : Comparaison des structures quaternaires des protéines Cre (en A, avec 10 pb du site <i>loxP</i> liées au cœur de liaison et au domaine catalytique) et XerD sans son site d'action (en B) .....12	12
Figure 5: Structure d'un protomère Cre lié à un bras de la jonction de Holliday formée de deux sites <i>loxP</i> . ....15	15
Figure 6: Organisation génomique des gènes <i>xerC</i> (en A) et <i>xerD</i> (en B) chez <i>E. coli</i> .17	17
Figure 7: Modèle de l'équilibre entre les interactions XerC-XerD.....27	27
Figure 8: Modèle du complexe formé entre deux sites <i>cer</i> par ArgR et PepA selon Alén <i>et al.</i> (1997). ....33	33
Figure 9 : Modèle du complexe Xer au site <i>cer</i> suite à l'obtention de la structure de PepA.....35	35
Figure 10: Test par l'excision d'un prophage .....46	46
Figure 11 : Modèle de la ségrégation des domaines Ter et de la résolution des dimères chromosomiques.....51	51
Figure 12: Biais d'éléments répétés et orientés riches en G flanquant de part et d'autre le site <i>dif</i> chez <i>E. coli</i> K12 .....53	53
Figure 13 : Schéma du cycle cellulaire d' <i>Escherichia coli</i> (AB1157) à un temps de génération de 50 min.....56	56
Figure 14: Schéma des différentes étapes du cycle d'essaimage chez <i>P. mirabilis</i> ...63	63
Figure 15: Schémas de l'inactivation de <i>xerDPm</i> par double recombinaison et de la création de <i>P. mirabilis xerC<sup>-</sup>xerD<sup>-</sup></i> . ....90	90
Figure 16 : Construction des souches de <i>P. mirabilis xerC<sup>-</sup></i> ou <i>xerD<sup>-</sup></i> par événement de simple recombinaison .....94	94
Figure 17 : Région entourant le site <i>dif</i> chez <i>E. coli</i> .....103	103
Figure 18: Comparaison des séquences en acides aminés des XerC d' <i>E. coli</i> (XerEc) et de <i>P. mirabilis</i> (XerCPm) .....107	107

Figure 19 : Organisation des gènes <i>xerC</i> et <i>xerD</i> de <i>P. mirabilis</i> , souche H14320 comparée à celle retrouvée chez <i>E. coli</i> .....	108
Figure 20: Tests de liaison aux sites <i>difEc</i> et <i>cer</i> .....	109
Figure 21: Réactions <i>in vitro</i> au site <i>cer</i> de ColE1..	112
Figure 22: Réactions <i>in vitro</i> au site <i>difEc</i> .....	114
Figure 23: Tests de complémentation dans <i>E. coli xer<sup>-</sup></i> : extractions plasmidiques.	116
Figure 24: Tests de stabilité plasmidique. ....	117
Figure 25 : Croissance de <i>P. mirabilis</i> sur milieu LA contenant du TTC.....	122
Figure 26 : Étude microscopique de diverses souches de <i>P. mirabilis</i> .....	125
Figure 27: Étude microscopique de diverses souches de <i>P. mirabilis</i> par coloration DAPI. ....	126
Figure 28 : Dosages $\beta$ -galactosidase de <i>P. mirabilis xerD<sup>-</sup></i> .....	128
Figure 29: Dosages $\beta$ -galactosidase des deux souches de <i>P. mirabilis xerC<sup>-</sup></i> .....	129
Figure 30 : Hybridation de type Northern avec la sonde d'ADN <i>xerDPm</i> contre l'ARN total de <i>P. mirabilis</i> .....	133
Figure 31: Hybridation de type Northern avec la sonde <i>xerCPm</i> sur l'ARN total de <i>P. mirabilis</i> . ....	134
Figure 32: Recherche d'un site <i>dif</i> chez <i>P. mirabilis</i> .....	135
Figure 33 : Multimérisation d'un vecteur contenant le site <i>dif</i> d' <i>E. coli</i> .....	137
Figure 34: Comparaison des régions incluant <i>dif</i> chez <i>E. coli</i> MG1655 et chez <i>P. mirabilis</i> 240-82..	140
Figure 35: Comparaisons entre les régions chromosomiques de <i>P. mirabilis</i> et d' <i>E. coli</i> où se trouve le site <i>dif</i> .....	142
Figure 36: Retards sur <i>difPm</i> . ....	144

## Liste des abréviations

### Acides aminés :

A : Ala : Alanine

C : Cys : Cystéine

D : Asp : Acide aspartique

E : Glu : Acide glutamique

F : Phe : Phénylalanine

G : Gly : Glycine

H : His : Histidine

I : Ile : Isoleucine

K : Lys : Lysine

L : Leu : Leucine

M : Met : Méthionine

N : Asn : Asparagine

P : Pro : Proline

Q : Gln : Glutamine

R : Arg : Arginine

S : Ser : Sérine

T : Thr : Thréonine

V : Val : Valine

W : Trp : Tryptophane

Y : Tyr : Tyrosine

\*\*\*\*\*

### Acides nucléiques :

A : Adénine

C : Cytosine

G : Guanine

T : Thymine

B : C ou G ou T

D : A ou G ou T

H : A ou C ou T

K : G ou T

M : A ou C

N : n'importe quel nucléotide (A ou C ou G ou T)

R : purine (A ou G)

S : C ou G

V : A ou C ou G

W : A ou T

Y : pyrimidine (C ou T)

\*\*\*\*\*

### Unités de mesure :

°C : degré Celsius

µl : microlitre

cm : centimètre

Da : Dalton

g : pour la centrifugation :  
représente  $9,81 \text{ m/s}^2$  (force  
gravitationnelle)

g : gramme

h : heure

kb : kilobase(s)

kDa : kilodalton

l : litre

M : molaire

min : minute

ml : millilitre

mM : millimolaire ou millimole/litre

ng : nanogramme

nm : nanomètre

pH : potentiel hydrogène

PM : Poids moléculaire

pmol : picomole

s : seconde

U : unité

v/v : volume par volume

V : volt

V/cm : volt par centimètre

µg : microgramme

µg/ml : microgramme par millilitre

µm : micromètre

\*\*\*\*\*

Général : $\beta$ -gal :  $\beta$ -galactosidase

3'OH : 3 prime hydroxyle

aa : acide aminé

ADN : Acide désoxyribonucléique

Ap : Ampicilline

APS : ammonium persulphate

ARN : Acide ribonucléique

BSA : Bovine serum albumine; albumine de sérum de boeuf

CDP-Star : Disodium 2-chloro-5-(4-methoxyspiro{1,2-dioxetane-3,2'-(5'-chloro)tricyclo[3.3.1.1.3,7]decan}-4-yl)-1phenyl phosphate

Cl<sup>-</sup> : ions chlore

Cm : chloramphénicol

Comm. Pers. : communication personnelle

CTAB : bromure d'ammonium cetyl-triméthyle (détergent cationique)

C-terminal : Carboxy-terminal

DAPI : 4',6-diamidino-2-phenylindole

ddH<sub>2</sub>O : eau déionisée ou dite MilliQ

DEPC : diethyl pyrocarbonate

dH<sub>2</sub>O : eau distillée

Dig : Digoxigénine

Dig-11-dUTP : Digoxigénine-11-2'-désoxy-uridine-5'- triphosphate

dNTP : désoxyribonucléotide

DNase : Désoxyribonucléase

DO : densité optique

DO<sub>600</sub> : densité optique prise à 600 nm

DTT : dithiothreitol

EDTA : Acide éthylène diamine tetra-acétique

*et al.* : et autres [abréviation de *et alii* (masculin pluriel) ou *et aliae* (féminin pluriel) ou *et alia* (pluriel neutre)]

EtOH : éthanol

Inc. : « Incorporated »

Int : Intégrase

IPTG : isopropyl-beta-D-thiogalactoside

Km : Kanamycine  
LA : LB-agar  
LB : Milieu Luria-Bertani  
LSW : « low swarming agar »  
MBP : maltose binding protein (protéine se liant au maltose)  
Mg<sup>++</sup> : ion magnésium  
Mn<sup>2+</sup> : Manganèse  
MOPS : acide morpholinéopropanesulfonique  
Na<sup>+</sup> : ion sodium  
NEB : New England Biolabs  
Ni<sup>2+</sup> : ion nickel  
Ni-NTA : Ni<sup>2+</sup>-acide nitrilotriacétique  
N-terminal : Amino-terminal  
OH<sup>-</sup> : ion hydroxyle  
ONPG : O-nitrophényl β-D-galactopyranoside  
ORF : open reading frame (cadre de lecture ouvert)  
PAGE : polyacrylamide gel electrophoresis; électrophorèse sur gel de polyacrylamide  
Par ex. : par exemple  
pb : paire de bases  
PBS : phosphate buffered saline (saline tamponnée au phosphate)  
PCR : Polymerase chain reaction (réaction de polymérisation en chaîne)  
PMSF : « phenylmethylsulphonyl fluoride »  
<sup>R</sup> : indique la résistance  
ref. : reference  
RNase : Ribonucléase  
SAP : « shrimp alkaline phosphatase » (phosphatase alcaline de crevette)  
SDS : Sodium dodecyl sulfate  
SDS-PAGE : Polyacrylamide gel electrophoresis in sodium dodecyl sulfate (gel d'électrophorèse de polyacrylamide contenant du SDS)  
SSC : Salt sodium citrate (sel et citrate de sodium)  
*sp.* : « species »  
Sm : streptomycine  
TAE : Tris-acétate EDTA  
TBE : Tris-borate EDTA  
Tc : Tétracycline

TE : Tris-EDTA (10 mM-0,1 mM)

TEMED : N, N, N', N'-tetramethylethylenediamine

TM : trademark

Tn : transposon

Tp : triméthoprime

TTC : 2,3,5-triphenyltetrazolium

UV : ultraviolets

vs : versus

X-gal : 5-bromo-4-chloro-3-indolyl-b-D-galactopyranoside

Zn<sup>2+</sup> : Zinc

*À François*

## Remerciements

Après toutes ces années passées au sein du laboratoire, je ne pourrais terminer ce document sans remercier la personne-ressource m'ayant accueillie dans son équipe : le professeur George Szatmari. Merci George pour tout, non seulement pour l'acquisition des connaissances techniques et théoriques au laboratoire, mais aussi pour la possibilité d'acquérir une certaine expérience de démonstratrice, d'avoir été mise en contact avec le monde de la vente de produits de laboratoire de biologie moléculaire, les responsabilités octroyées, etc.

Je tiens également à remercier les professeurs qui ont accepté de faire partie du comité de correction de cette thèse : Marc Drolet, Josée Harel, Paul Roy.

Je tiens à remercier tout particulièrement Micheline Fluet-Desmeules pour avoir donné les souches de *Proteus* avec lesquelles j'ai pu travailler tout au long de ces dernières années, pour avoir si gentiment donné de son temps pour l'ensemencement de souches en dehors de ses heures de travail et pour tous les documents qu'elle garde précieusement afin de les prêter à qui en aura besoin. Également Marielle Thivierge-Parent pour toutes les questions au sujet des milieux de cultures, les prêts de solutions à « usage unique »...

Je voudrais remercier chaleureusement tous les gens du département sans exception, pour toutes les questions auxquelles ils ont bien voulu répondre. Que ce soit d'un point de vue administratif (merci à Lise David pour m'avoir rappelée à l'ordre au moment des réinscriptions ou des lettres de prolongations !, merci à Nicole Fallecker pour toutes les bonnes discussions, merci à Carole Arsenault pour les arrangements de comptes biobars...) ou académique (tous les bons conseils de Marc Drolet qui m'a bien écoutée lorsque les doutes m'assaillaient au sujet de certains de mes résultats, ou encore Pierre Belhumeur pour tous ses partages d'idées... et oui, l'hybridation sur colonies a fini par fonctionner !).

Merci aussi à Annie-Claude Mialhe pour la chance offerte de découvrir un aspect du monde industriel avec Pierre Belhumeur, et aussi pour l'approfondissement de la systématique bactérienne. Merci à Daniel Morency pour tous ses essais photographiques afin d'observer adéquatement l'essaimage bactérien et à Monique Vasseur pour ses photos de cellules colorées au DAPI.

Je ne peux oublier non plus Julie Proulx qui a fait toutes les colorations de cellules de *Proteus*, et Valérie Haf-Gagné et Tien Lien Mi pour leur collaboration dans les premiers essais de complémentation.

Je voudrais faire un coucou tout particulier à Georges, Mélanie, Karine, Sonia et Jean-François, ainsi que tous les gens qui ont traversé le labo pendant mon doct : (procédons par ordre !: Nadine, Dimitri, Jean-Frédéric, Loubna, Hélène, et les nouveaux qui prennent la relève : Jérémy et Fuli, sans oublier bien sûr tous les stagiaires qui nous ont fait sourire tous les étés !).

Et enfin, je voudrais souligner l'effort incommensurable de François qui a tenté tant bien que mal de devenir un psychologue averti sur les sautes d'humeur d'une thésarde finissante et le papa qui m'a bien fait comprendre l'importance de toutes les racines des arbres.

# **INTRODUCTION**

# Chapitre I : Introduction

## 1 Les événements de recombinaison indépendants de RecA

### 1.1 La recombinaison spécifique de site

Dans les années 1950-1960, Allan Campbell concluait que le bactériophage  $\lambda$  s'insérait et s'excisait du chromosome d'*Escherichia coli* afin d'expliquer comment ce bactériophage attaquait son hôte. Ce phénomène a mené au concept de recombinaison spécifique de site et depuis lors, plusieurs systèmes similaires ont été étudiés (lire Nash, 1996).

Les recombinases spécifiques de site comme leur nom l'indique, agissent sur des sites précis, afin de remplir diverses fonctions biologiques. Ainsi, non seulement ce type de recombinaison peut-il servir à l'intégration et à l'excision de bactériophages et autres éléments génétiques dans le chromosome bactérien, mais il peut aussi permettre :

- la monomérisation des dimères chromosomiques et plasmidiques afin d'en faciliter la ségrégation aux cellules filles,
- le contrôle de l'expression génique,
- la réplication de certaines portions d'ADN par le repositionnement d'éléments de contrôle ou de segments de gènes et assemblage de gènes actifs par la délétion du matériel inclus entre les gènes, etc. (Craig, 2002).

En plus d'être spécifique (spécificité de liaison sur le site d'action), cette recombinaison, indépendante de RecA, est dite conservative, parce qu'il n'y a pas de dégradation ni de nouvelle synthèse d'ADN, elle ne nécessite pas de cofacteurs hautement énergétiques comme l'ATP et le nombre total de liens phosphodiester reste constant tout au long de la réaction (Mizuuchi et Baker, 2002).

### 1.2 Transposases et intégrases rétrovirales

Une autre catégorie de recombinases indépendantes de RecA, mais cette fois non spécifique de site et permettant la mobilisation de morceaux d'ADN, inclut les transposases et intégrases rétrovirales. Les éléments transposables sont définis comme des « segments discrets d'ADN (que ce soit directement ou via un intermédiaire d'ARN) capable de bouger d'un locus à l'autre dans le génome de leur hôte ou entre différents génomes » (Haren *et al.*, 1999). Ainsi, l'ADN génomique lui-même est l'intermédiaire mobile pour les

transposons (classe II), tandis que les intégrases rétrovirales (classe I) transposent via un intermédiaire d'ARN et encodent une transcriptase inverse (Haren *et al.*, 1999). Dans cette autre catégorie peuvent être retrouvés les transposons bactériens, les éléments d'insertion, les bactériophages se transposant tel Mu, la recombinase RAG permettant l'initiation de la recombinaison V(D)J afin d'assembler des segments de gènes pendant la maturation du système immunitaire chez les vertébrés, etc. (Mizuuchi et Baker, 2002).

## 2 Deux familles de recombinases spécifiques de site

Une définition simple de la recombinaison spécifique de site consiste à dire qu'« elle est une réaction dans laquelle les brins d'ADN sont brisés puis échangés en des positions précises sur deux cibles d'ADN et ce, afin de permettre une fonction biologique déterminée » (Hallet et Sherratt, 1997). Deux grandes familles ont ainsi été définies selon le résidu catalytique employé pour exercer l'activité de clivage. Il s'agit des résolvases-invertases (ou sérine recombinases) et des intégrases de la famille de l'intégrase du phage  $\lambda$  (maintenant nommées tyrosine recombinases).

### 2.1 Résolvases-Invertases

Les membres les plus connus de la famille des résolvases-invertases sont la résolvasse  $\gamma\delta$  et le Tn3 servant à la résolution de co-intégrats, l'invertase Hin servant à l'activation (« switching ») de l'antigène flagellaire chez *Salmonella* et l'invertase Gin employée pour l'inversion de la queue du phage Mu (voir Smith et Thorpe, 2002). Toutes ces enzymes clivent l'ADN par l'intermédiaire d'une sérine qui se trouve dans le domaine N-terminal de la protéine. Cette sérine va permettre le clivage concerté des quatre brins d'ADN à une séquence de 2 pb. Un lien phosphoséryle est ainsi créé du côté 5' récessif, ce qui libère un groupement 3'OH. La ligation se fait ensuite par le processus inverse : le groupe 3'OH va attaquer le lien 5'phosphoséryle afin de libérer l'enzyme et permettre la ligation de l'ADN pour obtenir les produits recombinants (Hallet et Sherratt, 1997).

### 2.2 Tyrosine recombinases

D'un autre côté, les membres de la famille des tyrosine recombinases exercent leur action par le biais d'une tyrosine située dans le domaine C-terminal de la protéine. Le clivage se fait d'un côté d'une région d'enjambement d'une longueur variant de 6 à 8 pb pour donner lieu à une jonction de Holliday. Deux brins sont ainsi clivés en même temps,

et non pas quatre, comme ce qui est vu avec les sérine recombinases. Il y a création d'un lien 3'phosphotyrosyle et le groupement 5'OH libéré pourra alors attaquer le lien 3'phosphotyrosyle du brin partenaire afin de créer la jonction de Holliday. Les produits recombinants proviennent ensuite de l'échange des deux autres brins par le même processus décrit précédemment (Hallet et Sherratt, 1997). Ce processus sera détaillé dans la section 3.3.

### 3 Les tyrosine recombinases

Il existe au-delà de 105 membres de la famille de l'intégrase du phage Lambda (Nunes-Düby *et al.*, 1998). Tous les membres possèdent une similitude au niveau des acides aminés employés pour exercer leur action de recombinase et ils exercent leur action de clivage de la même façon. Parmi tous ces membres, certains se recoupent suffisamment pour créer différentes sous-familles. Par exemple, la famille des Fim contenant les protéines FimB/FimE permettent le réarrangement du promoteur des gènes codant pour des *fimbriae* d'*E. coli* (Klemm, 1986), la famille de la protéine Flp contient des recombinases impliquées dans le maintien et le contrôle du nombre de copies de plasmides de levures (Sadowski, 1995). D'autres recombinases n'ont aucune similarité suffisante pour pouvoir être placées avec d'autres membres. Tel est le cas de la protéine Cre servant au maintien du prophage P1 à l'intérieur d'*E. coli*. Le regroupement de toutes ces sous-familles est retrouvé dans le site maintenu par Dominic Esposito : <http://mywebpages.comcast.net/domespo/tr/family.html> (cependant, la dernière mise à jour date de janvier 1999).

Dans les systèmes de recombinaison les plus simples de cette famille (tels Flp et Cre), seuls les cœurs des sites de recombinaison sont nécessaires pour une action de recombinaison. D'autres systèmes (comme le système Xer ou l'intégrase du phage  $\lambda$ ) ont besoin de sites accessoires afin de mettre en place un complexe synaptique fonctionnel.

#### 3.1 Quelques exemples :

##### 3.1.1 Les intégrons

Les intégrons contiennent minimalement un gène *intI*, qui encode la tyrosine recombinase, et un site adjacent de recombinaison, le site *attI*. Souvent, plusieurs cassettes sont intégrées en plus à *attI* ; ces cassettes contiennent un site *attC* et un ou plusieurs gènes

de résistance aux antibiotiques ( $\beta$ -lactames, chloramphénicol, triméthoprime, etc.) ou d'autres gènes détectables phénotypiquement (comme des facteurs pour la pathogénicité, des systèmes de restriction-modification, des toxines, etc.). Les cassettes seules sont des éléments considérés comme non-autonomes parce qu'elles requièrent la recombinase de l'intégron pour leur propre mobilité, et que cette recombinase est non mobilisable.

L'importance des intégrons a été révélée par la dissémination de la résistance aux antibiotiques au sein des populations bactériennes par transfert horizontal. Il est surtout retrouvé chez les bactéries gram-négatives pathogènes, comme les bactéries de la famille des *Vibrionaceae* ou des *Pseudomonas* (pour une revue de littérature sur ce sujet, lire Recchia et Sherratt, 2002).

Ces intégrons contenant des cassettes peuvent être retrouvés sur le chromosome, sur des éléments transposables ou encore sur des plasmides conjugatifs. Cependant, il n'est pas possible de les définir avec exactitude puisque les séquences qui les entourent n'ont pas encore été précisées. Les intégrons sont plutôt définis en termes fonctionnels : « les intégrons sont des éléments de capture de gènes encodant les déterminants pour la mobilisation et pour l'expression de cassettes de gènes » (Recchia et Sherratt, 2002).

### **3.1.2 Les cryptons**

Les différents facteurs impliqués dans la dynamique du génome ne sont pas toujours si aisément catégorisés. Ainsi, les cryptons sont un groupe récemment identifié de transposons mais qui encodent une recombinase de la famille des tyrosine recombinases, au lieu de la transposase classique (Goodwin *et al.*, 2003).

### **3.1.3 Les protéines Xer**

Les protéines Xer forment de leur côté le groupe le plus uni et le plus important (en nombre de représentants jusqu'à ce jour déterminés) de la famille des tyrosine recombinases. La suite de cette thèse servira d'ailleurs à tenter de mieux définir ce groupe et à mettre à jour les études faites sur ces recombinases.

## **3.2 Résidus catalytiques des tyrosine recombinases**

La famille des tyrosine recombinases a tout d'abord été établie par comparaisons de séquences et par des expériences de protéolyses limitées. Elle avait été définie par des

similitudes dans le domaine catalytique C-terminal. Ainsi, tous les membres de cette famille, malgré un degré de similitudes plutôt faible au niveau de la séquence globale, présentent des résidus conservés impliqués dans la catalyse de la réaction. Des études plus approfondies de comparaisons de séquences ont pu ensuite être réalisées à la suite de la cristallisation de certaines des recombinases et ont permis de révéler la présence de plusieurs motifs caractéristiques de cette famille (Esposito et Scocca, 1997, et Nunes-Düby *et al.*, 1998).

Les premières études avaient révélé quatre acides aminés absolument conservés dans la région C-terminale des protéines : une triade composée de deux arginines, puis d'une histidine (R-H-R) et un résidu tyrosine (Argos *et al.*, 1986, Abremski et Hoess, 1992). Ce résidu tyrosine fut déterminé comme le résidu nécessaire à la catalyse (tout d'abord chez l'intégrase du phage Lambda- Pargellis *et al.*, 1988). Alors que la première arginine se situe dans un premier motif hautement conservé (motif I, voir Figure 1), les résidus H-R-Y se situent dans un second motif, lequel est en extrémité C-terminale de la séquence protéique (motif II, voir Figure 1).

Plus récemment, un autre résidu impliqué dans la catalyse a été déterminé suite à l'analyse de la structure secondaire de plusieurs recombinases. Ce résidu, une lysine, est conservé avec la famille des topoisomérases de type IB eucaryotes (Cao et Hayes, 1999); il est situé dans un troisième motif retrouvé entre les deux premiers précédemment identifiés. Chez les protéines Xer, il ferait partie d'une séquence GKGXKXR (Cao *et al.*, 1997) placée dans le tour entre les feuillets  $\beta 2$  et  $\beta 3$  de la séquence (emplacement établi selon la structure de XerDEc). Bien que très conservé, ce résidu est remplacé par une arginine chez certaines recombinases (Nunes-Düby *et al.*, 1998).

Enfin, une seconde histidine retrouvée dans le motif II participerait aussi à la catalyse générale acide-base ; ce résidu est remplacé par l'acide aminé tryptophane chez la recombinase Cre (Sherratt et Wigley, 1998).

Les trois blocs de séquences présentées en Figure 1 montre l'emplacement des différents motifs chez certaines recombinases de type Xer, basé sur la structure de XerD d'*E. coli* (dans l'ordre : motif I, III et II).



**Figure 1 Alignement des domaines C-terminaux de quelques protéines de type Xer.** L'alignement a été fait grâce au programme pileup de GCG (Genetics Computer Group, Wisconsin package version 9.1, Madison, Wisconsin, © : 1982-1997), et le consensus avait été mis en valeur en 1998 par le programme BOXSHADE 3.21 (de K. Hofmann et M. Baron, anciennement retrouvé sur le site [http :\\www.isrec.isb-sib.ch/software/BOX\\_form.html](http://www.isrec.isb-sib.ch/software/BOX_form.html)).

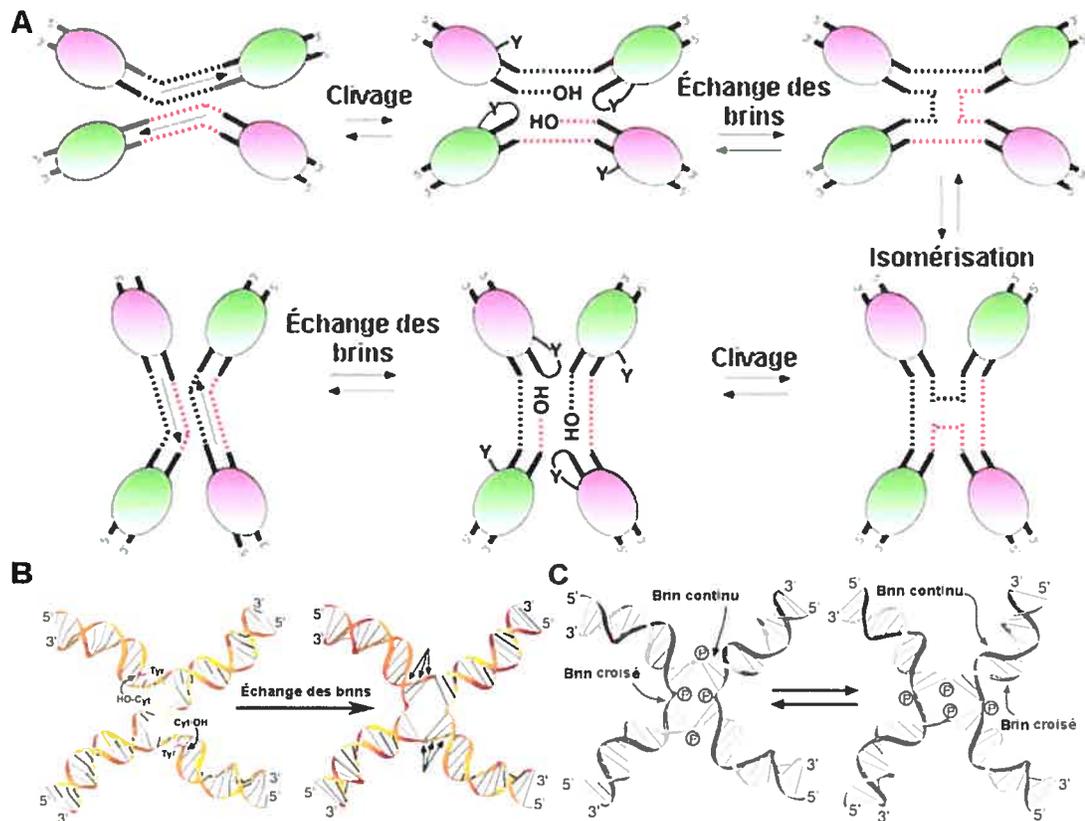
Les résidus conservés et formant le cœur catalytique des enzymes de la famille des tyrosine recombinases sont indiqués en violet. Les trois groupes de séquences correspondent respectivement aux motifs I, III et II tel que déterminés par Cao *et al.* (1997). Les résidus conservés sont en vert. Les hélices et feuillets indiqués au-dessus sont basés sur la structure de XerD d'*E. coli* (Subramanya *et al.*, 1997). Pa : *Pseudomonas aeruginosa*, Pf : *Pseudomonas fluorescens*, Ec : *E. coli*, St : *Salmonella typhimurium*, Hi : *Haemophilus influenzae*, Bs : *Bacillus subtilis*, Ll : *Leichmania lechmannii*, Ml : *Mycobacterium leprae*, Mt : *Mycobacterium tuberculosis* et Hp : *Helicobacter pylori*. « xC » et « xD » correspondent respectivement aux protéines XerC et XerD. Les séquences ont toutes été retrouvées grâce au site de Dominic Esposito : <http://mywebpages.comcast.net/domespo/trhome.html>.

### 3.3 Réaction de recombinaison

#### 3.3.1 « Strand-Swapping Isomerization model »

La réaction de recombinaison nécessite deux fragments d'ADN contenant chacun un site d'action, et quatre protomères. Le modèle proposé suite à diverses études biochimiques présente un échange des brins menant à la formation d'une jonction de Holliday, laquelle est par la suite isomérisée afin de permettre l'action de clivage sur les brins partenaires. Ce dernier échange mène donc aux produits recombinants.

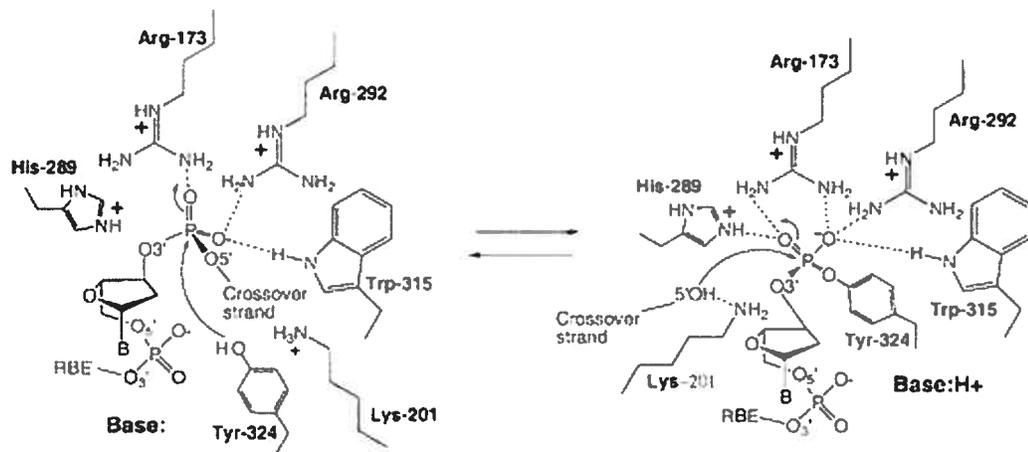
Ce modèle, connu sous le nom anglais de « Stand-Swapping Isomerization Model », est présenté en Figure 2A.



**Figure 2: Modèle de la réaction menée par les tyrosine recombinases : « Strand-Swapping Isomerisation model ».** A) Le modèle ici montré est basé sur l'ensemble des études de la protéine Cre sur son site d'action *loxP*. Les protomères présentés en vert sont catalytiquement actifs et exercent l'activité de clivage en premier (dans les trois premiers schémas de la réaction) tandis que l'activation du second protomère (en rose) est présentée dans les trois autres schémas. Après la première réaction de clivage, les brins sont échangés (« swapped ») pour mener lieu à la formation de la jonction de Holliday. Cette jonction est ensuite isomérisée (les brins initialement continus deviennent croisés), afin de permettre aux deux autres protomères d'exercer leur action par le même processus que précédemment. Il y a ainsi obtention des produits recombinants. B) La réaction d'échange des brins se fait par « swapping » (troc). La structure présentée ici est l'intermédiaire obtenu lors de la cristallisation de Cre liée de façon covalente à *loxP*. Ainsi, le fragment 5'-OH du résidu cytosine (ce résidu avait été clivé, ce qui le rendait libre d'exercer son action) « attaque » le lien 3'-phosphotyrosine pour donner lieu à la jonction de Holliday. Trois nouvelles paires de bases (montrées par les flèches à la suite de l'échange des brins) sont ainsi formées avec le substrat partenaire. La jonction est formée à la suite du mouvement du lien phosphodiester des bases échangées. C) Modèle pour l'isomérisation de la jonction de Holliday dans la recombinaison au site *loxP* par Cre. Ce sont les brins dits croisés qui sont prêts pour la réaction de clivage et l'échange entre les deux conformères. Entre ces deux conformères, il y a eu échange des points de branchement des groupements phosphates. Ces trois schémas ont été tirés et modifiés de Van Duyne (2001).

### 3.3.2 Réaction de clivage

Comme déjà mentionné au point 2.2 (voir plus haut), la région d'enjambement (autrement dit région du crossing-over) du site d'action sur lequel agissent les tyrosine recombinases, fait 6 à 8 pb. Cette réaction de clivage consiste en fait en l'attaque nucléophile du phosphate « scissile » via la tyrosine catalytique, suivi par l'expulsion du groupe 5' hydroxyle quittant de l'ensemble du groupe. La Figure 3 montre le modèle pour le clivage du site actif fait par la recombinase Cre, mais il pourrait également s'appliquer aux autres recombinases (avec les résidus équivalents).



**Figure 3: Modèle de la réaction de catalyse par Cre sur *loxP*.** Le résidu His-289 pourrait être le résidu qui accepte un proton provenant de la tyrosine nucléophile durant le clivage. L'acide responsable de donner un proton au groupement 5'OH quittant la réaction après le clivage est le résidu Lys-201. La réaction de ligation est l'inverse de ce qui est montré, avec une attaque nucléophile de l'intermédiaire 3'-phosphotyrosyle par le groupement 5'OH. Cette figure est tirée de VanDuyne (2001).

Ainsi, la triade R-H-R avec la chaîne contenant le résidu Trp (chez Cre et Flp, mais His pour les autres tyrosine recombinases, voir plus haut), procurent des liens donneurs d'hydrogène entourant le phosphate qui sera clivé. Ces résidus stabilisent l'état de transition pentavalent du phosphate et contribuent probablement à la catalyse de la réaction. Chez la topoisomérase du virus de la vaccine, le résidu lysine conservé a été montré comme étant l'acide catalytique donnant un proton au groupe 5'OH quittant la réaction (Krogh et Shuman, 2000). La base acceptrice du proton provenant de la tyrosine serait probablement le deuxième résidu histidine (du motif II) (VanDuyne, 2001).

### 3.3.3 Arrangement structural des résidus conservés

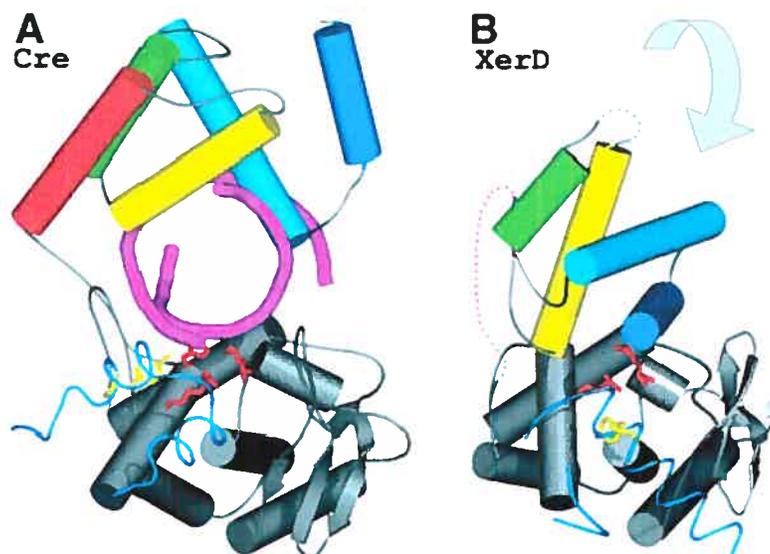
La détermination des structures de plusieurs tyrosine recombinases a permis de visualiser davantage le déroulement probable de la réaction de catalyse au niveau des sites d'action.

#### 3.3.3.1 *Les structures, un bref aperçu*

Jusqu'à présent, les structures complètes ont été établies pour les protéines XerD d'*E. coli* (Subramanya *et al.*, 1997) et Cre. Dans ce dernier cas, trois intermédiaires de la recombinaison ont été cristallisés avec un site *loxP* modifié (Gopaul *et al.*, 1998 et Guo *et al.*, 1997, 1999). La structure de Flp, impliquée dans l'amplification du plasmide 2  $\mu$  de *Saccharomyces cerevisiae*, a également été déterminée sur son site FRT (Chen *et al.*, 2000). Quant à l'intégrase de lambda ( $\text{Int}\lambda$ ), les domaines C- et N-terminaux ont été déterminés séparément (Kwon *et al.*, 1997 cristal du domaine C-terminal et Wojciak *et al.*, 2002 structure par RMN, résonance magnétique nucléaire, en solution). Enfin, seul le domaine catalytique de l'intégrase de HP1, laquelle sert à l'intégration de ce dernier dans *Haemophilus influenzae* Rd dans le locus *attB* (Goodman et Scocca, 1989), a été cristallisé (Hickman *et al.*, 1997).

Plus récemment, Cre a été cristallisée, cette fois avec son site naturel *loxP*. Le premier intermédiaire de cette cristallisation a montré la structure de la protéine avant qu'elle n'ait clivé l'ADN, tandis que le second a révélé la structure du complexe à la suite du premier échange des brins (Ennifar *et al.*, 2003).

La Figure 4 montre les structures quaternaires des protéines Cre et XerD. Tandis que les domaines C-terminaux des recombinases sont homologues les uns aux autres (à l'exception de l'hélice  $\alpha\text{M}$  qui contient le résidu catalytique ; voir plus loin), les domaines N-terminaux apparaissent plus divergents. Cependant, bien que les hélices  $\alpha\text{A}$  et  $\alpha\text{E}$  (cinquième hélice de Cre montrée en rouge sur la Figure 4, mais désordonnée chez XerD) ne peuvent se superposer, le repliement global de ces domaines apparaît relativement similaire (Swalla *et al.*, 2003).



**Figure 4 : Comparaison des structures quaternaires des protéines Cre (en A, avec 10 pb du site *loxP* liées au cœur de liaison et au domaine catalytique) et XerD sans son site d'action (en B).** Dans les deux cas, les hélices  $\alpha$  colorées représentent le cœur de liaison à l'ADN (domaine N-terminal : hélice A en bleu foncé, hélice B en bleu pâle, hélice C en vert et hélice D en jaune), tandis que les domaines C-terminaux sont représentés en gris. Le domaine N-terminal contient une cinquième hélice chez Cre (en rouge), tandis que cette région est désordonnée chez XerD (en pointillés rouges dans la figure B). XerD possède une seconde boucle désordonnée reliant les hélices  $\alpha$ C et  $\alpha$ D (en pointillés gris). La flèche de la figure B indique un changement possible de conformation pour XerD en l'absence d'ADN. Les domaines catalytiques sont ici représentés dans la même orientation chez les deux protéines : les résidus en rouge représentent le cœur R-H-R (chez XerD, l'His est cachée par l'hélice  $\alpha$ D), tandis que la tyrosine catalytique est représentée en jaune. Les hélices en bleu complétant le cœur catalytique en gris montrent les extrémités C-terminales différentes chez les deux protéines ; chez Cre, la toute fin de la protéine n'est pas représentée ici. En A, l'ADN est représenté en violet. Cette figure est tirée de Swalla *et al.* (2003).

### 3.3.3.2 *Emplacement des résidus catalytiques dans les structures*

D'après toutes les structures, le résidu catalytique se situe dans une région très flexible afin de permettre son action de clivage au niveau du phosphate « scissile » (phosphate devant être clivé. Pour une revue de l'organisation des sites actifs, voir Grainge et Jayaram, 1999). Cependant, la conformation de la région où se trouve cette tyrosine est différente dans chaque cas. Par exemple, chez Cre, la Tyr est placée à l'extrémité de l'hélice  $\alpha$ M, laquelle est placée de telle sorte de pouvoir bouger aisément à l'intérieur et à l'extérieur de la poche catalytique de son propre protomère (Guo *et al.*, 1997). Chez Flp, l'hélice  $\alpha$ M transportant la Tyr catalytique est désordonnée lorsqu'elle est inactive, tandis qu'elle atteint la poche catalytique de protomère partenaire lorsqu'elle est active (Chen et

Rice, 2003). D'un autre côté, alors que la tyrosine est située sur une hélice correspondant à l'hélice  $\alpha_M$  de Cre chez les autres recombinases, ce résidu est porté par un feuillet  $\beta$  situé près d'une boucle désordonnée chez Int $\lambda$  (Kwon *et al.*, 1997) ; il y aurait alors nécessité d'un réarrangement de la structure afin de permettre la réaction de clivage. La liaison à l'ADN par Int $\lambda$  sert d'ailleurs à stabiliser le repliement de la structure d'Int $\lambda$  (Kamadurai *et al.*, 2003), ce qui pourrait faciliter le placement adéquat du résidu Y afin qu'il effectue son action.

Tandis que les résidus RHR servent à la coordination des phosphates « scissiles » (liens qui seront clivés), la tyrosine catalytique est placée près du phosphate qui sera clivé. Chez Cre et Flp, un tryptophane forme une liaison hydrogène avec le phosphate « scissile » (Chen et Rice, 2003). Le résidu équivalent chez Int $\lambda$ , XerD et même chez les topoisomérases du virus de la vaccine et humaine est une deuxième histidine conservée servant à coordonner le phosphate sensible au clivage (Van Duyne, 2002). D'après ce qui fut révélé par la structure de Flp sur la jonction de Holliday (Chen *et al.*, 2000), ce résidu serait en plus important pour l'ancrage adéquat de l'hélice  $\alpha_M$ , compte tenu de ses interactions (en plus de l'ADN) avec les résidus avoisinants de la cavité catalytique (Chen et Rice, 2003).

Une revue complète des diverses structures des tyrosine recombinases, ainsi que les changements structuraux impliqués dans la réaction catalytique, a été écrite par Van Duyne, 2002.

### 3.3.4 Précisions sur le clivage

Comme mentionné plus haut, la tyrosine catalytique se situe sur l'hélice  $\alpha_M$  des différentes structures. Il y a donc quatre tyrosines potentielles dans le complexe créé au site d'action susceptibles d'effectuer la première réaction de clivage qui mènera à la formation de la jonction de Holliday (voir Figure 2A, 2<sup>e</sup> schéma).

Ainsi, le clivage pourrait se faire en *trans*, et alors le clivage serait effectué par le monomère situé à l'opposé du point de clivage. À l'opposé, le clivage en *cis* est fait par le monomère situé immédiatement à côté du site de clivage.

En échangeant son hélice  $\alpha M$ , il a été montré que la protéine Flp (ainsi que d'autres enzymes de la même famille retrouvées chez des levures) effectue un clivage en *trans* (Chen et Rice, 2003). Pour toutes les autres recombinases, le clivage se ferait en *cis*.

Certaines études faites avec la protéine Cre pouvaient laisser penser à un clivage en *cis* (Guo *et al.*, 1997, Tribble *et al.*, 2000), tandis que d'autres laissaient présager un clivage en *trans* (Shaikh et Sadowski, 1997). Ce paradoxe *cis-trans* fut largement discuté dans mon mémoire (Villion, 1998 et voir Jayaram, 1997), et ne sera donc pas repris ici. Cependant, les plus récentes études de conformation structurale par cristallographie ont pu éliminer cette ambiguïté quant à l'action de la protéine Cre. En effet, comme mentionné plus haut, un complexe de pré-clivage montre que la protéine Cre active est liée du côté gauche de la région où se fait le crossing-over sur *loxP* et induit une courbure de l'ADN. Une autre structure de l'intermédiaire de recombinaison, montre que c'est le lien phosphodiester sur le brin dit supérieur qui est clivé en premier; cette structure incluait le lien 3'-phosphotyrosyle de Cre sur son site. Ces intermédiaires ont ainsi montré que la courbure induite près du lien phosphodiester devant être clivé (« scissile ») activerait l'ADN pour le clivage par le résidu tyrosine de Cre située en *cis* (Ennifar *et al.*, 2003). C'est ce type de clivage qui est montré dans le modèle présenté en Figure 2A.

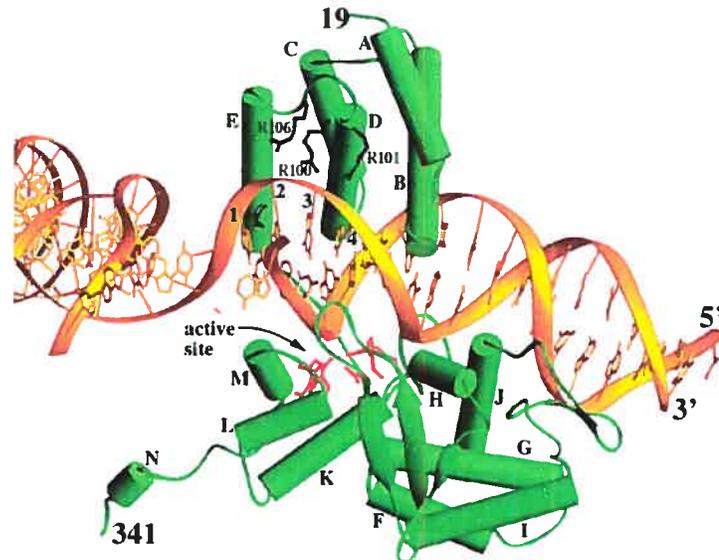
### 3.3.5 Interactions entre les sous-unités

#### 3.3.5.1 Révélation de la protéine Cre

Les structures de Cre en son site ont permis de révéler non seulement une interaction étendue de la protéine sur son site, mais également une interaction entre deux sous-unités protéiques. Les contacts se font aussi bien par le domaine N-terminal que par le domaine C-terminal via un échange d'hélices.

Chez Cre, l'échange des hélices du domaine C-terminal n'est pas réciproque (contrairement à ce qui est vu avec HP1) ; il est cyclique parmi les quatre sous-unités protéiques du complexe synaptique (Guo *et al.*, 1997). Le même type d'interactions se trouveraient chez XerC et XerD (voir le point 4), et chez Flp (Chen et Rice, 2003). De plus, compte tenu de la similarité entre les domaines N-terminaux de Cre et XerD, ils formeraient probablement des contacts à l'ADN, similaires (Figure 4 et voir VanDuyne, 2001).

En fait, la protéine Cre formerait une pince autour de son site de reconnaissance avec ses domaines N- et C-terminaux (comme le montre la Figure 5) et ce type de placement serait vu chez les autres recombinases.



**Figure 5: Structure d'un protomère Cre lié à un bras de la jonction de Holliday formée de deux sites *loxP*.** Le résidu catalytique est en rouge, tandis que les résidus R près du point de branchement sont en noirs. Le point rouge représente le groupe phosphate « attaqué » par la tyrosine catalytique. Cette figure est tirée de Gopaul *et al.*, 1998, montrant la structure des complexes formés de Cre sur la jonction de Holliday formée suite à l'action de clivage. Pour les détails, voir cet article.

### 3.3.5.2 Implications du modèle de la réaction

Le modèle d'échange des brins suivi de l'isomérisation (strand-swapping isomerisation model, présenté en Figure 2) implique des changements dans les interactions entre les diverses sous-unités pour que la réaction d'échange puisse avoir lieu : certaines sous-unités doivent être activées, alors que d'autres sont inactivées.

Chez Cre, ces deux formes sont induites par un placement différent des hélices  $\alpha$ M-lien- $\alpha$ N (voir la notation de ces hélices sur la Figure 5). Il en est de même pour la conformation de la boucle située entre les feuillets  $\beta$ 2 et  $\beta$ 3, même si celle-ci est moins bien définie. Le modèle implique en plus la production d'une synapse antiparallèle entre les sites pour que l'échange des brins soit productif. La coordination de la réaction permet l'activation de deux monomères à la fois via leur résidu catalytique. Grâce aux structures définies des intermédiaires de Cre sur son site d'action, l'étape de l'échange des brins (« swapping ») a pu être révélée (Figure 2B), et un modèle pour l'isomérisation de la

jonction de Holliday a pu être établi (Figure 2C). Le modèle global présenté ici (donc basé sur les structures de Cre sur son substrat) peut être appliqué à toutes les recombinases de la famille des Y-recombinases et sera précisé pour Xer dans la section suivante.

## 4 Les recombinases de type Xer

### 4.1 Découverte

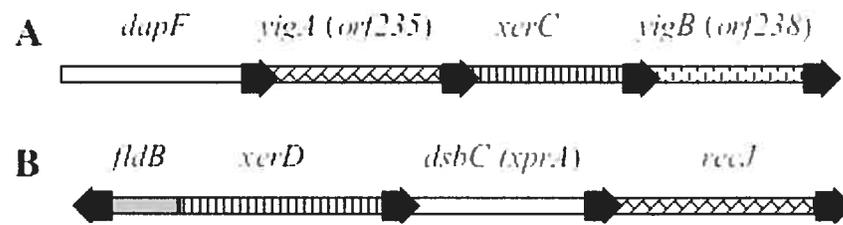
Le système de recombinaison Xer a été découvert tout d'abord par l'étude de la stabilité du plasmide ColE1 chez *E. coli*. En effet, ce plasmide ne possédait pas de système de partition, mais était malgré tout maintenu dans la souche, grâce à un site (le site *cer* ; « ColE1 resolution ») permettant la conversion des multimères en monomères (Summers et Sherratt, 1984). Ce site permettait alors la stabilité de ColE1.

Dans la quête de facteurs agissant au niveau du site *cer*, les gènes *xerA* et *xerB* ont d'abord été découverts [l'acronyme *xer* (pour : « **chromosomal ColE1 recombination function(s)** ») désigne les loci du chromosome permettant la recombinaison sur le plasmide ColE1, Stirling *et al.*, 1988a)]. Ces gènes *xerA* et *xerB* ont par la suite été définis comme des gènes codant pour des protéines accessoires mais nécessaires à la réaction de recombinaison (voir la section 5.1). En sélectionnant d'autres mutants déficients pour la recombinaison au site *cer*, l'équipe du laboratoire de Dr Sherratt a alors trouvé un troisième gène encodant une protéine agissant sur ce site (Colloms *et al.*, 1990).

Le gène *xerC* ainsi déterminé fait partie d'une unité multicistronique contenant le gène *dapF* (codant pour la diaminopimélate épimérase ; Richaud *et al.*, 1987 et Richaud et Printz, 1988), et il est flanqué des cadres de lecture ouverts (orfs), *yigA* et *yigB* (Blattner *et al.*, 1997).

Peu de temps après, un quatrième gène impliqué dans la recombinaison au site *cer*, le gène *xerD*, a été découvert par sa similarité de séquence avec le gène *xerC* (Blakely *et al.*, 1993). Il correspondait au 1<sup>er</sup> orf de la séquence nucléotidique d'un segment cloné contenant *recJ* (Lovett et Clark, 1985), et qui avait révélé une très forte similarité aux gènes *fimB* et *fimE* (Lovett et Kolodner, 1991). Toujours sur le même segment, un second gène, nommé alors *xprA*, s'est révélé être le gène *dsbC* codant pour l'isomérase disulfide impliquée dans la formation des ponts disulfures (Missiakas *et al.*, 1994).

La Figure 6 illustre l'organisation chromosomique des deux gènes *xer*.



**Figure 6: Organisation génomique des gènes *xerC* (en A) et *xerD* (en B) chez *E. coli*.**

## 4.2 Caractéristiques

Les protéines encodées par les gènes *xerC* et *D* ont 37% d'identité et font partie de la même famille que l'intégrase du phage Lambda (Colloms *et al.*, 1990, Blakely *et al.*, 1993). Tout comme les autres membres de la famille, elles possèdent donc les deux motifs hautement conservés contenant les résidus R-H-R et la tyrosine catalytique. Des mutants de *xerC* ont montré des problèmes dans la division cellulaire et dans la ségrégation des nucléoïdes, du même ordre que les mutants pour le site *dif* (voir la section 6 qui traite de ce site; Blakely *et al.*, 1991).

De plus, tout comme il avait été montré chez d'autres membres de la famille des tyrosine recombinases [par ex : Cre (Abremski *et al.*, 1986) et Int $\lambda$  (Kikuchi et Nash, 1979)], les recombinases de type Xer retrouvées chez *E. coli* montrent une activité topoisomérase de type I (Cornet *et al.*, 1997). Cette activité, dépendante de la tyrosine catalytique, serait spécifique de site puisque les recombinases n'ont pu relaxer que l'ADN d'un vecteur surenroulé contenant un site *dif*. Cette activité serait de plus mutuellement exclusive en ce sens que Cornet *et al.* ont montré que XerC ne pouvait relaxer l'ADN qu'en l'absence de XerD et inversement, XerD a montré cette activité topoisomérase qu'en l'absence de XerC (Cornet *et al.*, 1997).

Contrairement à la plupart des systèmes ne nécessitant qu'une seule recombinase, XerC et XerD doivent agir conjointement afin d'exécuter l'action de recombinaison. Cette caractéristique particulière a fait de ce système un excellent modèle afin d'étudier l'aspect moléculaire de la réaction de recombinaison spécifique. Il devenait alors plus aisé d'étudier, entre autres, les interactions protéine-protéine. Ainsi, plusieurs études de mutagenèse ont été réalisées afin de mieux cibler les points d'interaction entre les deux protéines, les points de contact avec l'ADN, de savoir quand et comment les protéines étaient activées pour réaliser leur activité de clivage.

### 4.3 Sites d'action des recombinaases de type Xer

Comme mentionné au début de cette section, les études ayant mené à la découverte de facteurs agissant en des sites spécifiques furent la découverte du locus *cer* sur le plasmide ColE1. En fait, plusieurs loci servant à la stabilité de divers plasmides ont été déterminés avant que l'on ne découvre les facteurs agissant sur ces sites. Par la suite, le même type de locus, retrouvé sur le plasmide R1, a pu s'intégrer à un locus retrouvé sur le chromosome d'*E. coli*. Ce site présentait une telle homologie avec le site *cer*, qu'il a été supposé qu'il devait remplir le même type de fonction, à savoir le maintien des chromosomes circulaires (Clerget, 1991). En parallèle, l'équipe de Kuempel montrait que ce site (nommé *dif*) était bel et bien un site de recombinaison indépendant de RecA (Kuempel *et al.*, 1991).

Tout au long des dernières années, les sites les plus étudiés furent *cer*, *dif* et *psi* [site retrouvé cette fois sur le plasmide pSC101 (Cornet *et al.*, 1994)]; ils donnaient chacun des produits de recombinaison différents lorsque portés par un vecteur de type pUC. Compte tenu de son importance durant le cycle cellulaire chez *E. coli*, une attention particulière sera portée sur le site *dif* dans une section à part.

#### 4.3.1 Liaison aux sites d'action

Ainsi, comme les autres membres des tyrosine recombinaases, les sites reconnus spécifiquement par les protéines Xer font 28 à 30 pb. Ces deux recombinaases se lient de façon coopérative et individuelle aux sites *dif* et *cer* à l'exception de XerC qui n'a pu se lier au site *cer* sans la présence de XerD. Toutes les deux se lient aux 11 pb flanquant la région d'enjambement variant de 6 à 8 pb. Malgré l'aspect presque palindromique de ces sites, il a été montré que XerC reconnaît spécifiquement la portion à gauche, tandis que XerD se lie à droite (Blakely *et al.* 1993). Comme on peut le voir par l'alignement de différents sites (voir Tableau I), le site de liaison de XerD est cependant beaucoup plus conservé que celui de XerC.

**Tableau I: Sites de recombinaison reconnus pas les protéines Xer.**

Sites plasmidiques	XerC	Espacement	XerD
CloDF13 ( <i>parB</i> )	GGTACCGATAA	..GGGATG	TTATGGTAAAT
ColA ( <i>car</i> )	GGTGCGTACAA	..CGGATG	TTATGGTAAAT
ColE1 ( <i>cer</i> )	GGTGCGTACAA	TTAAGGGA	TTATGGTAAAT
ColE3 ( <i>cer3</i> )	GGTGCGTACAA	..CGGGAG	TTATGGTAAAT
ColK ( <i>ckr</i> )	GGTGCGTACAA	TTAAGGGA	TTATGGTAAAT
ColN	GGTGCGTACAA	..TAAGGGA	TTATGGTAAAT
pHS-2 ( <i>hmr</i> )	GGTGCGCGCAA	..TAAAATG	TTATGTTAAAT
plasmide R1	GGTGCGCATAA	..TGTATA	TTATGTTAAAT
pMB1	GGTGCGTACAA	TTAAGGGA	TTATGGTAAAT
pNTP16 ( <i>nmr</i> )	GGTGCGCGTAA	..TGAGACG	TTATGGTAAAT
pSC101 ( <i>psi</i> )	GGTGCGCGCAA	..GATCCA	TTATGTTAAAC
pPvu1	GGTTCGATAA	..GACCCA	TTATGTTAAAC
<b>Sites chromosomiques (<i>dif</i>)</b>			
<i>E. coli</i>	GGTGCGCATAA	..TGTATA	TTATGTTAAAT
<i>S. typhimurium</i>	GGAGCGCATAA	..TGTATA	TTATGTTAAAT
<i>S. typhi</i>	GGAGCGCATAA	..TGTATA	TTATGTTAAAT
<i>H. influenzae</i>	ATTTTCGCATAA	..ATAAA	TTATGTTAAAT
<i>B. subtilis</i>	ACTTCCTAGAA	..TATATA	TTATGTTAACT

Les sites plasmidiques sont inscrits entre parenthèses à côté du plasmide sur lequel ils sont trouvés. La séquence consensus du site de liaison de XerC est : RBWDCSBDBAA (G/a;G/t/c;T/A;G/a/t;C;g/c;C/t/g;a/t/g;t/g/c;A;A), et celle de XerD est : TTATGKTA AAAAY (T;T;A;T;G;g/t;T;A;A;A/c;T/c).

Certains des nucléotides des sites d'action seraient responsables de cette spécificité. Ainsi, le nucléotide situé en position -10 sur le site *dif*<sup>1</sup> serait le plus important pour la liaison de XerC. D'ailleurs, ce nucléotide est conservé dans tous les sites analysés jusqu'à maintenant, et une mutation de cette position a empêché la liaison au site *dif in vitro* (Hayes et Sherratt 1997). Pour la liaison de XerD, les nucléotides +9 et +13 seraient les plus importants (Blakely et Sherratt, 1994 et Hayes et Sherratt 1997).

#### 4.3.2 Orientation de la réaction

La réaction de recombinaison pourrait théoriquement donner lieu à plusieurs produits différents. On parle de sélectivité pour la résolution lorsque la recombinaison a lieu préférentiellement dans le cas d'une réaction entre deux sites situés dans le même sens sur un même vecteur. Tel serait le cas lors de la formation d'un dimère plasmidique. À

<sup>1</sup> Les positions ont été données arbitrairement à partir de la région de crossing-over du site *dif*. Ainsi, la position +1 est le A central de la séquence TGTATA, tandis que la position -1 correspond au T central de cette même séquence.

l'opposé, une réaction de recombinaison intermoléculaire donnerait lieu à la formation de multimères. Un dernier type de réaction pourrait donner lieu à une inversion de la région comprise entre deux sites situés sur un même vecteur, mais alors, les sites devraient être en orientation inverse (voir Hallet et Sherratt, 1997).

Ainsi, même si la longueur des séquences (de 28 à 30 pb) est suffisante pour la reconnaissance des sites par les protéines Xer, des séquences accessoires sont nécessaires pour la recombinaison au site *cer* de ColE1. En réalité, ce site contient 160 à 180 pb en plus et ces séquences servent à la liaison de protéines dites accessoires qui permettent une sélection pour la résolution. Les protéines accessoires assurent une réaction de recombinaison intramoléculaire. Il y a alors résolution des multimères en monomères. (Ces protéines accessoires font l'objet de la section suivante).

#### 4.3.2.1 Une histoire de taille?

Les premières études sur les sites de recombinaison laissaient croire que la taille des régions centrales (zones de la région d'enjambement) avait un rôle à jouer dans les facteurs influençant la recombinaison.

En effet, dans une souche *recA<sup>-</sup>*, un vecteur contenant le site *dif* d'*E. coli* (la région centrale de ce site fait 6 pb) avait tendance à multimériser. Il n'y avait alors pas de sélectivité pour la résolution (Kuempel *et al.*, 1991 et revue dans Sherratt *et al.*, 1995). Par contre, lorsque la région centrale fut augmentée de 2 pb, ce site a nécessité des séquences et des protéines accessoires pour, non seulement que la recombinaison se produise, mais aussi qu'elle ne soit exercée que dans le sens d'une résolution (Blakely et Sherratt, 1996). D'un autre côté, un site *cer* dont la région centrale fut réduite de 2 pb, ne nécessitait plus de séquences ni de protéines accessoires pour que la recombinaison ait lieu, et alors, il n'y avait plus de sélectivité pour la résolution (Summers, 1989).

Cependant, d'autres sites plasmidiques ont montré une région centrale de 6 pb (voir le Tableau I). Les sites *psi* (de pSC101) et *parB* (de CloDF13), tout comme *cer*, ont montré une préférence pour la recombinaison intramoléculaire lorsque des séquences et protéines accessoires sont présentes (Cornet *et al.*, 1994 et Colloms *et al.*, 1996). La taille de la région centrale ne pouvait donc plus servir seule à expliquer « l'orientation » de la recombinaison.

Par la suite, il a été déduit que la sélectivité pour la résolution était influencée par la séquence de liaison des recombinases, et ce, grâce à la construction d'hybrides entre les sites *dif* et *psi* (leur site de liaison de XerC n'est différent que par deux nucléotides; voir le Tableau I). L'orientation avait alors pu être modifiée par le changement d'un seul nucléotide dans le site de liaison de XerC (Blake *et al.*, 1997). Il a été alors supposé que les protéines accessoires étaient nécessaires pour compenser le fait que la liaison des recombinases était plus faible aux sites *psi* et *cer* qu'elle ne pouvait l'être au niveau du site *dif*.

#### 4.4 La réaction de recombinaison de type Xer

Les protéines Xer, par le fait qu'elles sont deux, ont permis l'étude de plusieurs aspects de la réaction de recombinaison : l'action de clivage, l'activation des protéines, les contacts entre les protéines, etc. Par la suite, la visualisation de l'organisation de la poche catalytique a été rendue possible par la cristallisation de la protéine XerD (Subramanya *et al.*, 1997).

Brièvement, la protéine XerD peut être divisée en deux domaines (voir la Figure 4B), comme c'est le cas de la protéine Cre (Figure 4A). Le domaine N-terminal (résidus 1 à 107) est composé de quatre hélices  $\alpha$  (A à D), tandis que le domaine C-terminal (résidus 108 à 298), également appelé domaine CAT (catalytique), est constitué de 9 hélices  $\alpha$  (E à N) et de trois feuilletts  $\beta$ . Même si la structure n'a pas été définie avec son substrat, il est supposé que XerD formerait une pince autour de l'ADN, comme ce qui a été montré avec la protéine Cre (Gopaul *et al.*, 1998 et revoir Figure 5). Les contacts se feraient alors par l'intermédiaire des hélices  $\alpha$ B et  $\alpha$ D du premier domaine, et aussi par l'hélice  $\alpha$ J du second (Subramanya *et al.*, 1997).

##### 4.4.1 Petit historique sur le modèle

Afin d'expliquer comment pouvait s'effectuer la recombinaison spécifique de site, il avait été supposé au départ un modèle de migration des branches, un peu comme ce qui se produit au niveau de la recombinaison homologue. Cependant, les études de Nunes-Düby *et al.* (1995), portant sur la recombinaison avec l'intégrase du phage  $\lambda$ , ne permettaient plus d'étayer cette hypothèse puisque la position optimale de la jonction n'était pas aux extrémités de la région centrale du site *att*, mais près du milieu. Les jonctions synthétiques

étaient résolues malgré une possibilité de mobilité réduite des branches. Leurs résultats leur ont permis de soumettre un autre modèle alors nommé « Strand-swapping isomerisation model » (voir point 3.3.1). Selon ce modèle, la jonction de Holliday provient du clivage des brins dits supérieurs<sup>2</sup> des sites d'action, suivi par la « fonte » de deux ou trois nucléotides de ces brins dans la région d'enjambement (« crossing-over »). Ces nucléotides sont ensuite échangés entre les poches du site catalytique des deux protomères partenaires avant qu'il n'y ait la liaison des brins pour générer la jonction de Holliday (Nunes-Düby *et al.*, 1995 et voir le schéma du modèle amélioré pour la réaction de recombinaison réalisée par Cre, Figure 2).

Les diverses études portant sur le système Xer ont permis de supporter ce modèle par la suite. Tout d'abord, la jonction de Holliday a été montrée comme l'étape intermédiaire de la recombinaison au site *cer in vivo* par microscopie électronique, et ces jonctions étaient le fruit d'un échange des brins dits supérieurs (McCulloch *et al.*, 1994b). Par la suite, la construction de jonctions de Holliday synthétiques et de substrats suicides a permis d'étudier les diverses étapes de cette réaction, *in vitro*.

#### 4.4.2 Les développements

##### 4.4.2.1 Formation de la jonction de Holliday chez Xer

Ainsi, lors d'une première étude avec une jonction de Holliday contenant un site *cer* modifié, il a été montré que cette jonction n'avait pas besoin d'être au point de clivage pour être formée; la position optimale de la jonction était située à au moins deux nucléotides en 3' du site de clivage (Arciszewska *et al.*, 1995).

En fait, une étude subséquente a permis de déduire que le point de branchement de la jonction serait exactement au centre du site de recombinaison; la liaison des recombinases avait induit le détachement des bases situées au centre du site (Arciszewska *et al.*, 1997).

Afin de réaliser cette jonction, il a été montré que la première réaction de clivage était réalisée par XerC au niveau d'un site *cer*. Le point de clivage de cette protéine était

---

<sup>2</sup> Les notions de brins « inférieurs » et « supérieurs » ont été déterminés arbitrairement. Le brin supérieur correspond celui qui contient le site d'action tel que montré dans l'alignement des sites au Tableau I, et le brin inférieur correspond à la séquence complémentaire de ce brin.

situé sur le brin supérieur, entre le dernier nucléotide de son site de liaison et le premier nucléotide de la région d'espacement, afin de générer un complexe 3'-phosphotyrosyle et une extrémité 5'-hydroxyle (Arciszewska et Sherratt, 1995). Pour réaliser cette action, l'activité catalytique de XerD n'était pas nécessaire, même si sa « présence » l'était.

Enfin, la jonction de Holliday est complétée par la liaison entre l'extrémité 5'OH et le lien 3'-phosphotyrosyle du protomère associé à son site de liaison. Cette liaison se fait donc en *cis* par la protéine XerC liée de façon covalente à son site de liaison, mais cette fois, sans la nécessité de la protéine partenaire XerD (Grainge et Sherratt, 1999).

#### 4.4.2.2 *Résolution de la jonction de Holliday*

D'autres réactions *in vitro*, cette fois sur des vecteurs surenroulés contenant deux sites de recombinaison directement répétés flanquant une certaine région d'ADN ont également confirmé l'action de clivage de la protéine XerC (Colloms *et al.*, 1996). Par contre, la réaction de recombinaison au site *cer* s'est arrêtée à l'obtention de la jonction de Holliday. Ainsi, bien que XerD soit nécessaire à cette recombinaison, le facteur permettant la résolution de la jonction de Holliday était inconnu et l'est toujours aujourd'hui (octobre 2004).

Cependant, d'autres réactions *in vitro*, réalisées cette fois sur un vecteur surenroulé contenant deux sites *psi*, ont donné lieu à la production de caténanes. La résolution de la jonction de Holliday avait donc eu lieu et il a été montré que la réaction de clivage (sur les brins inférieurs) l'ayant permis avait été réalisée par XerD (Colloms *et al.*, 1996).

Le clivage des brins dits inférieurs par XerD a également pu être montré grâce à l'emploi de substrats suicides. Ces substrats sont des substrats linéaires contenant un bris dans la région d'enjambement du site d'action des recombinases. Ainsi, l'extrémité 5' hydroxyle libre engendrée après la réaction de clivage par la recombinase (qui elle est liée à son site de liaison), ne peut plus compléter la réaction. La recombinase reste donc liée de façon covalente à son site; les intermédiaires de recombinaison s'accumulent et peuvent ainsi être étudiés plus aisément (Nunes-Düby *et al.*, 1987).

Le substrat suicide contenant le site *dif* d'*E. coli* a ainsi permis de montrer que le point de clivage de XerC était situé entre le dernier nucléotide de son site de liaison et le premier nucléotide de la région où a lieu le crossing-over. Il en était de même pour la protéine XerD (Blakely *et al.*, 1997). L'efficacité du clivage par XerC a été plus

importante que celle de XerD; ce clivage a même été possible sans la présence de XerD (mais alors très faiblement).

#### 4.4.2.3 *Ordre d'échange des brins*

Selon les résultats des études présentées sur les sites *cer*, *dif* et *psi*, l'ordre d'échange des brins des réactions menées par les recombinases Xer était : 1) brins supérieurs clivés par XerC, 2) formation de la jonction de Holliday, puis 3) résolution par XerD (à l'exception du site *cer*) par le clivage des brins inférieurs. Cependant, c'est une série de paramètres qui permet la détermination de l'ordre d'échange des brins.

Lors de la formation de la jonction de Holliday, il peut exister deux isoformes suivant les brins qui sont croisés ou continus (voir Figure 2C). Il avait été montré que la richesse en purines des brins pouvait influencer la forme de la jonction de Holliday. En effet, Azaro et Landy (1997) avaient montré que la forme isomérique observée était celle dans laquelle les brins croisés avaient un maximum de purines.

Par l'utilisation de jonctions de Holliday synthétiques contenant le site *dif* et dont deux des quatre bras étaient attachés entre-eux, il a été montré que XerD ne clivait efficacement que les brins inférieurs initialement croisés. Ce sont les brins inférieurs de cette jonction qui étaient liés. De la même façon, XerC a clivé plus efficacement les brins supérieurs qui étaient initialement croisés (mais elle a aussi pu cliver, moins efficacement cependant, des brins supérieurs continus; Arciszewska *et al.*, 1997). Contrairement à ce qui est retrouvé chez *dif* ou *psi*, le brin supérieur de la région centrale du site *cer* est riche en purines, ce qui semble imposer une conformation dans laquelle les brins supérieurs sont croisés. Ainsi, le clivage par XerC est favorisé.

La configuration du site d'action pouvait donc avoir son importance pour déterminer la protéine qui devait cliver en premier. Selon Arciszewska *et al.* (2000), l'échange médié par XerC était plus le résultat d'une conformation favorable du complexe tétramérique au site d'action pour l'échange médié par XerC, que le fait de son activité de clivage plus élevée que celle de XerD.

Par la suite, il a été montré que l'ordre d'échange des brins était déterminé par l'architecture créée par les facteurs accessoires dans le complexe synaptique formé pour la recombinaison au site *psi* (Bregu *et al.*, 2002). En inversant l'orientation du cœur du site *psi* par rapport aux séquences accessoires dans ce site, la jonction de Holliday a été formée

par XerD et non pas par XerC, et ce, sans que la topologie du produit final ne soit changée (Bregu *et al.*, 2002). Au site *dif*, la réaction de recombinaison est régulée par une protéine impliquée dans la division cellulaire, la protéine FtsK (les rôles de cette protéine seront discutés plus loin dans l'introduction, au point 6.2). Entre autres choses, cette protéine renverse l'état catalytique des recombinases : la jonction de Holliday est formée par XerD, puis résolue par XerC. L'ordre d'échange des brins est donc modifié en sa présence (Aussel *et al.*, 2002).

#### 4.4.2.4 *Études des interactions protéiques*

Par la délétion d'une certaine portion de séquence du domaine C-terminal des protéines Xer, il a été montré que 36 résidus étaient importants pour la coopérativité entre les protéines XerC et XerD (Spiers et Sherratt, 1997). Ces résidus étaient inclus dans ce qui serait l'hélice  $\alpha$ M de XerD (Subramanya *et al.*, 1997). De plus, la délétion des cinq derniers résidus de la protéine XerD était suffisante pour empêcher la recombinaison *in vivo* (et ce peu importe le site employé) alors que les activités de liaison et de catalyse étaient encore fonctionnelles (Spiers et Sherratt, 1997). Ensuite, il a été montré que la délétion des résidus de l'extrémité C-terminale de XerD n'avait pas d'effets directs sur les interactions coopératives avec XerC pour la liaison au site *dif*. Cependant, la coopérativité de la liaison était affectée par une délétion des 5 à 10 derniers résidus de XerC. Lorsque les deux protéines étaient tronquées de leur extrémité C-terminale, la coopérativité pour la liaison se voyait diminuer davantage (Spiers et Sherratt, 1999).

En fait, la délétion de ces derniers résidus de XerC a montré une inactivation du soi, mais une stimulation de la protéine partenaire XerD ; le phénotype engendré a alors été nommé SPIS (stimulation of partner and impairment of self) (Hallet *et al.*, 1999). Ce phénotype a également été observé lorsque le résidu conservé « arginine » du premier motif (voir la Figure 1) a été muté dans chacune des protéines. Il avait alors été conclu que la stimulation catalytique du partenaire au moment de l'inhibition du soi devait se produire avant ou au moment du clivage (Arciszewska *et al.*, 2000).

Ce phénotype SPIS laissait supposer des changements dans la conformation de chacune des protéines, et ce, afin de permettre la catalyse de telle protéine à un moment donné. Ces résultats venaient de nouveau corroborer le modèle demandant une

isomérisation de la jonction afin de permettre à la deuxième série de brins d'être clivés (Arciszewska *et al.*, 2000).

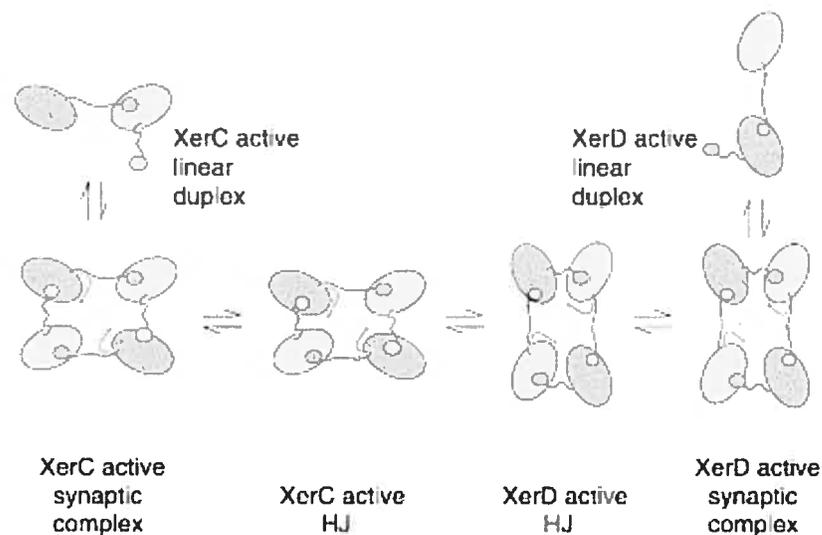
Dans la même étude découvrant le phénotype SPIS, d'autres résidus se sont révélés importants pour le contrôle de l'activité catalytique de chacune des protéines (Hallet *et al.*, 1999). Par l'alignement de plusieurs séquences de XerC et de XerD de différentes espèces, il avait été retrouvé un motif exclusivement réservé pour les protéines XerC, alors qu'au même endroit se trouvait un motif différent pour les protéines XerD (ce « motif » est constitué des acides aminés ESS chez XerC et de NHG chez XerD, retrouvés à la fin de ce qui correspond à l'hélice  $\alpha$ L de XerDEC; voir l'alignement de la Figure 1). L'échange de ces résidus entre les protéines partenaires a engendré un phénotype inverse à SPIS et fut nommé IPSS (impaired partner/stimulated self), c'est-à-dire une stimulation de la protéine possédant le motif et une inactivation de l'autre recombinase (Hallet *et al.*, 1999). Ces résultats laissaient suggérer qu'il pouvait y avoir un contrôle réciproque de l'activité catalytique de chacune des protéines XerC et XerD. Ce contrôle impliquerait la région C-terminale de l'une des recombinases et la région définie par les trois acides aminés de la recombinase partenaire (Hallet *et al.*, 1999).

Par la suite, d'autres mutations ont permis de montrer que le domaine N-terminal avait un rôle à jouer dans le contrôle de l'échange des brins par les recombinases Xer, et ce, même s'il n'est pas nécessaire à l'activité catalytique. Ainsi, l'hélice  $\alpha$ B du domaine N-terminal de XerD serait impliquée dans le maintien de la conformation de l'hétérotétramère nécessaire à la catalyse par XerC. De la même façon, lorsqu'une jonction de Holliday est organisée de telle sorte de permettre le clivage par XerD, le domaine N-terminal de XerC doit être présent pour ne pas entraîner un renversement de la catalyse parce qu'alors, c'est XerC qui exercerait la réaction de clivage (Feirrer *et al.*, 2001).

Enfin, l'échange des résidus du motif II de XerD (motif où se trouve la tyrosine catalytique; Figure 1) par ceux de XerC a modifié l'activité catalytique de XerC sur un substrat permettant normalement l'échange par XerD. Il a donc été déterminé que la région réceptrice de la recombinase se situait, du moins en partie, dans ce motif II (Ferreira *et al.*, 2003).

#### 4.4.3 Modèle final pour l'action des protéines Xer

Ainsi, comme ce qui avait été vu pour la protéine Cre (Guo *et al.*, 1997, Gopaul *et al.*, 1998, et voir point 3.3.5.1), toutes ces études ont permis de renforcer le modèle selon lequel il y aurait contrôle réciproque et cyclique des deux recombinases impliquées dans la réaction de recombinaison. Le fait que XerC soit active pendant que XerD est inactive dépendrait des interactions entre l'extrémité C-terminale de XerC et la région réceptrice de XerD située à l'intérieur du domaine C-terminal tel qu'il est montré sur le modèle présenté en Figure 7. Il y aurait un certain équilibre entre les interactions XerC et XerD. L'isomérisation de la jonction de Holliday serait alors nécessaire afin de permettre une configuration adéquate pour la réaction catalytique de la protéine partenaire.



**Figure 7: Modèle de l'équilibre entre les interactions XerC-XerD.** Puisque la structure de la protéine XerD est similaire à la protéine Cre (malgré la divergence au niveau de la séquence en acides aminés), un modèle de l'activation de l'activité catalytique de XerC vs XerD à l'intérieur du complexe protéines-ADN a été dérivé des structure de Cre au site *loxP*. Dans ce modèle, l'extrémité C-terminale du bras de XerC (ici représentée en bleu-vert) est étendue de telle sorte que la tyrosine nucléophile puisse attaquer le groupement phosphate du site d'action devant être clivé (« scissile phosphate »). Cette extrémité C-terminale interagit alors avec le domaine receveur de la protéine partenaire XerD. La région C-terminale de XerD (ici en mauve) est alors en conformation non-étendue (forme compacte), mais aide à la synapse. Pour que XerD puisse être à son tour activée, il doit y avoir isomérisation de la jonction de Holliday. Les extrémités C-terminales de XerD sont alors en conformation étendue et XerC est inactivée. Les attaques nucléophiles sont représentées par les flèches rouges. HJ : Jonction de Holliday. Cette figure est tirée de Ferreira *et al.*, 2003.

## 5 Les sites de recombinaison plasmidiques

Le maintien du plasmide ColE1 se fait donc grâce au site *cer* sur lequel agissent les recombinases de type Xer (voir la section précédente). En réalité, ce type de locus retrouvé en *cis* a été révélé non seulement sur d'autres plasmides de la série des « Col » tels ColE2 et ColE3 (Yasueda *et al.*, 1989), ColK (site *ckr* ; Summers *et al.*, 1985), ColA (Morlon *et al.*, 1988) et ColN (Kolot, 1990), mais aussi sur des plasmides multicopies d'autres familles, tels CloDF13 (site *parB* ; Hakkart *et al.*, 1984) et pSC101 (site *psi* ; Cornet *et al.*, 1994). Ce type de site est retrouvé également sur des plasmides d'autres espèces, comme pHS-2 de *Shigella flexneri* (Réhel et Szatmari, 1995) ou pNTP16 de *Salmonella typhimurium* (Cannon et Strike, 1992).

### 5.1 Nécessité de protéines accessoires pour la recombinaison aux sites plasmidiques

Avant la découverte des protéines XerC et XerD, deux facteurs (XerA et XerB) se sont révélés nécessaires pour que la recombinaison au site *cer* ait lieu (Stirling *et al.*, 1988a). Par la suite, ces protéines ont été identifiées, dans l'ordre, comme le répresseur de la biosynthèse de l'arginine, ArgR (Stirling *et al.*, 1988b) et l'aminopeptidase de leucine, PepA (Stirling *et al.*, 1989).

Au niveau du site *psi*, la protéine PepA semblait être le seul facteur accessoire requis *in vitro* (Colloms *et al.*, 1996), mais par la suite, il a été montré que la protéine ArcA jouait un rôle similaire à celui de ArgR sur *cer* (Colloms *et al.*, 1998).

Ainsi, afin de permettre la réaction catalytique aux sites plasmidiques, les contacts entre les recombinases ne suffisent pas. Les facteurs accessoires qui seront traités dans cette section ont alors un rôle architectural au niveau de la réaction de recombinaison.

#### 5.1.1 La protéine ArgR

La stabilité du plasmide ColE1 dépend de la protéine ArgR (originellement nommée XerA; Stirling *et al.*, 1988b). Les multimères de ce plasmide sont générés par RecA, mais résolus en monomères grâce au locus *cer* (Summers et Sherratt, 1984).

ArgR est le répresseur de la biosynthèse de l'arginine (pour une revue de littérature, lire Maas, 1994). Ce répresseur reconnaît et se lie à deux boîtes ARG de 18 pb chacune

(deux fois 9 pb en palindrome) séparées par 2 à 3 pb (Tian *et al.*, 1992). Une boîte ARG est retrouvée dans la région promotrice des gènes de la biosynthèse de l'arginine.

ArgR est une protéine hexamérique qui se lie à l'ADN en présence de L-arginine (Lim *et al.*, 1987). Cette protéine montre une haute affinité pour son site lorsque les deux boîtes ARG sont séparées par 3 pb (Chen *et al.*, 1997). En absence de L-arginine, la liaison se fait en tant que trimère.

La L-arginine servirait à deux choses : elle permet la stabilisation des hexamères et elle est nécessaire à la liaison à l'ADN (Van Duyne *et al.*, 1996). En son absence, ArgR se lie à l'ADN en tant que trimère (Burke *et al.*, 1994). Ce répresseur, en se liant à l'ADN induit une courbure d'environ 70° dont le centre se situe à l'intérieur des deux boîtes ARG adjacentes (Maas, 1994).

La protéine ArgR possède deux domaines dont les structures ont été déterminées séparément pour *E. coli* (Van Duyne *et al.*, 1996, pour la portion N-terminale, et Sunnerhagen *et al.*, 1997, pour la portion C-terminale). Le domaine N-terminal est le domaine de liaison à l'ADN, tandis que le domaine C-terminal sert à la liaison de la L-arginine (pour la stabilisation de la formation de l'hexamère) et est le domaine d'oligomérisation (Van Duyne *et al.*, 1996).

Une boîte ARG a été retrouvée à environ 100 pb en amont de la région d'échange du site *cer* (Summers, 1989) et ArgR a pu s'y lier (Colloms *et al.*, 1996). D'ailleurs la protéine ArgR en conjonction avec PepA (voir la section suivante) sont des protéines structurales autour desquelles deux segments d'ADN sont interliés dans une géométrie précise (Alén *et al.*, 1997). L'imperfection des boîtes ARG serait importante *in vivo* pour un meilleur contrôle de la répression par ArgR puisque la liaison à une boîte ARG dite « parfaite » est aussi forte que la liaison à deux boîtes « naturelles » espacées par 2 ou 3 pb, alors que normalement, la liaison à une seule boîte est 100 fois plus faible (Chen *et al.*, 1997).

La protéine AhrC de *Bacillus subtilis* présente 27% d'identité à la protéine ArgR d'*E. coli*. Elle a d'ailleurs pu remplacer ArgR dans la recombinaison au site *cer* (Stirling *et al.*, 1988b) et elle est parfaitement fonctionnelle dans les autres fonctions de ArgR dans *E. coli* (homologie fonctionnelle ; Czaplewski *et al.*, 1992).

La structure complète de ArgR a été ensuite déterminée, mais chez *Bacillus stearothermophilus*. L'hexamère est formé de six sous-unités identiques où les 6 domaines de liaison à l'arginine entourent le cœur central oligomérique (Ni *et al.*, 1999). La différence principale révélée par les structures déterminées avec et sans arginine est la formation d'une interface étroite ayant une orientation précise entre les trimères (Ni *et al.*, 1999).

### 5.1.2 PepA

Précédemment nommée XerB, la protéine PepA est également nécessaire en tant que facteur accessoire dans la recombinaison au site *cer* (Stirling *et al.*, 1989).

PepA est une protéine hexamérique dont le monomère fait 55,3 kDa. Elle requiert des ions  $Mn^{2+}$  pour son activité et clive les résidus N-terminaux de certains peptides : c'est une aminopeptidase de leucine (Taylor, 1993). Cette activité aminopeptidase n'est cependant pas requise pour la recombinaison au site *cer* (McCulloch *et al.*, 1994a).

D'un autre côté, PepA, alias CarP à ce moment, a pu se lier à la région contrôle de la région *carAB* de *S. typhimurium* (Charlier *et al.*, 1995a) et d'*E. coli* (Charlier *et al.*, 1995b). En agissant sur la transcription de *carAB*, PepA est donc aussi impliquée dans la régulation des pyrimidines par la carbamoylphosphate synthétase d'*E. coli* et alors, l'activité peptidase de la protéine n'est pas requise (Charlier *et al.*, 1995b).

La structure de PepA a été déterminée à 2,5 Å : PepA est organisée comme un dimère de trimères de symétrie 3 2. Chacune des sous-unités monomériques comprend deux domaines : le domaine N-terminal (les 166 premiers acides aminés) est lié au domaine C-terminal (à partir du 193<sup>e</sup> acide aminé) par une longue hélice  $\alpha$  formée des résidus 167 à 192. Tandis que l'activité peptidase est retrouvée dans le domaine C-terminal, il a été supposé que le domaine N-terminal contiendrait le domaine de liaison à l'ADN (Sträter *et al.*, 1999).

Peu après, la comparaison de la structure de PepA à la structure de la LAP (bovine lens aminopeptidase ; structure déterminée par Sträter et Lipscomb, 1995), a permis de renforcer cette idée puisque la LAP n'avait pas montré d'activité de liaison à l'ADN. De plus, la plupart des mutants déficients pour la liaison, mutants obtenus par mutagenèse aléatoire, étaient touchés dans le domaine N-terminal ou alors dans l'extrémité la plus C-terminale de PepA (Charlier *et al.*, 2000).

### 5.1.3 La protéine ArcA

Contrairement à ce qui est vu pour le site *cer*, ArgR n'a pas été montré comme ayant un rôle dans la recombinaison au site *psi*. Seule PepA était alors essentielle (en plus, bien sûr des recombinases XerC et XerD ; Colloms *et al.*, 1996). Cependant, l'efficacité de la recombinaison au site *psi in vitro* a été améliorée par l'ajout de la protéine ArcA et cette protéine s'est montrée essentielle *in vivo* (Colloms *et al.*, 1998).

Cette protéine fait partie du système ArcAB (pour aerobic respiration control ; contrôle de la respiration en aérobiose) dans lequel la protéine ArcB est une protéine senseur et ArcA la protéine régulatrice qui se lie à l'ADN. Ce système à deux composantes régule plus de 30 opérons en conditions d'anaérobiose. Tandis que *arcB* est exprimé constitutivement, *arcA* est exprimé quatre fois plus dans les conditions d'anaérobiose (Iuchi et Weiner, 1996). ArcB est en fait une kinase autophosphorylée liée à la membrane et qui phosphoryle à son tour la protéine ArcA afin de l'activer via un relais de transmission du phosphate à l'intérieur de ArcB (Kwon *et al.*, 2000).

La protéine ArcA-P (ArcA phosphorylée) serait en fait un régulateur plus global, impliqué dans des mécanismes autres que celui de la respiration. Il a été montré qu'elle s'était liée à gauche de la région d'*oriC* et avait inhibé la réplication du chromosome d'*E. coli in vitro* (Lee *et al.*, 2001). De plus, un site de liaison de ArcA est retrouvé sur le plasmide R1 dans la région de transfert *traJ* (Strohmaier *et al.*, 1998). Récemment, en établissant le profil de transcription dépendant de ArcA-P puis en développant une carte d'affinité de séquence d'ArcA-P, 55 nouveaux opérons se sont révélés être régulés par cette protéine (Liu et De Wulf, 2004).

Également, une boîte ayant un consensus de 8 sur 10 pb d'une boîte ArcA a été retrouvée à environ 100 pb en amont de la région d'échange du site *psi*. ArcA, phosphorylée ou non, a d'ailleurs pu s'y lier. Ainsi, il a été supposé qu'ArcA pourrait jouer un rôle similaire à celui rempli par ArgR au niveau du site *cer* (Colloms *et al.*, 1998).

## 5.2 Implication topologique des facteurs accessoires

Afin d'expliquer comment les protéines ArgR et PepA faisaient en sorte d'orienter la réaction de recombinaison au site *cer in vivo* vers une réaction intramoléculaire et non pas intermoléculaire, la notion de filtre topologique avait été employée (Summers, 1989).

Cette sélectivité topologique serait expliquée par la formation de synapses ayant une géométrie protéine/ADN fixe (Colloms *et al.*, 1996). En effet, après avoir constaté que la recombinaison entre deux sites *psi* sur un même vecteur donnait lieu à la production de caténanes (Colloms *et al.*, 1996), il a été par la suite montré que ces caténanes avaient une topologie fixe : la synapse productive a été formée en entourant les séquences accessoires autour des protéines accessoires de telle sorte d'attraper 4 nœuds (« nodes »; le terme « node » en anglais signifie le point où deux segments d'ADN se croisent -Dröge et Cozzarelli, 1992). Les analyses de Colloms *et al.* ont laissé suggérer que la synapse, en se créant, avait attrapé trois supertours négatifs entre les sites de recombinaison placés de façon anti-parallèle et que la réaction d'échange avait introduit un surenroulement négatif supplémentaire (Colloms *et al.*, 1997). Les caténanes ainsi formés étaient enroulés dans un sens « tourné vers la droite ».

D'un autre côté, bien que la recombinaison *in vitro* entre deux sites *cer* sur un vecteur surenroulé s'arrête après la première réaction d'échange (Colloms *et al.*, 1996), la formation de la synapse se serait déroulée de la même façon (Colloms *et al.*, 1997).

La sélectivité pour la recombinaison intramoléculaire au site *psi* n'a pu être contournée qu'en plaçant deux sites de recombinaison sur des caténanes plusieurs fois reliés (multiply interlinked catenanes). Tout comme il avait été montré précédemment, la recombinaison n'est alors survenue efficacement que lorsque les sites étaient antiparallèles l'un à l'autre et alors, il y a eu introduction d'un nouveau croisement (Bath *et al.*, 1999). Ainsi, la notion de filtre topologique exercé par les protéines accessoires restait cohérente, puisque la recombinaison n'était survenue que dans une synapse définie productive (Bath *et al.*, 1999).

### 5.3 Modèles du complexe synaptique au site *cer*

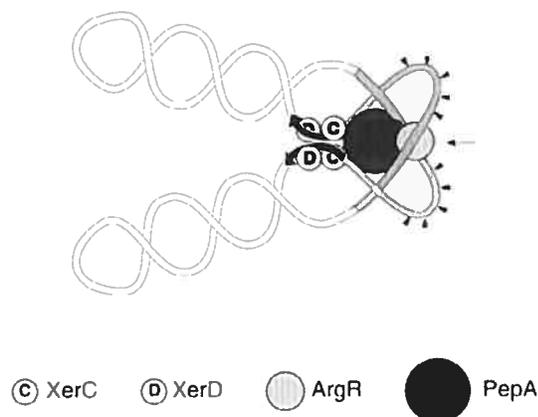
Les protéines ArgR et PepA sont toutes deux des protéines hexamériques pouvant se lier à l'ADN. Elles ont montré deux fonctions bien distinctes au sein de la cellule bactérienne : une fonction dans la recombinaison spécifique au site *cer* et une autre dans la modulation de l'expression de certains gènes.

Comme il a été mentionné plus haut, dans la recombinaison au site *cer*, ArgR et PepA feraient en sorte d'assurer la sélectivité pour la résolution. En effet, la recombinaison à ce site est exclusivement intramoléculaire et survient entre des sites directement répétés.

PepA serait essentielle pour le pairage des sites de recombinaison (Guhathakurta et Summers, 1995). Par la suite, la même équipe a montré que les facteurs accessoires servaient à bloquer la réaction de recombinaison en *trans*. Dans cette étude, une nouvelle classe de sites de recombinaison a montré une dépendance aux facteurs accessoires pour devenir monomères; en leur absence, des multimères étaient formés (Guhathakurta *et al.*, 1996).

Un premier modèle de l'architecture du complexe au site *cer* a alors été créé et qui montrait la protéine PepA en contact avec les recombinases (Guhathakurta *et al.*, 1996). Le modèle a, bien entendu, été étoffé par la suite par toutes les études biochimiques qui ont suivi.

Le second modèle a pris alors en compte les produits topologiques obtenus suite aux interactions des protéines accessoires et des recombinases. Tout d'abord, les études de Colloms *et al.* (1997, et voir le point précédent) ont montré que le produit obtenu de la recombinaison au site *cer* avait une topologie fixe. Effectivement, les deux sites *cer* sont enroulés de façon plectonémique dans le complexe nucléoprotéique (Alén *et al.*, 1997).



**Figure 8: Modèle du complexe formé entre deux sites *cer* par ArgR et PepA selon Alén *et al.* (1997).** Les deux sites *cer* antiparallèles sont enroulés dans un plectonème tourné vers la droite par ArgR et PepA, lesquelles attrapent alors trois supertours négatifs. Les petites flèches indiqués sur les sites d'action montrent les zones de flexibilité (non protégées) révélées par un test à la DNase I. Ici, seul le modèle dans lequel une molécule de PepA est impliquée est montré. Deux molécules de PepA pourraient également être impliquées. Cette figure est tirée de Alén *et al.* (1997).

Enfin, la démonstration de la liaison potentielle de PepA par des tests de protection à la DNase I en trois zones du site *cer* (Alén *et al.*, 1997), le fait que ArgR induise une courbure de 65° lorsqu'elle se lie à une seule boîte ARG (Burke *et al.*, 1994) et la définition topologique des produits obtenus (Colloms *et al.*, 1997, Alén *et al.*, 1997) ont permis de

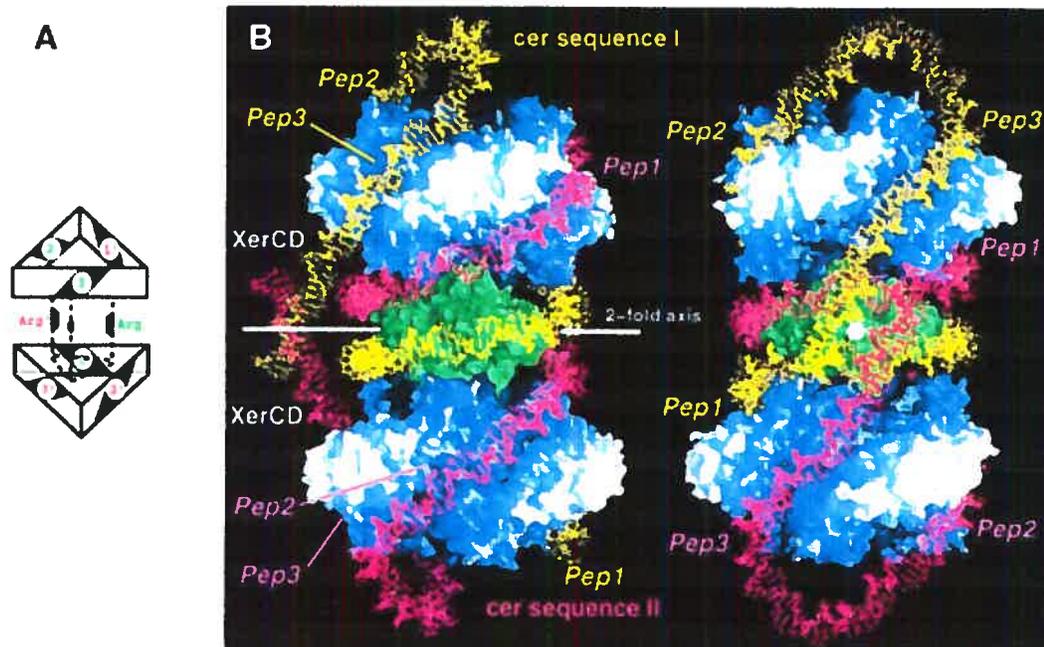
préciser un second modèle (Figure 8). À ce stade, on ne pouvait déterminer si un seul ou deux hexamères de PepA étai(en)t impliqué(s), et c'est pourquoi deux options restent possibles. Dans les deux situations, les séquences accessoires sont attrapées par PepA et ArgR de telle sorte que trois surenroulements négatifs sont captés pour donner un plectonème tourné vers la droite dont les sites sont anti-parallèles. Dans ce modèle, le complexe synaptique semble donc se former au moment de la formation d'un dimère plasmidique (assemblage *de novo*) (Alén *et al.*, 1997).

Peu de temps après cette étude, une autre équipe a proposé un modèle de la formation du complexe synaptique en se basant principalement sur des analyses informatiques et sur les structures existantes (XerD et ArgR). Ce modèle mentionne le complexe obtenu suite à la formation d'un dimère, et aussi le complexe pouvant se former au niveau d'un seul site *cer*. Dans ce cas, tout comme il avait été proposé précédemment par la même équipe (Guhathakurta *et al.*, 1996, voir plus haut pour le premier modèle), PepA agirait en tant qu'adaptateur pour connecter les recombinases XerC et XerD à l'hexamère ArgR. Ce modèle présume alors d'un contact entre les protéines accessoires ; bien que cela n'ait pas été encore prouvé, ce contact est soupçonné par les surfaces et charges potentielles entre les protéines. Ce modèle prédit aussi un contact probable entre PepA et les recombinases.

En fait la plus grande variation avec le modèle précédent est que le modèle suppose la formation d'un complexe à un site unique *in vivo* et que le complexe synaptique « complet » survient à la suite de la rencontre de deux complexes lors de la formation d'un dimère. Il faut donc au système un moyen de distinguer les sites qui seraient en *cis* de ceux situés en *trans*. L'équipe de Summers a alors proposé que le niveau de surenroulement de l'ADN pouvait servir de « ressort » pour renforcer le pont créé par ArgR (si nécessaire) et permettre un changement de niveau de l'ADN (« gearshift ») ; le contact résulterait donc de ce « glissement » dû au ressort d'ADN. Cette isomérisation permet ainsi le rapprochement de chacun des sites complexés avec XerCD et un hexamère de PepA, pour former le complexe synaptique (Hodgman *et al.*, 1998).

Enfin, le dernier modèle proposé jusqu'à maintenant pour la formation du complexe synaptique au site *cer* a consisté à intégrer de nouvelles données concernant la structure de PepA d'*E. coli*, laquelle présente trois sillons qui pourraient être les sites de liaison potentiels à l'ADN (Sträter *et al.*, 1999). Le modèle retenu est montré en Figure 9. D'après

la structure de PepA et les contacts possibles déterminés précédemment (Alén *et al.*, 1997), deux hexamères de PepA sont supposé permettre l'enroulement des séquences accessoires en conjonction avec un seul hexamère de ArgR et former un genre de sandwich le long de leur axe moléculaire à trois plis (ce modèle intègre donc le modèle de la Figure 8, mais contenant deux molécules de PepA au lieu d'une seule). Comme démontré auparavant, cet alignement permet d'entourer l'ADN tourné vers la droite afin d'attraper les trois surenroulements négatifs (Sträter *et al.*, 1999).



**Figure 9 : Modèle du complexe Xer au site *cer* suite à l'obtention de la structure de PepA.** A) L'orientation des molécules PepA et ArgR n'étant pas encore connue, l'un des modèles les plus probables du complexe Xer au site *cer* est montré. En noir sont indiqués les sites de liaison possibles de PepA (représentée en triangle). Ce modèle contient l'axe symétrique à deux plis reliant deux molécules PepA aux sites *cer* et qui correspondent à l'axe de symétrie de ArgR. B) Modèle moléculaire du modèle présenté en A. Les deux molécules de PepA sont représentées en bleu tandis que ArgR est en vert. Les sites *cer* sont ici en pourpre et en jaune. Cette figure est tirée de Sträter *et al.* (1999).

La structure des recombinases au site d'action, la connaissance des points de liaison exacts de PepA, ainsi que la vérification d'une liaison entre les protéines recombinases et accessoires, permettraient de confirmer ou d'infirmer certains aspects des modèles proposés. Selon les données de Gourlay et Colloms (2004) apportées par l'emploi des protéines ArgR et PepA pour la recombinaison avec Cre sur *loxP* (voir la section suivante), les modèles pourraient être légèrement modifiés afin de montrer l'absence de nécessité de

contact entre les protéines accessoires pour que le complexe soit formé et la recombinaison avoir lieu.

#### **5.4 Rôle des protéines accessoires sur le prophage P1**

Dans l'attente de savoir comment les protéines accessoires pouvaient activer la recombinaison spécifique au site *cer*, Gourlay et Colloms (2004) ont tenté d'étudier la recombinaison sur des sites hybrides contenant les séquences accessoires de *cer* ou *psi* avec la région centrale du site *loxP*, *in vitro*. L'ajout des séquences accessoires a fait en sorte que PepA a modifié la topologie normalement engendrée par Cre au niveau de *loxP*, pour obtenir des produits similaires à ceux obtenus pour *cer* ou *psi* (caténanes à 4 nœuds). ArgR ou ArcA n'auraient qu'un rôle secondaire, mais lorsque ajoutés, ont permis de réduire la quantité de PepA nécessaire. Enfin, les résultats obtenus laissent suggérer qu'il n'y aurait pas nécessité d'interactions entre les protéines accessoires et les recombinases (Gourlay et Colloms, 2004).

Les protéines accessoires ont donc maintenant été testées sur d'autres sites et il s'avère qu'elles peuvent jouer un rôle dans la recombinaison spécifique au site *loxP* (Gourlay et Colloms, 2004). *In vivo*, ces deux protéines se sont récemment révélées importantes pour la stabilité du prophage P1 dans les cellules d'*E. coli* (Paul et Summers, 2004). Une séquence pouvant peut-être correspondre à une boîte ARG a d'ailleurs été trouvée en amont de la région d'échange du site *loxP*, mais elle n'a pas encore été testée. Par contre, alors que les deux protéines sont tout aussi importantes pour la recombinaison au site *cer*, il a semblé ici que PepA l'était davantage que ArgR pour le maintien de P1 (Paul et Summers, 2004), ce qui a laissé suggérer aux auteurs qu'un complexe synaptique pourrait éventuellement se former en l'absence de ArgR, même si moins stable qu'en sa présence.

#### **5.5 Stabilité plasmidique de ColE1 et le transcrit Rcd**

D'après ce qui a été vu précédemment, la stabilité plasmidique de ColE1 est maintenue grâce à la recombinaison spécifique au site *cer* par les recombinases XerC et XerD et les facteurs accessoires ArgR et PepA. Cependant, cette stabilité n'est pas due uniquement au fait de cette recombinaison.

En amont de la région d'échange du site *cer* se trouve un promoteur (*P<sub>cer</sub>*) orienté vers la région d'échange de *cer* (Summers et Sherratt, 1988). Une mutation stimulant le promoteur (« up-mutation ») n'a pas eu d'effet sur la résolution des dimères en monomères (Summers et Sherratt, 1988), mais une mutation diminuant l'activité du promoteur (« down-mutation ») a eu pour effet de réduire la stabilité plasmidique de ColE1, sans que la recombinaison ne soit touchée (Patient et Summers, 1993).

Ce promoteur permet la synthèse d'un petit ARN nommé Rcd (pour régulateur de la division cellulaire ; regulator of cell division) d'environ 70 nucléotides qui empêcherait la division cellulaire (Patient et Summers, 1993). Il pourrait ainsi servir de point de vérification et établir le lien entre la résolution des multimères et la division cellulaire. En effet, le processus de résolution par la recombinaison spécifique est un processus relativement lent ; seul, ce processus ne suffirait pas à empêcher l'accumulation de dimères plasmidiques et ainsi courir le risque de ce qui a été nommé la « catastrophe dimérique » (Summers, 1998).

Cette expression, la « catastrophe dimérique », a été attribuée au fait que les dimères plasmidiques (générés par la recombinaison homologue) ont un potentiel deux fois plus grand d'être répliqué compte tenu de leurs deux origines de répllication. Si tel était le cas, après « x » générations, la population bactérienne finirait par ne contenir que des dimères, trimères, etc... puis il y aurait éventuellement perte du plasmide (Summers *et al.*, 1993). Heureusement cette « catastrophe » est évitée puisque les cellules contenant des multimères poussent plus lentement et donnent ainsi le temps aux recombinases de type Xer d'exercer leur action (Summers 1998).

Puisque l'activité métabolique des cellules surexprimant Rcd n'était pas empêchée, mais que le taux de croissance des cellules était réduit et éventuellement arrêté dans le cas de cellules dites quiescentes<sup>3</sup> (Rowe et Summers, 1999) et compte tenu de la structure secondaire de cet ARN, supposée par informatique, il était pensé que Rcd pouvait agir en

---

<sup>3</sup> Les cellules dites quiescentes sont des cellules mutantes pour *hns* et qui surexpriment Rcd. Dans ces cellules, la division cellulaire est bloquée et l'activité métabolique n'est concentrée que sur l'expression des gènes plasmidiques (Rowe et Summers, 1999). D'ailleurs, l'intérêt pour ces cellules et l'induction de Rcd a permis la mise au point d'un système pour exprimer abondamment des protéines exogènes à partir de plasmides (Summers 2002 *Microbiology Today*, 29 : 76-78), comme un fragment d'anticorps à une chaîne (scFV) employé en médecine (Mukherjee *et al.* 2004).

tant qu'ARN anti-sens. Cependant, aucune cible n'a pu être déterminée (Sharpe *et al.*, 1999). Il a alors été supposé que la cible serait l'ARNm codant pour une protéine impliquée dans la division cellulaire, mais rien n'a été prouvé jusqu'à maintenant.

L'activation de *P<sub>cer</sub>* se ferait par la multimérisation des plasmides, compte tenu du niveau élevé d'expression dans des cellules (*recF*<sup>+</sup> et *sbcA*<sup>-</sup>) contenant des multimères (Patient et Summers, 1993). La longueur réelle de l'espace entre la région -35 et -10 du promoteur *P<sub>cer</sub>* est de 15 pb. Or, il a été montré que la longueur optimale de cet espace pour l'expression de Rcd (et d'autres transcrits) était de 17 pb (Chatwin et Summers, 2001). Cet espacement de deux bases supplémentaires correspondrait à une rotation de 69° de l'ADN. Il a ainsi été supposé que la formation du complexe synaptique permettait un réalignement du promoteur grâce à une torsion formée par ce complexe au site *cer* (hypothèse de la torsion locale). L'activation de *P<sub>cer</sub>* pour la production de Rcd nécessiterait donc la formation d'un complexe nucléoprotéique complet obtenu sur un dimère plasmidique (Chatwin et Summers, 2001). (L'organisation de ces complexes a été montrée à la section précédente).

## 6 Un autre site d'action pour le système Xer : le site *dif* chromosomique

### 6.1 Généralités

En 1991, un locus adjacent au gène *hipA* était identifié. [Ce gène *hipA* (pour « high persistence locus ») est situé près de *terC3* dans la région de terminaison de la fourche de réplication du chromosome d'*E. coli* (Black *et al.*, 1991)]. La délétion de ce locus causait un phénomène de filamentation, une structure aberrante du nucléoïde, et l'induction du système SOS chez un certain pourcentage de cellules (environ 15%). Ce locus, dont une fraction présentait une très forte similarité au site *cer*, a été nommé *dif* pour deletion-induced filamentation (filamentation induite par la délétion) (Kuempel *et al.*, 1991). Le retour au phénotype sauvage par complémentation n'a pu se faire qu'en « *cis* » (avec intégration dans le chromosome) mais alors, l'orientation du site *dif* n'avait pas d'importance.

Il a été montré par la suite que seules les 33 pb de ce locus contenant la séquence similaire à la région d'action du site *cer*, sont suffisantes et nécessaires pour contrecarrer le phénotype de filamentation engendré par la délétion de la zone plus large entourant *dif*.

Cependant, le site *dif* doit se situer alors dans la zone de terminaison de réplication (Leslie et Sherratt, 1995, Tecklenburg *et al.*, 1995). Malgré cette localisation de *dif*, il a été montré que la terminaison de la réplication (et d'où la rencontre avec les fourches de réplication) n'était pas nécessaire pour la fonctionnalité du site *dif* (Cornet *et al.*, 1996, Kuempel *et al.*, 1996).

Le site *dif*, dont la recombinaison était indépendante de RecA (Kuempel *et al.*, 1991) a bel et bien été montré comme un site de recombinaison spécifique reconnu par les recombinases XerC (Blakely *et al.*, 1991) et XerD (Blakely *et al.*, 1993) et il semblait impliqué dans la bonne mise en œuvre de la ségrégation ou de la division cellulaire (Leslie et Sherratt, 1995). Par un test de marquage par densité, Steiner et Kuempel (1998a) ont pu ensuite confirmer que la recombinaison au site *dif* d'*E. coli* nécessite les recombinases XerC et XerD *in vivo*. D'un autre côté, la recombinaison au site *dif* n'a pas eu lieu en absence de division cellulaire, c'est-à-dire qu'aucune résolution n'a été observée lorsque la formation initiale du septum fut inhibée (Steiner et Kuempel, 1998a).

Les dimères chromosomiques, qui sont éventuellement résolus par la recombinaison au site *dif*, proviennent des échanges entre les chromatides-sœurs (SCE ; sister chromatide exchange). Cette formation de dimères chromosomiques est elle-même dépendante de RecA. Toujours grâce à un test de marquage par densité, il a été confirmé que la recombinaison au site *dif* était éliminée dans une souche *recA*<sup>-</sup> (Steiner et Kuempel, 1998a). De plus, alors que des mutations individuelles de *recB* et *recF* ont réduit de 50% les événements de recombinaison au site *dif*, une mutation dans les deux gènes a pratiquement éliminé la recombinaison à ce site. Ainsi, les deux voies de recombinaison (RecBCD et RecFOR) seraient impliquées dans les événements d'échanges entre les chromatides-sœurs (SCE ; sister chromatide exchange) (Steiner et Kuempel, 1998b).

De la même façon, des mutations hyper recombinantes [*polA* (codant pour la polymérase I, *dut* (codant pour la dUTPase) et *uvrD* (codant pour une hélicase)], ainsi qu'une mutation dans *fur* (codant pour une protéine impliquée dans la prise du fer dans la cellule), ont montré une augmentation du nombre d'événements de recombinaison au site *dif*. Il a été alors supposé que les échanges entre chromatides-sœurs pouvaient provenir de bris simple brins, causés par le processus de réplication en tant que tel, ou encore de dommages oxydatifs à l'ADN (Steiner et Kuempel, 1998b).

Enfin, le site *dif* fut montré comme un site préférentiel pour la topoisomérase IV ; les recombinases XerC/D seraient nécessaires pour son activité de clivage, mais pas FtsK, ni la division cellulaire. Cette affinité serait peut-être due à une structure particulière créée par les recombinases XerC/D puisque la séquence n'a pas d'importance. Il est à noter cependant que le site *dif* n'est pas essentiel pour la décaténation, et que, s'il est absent, d'autres sites pourraient probablement jouer ce rôle afin de veiller à la bonne complétion de la réplication (Hojgaard *et al.*, 1999).

## 6.2 La protéine FtsK

Avant que son rôle dans la recombinaison au site *dif* n'ait été révélé, la protéine FtsK avait été découverte comme une protéine importante impliquée dans la division cellulaire (Begg *et al.*, 1995). En effet, une mutation par substitution non-sens dans le domaine N-terminal (*ftsK44*) a empêché la division cellulaire. Cependant, la ségrégation de l'ADN n'a été nullement affectée dans la souche contenant cette mutation (Begg *et al.*, 1995). Une autre équipe a montré que cette mutation réduisait la capacité de localisation au septum de FtsK (Yu *et al.*, 1998a).

Par contre, une certaine proportion des cellules ne contenant que la portion N-terminale de FtsK a vu leur nucleoïde mal placé et le système SOS induit lorsqu'il y a eu division cellulaire, probablement suite à un endommagement de l'ADN. Il a donc été supposé que le reste de la protéine servirait à la localisation du chromosome dans la cellule (Liu *et al.*, 1998). D'un autre côté, une surexpression de la protéine FtsK nuit à la division cellulaire en inhibant l'assemblage de l'anneau créé par FtsZ (Draper *et al.*, 1998).

### 6.2.1 Caractéristiques

Cette protéine, d'environ 147 kDa (1329 aa), contiendrait trois domaines distincts : un domaine N-terminal d'environ 220 aa, une zone centrale d'environ 500 aa riche en PQ (proline et glutamine), et un domaine C-terminal d'environ 470 aa ayant une très forte similitude avec des protéines transférant l'ADN chez d'autres espèces (par exemple : la protéine SpoIIIE de *Bacillus subtilis* (Begg *et al.*, 1995). Toutes les protéines présentant ce domaine hydrophile ont une séquence potentielle très conservée de liaison à l'ADN. Il a par la suite été déterminé que le domaine N-terminal était le domaine essentiel impliqué

dans la division cellulaire en tant que telle et que le reste de la protéine n'était pas nécessaire à la viabilité (Draper *et al.*, 1998).

La caractérisation plus avancée de cette protéine a révélé que, non seulement les 202 premiers acides aminés sont essentiels pour la division cellulaire et la viabilité, mais aussi que cette protéine était associée à la membrane par l'intermédiaire de ce domaine (Wang et Lutkenhaus, 1998 et Yu *et al.*, 1998a). De son côté, le domaine cytoplasmique de FtsK, qui est le domaine C-terminal de la protéine, sert à la ségrégation chromosomique (Yu *et al.*, 1998b). Plus précisément, il sert à la résolution des dimères chromosomiques puisque la recombinaison spécifique de site au site *dif* nécessite ce domaine (Steiner *et al.*, 1999). Enfin, une série de mutants de délétion a permis de montrer que la région centrale riche en PQ avait aussi un rôle à jouer dans cette résolution (Boyle *et al.*, 2000).

### 6.2.2 Domaine C-terminal de FtsK

Le domaine C-terminal de FtsK a pu ensuite être subdivisé en trois parties (Yates *et al.*, 2003) :

1. la portion la plus N-terminale ( $\alpha$ ) serait peut-être pour la modulation de la réaction,
2. la portion centrale ( $\beta$ ) est très conservée et possède le motif des protéines de la famille des AAA (ATPases ayant des activités diverses) et enfin,
3. la portion la plus C-terminale contenant les 140 derniers aa ( $\gamma$ ) est responsable de la spécificité d'espèce.

Ainsi, la portion  $\beta$  montre que la protéine fait partie de la superfamille des protéines AAA (pour ATPases associées à différentes activités), et dans le sous-groupe associé à SpoIIIE. D'ailleurs, l'activité de FtsK au niveau de la résolution des dimères chromosomiques nécessite l'hydrolyse de l'ATP (Aussel *et al.*, 2002).

De plus FtsK s'est révélée être spécifique d'espèce puisque que la protéine d'*E. coli* n'a pu remplacer celle d'*Haemophilus influenzae* pour la recombinaison au site *dif*, et vice-versa ; cette spécificité d'espèce fut octroyée au domaine  $\gamma$  du domaine C-terminal de FtsK (Yates *et al.*, 2003).

### 6.2.3 Similarités de FtsK avec SpoIIIE

Comme mentionné, la région C-terminale de FtsK d'*E. coli* est très similaire à la protéine SpoIIIE. Cette protéine sert au transport de l'ADN de la cellule-mère à la présore lors de la division effectuée au cours de la sporulation chez *Bacillus subtilis* (Errington *et al.*, 2001 et Bath *et al.*, 2000).

En réalité, plusieurs similitudes peuvent être recoupées chez les deux espèces quant aux fonctions de ces deux protéines. Ainsi, le domaine N-terminal de la protéine FtsK est nécessaire à la localisation au septum de division cellulaire ; chez SpoIIIE, ce même domaine sert à l'ancrage, mais la localisation est plus aléatoire. Alors que le domaine C-terminal de FtsK est impliqué dans la formation du septum en plus de la résolution des dimères chromosomiques, le domaine C-terminal de SpoIIIE est requis lors de la sporulation, et sert aussi à la distribution adéquate du chromosome (pour une revue : Sciochetti et Piggot, 2000).

### 6.3 Études de la recombinaison au site *dif* d'*E. coli* avec la protéine FtsK

Ainsi, après avoir pensé que la recombinaison au site *dif* ne nécessitait pas de protéines accessoires, comme ce qui avait été vu pour les recombinaisons aux sites *cer* ou *psi* par exemple, le domaine cytoplasmique de FtsK a été montré comme ayant un rôle à y jouer. La recombinaison entre deux sites *dif* portés par un plasmide ou entre un site *dif* porté par le chromosome et un autre porté par un plasmide ne survient efficacement que lorsque ce domaine est présent et en présence de RecA (il ne faut pas oublier qu'en absence de RecA, il y a moins de formation de multimères, et alors, la nécessité de recombinaison de type Xer au site *dif* n'est plus). Cette protéine « accessoire » n'est cependant pas nécessaire à la recombinaison aux sites *cer* ou *psi* (Recchia *et al.*, 1999).

Cette « nécessité » de la présence de FtsK pour la recombinaison au site *dif in vitro* n'a pu être évitée que lorsque la recombinaison fut effectuée avec les protéines XerC d'*Haemophilus influenzae* et XerD d'*E. coli* (Neilson *et al.*, 1999).

Par la suite, il a été montré que la recombinaison au site *dif* était activée par FtsK et qu'une fois lancée, la réaction avait lieu, même en l'absence de formation de dimères chromosomiques (Barre *et al.*, 2000). Cependant, compte tenu de la toxicité d'une

surexpression de la protéine entière et de la non activité d'une protéine tronquée ne contenant que strictement le domaine C-terminal, une étude subséquente a montré que le domaine C-terminal combiné à 50 aa situés plus en amont (résidus 179 à 230) suffisait pour produire une protéine fonctionnelle pour la recombinaison spécifique de site (Aussel *et al.*, 2002). Ainsi, des tests *in vitro* ont montré que cette protéine, combinée aux recombinases Xer et à l'hydrolyse de l'ATP, était suffisante pour effectuer une réaction d'échange à *dif*. La même étude a pu démontrer que cette portion de FtsK formait un hexamère en forme d'anneau, qu'elle possédait une activité translocase d'ADN (elle génère des supertours positifs en présence de la Topo I sur des cercles d'ADN relaxés), mais que la réaction d'échange des brins à *dif* était indépendante de la topologie du substrat.

Cette activité translocase, permettant la formation de boucles dans l'ADN, s'est par la suite révélée indépendante des protéines Xer (Saleh *et al.*, 2004).

#### 6.4 Lien avec XerC et XerD

À la suite de la construction de plusieurs dérivés de FtsK, il a été montré que le domaine C-terminal était seul suffisant pour qu'une recombinaison complète puisse être effectuée au site *dif in vitro*. L'échange des premiers brins se ferait par XerC seule, mais cet échange est non stable et converti immédiatement en substrat original. FtsK viendrait alors stabiliser la jonction de Holliday formée et activerait la deuxième réaction d'échange par XerD (Barre *et al.*, 2000).

Ainsi, comme le faisait remarquer Donachie : « La question demeurerait pourtant à savoir la raison d'être réelle de FtsK au niveau de la recombinaison au site *dif*, surtout que des homologues de FtsK chez d'autres espèces bactériennes ne servent pas à la monomérisation des dimères chromosomiques, mais servent plutôt au transfert d'ADN entre les cellules » (Donachie, 2002).

C'est ainsi qu'Aussel *et al.* ont montré que FtsK active un changement dans l'ordre d'échange des brins au moment de la recombinaison au site *dif*. Ainsi, contrairement à ce qui est vu au niveau du site *cer* (ou au site *dif* en l'absence de FtsK), XerD clive le premier brin et XerC complète la réaction en résolvant la jonction de Holliday formée. Cette résolution se ferait par contre sans la nécessité de FtsK (Aussel *et al.*, 2002).

Par la suite, il a été proposé que FtsK servait à positionner adéquatement les régions contenant *dif* pour que la synapse entre les deux sites *dif* soit productive. En fait, le domaine C-terminal de FtsK est nécessaire à la résolution des dimères chromosomiques, mais n'est pas nécessaire à l'action de recombinaison. Ceci a été vu lorsque le site *dif* a été remplacé par le site *loxP*, et l'action de recombinaison effectuée par Cre. De plus, comme vu chez la souche *dif*, le phénotype de filamentation engendré par l'enlèvement de la portion C-terminale de FtsK disparaît quand la souche est également *recA* (Capiaux *et al.*, 2002).

Le phénotype de filamentation engendré par la délétion du site *dif* avait pu être éliminé lorsque le site *loxP* avait été mis à la place (en fournissant la recombinase Cre ; Leslie et Sherratt, 1995) et même par *psi* (Cornet *et al.*, 1994). Ce remplacement du site *dif* par d'autres sites de recombinaison [*psi*, *loxP*, et aussi *cer3* (site *cer* modifié)] a ensuite permis de séparer l'événement de résolution des dimères chromosomiques (ou CDR) de l'événement de recombinaison. En effet, dans le cas des trois sites mentionnés, la recombinaison pouvait avoir lieu en absence de FtsK, mais pas la résolution des dimères chromosomiques (Capiaux *et al.*, 2002).

Plus récemment encore, Bigot *et al.* ont tenté de mieux distinguer les deux rôles de la portion C-terminale de FtsK, c'est-à-dire l'activation directe de la recombinaison par XerC/D sur le site *dif*, et l'activité de translocation d'ADN afin d'amener les sites ensemble et/ou de les éloigner du septum de division cellulaire en train de se refermer (Bigot *et al.*, 2004). En parallèle, ils sont parvenus à mieux définir l'avantage possible de la région « linker » riche en PQ. Cette région est très divergente d'une espèce à l'autre, contrairement à la région C-terminale. Elle n'est pas nécessaire à l'activation de la recombinaison par les protéines Xer, mais elle est nécessaire à la bonne mise en place de la division cellulaire. L'hypothèse retenue par l'étude est qu'elle aiderait à la stabilisation de l'interaction de FtsK avec les autres protéines impliquées dans la division cellulaire, ou encore que sa délétion modifierait une structure d'ordre complexe des autres protéines du septum (Bigot *et al.*, 2004).

### **6.5 Zone d'activité du site *dif***

Comme mentionné au point 6.1, la position du site *dif* dans le chromosome d'*E. coli* est importante pour que ce site puisse remplir sa fonction. En fait, le site *dif* ne peut être

déplacé qu'à l'intérieur d'une zone restreinte équivalent à environ 30 kb sans que l'on ne voit apparaître le phénotype de filamentation. Cette zone a alors été dénommée DAZ pour zone d'activité du site *dif* (*dif* activity zone ; Kuempel *et al.*, 1996, Cornet *et al.*, 1996).

D'un autre côté, un site *dif* minimal de 33 pb reste fonctionnel dans une souche présentant une délétion d'une certaine portion de la région entourant normalement *dif* (Tecklenberg *et al.*, 1995, Kuempel *et al.*, 1996). Dans ce cas, il a été supposé que des réarrangements chromosomiques devaient permettre la construction d'une nouvelle DAZ et ainsi éliminer le phénotype de filamentation (Kuempel *et al.*, 1996). Ces phénomènes seront expliqués un peu plus loin dans le texte.

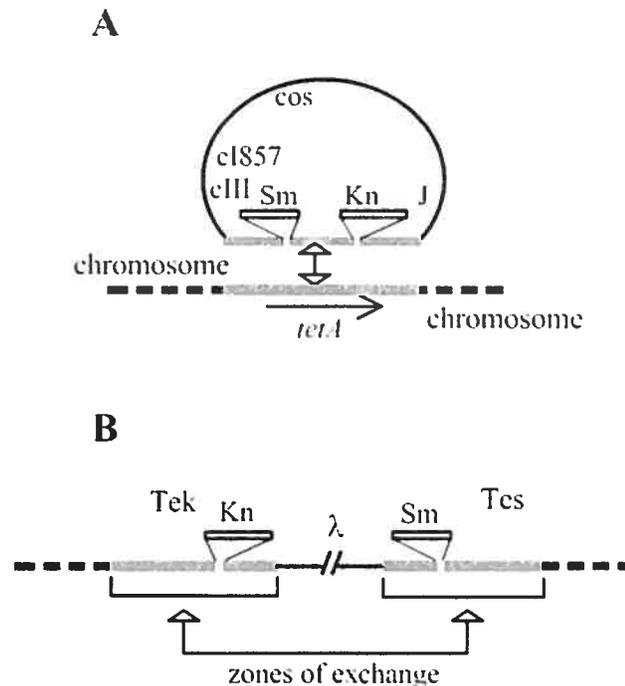
## 6.6 Phénomène d'hyper-recombinaison

Également dans la région terminale de réplication, à l'époque de la découverte du site *dif*, un phénomène de recombinaison excessive par RecA a été observé entre le site de pause le plus employé, TerC (l'un des sites sur lequel agit la protéine Tus requise pour arrêter l'avancée de la fourche de réplication –lire Hill et Marians, 1990) et le site nouvellement défini, *dif* (Louarn *et al.*, 1991). La région a été définie par la suite comme débutant aux alentours de la 23,3<sup>e</sup> min (près de *terE*) et se terminant entre la 39,5<sup>e</sup> min et 40,6<sup>e</sup> min (soit peu après la position du gène essentiel *nadE* codant pour la NAD synthétase). Elle fut alors nommée TRZ (pour « terminal replication zone » ; zone terminale de recombinaison) (Louarn *et al.*, 1994<sup>4</sup>).

Ce phénomène de recombinaison terminale a été mesuré par la fréquence d'excision d'un prophage  $\lambda$ TSK dont le principe est illustré à la Figure 10. Il a été déterminé que cet événement n'est pas (ou très peu) influencé par la recombinaison au site *dif*, ni ne nécessite des pauses de la fourche de réplication (Louarn *et al.*, 1994). Par contre, il est dépendant du complexe RecBCD, ce qui laissait suggérer que cette recombinaison provenait de bris double-brins ou encore de bris menant à la formation de bris double-brins. De plus, tout comme ce fut le cas pour la recombinaison au site *dif*, il n'y avait pas de relation directe entre la recombinaison terminale et la terminaison de la réplication puisque l'hyper-recombinaison est survenue dans la TRZ dans une souche Tus<sup>-</sup>, et ce, même quand cette zone avait été déplacée près d'oriC (Corre *et al.*, 1997).

---

<sup>4</sup> Les définitions des positions entre parenthèses ont été retrouvées sur le site [<http://www.shigen.nig.ac.jp/ecoli/pec/index.jsp>] consulté le 24 septembre 2004.



**Figure 10: Test par l'excision d'un prophage.** A) Intégration du prophage  $\lambda$ TSK par recombinaison homologue. Le dérivé du phage  $\lambda$  ( $\lambda$ TSK) contient les gènes *tet* du *Tn10* au lieu de la région *att int*. Les gènes conférant les résistances à la kanamycine ( $Km^r$ ) et à la streptomycine-spectinomycine ( $Sm^r$ ) ont été insérés dans *tetA*. Ce prophage a donc la capacité de s'intégrer dans un chromosome d'*E. coli* contenant un *Tn10* résident, par le biais de la recombinaison homologue, mais il ne peut plus s'intégrer dans le site chromosomique *attλ*. B) Excision du prophage. La souche bactérienne est donc devenue  $Tc^s Km^r Sm^r$  et le prophage intégré est alors flanqué de séquences répétées. De plus, compte tenu de la thermosensibilité du prophage intégré (il contient le répresseur *cI* muté : *cI857*), il sera conservé à 30°C. Après une induction à 42°C, les bactéries survivantes seront essentiellement le résultat d'une recombinaison homologue excisive entre les séquences répétées. La souche pourra redevenir thermorésistante (Tr) et  $Tc^r$  ou alors  $Tr Km^r$  (Tek) ou  $Tr Sm^r$  (Tes) selon comment se retirera le prophage. Cette figure est tirée de Corre et Louarn, 2002.

Ainsi, cette hyper-recombinaison laissait supposer l'existence d'une haute fréquence de séquences répétées dans cette région chromosomique (Louarn *et al.*, 1994). D'ailleurs, la fréquence d'excision du prophage  $\lambda$ TSK a été influencée par l'orientation de séquences répétées (Corre *et al.*, 1997) et ce, de façon indépendante de Xer (Corre *et al.*, 2000). Cependant, il n'a pas été observé dans une souche sauvage quand le prophage n'était pas inséré près du site *dif* (Corre *et al.*, 2000). Enfin, lorsque le prophage était inséré très près de *dif* dans l'orientation permettant la recombinaison terminale, la résolution des dimères chromosomiques était inhibée (phénomène appelé atténuation du CDR– « chromosome

dimer resolution ». Corre *et al.*, 2000). Ainsi, Corre *et al.* (2000) en ont supposé que le prophage  $\lambda$  modifié contenait lui-même suffisamment d'éléments orientés dans une direction pour qu'ils soient reconnus et qu'ils puissent former un côté de la DAZ. La fréquence d'excision du prophage était contrôlée par des éléments portés par le prophage lui-même et elle dépendait de l'orientation du prophage par rapport à *dif*.

## 6.7 Des éléments polaires dans la DAZ

La présence d'éléments DOPE (DAZ organizing polar elements ; éléments polaires organisant la DAZ), pourraient expliquer les comportements obtenus suite à certaines délétions ou inversions réalisées au niveau de la DAZ (Pérals *et al.*, 2000). Pour tester ces délétions-inversions, deux méthodes ont été employées. La première, mesure indirecte de la résolution des dimères chromosomiques, consistait en un test de co-culture : dans ce test, une souche contenant une délétion du site *dif* était mise en culture avec la souche à tester dont le site *dif* avait été déplacé. On voyait donc par ce test, si la souche testée croissait mieux que la souche *dif*. Le second test consistait en une mesure directe de la résolution par l'insertion d'une cassette Km<sup>R</sup> flanquée de deux sites *dif* en tandem (bien entendu, le gène *xerC* était amené par un vecteur inductible dans la souche construite avec cette cassette).

Ainsi, certaines inversions ont montré une inactivation de la DAZ, tandis que d'autres non, et des délétions ont permis une reformation de la DAZ. Ces délétions « autorisées » permettraient une réorganisation de la DAZ puisqu'elles n'empêcheraient pas aux éléments polarisés de converger vers le site *dif*. Les éléments polarisés situés de part et d'autres du site *dif*, mais un peu plus éloignés du centre, serviraient alors à la réorganisation de la DAZ.

Une délétion à l'extérieur du site *dif* a permis aussi de conserver l'activité au site *dif*. De la même façon, le segment chromosomique contenant *dif* a pu être inversé sans voir modifiée l'activité au site *dif*. Le site *dif* serait alors toujours à la rencontre des éléments polarisés, même si l'ensemble de la séquence fut inversé. Par contre, une inversion d'un segment chromosomique d'un seul côté du site *dif*, n'a pas permis l'activité du site. Cela serait expliqué par le fait que le site *dif* ne serait plus au lieu de convergence des éléments polaires, la DAZ s'étant reformée en excluant *dif*. Ainsi, ces éléments DOPEs sont

interchangeables, leur polarité peut être inversée, et ils peuvent même se retrouver un peu plus loin de la région terminale (Pérals *et al.*, 2000).

Toujours grâce à la cassette construite avec les deux sites *dif* flanquant le gène de la résistance à la Km, Pérals *et al.* ont observé que la recombinaison entre les sites *dif* avait lieu aussi bien de façon intra- qu'intermoléculaire, en autant que la cassette se situait au niveau de la DAZ précédemment déterminée (voir le point 6.5). De plus, ils ont montré qu'une surexpression de FtsK augmentait la recombinaison au niveau des sites *dif* déplacés, mais pas la résolution des dimères chromosomiques. Il semblerait donc que la quantité intracellulaire de FtsK soit un facteur limitant pour la recombinaison au site *dif*, ce qui serait expliqué par sa position au septum de division cellulaire (Pérals *et al.*, 2001 et voir le modèle de co-localisation, point 6.9).

## 6.8 Fragilité du terminus

La région du terminus est une région fragile où des événements intenses de recombinaison homologue surviennent quand la résolution des dimères chromosomiques est inactivée (souche *dif*, *xerCD*<sup>-</sup>) effet qui est rehaussé par une mutation dans *recD* (Corre *et al.*, 2000). Elle serait également rendue fragile par l'inactivation du domaine C-terminal de FtsK (Corre et Louarn, 2002). La zone fragile est plus étendue dans la souche mutante pour ce domaine de FtsK que dans les mutants *xer*, ce qui a laissé suggérer que cette région était moins bien déterminée. En effet, dans les mutants *xer*, cette zone se confine au voisinage de *dif* (Corre et Louarn, 2002).

Comme vu au point 6.6, le prophage  $\lambda$ TSK (revoir la Figure 10) inséré près de *dif* dans l'orientation permettant son excision (recombinaison hyper-terminale) ne permet plus la résolution des dimères chromosomiques (Corre *et al.*, 2000). La fréquence d'excision du prophage est alors un indicatif du degré de fragilité de la zone. En effet, lorsque les dimères chromosomiques ne peuvent plus être résolus, il y a un processus de dégradation de l'ADN qui est enclenché (déchirement de l'ADN au terminus ; « effet guillotine », voir la section suivante), ou encore, de recombinaison homologue pour la réparation. Ainsi, la recombinaison effectuée entre les éléments répétés flanquant la séquence du prophage induit son excision, phénomène qui peut être mesuré.

Un fait intéressant découlant de ces études (Corre et Louarn, 2002) est que le rôle de la polarité des séquences n'a pas été vu dans une souche *ftsKc*<sup>-</sup> alors qu'il est vu dans une

souche *ftsK*<sup>+</sup> ; les bris d'ADN (vus par la fréquence d'excision du prophage) sont répartis alors de façon beaucoup plus homogène et plus étendue dans le chromosome d'*E. coli*. Avec ce qui avait été vu précédemment, il a été supposé que le positionnement du terminus serait le résultat de deux processus consécutifs : un premier positionnement fait grossièrement et indépendamment de FtsK qui accompagne ou suit la reconstruction post-répllicative du nucléoïde, et un second plus précis (« fine tuning ») grâce à FtsK qui ferait en sorte de placer les sites *dif* près l'un de l'autre sous le septum (Corre et Louarn, 2002).

Il a été alors supposé que FtsK lisait en quelque sorte l'ADN, que les éléments polaires étaient reconnus d'une quelconque façon.

## 6.9 Modèle de co-localisation

Deux phénomènes, ont été observés au niveau de la région terminale de réplication, lesquels pouvaient être expliqués par la présence de multiples éléments polaires : 1) le contrôle de l'activité du site *dif* et de la recombinaison terminale dans une région restreinte, et 2) l'existence de zones non divisibles (NDZ pour « non-divisible zones ») centrées sur *dif* dans le chromosome (voir la revue de Capiaux *et al.*, 2001).

En fait, le macrodomaine Ter (décrit par Niki *et al.*, 2000), qui inclut la DAZ, coïncide avec une région où ces zones non divisibles (NDZ) sont retrouvées. Ces zones (délétères si inversées) seraient en fait une succession de plusieurs petites zones dans lesquelles on retrouve des séquences dont l'orientation est physiologiquement importante. La séparation de certains groupes de séquences ont d'ailleurs causé des anomalies dans la forme et dans la distribution des nucléoïdes (Guijo *et al.*, 2001). À l'opposé sur le chromosome d'*E. coli* se trouve un macrodomaine Ori qui contient *oriC* (Niki *et al.*, 2000).

Il a été observé que l'ADN était « déchiré » par le septum au moment de la division cellulaire. Cet effet, nommé « effet guillotine », a été vu tout d'abord chez les mutants *mukB* d'*E. coli* (Niki *et al.*, 1991), et ensuite dans les mutants montrant le phénotype Dif (Hendricks *et al.*, 2000). L'induction du système SOS survenait au moment de la division cellulaire dans les mutants affectés dans la recombinaison Xer au site *dif*, mais n'apparaissait pas lorsque la division était bloquée. Comme déjà mentionné, le site *dif* est situé en aval de *hipA*. Cette étude a montré que la longueur des filaments augmentait dans une souche double mutante *xerC*<sup>-</sup> *hipA*<sup>-</sup>. Cet effet serait expliqué par le fait que, le phénomène « guillotine » survenant dans les mutants *xer*<sup>-</sup>, il y aurait dégradation du locus

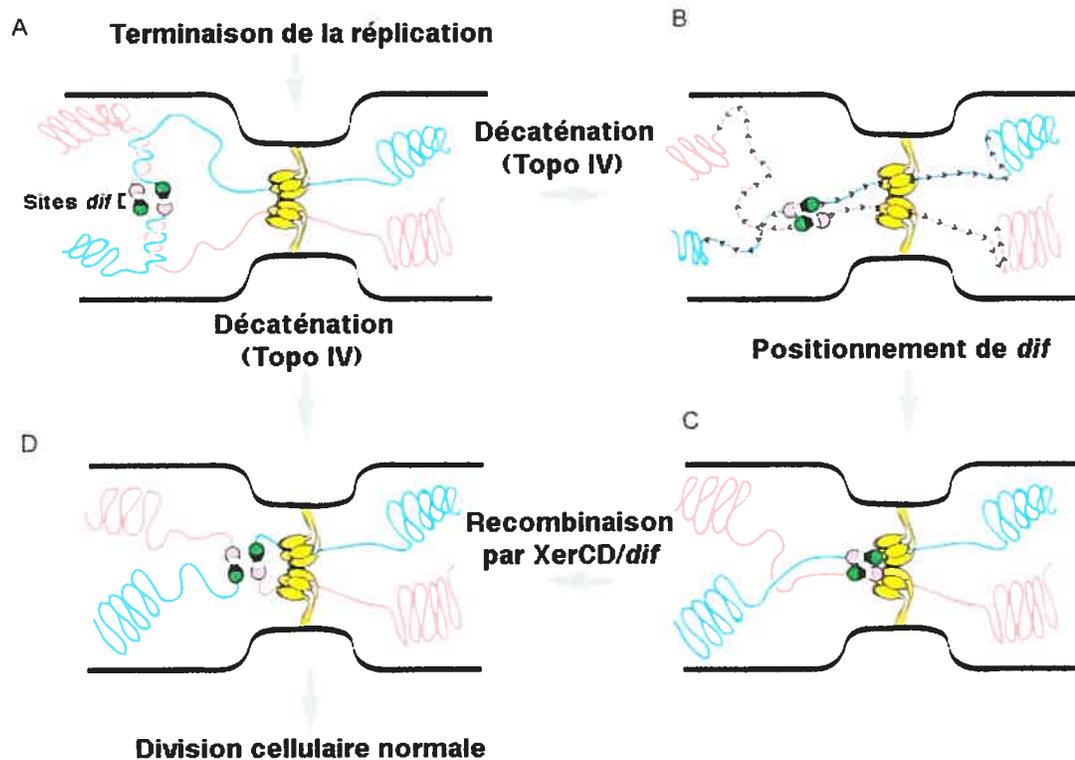
*hipAB* au moment de la formation du septum, et alors, le manque de production de HipB mènerait à la relâche (activation) des protéines HipA liées à l'ADN, ce qui inhiberait ensuite la croissance cellulaire et provoquerait la mort (Hendricks *et al.*, 2000).

Par hybridation d'ADN, il a ensuite été montré que la dégradation de l'ADN survenait effectivement dans la région du site *dif* et qu'elle était dépendante de RecD (ayant l'activité exonucléase V du complexe RecBCD) (Prikryl *et al.*, 2001).

Suite à ces résultats et à ceux des études sur le site *dif*, un modèle de co-localisation a été présenté (par Capioux *et al.* 2001, puis raffiné par Louarn *et al.*, 2004), postulant que la structure joue un rôle majeur dans la résolution des dimères chromosomiques (CDR). Selon ce modèle, les complexes XerCD/*dif* et FtsK (par son domaine N-terminal) se logent conjointement au septum de division cellulaire au moment de la septation, par le processus de la polarisation de l'ADN. La recombinaison au site *dif* ne peut alors se faire qu'au septum. Ainsi :

1. La polarisation des régions entourant le site *dif* est nécessaire à la localisation des sites *dif* d'un dimère chromosomique près du septum. Les DOPEs pourraient être impliqués dans le processus de mobilisation du terminus par le système de partition, afin de retirer le domaine TER du plan de division cellulaire [à ce moment, les chromosomes-fils peuvent être monomériques ou dimériques. C'est dans le cas de la présence d'un dimère seulement qu'il y aurait un genre de lutte à la corde (« tug-of-war ») qui finirait par placer les sites *dif* sous le septum].
2. Dans le cas de la formation d'un dimère chromosomique, les deux sites *dif* d'un dimère chromosomique seraient donc placés près du domaine C-terminal de FtsK, laquelle est ancrée au septum.
3. Ce positionnement faciliterait donc l'action de FtsK en conjonction avec les protéines Xer. Lorsque l'échange est complet et qu'il y a résolution des dimères chromosomiques, le processus de partition entre immédiatement en action et tire le domaine TER loin du septum pour faire en sorte qu'aucune autre activité de recombinaison n'ait lieu.
4. Enfin, dans le cas où la résolution n'aurait pas été possible, les régions contenant les sites *dif* sur le dimère se retrouvent prises au septum et subissent alors de la dégradation ou des altérations recombinantes.

La Figure 11 retrouvée dans la revue de Lesterlin *et al.* (2004) reprend les grandes lignes de ce modèle en y incluant la procédure de décaténation devant survenir pour que tous ces événements aient lieu.



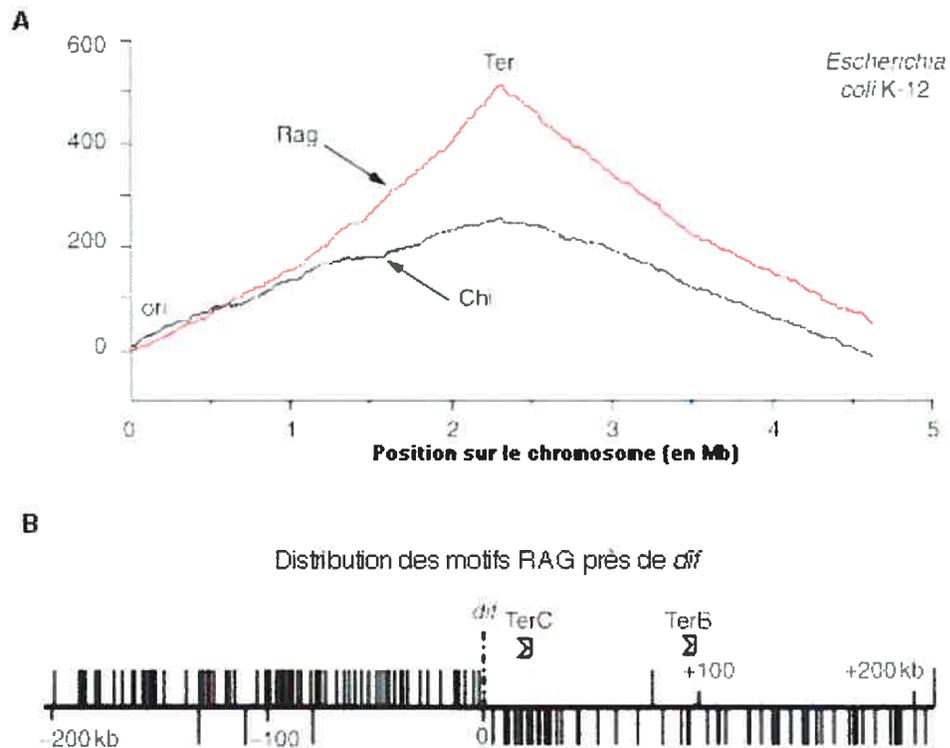
**Figure 11 : Modèle de la ségrégation des domaines Ter et de la résolution des dimères chromosomiques.** Les hexamères de FtsK sont représentés en jaune (au centre de chacun des schémas), les protéines XerC et XerD sont en rose et vert, les sites *dif* sont les demi-cercles noir et blanc, et enfin, les lignes bleues et rosâtres schématisent les chromosomes. A) À la fin de la réplication, les deux sites *dif* sont en quelque sorte stabilisés par les liens d'intercaténation situés entre eux. Deux situations peuvent alors se produire : lorsqu'il y a eu monomérisation, la décaténation suffit pour séparer les deux chromosomes nouvellement répliqués (Processus de A vers D). B) Lorsqu'il y a présence de dimères, par contre, les sites *dif* sont attrapés par le septum de division suite à la décaténation, et FtsK serait transloquée vers les sites *dif* grâce à la polarisation (représentée ici par les flèches noires). En parallèle, la topo IV interagit avec FtsK (contact prouvé par Espeli *et al.*, 2003) et retire les liens de caténation. C) Dans cette situation, le positionnement adéquat de *dif* serait assuré par la polarisation de l'ADN dans cette région chromosomique, et alors, la synapse requise pour permettre la réaction de clivage par XerD serait réalisée par le contact avec FtsK. D) Et enfin, la résolution des dimères chromosomiques permet la séparation des deux chromosomes-fils; la division cellulaire peut ensuite être complétée normalement. Cette figure a été modifiée de Lesterlin *et al.* (2004).

## 6.10 Les éléments polarisés au terminus : les éléments RAG

En 1998, Salzberg *et al.* ont retrouvé des oligomères dont l'orientation a été qualifiée de biaisée, particulièrement autour de l'origine de réplication et de la région du

terminus; cela signifiait que l'octamère étudié se retrouvait davantage sur le brin continu (« leading strand ») que sur le brin discontinu (« lagging strand »). Ainsi, en plus de l'élément CHI (TGTGGCG) déjà déterminé et étudié (Blattner *et al.*, 1997), cette équipe avait retrouvé un autre motif riche en purines (RRCAGGG où R désigne une purine) qui ne pouvait être simplement expliqué par les taux de mutation ou de réparation survenant régulièrement dans la cellule. Le terme chirochore avait été donné à tout segment individuel dans le génome d'*E. coli* qui ne respectait pas la règle de la parité, c'est-à-dire, la règle disant que si les événements de sélection et le nombre de mutations étaient équivalents pour chacun des deux brins d'ADN, un équilibre devait être atteint entre le nombre d'A versus T et de G versus C (Lobry, 1996). On désigna ensuite chacun des bras de réplication comme des réplichores, puisqu'un chirochore y était retrouvé (Blattner *et al.*, 1997). Le biais retrouvé avec le second motif était maximal avec l'octamère RRNAGGGS (où N représente n'importe quel nucléotide et S : G ou C) qui fut nommé RAG.

Ces éléments RAG pourraient représenter ces DOPEs compte tenu de leur représentativité autour de *dif* (Lobry et Louarn, 2003). La Figure 12B montre la distribution de ces éléments entourant *dif*. Comme il est relevé, le changement de polarité est presque de 100% lorsque l'on atteint Ter (ou *dif*. Les éléments RAG sont représentés par les barres verticales). Dans cette zone de 400 kb, nous pouvons constater que seuls quelques éléments ne sont pas polarisés dans la même orientation. Par comparaison, le biais du motif CHI est moins important, en ce sens que l'endroit de changement de polarité est moins clairement défini et que la différence entre une orientation versus l'autre est moins prononcée que pour les éléments RAG (voir la Figure 12A, comparez les deux courbes. Les valeurs de l'ordonnée représentent le nombre de fois qu'un motif est représenté dans la même orientation sur le 1<sup>er</sup> réplichore, par rapport à l'orientation inverse).



**Figure 12: Biases d'éléments répétés et orientés riches en G flquant de part et d'autre le site *dif* chez *E. coli* K12.** A) « Marche informatique » le long du brin chromosomique publié. Au fur et à mesure de la promenade, selon le motif analysé, il peut y avoir une unité franchie vers le haut lorsque le motif se présente dans le même sens que l'analyse, et une unité en moins lorsque le complément du motif est rencontré. Sur cette figure, la marche a permis de montrer que l'on retrouvait environ 500 éléments RAG (RGNAGGGS) de plus dans une orientation, lorsque l'on se dirige de l'origine de réplication vers le terminus. Lorsque l'on poursuit la marche après le terminus, le biais rencontré est inversé. Le changement du biais se révèle précisément au site *dif*, et, comme on le voit par la courbe obtenue, le biais augmente plus on s'approche de cette région. De la même façon, on peut rencontrer les motifs Chi (GCTGGTGG) de façon orientée, mais le biais est moins prononcé et la coupure est moins bien révélée que ce qui est vu pour les éléments RAG. B) Distribution des motifs RAG près du site *dif*. Les éléments RAG sont indiqués par les barres verticales, et comme mentionné en A, on peut voir que le site *dif* est logé à la jonction de deux régions fortement polarisées par ces éléments. Cette figure fut modifiée de Lobry et Louarn, 2003.

### 6.11 Orientation de l'activité de FtsK

La polarité remarquée du chromosome d'*E. coli* n'est pas unique. Il a été remarqué que les séquences entourant les sites *dif* de *H. influenzae* et de *S. typhimurium* sont très différentes de celles d'*E. coli* ; pourtant, le même biais d'éléments polarisés de type RAG y ont été retrouvés (Lobry et Louarn, 2003). D'un autre côté, le domaine C-terminal de la

protéine FtsK d'*E. coli* a pu être remplacé par le domaine correspondant de la protéine FtsK d'*H. influenzae*, sans que la « lecture » de la polarité ne soit modifiée (Bigot *et al.*, 2004).

Cependant, une étude récente faite sur une molécule simple (« single molecule experiment ») a permis de montrer que la protéine FtsK tronquée (ne contenant que le domaine C-terminal) créait des boucles dans l'ADN, peu importe la séquence (Saleh *et al.*, 2004).

Afin d'essayer de comprendre comment la polarité de l'ADN chromosomique pouvait alors influencer l'activité translocase de FtsK, l'équipe de Bigot *et al.* (2004) a ensuite émis plusieurs hypothèses :

1. L'influence des séquences d'ADN pourrait être directe. En effet, certaines séquences orientées sont également retrouvées chez *H. influenzae*, et il s'agirait alors de les déterminer précisément. Cependant, comme mentionné plus haut, des expériences *in vitro* n'ont pas montré d'influence de régions précises, du moins avec la protéine FtsK<sub>50C</sub> (Massey *et al.*, 2004, Saleh *et al.*, 2004).
2. L'influence pourrait être indirecte par l'intermédiaire d'une protéine autre non encore déterminée, laquelle influencerait par la suite FtsK par son action sur les séquences orientées.
3. Et enfin, comme il a déjà été proposé par l'étude de Saleh *et al.* (2004), la translocation de l'ADN par FtsK serait dictée par les tensions créées dans l'ADN, tensions qui forceraient FtsK à tirer l'ADN dans une direction précise.
4. Ou alors, un domaine autre que le domaine C-terminal de la protéine FtsK pourrait être impliqué...

Au moment de la rédaction de cette thèse, aucun autre article n'est venu compléter cette étude afin de déterminer comment la polarité de l'ADN pouvait influencer l'activité de FtsK au niveau de la terminaison de la réplication afin d'assurer la résolution des dimères chromosomiques (automne 2004).

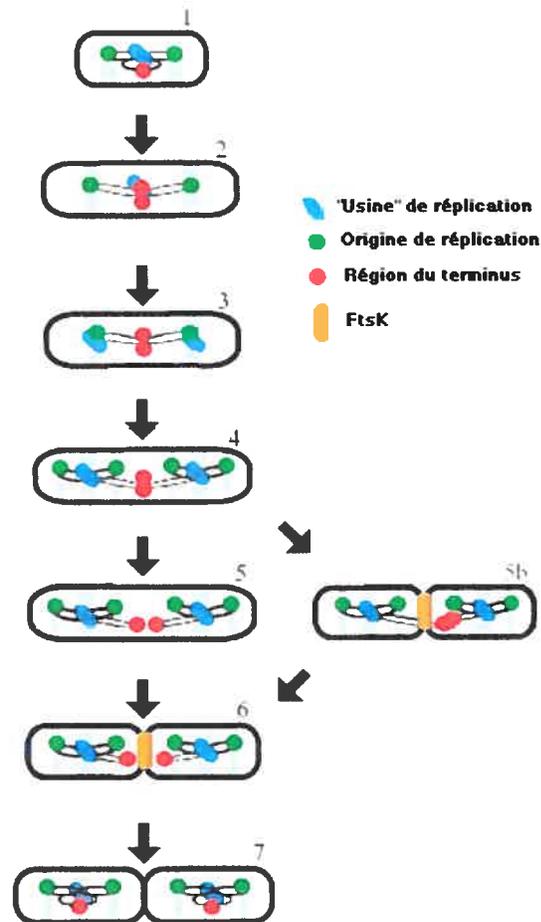
## 6.12 Lien entre la recombinaison et la ségrégation chromosomique

Depuis quelques années, plusieurs études tendent à montrer les liens entre la recombinaison homologue, la réplication, la recombinaison spécifique de site lors de la terminaison de la réplication, la ségrégation chromosomique et la division cellulaire. Tous ces événements, originellement étudiés séparément, ont vu graduellement se tisser une toile

qui les reliait entre eux. Les modèles se sont raffinés au fur et à mesure des études et un schéma plus global a pu être tiré.

On peut maintenant parler de macrodomaines et d'organisation spatiale du chromosome (voir Niki *et al.*, 2000). Selon les études génétiques de Valens *et al.*, (2004), le chromosome serait organisé comme un anneau composé en fait de quatre macrodomaines [deux de ces macrodomaines, Ter et Ori, avaient été précédemment définis par FISH (« fluorescence *in situ* hybridization »; Niki *et al.*, 2000)] et de deux régions moins bien structurées flanquant le macrodomaine Ori. Le chromosome peut alors être vu comme une subdivision de larges régions dans lesquelles les interactions d'ADN sont favorisées (Valens *et al.*, 2004).

Dans *E. coli*, le positionnement précis de la machinerie impliquée dans la réplication et la recombinaison a pu être déterminé, grâce surtout, à la visualisation de certaines molécules par marquage par fluorescence. Il a été ainsi montré que les origines et terminaisons de réplication ont une position fixe. Au fur et à mesure de la progression du cycle cellulaire, cette position est modifiée, comme on peut le voir à la Figure 13 (Lemon et Grosman, 2001, Sherratt *et al.*, 2001 et Lau *et al.*, 2003). Cette figure montre le schéma du cycle cellulaire d'*E. coli* que l'on a pu tirer des différents résultats de marquages et colocalisation.



**Figure 13 : Schéma du cycle cellulaire d'*Escherichia coli* (AB1157) à un temps de génération de 50 min.** Les lignes noires représentent l'ADN et les espaces grisâtres situent l'endroit des futures septation. En 1 et 7, les nouvelles cellules ont en leur centre, leur usine de réplication et la région terminale de réplication et les origines se situent près du quart de la cellule ; les chromosomes sont répliqués à 50%. À peu près à la moitié du cycle cellulaire, la réplication cesse, l'usine de réplication se sépare et de nouvelles usines sont assemblées aux origines (montré en 3). Peu après l'initiation de la réplication à la position 1/4 de la cellule, les origines se séparent et se placent dans ce qui deviendra la position 1/4 des cellules-filles après la division cellulaire (en 4). Les régions terminales de réplication peuvent alors se séparer avant l'invagination (en 5) ou après que l'invagination n'ait débuté (5b puis vers 6) pour ensuite atteindre la position centrale des nouvelles cellules-filles (en 7). Ces étapes ont été démontrées par des prises de photos séquentielles des cellules contenant diverses protéines (impliquées dans le processus de ségrégation chromosomique) marquées, en train de se diviser. Les protéines ou loci marqués par différents dérivés de la GFP (Green Fluorescent protein) étaient : RarA (pour représenter l'usine de réplication ; il avait été montré qu'elle se co-localisait avec SeqA), l'origine de réplication (une insertion de plusieurs opérateurs *lacO* en tandem dans la région de l'origine était détectée par un LacI marqué), le terminus (une insertion de plusieurs opérateurs *tetO* en tandem dans la région du terminus était détectée par un TetR marqué) et enfin, FtsZ. Cette figure fut adaptée de Sherratt *et al.* (2004) à partir des résultats présentés dans Lau *et al.* (2003).

Les études de Lau *et al.* (2003) ont permis de visualiser les régions du terminus et de l'origine de réplication grâce à un système ingénieux permettant de contourner le problème de l'instabilité de ces régions. Par le système de recombinaison homologue médiée par  $\lambda$ Red (Datsenko et Wanner, 2000), plusieurs opérateurs de l'opéron *lac* (*lacO*) en tandem ont pu être insérés dans la région du terminus de réplication, afin de détecter cette dernière par le LacI exprimé sous un promoteur inductible à l'arabinose et fusionné à la protéine fluorescente cyan (ECFP). De la même façon, plusieurs opérateurs de l'opéron *tet* (*tetO*) en tandem ont pu être insérés dans la région de l'origine de réplication, afin d'être détectés par TetR exprimé sous le même type de promoteur, mais marqué cette fois par la protéine fluorescente jaune (EYFP). Les différents résultats obtenus de cette étude sont résumés à la Figure 13. Une caractéristique de l'étude fut l'obtention du terminus de réplication logé de façon asymétrique d'un côté du futur site de division cellulaire, que ce soit dans des cellules jeunes ayant deux *ori* ou dans des cellules plus âgées ayant 4 *ori*. Pour une portion de cellules, un focus simple du terminus était retrouvé d'un côté du septum en train de s'invaginer, se divisait ensuite en deux pour que l'un d'eux passe au travers du septum et se rende dans l'autre cellule-fille. Dans l'autre moitié de cellules, le focus du terminus s'est divisé en deux avant que l'invagination ne s'enclenche. Cette asymétrie n'a pas été observée pour les foci de l'origine de réplication.

Un autre article a montré la ségrégation du terminus de réplication chez *E. coli* : cette fois, ce fut par l'insertion du site *parS* à plusieurs endroits dans la région terminale, afin que cette dernière puisse être visualisée avec la protéine ParB marquée à la GFP (Li *et al.*, 2003). Le positionnement central de la région terminale n'était pas lié à la recombinaison au site *dif* via la recombinaison XerC, mais c'est plutôt la portion C-terminale de FtsK qui aurait un rôle direct ou indirect à y jouer. Le transfert de l'un des termini nouvellement répliqués pourrait être facilité par l'activité translocase de cette protéine.

### 6.12.1 Modèle d'extrusion-capture

Les résultats de Lau *et al.* (2003) venaient corroborer le modèle d'« extrusion-capture » de Lemon et Grossman (2001, et voir la Figure 13 qui représente le modèle modifié selon Lau *et al.*, 2003). Selon ce modèle, plusieurs espèces bactériennes contrôleraient l'énergie produite par l'usine de réplication durant la polymérisation de

l'ADN pour faciliter le processus de la partition du chromosome nouvellement répliqué. Cette partition des chromosomes circulaires se ferait en plusieurs étapes :

1. les origines de réplication s'éloignent l'une de l'autre vers les pôles opposés de la cellule (et donc du réplisome central) sitôt cette portion répliquée, et sont maintenues d'une quelconque façon aux 1/4 de la cellule.
2. L'ensemble de l'ADN chromosomique situé entre l'origine et le terminus est ensuite répliqué, replié et alors partitionné.
3. Enfin, la réplication se terminant, les chromatides sœurs sont alors décaténées, les dimères résolus en monomères (si nécessaire), et les deux régions terminales se placent de part et d'autre du septum de division cellulaire. Peu de temps après la division cellulaire, le terminus se replace au centre de la cellule.

Selon le modèle, le réplisome serait le moteur permettant de tirer le chromosome au centre de la cellule à travers l'usine de réplication, pour ensuite exécuter la réplication et éloigner les produits l'un de l'autre en maintenant les origines aux deux extrémités de la cellule. L'ADN polymérase ne serait donc pas une locomotive roulant sur l'ADN (Lemon et Grossman, 2001).

Ainsi la position de RarA, protéine stimulant les ATPases de la classe des AAA (adénosines tri-phosphatases-ATPases- associées à diverses activités ; classe dans laquelle fait partie FtsK), a montré qu'elle était associée au réplisome ; elle est co-localisée avec la protéine SeqA (Hiraga *et al.*, 2000 et Lau *et al.*, 2003).

Une étude plus récente a d'ailleurs précisé le positionnement de SeqA (Molina et Skarstad, 2004). SeqA est une protéine qui se lie aux sites hémi-méthylés GATC de la région d'*oriC*, séquestrant ainsi les sites de liaison de DnaA et empêchant toute nouvelle initiation de réplication (Messer, 2002). Cette protéine serait présente sur plusieurs foci montrant plusieurs fourches de réplication. Le modèle de l'usine de réplication se voit alors modifié afin de pouvoir intégrer 4, 8, ou 12 usines de réplication (Molina et Skarstad, 2004).

La protéine RarA a également été covisualisée avec la protéine RecQ (Sherratt *et al.* 2004). RecQ est une hélicase 3'→5' impliquée dans la réparation des fourches de réplication bloquées ou brisées. Cette colocalisation de RarA (associée à la fourche de réplication d'après le co-marquage avec SeqA) et de RecQ (représentant de l'activité de

recombinaison homologue), montre à quel point la réplication et la recombinaison homologue sont intimement liées. Les protéines associées à la recombinaison homologue doivent pouvoir rapidement agir afin de répondre à tout problème que la fourche de réplication pourrait rencontrer (Sherratt *et al.*, 2004).

Le processus de la ségrégation des chromosomes nouvellement dupliqués implique en fait plusieurs protéines de différentes fonctions (structurale, de partition, pour le surenroulement, etc.) pour aider à la condensation et à toute l'organisation des nouveaux chromosomes une fois qu'ils ont été bien mis en place (revue dans Sherratt, 2003). De leur côté, les événements de recombinaison entre des régions homologues du chromosome peuvent mener à la dimérisation des chromosomes. C'est à ce moment que le système de recombinaison spécifique de site Xer entre en action : il permet alors d'assurer la monomérisation des chromosomes afin qu'ils soient ségrégués adéquatement aux cellules-filles.

De par l'importance de son action, on pouvait donc s'attendre à ce que ce système soit ubiquitaire, du moins pour toutes les espèces bactériennes ayant un chromosome circulaire.

## 7 *Proteus mirabilis*

### 7.1 Ubiquité des recombinases de type Xer

L'avancée récente des projets de séquençage de divers génomes bactériens a permis de déterminer que le système Xer se retrouvait dans presque toutes les espèces ayant un chromosome circulaire, que l'espèce soit gram-positive ou gram-négative. Ce système est absent des génomes de *Mycobacterium genitalium* et de *M. pneumoniae* (Himmelreich *et al.*, 1997), probablement parce que le système de recombinaison homologue est absent également (Recchia et Sherratt, 1999), mais un gène similaire est retrouvé dans *Mycoplasma agalactiae* (Glew *et al.*, 2002).

Il suffit de faire une courte recherche des protéines dans Pubmed de NCBI pour se rendre compte de l'ubiquité de ces gènes dans le monde bactérien (<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi>). Le système a même été retrouvé chez des archéobactéries (Recchia et Sherratt, 1999). Le Tableau II présente une liste, loin d'être

exhaustive, de quelques-unes des espèces dans lesquelles les gènes *xerC* et *xerD* ont été retrouvés.

**Tableau II : Quelques exemples d'espèces bactériennes possédant les gènes *xer*.**

Espèce	Gène	Numéro d'accèsion	Taille (en aa <sup>1</sup> )	Référence <sup>2</sup>
<i>Bacillus subtilis</i>	<i>codV</i> (≡ <i>xerC</i> )	U13634	304	Slack <i>et al.</i> 1995
	<i>ripX</i> (≡ <i>xerD</i> )	Z99116	296	Sciochetti <i>et al.</i> 1999
<i>Caulobacter crescentus</i>	<i>xerC</i>	AE005708	304	Jouan et Szatmari, 2003
	<i>xerD</i>	AE005964	305	
<i>Escherichia coli</i>	<i>xerC</i>	M38257	298	Colloms <i>et al.</i> 1990
	<i>xerD</i>	M54884	298	Blakely <i>et al.</i> 1993
<i>Haemophilus influenzae</i>	<i>xerC</i>	U32750	295	Neilson <i>et al.</i> 1999
	<i>xerD</i>	U32716	297	
<i>Helicobacter pylori</i>	<i>xerC</i>	AE000580	362	
	<i>xerD</i>	AE000608	355	
<i>Lactobacillus leishmannii</i>	<i>xerC</i>	X84261	295	Becker et Brendel, 1996
<i>Methanobacterium thermoautotrophicum</i>	<i>xerC</i>	AE000666	311	
<i>Mycobacterium leprae</i>	<i>xerC</i>	Z97369	302	
	<i>xerD</i>	U00021	316	
<i>Mycobacterium tuberculosis</i>	<i>xerC</i>	BX842581	315	
	<i>xerD</i>	BX842577	311	
<i>Pseudomonas aeruginosa</i>	<i>sss</i> (≡ <i>xerC</i> )	X78478	302	Höfte <i>et al.</i> 1994
<i>P. fluorescens</i>	<i>sss</i> (≡ <i>xerC</i> )	Y12268	299	Dekkers <i>et al.</i> 1998
<i>Proteus mirabilis</i>	<i>xerC</i>	AF033498	307	Villion et Szatmari, 2003
	<i>xerD</i>	AF033497	313	
<i>Serratia marcescens</i>	<i>xerC</i>	AF028736	303	Guynn <i>et al.</i> 1998
<i>Salmonella typhimurium</i>	<i>xerC</i>	U92524	300	Hayes <i>et al.</i> 1997
	<i>xerD</i>	U92525	298	
<i>Staphylococcus aureus</i>	<i>xerD</i>	AF093548	295	Chalker <i>et al.</i> 2000
	<i>xerC</i>	AF173870		
	<i>xerC</i>	AE007417	356	
<i>Streptococcus pneumoniae</i>	<i>xerD</i>	AJN277766	244	Reichmann et Hakenbeck, 2002
<i>Vibrio cholerae</i>	<i>xerC</i>	AE004103	311	Huber et Waldor, 2002
	<i>xerD</i>	AE004311	302	

≡: signifie "correspond à".

Les numéros d'accèsion sont ceux de la séquence nucléotidique

1: La taille est donnée en nombre d'acides aminés

2: Ces références montrent la première étude concernant les gènes *xer* mentionnés

### 7.1.1 D'autres fonctions pour Xer ?

Certaines de ces recombinases de type Xer ont montré d'autres fonctions que la fonction de recombinase en tant que telle. Ainsi, chez *Pseudomonas aeruginosa*, la protéine Sss, homologue de XerC, joue un rôle dans la production de la pyoverdine (Höfte

*et al.*, 1994) et chez *P. fluorescens*, cette même protéine est requise pour la colonisation de l'espèce au niveau des racines de pommes de terre (Dekkers *et al.*, 1998).

Chez *S. aureus*, une étude a laissé entendre que le gène *xerD* serait un gène essentiel chez cette espèce, tandis que le gène *xerC* serait impliqué dans la virulence, puisque la souche mutante pour ce gène a montré une aptitude réduite à établir l'infection dans des modèles de souris, malgré une croissance normale *in vitro* (Chalker *et al.*, 2000). Chez *Streptococcus pneumoniae*, tout comme chez plusieurs autres espèces de streptocoques, le gène *xerD* est retrouvé, mais possède des caractéristiques particulières. En effet, les résidus conservés de la poche catalytique ne sont pas les mêmes. Cependant, les études de fonctionnalité restent à faire (Reichmann et Hakenbeck, 2002). Par comparaison de séquences, il a été défini que le gène *xerC* est également présent dans une vingtaine de souches de *S. pneumoniae*, tandis que deux autres gènes similaires à *xerC* seraient d'origine phagique ou de transposon ; ces derniers ne sont pas retrouvés dans toutes les souches de cette espèce (Hakenbeck *et al.*, 2001). Le gène *xerC* (auparavant dit *xerI*), serait également impliqué dans la virulence de l'espèce (Chalker *et al.*, 2000).

Chez *Vibrio cholerae*, les recombinases Xer sont impliquées dans l'intégration du bactériophage filamenteux CTX $\phi$  encodant la toxine cholérique (Waldor et Mekalanos, 1996) dans le chromosome I, le plus large des deux chromosomes de l'espèce. De plus, cette intégration se fait dans un site qui chevauche le site *dif* de ce chromosome (Huber et Waldor, 2002). Cependant, contrairement à la recombinaison au site *dif* qui nécessite le domaine C-terminal de FtsK et la localisation du site dans la région terminale de réplication, l'intégration de CTX $\phi$  en son site *attB* ne nécessite ni l'un, ni l'autre (Huber et Waldor, 2002).

## **7.2 *Proteus mirabilis***

Dans notre laboratoire, nous avons identifié et séquencé les gènes encodant les recombinases de *Proteus mirabilis*. Cette espèce est un pathogène opportuniste qui peut causer des infections urinaires en particulier chez les individus ayant des malformations du tractus urinaire ou encore chez ceux à qui un cathéter urinaire a été mis en place pour une longue durée. Dans ces derniers cas, des complications peuvent survenir suite à la formation d'un biofilm sur le cathéter : le biofilm peut éventuellement cristalliser et

boucher le cathéter, ce qui entraîne une cascade de réactions pouvant mener à une pyélonéphrite (Rózalski *et al.*, 1997).

Une caractéristique importante de l'espèce est sa capacité d'essaimage sur un milieu semi solide. Cette caractéristique étant la plus étudiée chez cette espèce, le prochain chapitre traitera brièvement de ce mouvement bactérien permettant à l'espèce de coloniser le tractus urinaire.

### **7.2.1 La mobilité bactérienne de surface**

Plusieurs espèces bactériennes peuvent coloniser différents milieux en étant mobiles. En 1972, Henrichsen classait les différentes formes de mobilité bactérienne en 6 catégories : la nage (swimming), l'essaimage (swarming), le mouvement saccadé (twitching), le flottement (gliding), le glissement-étalement (sliding) et l'élan (darting). Tandis que les deux premiers mouvements sont flagellaires, le mouvement par saccades nécessite des pili de type IV, tandis que le glissement-étalement est une forme passive de translocation. La cause du dernier type de mouvement n'est pas encore connue.

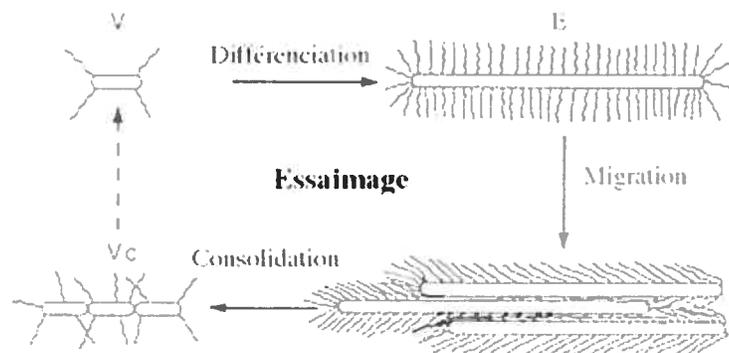
La nage, comme son nom l'indique, est un mouvement en milieu liquide. Au contraire de la nage qui est essentiellement un mouvement individuel, l'essaimage est un mouvement de groupe, sur une surface solide qui nécessite un plus grand nombre de flagelles. La mobilité par saccades est caractérisée par des secousses par intervalles, des sursauts dans le mouvement. Tout comme l'essaimage, la mobilité par saccades est un mouvement de groupe, bien que des cellules individuelles puissent l'exécuter. Le flottement est un mouvement doux, un peu comme le mouvement d'un vol plané. Enfin, le glissement ou étalement est un mouvement passif par lequel des ensembles de cellules allant dans différentes orientations d'étalement comme un tout (pour une description de chacun de ces mouvements, lire Harshey, 2003). Dans tous les cas, le niveau d'humidité est important pour la mobilité. Par exemple, pour pouvoir se mouvoir, les cellules d'essaimage excrètent des polysaccharides pour avoir un niveau d'humidité adéquat (cela forme une substance visqueuse observée à la suite du mouvement de migration).

### **7.2.2 L'essaimage chez *P. mirabilis***

La mobilité par essaimage, mieux connu sous son nom anglais de « swarming », est un processus multicellulaire qui permet à certaines espèces bactériennes de se mouvoir sur

une surface solide (Fraser *et al.*, 2000). L'essaimage a été décrit pour la première fois en 1885 par Hauser et c'est de cette observation qu'il a donné le nom de *Proteus* au genre bactérien capable d'une telle différenciation : ce nom appartenait au Dieu grec (Protée en français), fils de Poséidon, qui avait la capacité de se métamorphoser lorsqu'il désirait éviter de répondre aux questions qu'on lui posait.... Ce mouvement, dépendant des flagelles, a été retrouvé, également chez les genres *Vibrio*, *Bacillus*, *Clostridium*, *Escherichia*, *Salmonella*, etc.

L'essaimage est caractérisé par la production de zones concentriques (terrasses) provenant de la migration de l'ensemble de la population cellulaire entrecoupées par la croissance en bordure de la colonie, sans migration cette fois.



**Figure 14: Schéma des différentes étapes du cycle d'essaimage chez *P. mirabilis*.** V et Vc : formes végétatives de la cellule et E : cellule d'essaimage hyperflagellée (« swarmer »). Les détails de l'essaimage sont donnés dans le texte. Cette figure fut modifiée à partir de Fraser *et al.*, 2000).

L'essaimage peut se diviser en trois étapes apparentes (Rauprich *et al.*, 1996 et Fraser *et al.*, 2000 ; chacune des étapes est schématisée en Figure 14):

1. Phase de différenciation : après une phase de latence plus ou moins longue pendant laquelle des cellules végétatives finissent par former une colonie sur le milieu sur lequel elles viennent d'être inoculées, les cellules se trouvant en bordure de cette colonie se différencient en de longues cellules d'essaimage polyploïdes et hyperflagellées (l'ARNm de l'opéron flagellaire est exprimé jusqu'à 30 fois plus dans les cellules d'essaimage que dans les cellules végétatives). En fait, de cellules nageuses mesurant de 1,5 à 2  $\mu\text{m}$  en longueur et possédant 6 à 10 flagelles péritriches, elles deviennent des cellules

d'essaimage ; elles s'allongent jusqu'à 60 à 80  $\mu\text{m}$  et possèdent près de 50 fois plus de flagelles.

2. Phase de migration : les cellules d'essaimage se rassemblent alors en formant des genres de radeaux (« rafts ») lesquels migrent en s'éloignent de la colonie. Durant tout le temps de cette phase, il y a un apport constant de nouvelles cellules différenciées qui suivent le flot du front de migration.
3. Phase de consolidation : les cellules finissent par cesser de se mouvoir et il y a consolidation. Les cellules se divisent alors et redeviennent végétatives, dans leur forme courte. Tant que la seconde phase de migration n'est pas enclenchée, la première vague de migration continue de s'étendre et il y a croissance et division de cellules différenciées aussi bien que de non-différenciées. La première vague de consolidation ne cesse que par la suite pour former la bordure extrême de la première « terrasse ».

La durée totale de la période incluant la migration et la consolidation a été déterminée et durerait environ 4 h sur plusieurs types de milieux. Cependant, la taille et la durée de la migration seraient influencées par la concentration en agar du milieu (pour un modèle mathématique de l'essaimage, voir Czirók *et al.*, 2001).

L'initiation de la différenciation en cellules d'essaimage est induite lorsque la bactérie entre en contact avec un milieu solide, ou sur un milieu à viscosité élevée. Ce changement environnemental serait détecté par les flagelles. D'autres stimuli physiologiques et environnementaux induisent ce changement morphologique. Cependant, ce changement morphologique n'est pas suffisant pour engendrer la migration. Il faut encore que les cellules se réunissent en radeaux (rafts) et la densité cellulaire en serait le facteur principal. Le contact entre les cellules serait stabilisé par la formation d'une couche visqueuse formée par un polysaccharide capsulaire acide. Cette couche permettrait aussi la création d'un microenvironnement liquide permettant à la bactérie de se mouvoir en essaim.

Suite à des mutations aléatoires ayant pour but de détecter les gènes impliqués dans le phénomène d'essaimage, 5 classes phénotypiques ont été déterminées (Fraser *et al.*, 2000):

1. Il existe des mutants non mobiles et non formateurs d'essaims (nonswarming). Ces mutants ne peuvent alors ni nager dans un milieu liquide, ni progresser sur la surface d'un milieu solide.
2. D'autres mutants sont mobiles, mais non formateurs d'essaims. Ils ont donc conservé leur mobilité en milieu liquide, mais ont perdu leur capacité de migration sur milieu solide.
3. D'autres mutants, ayant conservé leur capacité de migration, présentent cependant un patron de migration en dendrites (dendritic swarming), mais la fréquence des terrasses formées reste la même que dans la souche sauvage.
4. D'autres encore présentent un nombre de terrasses (consolidation) différents : soit moins fréquemment que la souche sauvage (l'apparition des vagues est très rare),
5. soit plus fréquemment que la souche sauvage.

Il est à noter, que bien que la conversion des cellules dites végétatives (cellules mobiles courtes ayant peu de flagelles péritriches) en cellules d'essaimage ne soit pas un processus nécessaire à la survie de l'espèce, il reste lié à l'environnement.

### 7.2.3 L'essaimage et la virulence de *P. mirabilis*

Lors de sa colonisation du tractus urinaire, *P. mirabilis* exprime plusieurs facteurs de virulence tels que l'hémolysine, les flagelles, des fimbriae et des protéases. Tandis que les fimbriae (MR/P) servent à l'adhérence de *P. mirabilis* aux cellules épithéliales, les flagelles permettent la migration de l'uretère aux reins. La production d'uréase contribue de son côté à augmenter le pH de l'urine, ce qui favoriserait la formation de calculs dans la vessie (une revue des différents facteurs de virulence est présentée par Rózalski *et al.*, 1997).

Le processus de différenciation cellulaire est lié à la virulence de *P. mirabilis* par le fait que plusieurs facteurs de virulence sont plus exprimés dans les cellules d'essaimage que dans les cellules dites végétatives. Tel est le cas de l'uréase, de la métalloprotéase ZapA et l'hémolysine (Rather, 2005). La formation du biofilm serait également dépendante de la différenciation cellulaire.

L'expression de ces facteurs de virulence et la différenciation cellulaire sont toutes deux régulées négativement par les protéines RsbA et RsmA. RsmA est un régulateur

global servant à contrôler la stabilité des ARNm ; l'expression des facteurs de virulence serait inhibée via la dégradation de leurs ARNm (Liaw *et al.*, 2003). D'un autre côté, il a été montré que la protéine RsbA était impliquée dans la formation de biofilm par *P. mirabilis* et dans la production d'exopolysaccharides (EPS). Cependant, pour des raisons encore inconnues, alors que RsbA agit comme un répresseur dans le premier cas, elle serait un activateur dans le second (Liaw *et al.*, 2004). Par l'intermédiaire d'une voie dépendante de RsbA, l'essaimage et l'expression de l'hémolysine seraient aussi régulés par certains acides gras ; ces acides pourraient alors agir comme signaux extracellulaires (Liaw *et al.*, 2004).

## 8 Projet

Comme nous l'avons vu précédemment, le système Xer est retrouvé dans une panoplie d'espèces bactériennes. Notre projet s'est donc intéressé à l'une d'entre-elles, *Proteus mirabilis*. Nous avons d'ailleurs commencé à étudier la recombinase XerD lors d'une étude précédente (Villion et Szatmari, 1998).

Par ce projet, nous avons voulu poursuivre la caractérisation des recombinases Xer de *P. mirabilis*. En effet, il avait été montré que le plasmide ColE1 formait des multimères dans cette espèce (Goebel et Helinski, 1968). Des multimères d'un vecteur dérivé de pUC contenant les sites *cer* et *nmr* (de pNTP16) flanquant le gène de résistance à la kanamycine ont également été observés dans *P. mirabilis* (Sirois et Szatmari, 1995). Nous étions donc intéressés à savoir si le système Xer que nous avons découvert était réellement fonctionnel.

Dans une première partie, nous montrons la reconnaissance puis le clivage de sites spécifiques par ces protéines (les sites *dif* d'*E. coli* et *cer* de ColE1), par des réactions *in vitro*. Des tests de complémentation ont ensuite été réalisés dans une souche d'*E. coli* *xerC*<sup>-</sup> *xerD*<sup>-</sup> afin de vérifier la fonctionnalité des protéines de *P. mirabilis*, *in vivo*. Cependant, alors que les deux protéines peuvent agir ensemble pour résoudre et maintenir un vecteur contenant deux sites *cer*, la collaboration entre la protéine XerC de *P. mirabilis* et XerD d'*E. coli* n'a pas été possible. Ces premiers résultats ont pu être publiés dans FEMS Microbiology Letters en septembre 2003 (Villion et Szatmari, 2003).

Afin de poursuivre cette étude et de tenter de comprendre ce qui a pu se produire lors de ces tests de complémentation, nous avons construit des souches mutantes de *P. mirabilis* pour ces gènes. Les expériences réalisées avec ces souches constituent en grande

partie la seconde moitié des résultats. Des tests de complémentation *in vivo* ont de nouveau été tentés, mais le plasmide s'est avéré trop instable dans ces souches, pour que nous puissions poursuivre. Nous avons ensuite brièvement étudié l'expression des promoteurs potentiels de XerC et de XerD. En effet, mis à part une étude faite sur l'expression du gène *xerC* d'*E. coli* (Aller et Szatmari, non publié), aucune recherche ne semble avoir été faite sur les niveaux d'expression de ces gènes ou sur les promoteurs potentiels. En parallèle, nous avons remarqué que le phénotype engendré par les mutations Xer sur la capacité de migration par essaimage de *P. mirabilis* avait été légèrement modifié. La raison exacte demeure cependant inconnue et différentes possibilités restent à explorer.

Finalement, nous avons pu identifier le site *dif* de *P. mirabilis* sur lequel les recombinaisons Xer ont pu se lier. Une brève analyse en est présentée.

# **MATÉRIEL ET MÉTHODES**

## Chapitre II : Matériel et méthodes

### 1 Souches bactériennes et plasmides employés

Les souches d'*E. coli* employées au cours de cette étude sont indiquées dans le Tableau III. La construction des souches de *P. mirabilis* est indiquée plus loin dans le texte ; la souche de base employée est un isolat clinique qui avait été donné au département de microbiologie et immunologie de l'Université de Montréal (UM240-82, et voir Sirois et Szatmari, 1995).

Tous les plasmides sont indiqués dans le Tableau IV. La construction de certains de ces plasmides est détaillée plus loin dans le texte.

#### 1.1 Milieux de croissance

Les diverses souches (*E. coli* et *P. mirabilis*) employées dans cette étude ont été mises en culture dans du LB [bouillon Luria-Bertani (par litre : 10 g tryptone, 10 g d'extraits de levures (« yeast extract ») et 5 g NaCl] avec ou sans antibiotiques, selon le cas.

*P. mirabilis* est reconnue pour son mouvement d'essaimage mieux connu sous son terme anglais de « swarming », lorsqu'elle est étalée sur un milieu LA ordinaire (LB + 15 g/l d'agar). Ainsi, pour obtenir des colonies isolées, les cultures de cette espèce ont été étalées sur le milieu LSW [Low Swarming Agar (par litre): 10 g tryptone, 5 g d'extraits de levures, 5 ml glycérol, 0,4 g NaCl et 20 g d'agar) (Belas *et al.*, 1991). Pour *E. coli*, l'obtention de colonies isolées s'est faite sur LA, sauf lorsque indiqué.

Les antibiotiques ont été employés aux concentrations suivantes : Ap (100 µg/ml dans tous les cas), Km (50 µg/ml), Tc (6,25 µg/ml pour les cultures qui contenaient le pCS210 et 12,5 µg/ml pour les cultures de *P. mirabilis*), Sm (100 µg/ml) et Cm (34 µg/ml pour *E. coli* et 100 µg/ml pour *P. mirabilis*).

**Tableau III: Souches bactériennes**

Souches	Description	Références
<b><i>E. coli</i></b>		
DS941	AB1157 <i>recF lac<sup>R</sup> lacZΔM15</i>	Colloms <i>et al.</i> , 1990
DS9029	DS941 <i>xerD::Tp<sup>R</sup> xerC::miniMu PR13</i>	Colloms <i>et al.</i> , 1996
DS955 <i>recA<sup>-</sup></i>	DS941 <i>argR::Tp<sup>R</sup> pepA::Tn5 Δ(srl-recA)306::Tn10</i>	G. Szatmari
BL21 pREP4	F <sup>-</sup> <i>ompT hsdS<sub>B</sub> (r<sub>B</sub>m<sub>B</sub>) gal dcm</i> contenant le vecteur pREP4 possédant <i>lacI<sup>q</sup></i>	Studier et Moffatt, 1986
HMS174(DE3)	F <sup>-</sup> <i>recA hsdR(r<sub>K12</sub>m<sub>K12</sub><sup>+</sup>) Rif<sup>R</sup> (DE3)</i>	Campbell <i>et al.</i> , 1978
DH5α	F <sup>-</sup> <i>Φ80ΔlacZΔM15 Δ(lacZYA-argF) U169 endA1 recA1 hsdR17(r<sub>K12</sub>m<sub>K12</sub><sup>+</sup>) deo<sup>R</sup> thi-1 supE44 λ- gyrA96 relA1</i>	Grant <i>et al.</i> 1990
SM10λ <i>pir</i>	λ <i>pir thi thr leu tonA lacY supE recA::RP4-2-Tc::Mu Km<sup>R</sup> NaI<sup>R</sup></i>	Miller et Mekalanos, 1988
CC118λ <i>pir</i>	Δ( <i>ara-leu</i> ) <i>araD ΔlacX74 galE phoA20 thi-1 rpsE rpoB argE (am) recA1 λpir</i>	Herrero <i>et al.</i> , 1990
S17.1λ <i>pir</i>	λ <i>pir recA thi pro hsdR Mu+ RP4: 2-Tc::Mu Km::Tn7 Tp<sup>R</sup> Sm<sup>R</sup></i>	Simon <i>et al.</i> 1983 et James <i>et al.</i> , 1995
BW25113	Δ( <i>araD-araB</i> )567, Δ <i>lacZ4787(::rrnB-3)</i> , <i>lacIp-4000(lacI<sup>q</sup>)</i> , λ <sup>-</sup> , <i>rph-1</i> , Δ( <i>rhaD-rhaB</i> )568, <i>hsdR514</i>	Datsenko et Wanner, 2000
BW25141	Δ( <i>araD-araB</i> )567, Δ <i>lacZ4787(::rrnB-3)</i> , <i>lacIp-4000(lacI<sup>q</sup>)</i> , Δ( <i>phoB-phoR</i> )580, λ <sup>-</sup> , <i>galU95</i> , Δ <i>uidA3::pir<sup>+</sup></i> , <i>recA1</i> , <i>endA9(del-ins)::FRT</i> , <i>rph-1</i> , Δ( <i>rhaD-rhaB</i> )568, <i>rrnB-3</i> , <i>hsdR514</i>	Datsenko et Wanner 2000
BWΔ <i>dif</i>	BW25113 <i>dif(del-ins) ::FRT</i>	Cette étude
<b><i>Proteus mirabilis</i></b>		
Pm 240-82	Souche sauvage provenant d'un isolat clinique	Université de Montréal
PmC8	Pm 240-82 <i>xerC::lacZYA-Cm<sup>R</sup></i>	Cette étude
PmD1	Pm 240-82 <i>xerD::lacZYA-Cm<sup>R</sup></i>	Cette étude
Pm2DC <i>lacZ</i>	Pm 240-82 <i>xerD::Cm<sup>R</sup> xerC::lacZYA-Km<sup>R</sup></i>	Cette étude

Note: Toute les souches d'*E. coli*, à l'exception de la BL21, sont dérivées d'*E. coli* K-12.

Tableau IV: Plasmides

Plasmides	Description	Références
pAL101	pKNG101 Str <sup>S</sup> Cm <sup>R</sup> (Cm <sup>R</sup> provient de pCAT19)	F. Aller
pCAT19	pUC19 modifié contenant le gène <i>cat</i> de Tn9 extrait de pBR325	Fuqua, 1992
pCP20	Plasmide Cm <sup>R</sup> Ap <sup>R</sup> <i>ts-rep</i> permettant la synthèse de FLP par induction thermique avec [ <i>cl857</i> ](lambda)( <i>ts</i> )	Cherepanov et Wackernagel, 1995
pCS126	Vecteur permettant la surexpression de <i>pepA</i> à partir d'un promoteur <i>tac</i>	Stirling <i>et al.</i> , 1989
pCS210	Dérivé de pACYC184 ayant le gène <i>lacZ</i> entre deux sites <i>cer</i>	Stirling <i>et al.</i> , 1989
pCS210Km <sup>R</sup>	pCS210 dont la résistance à la Tc a été remplacée par la Km <sup>R</sup>	Cette étude
pCS211(Km <sup>R</sup> )	Produit de la résolution de pCS210 ou de pCS210Km (avec un site <i>cer</i> )	
pDB169	Vecteur permettant la surexpression d' <i>argR</i> à partir d'un promoteur <i>tac</i>	Lim <i>et al.</i> , 1987
pDEcK	pKNG101 contenant le site <i>dif</i> d' <i>E. coli</i> (33pb)	Cette étude
pEcCD	pRM130 contenant le gène <i>xerC</i> d' <i>E. coli</i> provenant de pSDC105	Cette étude
pFUSE	Vecteur suicide pour fusion d'opéron avec <i>lacZ</i> : <i>oriR6K mobRP4 cat lacZYA</i>	Bäumler <i>et al.</i> , 1996
pGB500	pMAL-c2 contenant <i>xerC</i> d' <i>E. coli</i>	Arciszewska et Sherratt, 1995
pHP45ΩKm	pHP45:Ω dont l'interposon original a été remplacé par un interposon Ω-Km (Km <sup>R</sup> provient de Tn5)	Fellay <i>et al.</i> , 1987
pKD3	Plasmide matrice <i>ori R6K Ap<sup>R</sup></i> contenant Cm <sup>R</sup> flanqué de deux sites <i>FRT</i>	Wanner et Datsenko, 2000
pKD46	Vecteur "helper" permettant l'expression de la recombinase Red sous promoteur inductible au galactose: <i>ori R101 repA101(ts) araBp gam bet exo</i>	Wanner et Datsenko, 2000
pKΔSacKm <sup>R</sup>	pKNG101 dont le fragment <i>XmnI-FspI</i> contenant <i>sacBR</i> a été délété + ajout de la Km <sup>R</sup> dans le site <i>SbfI</i>	Cette étude
pKNG101	Vecteur suicide: <i>strAB ori R6K sacBR mob</i>	Kaniga <i>et al.</i> , 1991
pKNG-Km	pKNG101 avec insertion de Km <sup>R</sup> dans le site <i>BamHI</i>	J. Harel
pKS492	Fragment <i>HpaII-TaqI</i> de ColE1 (280 pb correspondant au site <i>cer</i> ) dans pUC-18	Stirling <i>et al.</i> , 1988
pMIN33	Vecteur pUC18 contenant un seul site <i>dif</i> (33 pb)	Blakely <i>et al.</i> , 1991
pMV17C	<i>xerCPm</i> dans pQE-31 (QIAGEN)	Cette étude
pMVC6	<i>xerCPm</i> dans pTZ-PC	Cette étude
pMVC <sub>mbp</sub>	pMAL-c2 (NEB) ayant le gène <i>xerC</i> de <i>P. mirabilis</i>	Cette étude
pMVD37	<i>xerDPm</i> dans pTZ-PC	Villion et Szatmari, 1998
pMVD72	pET-14b (Novagen) contenant le gène <i>xerD</i> de <i>P. mirabilis</i>	Villion et Szatmari, 1998
pPmCD	pUC19 avec les gènes <i>xerC</i> et <i>xerD</i> de <i>P. mirabilis</i>	Cette étude
pPmCEcD	pRM130 contenant le gène <i>xerC</i> de <i>P. mirabilis</i>	Cette étude
pPmDEcC	pUC19 contenant les gènes <i>xerC</i> d' <i>E. coli</i> de pSCD105 et <i>xerD</i> de <i>P. mirabilis</i>	Cette étude

### Suite de Tableau IV: Plasmides

pRM130	Vecteur pMB1 (Ap <sup>R</sup> ) surexprimant <i>xerD</i> d' <i>E. coli</i>	Blakely <i>et al.</i> , 1993
pSDC105	Vecteur surexprimant le gène <i>xerC</i> d' <i>E. coli</i> à partir d'un promoteur <i>tac</i>	Colloms <i>et al.</i> , 1990
pSDC124	Plasmide rapporteur Ap <sup>R</sup> contenant deux sites <i>dif</i> flanquant Km <sup>R</sup>	Blakely <i>et al.</i> , 1991
pTZ-PC	Dérivé de pTZ-19R contenant deux sites <i>XcmI</i> pour le clonage de produits de PCR générés par la polymérase Taq	D.C. Tessier et D.Y. Thomas, comm. pers.
pUC119-K	pUC19 contenant Km <sup>R</sup> de Tn903	G. Szatmari
pUC19	Vecteur de base	Yanish-Perron <i>et al.</i> , 1985
pUC57	Vecteur dérivé de pUC19 ayant des sites supplémentaires dans le site de multiclonage.	Fermentas
pVi11	site <i>dif</i> d' <i>E. coli</i> (103 pb) dans pKNG101	Cette étude

### Tableau V: Amorces pour les diverses réactions de PCR

Nom	Séquence (5' vers 3')	Site de restriction	Séquence amplifiée ou emploi
A1 C	GNG AYM RNG CNH TNY TNG ARB T RTG NGT RTA DAT YTG NGT		Recherche de la portion C-terminale conservée des gènes <i>xer</i> d'entérobactéries.
PrC5' PmC3'	GAT CGC AAG CAT ATG AGC CA GTG GGA TCC GAT TTT TCT CTC	<i>NdeI</i>	Gène <i>xerCPm</i> entier
PrC2 mirC2	CTT TAC CTA ATA CTC GGA CTT C CTC TTT TGC CAC TCA TTT AC		Amorces servant à rechercher le reste du gène <i>xerCPm</i> par PCR inverse à partir de la séquence C-terminale.
PmirxerCF PmirxerCR	GGA TAA ACT CGC GCA G CCA TCA CAG GAA CAT TGT C		Confirmation de la séquence de <i>xerCPm</i>
PmxD-IntF PmXD-IntR	ACA CGC TTT TCT TGC CAC TC GCA ACA CAT GGG GTG ATA AT		Séquence interne de <i>xerDPm</i>
XDL-R	CAG GGT TTT CCC AGT CAC		Vérification des fusions <i>lacZ</i>
PmC-F PmC-R	CAA GAA GGT ATG GGA ACG CGG ATA AAC GCT AAG GTC		Gène <i>xerCPm</i> entier
PmDNdeI PmD-F PmD-R	GAG TAA ACA TAT GAC CCA GAC C GCC GCC ATT TCG GTA TAA CAG CCC ACT ATC GGT GTC GGT TG	<i>NdeI</i>	Amorce employée avec PmD-R pour obtention de <i>xerDPm</i> à mettre dans pET14B Gène <i>xerDPm</i> entier
PrD MirD	GTT TTC CCT TGC ATA AGG GAT GTT ATT GGG GCA CAG		Amorces servant à rechercher le reste de <i>xerDPm</i> par PCR inverse à partir de la séquence C-terminale
PmXDFseq PmXDRseq	GCC AGT TGC GCT TTA ATA GC GGT CAC TAT TGT ACC AG		Confirmation de la séquence de <i>xerDPm</i>

**Suite de Tableau V: Amorces pour les diverses réactions de PCR**

EccF-blunt	CGT ATG ACC GAT TTA CAC ACC		
EccF-Kpn	GAA GGT ACC ACG CGT ATG AC	<i>KpnI</i>	Amplification de <i>xerCEc</i>
EccR-Pst	GAT CTG CAG AAA GGT GAG CG	<i>PstI</i>	
EcdF-Kpn	TAG GGT ACC GCA AGT GAA ACA G	<i>KpnI</i>	Amplification de <i>xerDEc</i>
EcdR-Pst	GAA CTG CAG ATC AGC CTG AGC	<i>PstI</i>	
PmxC-IntF	TCG CAG TGG ATT ACA ACC TG		Séquence interne de <i>xerCPm</i>
PmxC-IntR	TGT ACT GCT CGC AAA TCA CC		
PmxCN5'	GAT GCA TAA ACA CTC GCA CT	<i>NsiI</i>	Amplification d'un fragment en amont de <i>xerCPm</i> (incluant une portion de la région N-terminale)
PmxCN3'	CGG AAT TCG CCT TCT TGG AG	<i>EcoRI</i>	
PmxCC5'	GAT GCT GCA GAT CCA AGA G	<i>PstI</i>	Amplification d'un fragment en aval de <i>xerCPm</i> (incluant une portion de la région C-terminale)
PmxCC3'	TGC AAA ACT AGT CCC ACT GG	<i>SpeI</i>	
PmxDN5'	GCA TGC ATC CAT TTC GGT ATA ACA	<i>NsiI</i>	Amplification d'un fragment en amont de <i>xerDPm</i> (incluant une portion de la région N-terminale)
PmxDN3'	GCG AAT TCA TAA GGG CTG GTC	<i>EcoRI</i>	
PmxDC5'a	GCC GAC TGC AGG GTG TTA AAT AG	<i>PstI</i>	Amplification d'un fragment en aval de <i>xerDPm</i> (incluant une portion de la région C-terminale)
PmxDC3'	GCC TCA ACT AGT AAA GGG ATA GC	<i>SpeI</i>	
DifF22	CAG AAA AGC ACT TCG CAT CAC		Amplification du site <i>difEc</i> , région de 103 pb
DifR4	CAA TCA TGA CCG CCA ACG AC		
dif-KpnF	GTG GTA CCA TTC CCG CTT GAT C	<i>KpnI</i>	Amplification d'une zone de 350 pb entourant <i>difEc</i>
dif-EcoR	CCA GAA TTC TTC CAG CAG CAC T	<i>EcoRI</i>	
dif31	VHD RBW KYR YAT WAT VTA TVT TAT GTT AAA T		Séquence dégénérée représentant un site <i>dif</i> potentiel
dif31Inv	ATT TAA CAT AAB ATA BAT WAT RYR MWV YHD B		Inverse de l'amorce <i>dif31</i>
difPm5'	GAT TTC AGC GCC TAT CCG TA		Amplification du site <i>difPm</i> , région d'une centaine de pb.
difPm3'	GGC TTG AAG CGC TAA ATC AT		
difEc-totF	AGC GAT CTC AAA CTG GCA AT		Amplification de <i>difEc</i> avec les séquences environnantes
difEc-totR	TAG GCT GCT GTG GTG ACA TC		
difF1-P1	CCG ACA GAT TTT CCG GAG AAC GTG GTG ACG GCA GTT GAA AGC AAT <b>GTG TGT AGG CTG GAG CTG</b> <b>CTT C</b>		Obtention de la séquence linéaire permettant la délétion du site <i>difEc</i> (en gras: portion des amorces permettant l'amplification de Cm <sup>R</sup> de pKD3).
difR3-P2	CGA CAG ACA CTG CGC TCT TAG CAT GGG GCA GGG AAT AAC CAG GCT ATC <b>ATA TGA ATA TCC TCC</b> <b>TTA G</b>		

## 1.2 Enzymes employées :

Toutes les enzymes de restriction employées, ainsi que le fragment Klenow de la polymérase d'*E. coli* et la ligase T4 d'ADN provenaient de New England Biolabs (sauf lorsque indiqué) et furent employées selon les recommandations d'usage. La phosphatase alcaline de crevette (SAP) provenait de Roche Diagnostics, Inc.

Les enzymes de polymérisation en chaîne étaient la polymérase Taq de NEB ou de QIAGEN ou la polymérase Vent<sup>TM</sup> de NEB.

## 1.3 Réactions de PCR :

Toutes les réactions de PCR (réactions de polymérisation en chaîne), ont été réalisées dans des tubes de 200 µl à paroi mince à l'aide d'un thermocycleur Biometra (à l'exception du programme employé pour la recherche de *xerCPm*, voir plus loin). Sauf exceptions, elles ont contenu, en plus de l'ADN matrice, 4 µmol de chacun des dNTPs, 100 pmol de chacune des amorces, 1/5 du volume final en tampon Q (QIAGEN), 1/10 du volume final en tampon de réaction contenant le Mg<sup>++</sup> (de QIAGEN ou de NEB) et 1 U d'enzyme (soit la Taq provenant de QIAGEN ou de NEB, respectivement). Les programmes débutaient tous par un cycle à 95°C/2 min, et se terminaient par un cycle à 72°C/5 min. Les 30 cycles centraux se répartissaient ainsi : 95°C/ 15 s, température d'hybridation (variant selon les amorces employées) / 30 s et 72°C/ temps voulu (selon la longueur de l'échantillon d'ADN à amplifier sur la base de 1 kb/ min). Toutes les séquences des amorces mentionnées dans le reste de cette section sont inscrites dans le Tableau V et ne seront donc pas reprises plus loin.

## 2 Extractions d'ADN et purifications

### 2.1 Méthode de « craquelage (cracking) »

La méthode de craquelage (Sambrook *et al.*, 1989) est une méthode d'extraction rapide de l'ADN qui nous permet de voir si les clones obtenus suite à une transformation d'un produit de ligation dans une souche d'*E. coli* (par ex.) contiennent l'insert désiré. Une fois la méthode de craquelage réalisée, il suffit de mettre en culture le clone voulu pour en extraire et purifier l'ADN plasmidique par une méthode standard.

Pour réaliser le craquelage, il faut tout d'abord étaler sur une boîte de gélose plusieurs des clones obtenus suite à la transformation sur une surface carrée d'environ

1 cm<sup>2</sup>. On laisse ensuite croître les divers clones pendant la nuit à 37°C. Le lendemain, une portion de la surface étalée d'un des clones est prélevée à l'aide d'un cure-dent préalablement stérilisé, puis mise dans un tube Eppendorf contenant 50 µl de la solution de lyse (0,2N NaOH, 0,5% SDS, 20% saccharose, 5 mM EDTA, 0,025% bleu de bromophénol). On répète ainsi pour tous les clones étalés. Tous les tubes sont ensuite centrifugés pendant 15 min à haute vitesse (12 000 g) et environ 20-25 µl du surnageant sont prélevés et migrés sur gel d'agarose.

## **2.2 Minipreps :**

Toutes les extractions d'ADN plasmidique ont été faites avec la trousse de QIAGEN : QIAprep Spin Miniprep kit en suivant les recommandations de la compagnie.

Dans certaines situations, l'ADN plasmidique a été extrait par la méthode de lyse alcaline standard (Sambrook *et al.*, 1989). Ainsi, après avoir centrifugé en tout 3 ml de la culture de nuit d'intérêt dans un tube Eppendorf de 1,5 ml, le culot fut suspendu par Vortex dans 100 µl du tampon de suspension (0,1M Tris-Cl, pH 8,0 + RNase A). La lyse est ensuite réalisée en ajoutant 250 µl de solution de lyse fraîche (0,2 N NaOH + 1% SDS) ; le mélange se fait délicatement par inversion du tube 4 ou 5 fois. Après un maximum de 5 min, il faut arrêter la réaction avec 350 µl de solution de neutralisation (3M acétate de potassium, 11,5% acide acétique v/v) en inversant 4 ou 5 fois, puis en laissant le tube sur glace pendant 5 min. Le tube est ensuite centrifugé à haute vitesse (12 000 g), puis le surnageant est récupéré délicatement en faisant attention de ne prélever aucun débris du culot. Ce surnageant est ensuite traité avec l'ajout d'un volume de phénol-chloroforme-alcool isoamyle (ratio de 25 :24 :1), puis au chloroforme-alcool isoamyle (ratio de 24 :1). Enfin, la précipitation du surnageant est faite par l'ajout de 0,6 volume en isopropanol, et le culot nettoyé avec de l'éthanol à 70%. Le culot final contenant l'ADN plasmidique pur est remis en solution dans 100 µl de tris (10 mM, pH 8,0).

## **2.3 Purification de fragments d'ADN extraits de gels :**

### **2.3.1 Gels d'agarose**

Pour l'ADN provenant de gel d'agarose, le kit d'extraction rapide d'ADN de QIAGEN (QIAquick Gel Extraction kit) a été employé tel qu'indiqué par le manufacturier.

Au tout début des expériences, lors de la recherche de *xerCPm*, le kit alors employé provenait de la compagnie Amersham Pharmacia Biotech (Sephaglas BandPrep kit).

### 2.3.2 Gels de polyacrylamide

Pour l'ADN provenant de gel de polyacrylamide non dénaturant, l'extraction a d'abord débuté par l'incubation pendant 1 h/ 50°C d'un tube Eppendorf contenant la portion découpée du gel et 2 volumes de tampon de diffusion (0,5 M acétate d'ammonium, 10 mM acétate de magnésium, 1 mM EDTA pH 8,0 et 0,1% SDS). Avant la poursuite de l'extraction avec le kit d'extraction rapide d'ADN (QIAGEN) tel qu'indiqué, 6 volumes de tampon QG + 1 volume d'isopropanol ont été ajoutés au surnageant récupéré sans les débris de polyacrylamide.

## 3 Préparation de cellules compétentes et transformations:

Deux types de protocoles ont été suivis lors de cette étude pour rendre les cellules bactériennes compétentes à la transformation : un protocole pour faciliter l'incorporation d'ADN suite à un choc thermique et un autre pour incorporer l'ADN par électroporation.

### 3.1 Chimiocompétence :

Un bouillon LB de 50 ml estensemencé avec 500 µl d'une culture en bouillon ayant poussé toute la nuit. Lorsque la culture atteint une  $DO_{600}$  de  $\approx 0,3$ , les cellules sont mises immédiatement sur glace pendant 5 à 10 min, puis centrifugées à 4°C à 1600g/10 min. Le culot est ensuite lavé deux fois dans du tampon  $CaCl_2$  froid (60 mM  $CaCl_2$ , 10 mM HEPES pH7,0, 15% glycérol). Le culot final est resuspendu dans 1 ml du même tampon, et les cellules sont transférées à coup de 100 µl dans des tubes vissés de 2 ml en plastique et conservées à -70°C jusqu'à leur utilisation.

Pour la transformation, l'ADN à incorporer dans les cellules préparées (plasmide ou produit de ligation) est mis en contact avec 100 µl de cellules chimiocompétentes pendant 15 à 30 min, sur glace. Un choc thermique fut ensuite réalisé à 42°C pendant 40 s, puis les cellules remises sur glace pendant 1 min. 1 ml de LB ou de SOC fut ensuite ajouté, et le tube Eppendorf fut incubé pendant 1 h à 37°C.

## 3.2 Électrocompétence :

### 3.2.1 Pour *E. coli*

Le protocole est similaire à celui employé pour rendre les cellules bactériennes chimiocompétentes, en ce sens que les cellules sont mises en culture dans du LB frais à partir d'une culture de nuit. Cependant, la densité optique atteinte est de  $\approx 0,6$ . De plus, il s'agit ici de laver les cellules bactériennes avec de l'eau MilliQ stérile afin d'enlever toute trace de sels. Ainsi, une fois la densité optique atteinte, les cellules refroidies sur glace et centrifugées, sont lavées trois fois avec de l'eau MilliQ stérile contenant 10% de glycérol. Après la dernière centrifugation, le culot est resuspendu dans la goutte finale de surnageant avant d'être réparti à coup de 50  $\mu\text{l}$  dans des tubes afin d'être conservé à  $-70^\circ\text{C}$  jusqu'à utilisation.

L'électroporation a été ensuite réalisée de façon standard. Si le produit à transformer provenait d'une ligation, il était préalablement précipité au N-butanol et resuspendu dans un volume minimal d'eau déionisée. Si le produit employé était un plasmide préalablement purifié et resuspendu dans de l'eau déionisée, il était employé tel quel.

Dans les deux cas, une fois l'ADN mis en contact avec le 50  $\mu\text{l}$  de cellules électrocompétentes, le tout était transféré dans une cuvette à électroporation de 0,2 cm de largeur gardée au froid. Le choc fut ensuite réalisé à 2,5 kV pendant 4 à 5 ms. Immédiatement après le choc, 800  $\mu\text{l}$  de SOC furent ajoutés dans la cuvette, et ensuite transférés dans un tube Eppendorf de 1,5 ml. Les cellules furent ensuite incubées à  $37^\circ\text{C}$  pendant au moins 1 h, avant d'être étalées sur le milieu sélectif approprié.

### 3.2.2 Pour *P. mirabilis*

Dans le cas de cellules de *P. mirabilis*, l'électroporation a été réalisée après un contact ADN-cellules d'une heure sur glace, tel qu'il avait été suggéré par Fernandes et da Costa (1996). Le reste de la transformation a été réalisé tel qu'indiqué ci-haut. Il est à noter cependant que, bien que ce fut le protocole nous ayant donné les meilleurs résultats, le taux d'efficacité fut très minime comparativement à ce que l'on pouvait obtenir avec des cellules d'*E. coli* (une dizaine de transformants pour la moitié de la transformation de *P. mirabilis* comparativement à un tapis de colonies pour 50  $\mu\text{l}$  étalés d'*E. coli*).

Bien entendu, dans le cas de la transformation d'un vecteur thermosensible, l'incubation a été de 2 à 3 h à 28°C au lieu de 37°C, et ce, pour les deux types de transformations.

## **4 Sondes d'ADN (pour hybridation ou pour retards sur gel) :**

### **4.1 Marquage**

Les sondes ont été marquées à la digoxygénine soit par une réaction de PCR normale, soit par l'incorporation de nucléotides par la Klenow.

#### **4.1.1 Marquage aléatoire par PCR**

Pour les réactions de PCR, 1 µl de Dig-mix [contenant 1 µl dATP 10mM, 1 µl dGTP 10 mM, 1 µl dCTP 10 mM, 0,65 µl dTTP 10 mM et 3,5 µl Digoxigenin-11-dUTP 25 nM (Roche) dans une solution de 10 µl complétée avec ddH<sub>2</sub>O] + 3 µl de dNTPs (2 mM chacun) ont été ajoutés au lieu des 4 µmol de dNTPs mis normalement. Le reste de la réaction se fait normalement comme indiqué à la section des réactions de PCR (point 1.3 de ce chapitre).

#### **4.1.2 Marquage aléatoire par polymérisation par le fragment Klenow**

Pour le marquage par la Klenow, le fragment d'ADN d'intérêt purifié est dénaturé à 100°C pendant 15 min avant l'ajout d'un mélange contenant des hexanucléotides, les dNTPs (incluant de la DIG-dUTP dans le ratio d'une unité pour 19 de dTTP), le tampon pour la Klenow (Ecopol buffer à mettre à 1X ; NEB) et 5U de Klenow (NEB). La réaction est laissée environ 18 h à 37°C avant d'être arrêtée par l'ajout d'EDTA (1 µl de 0,5 M dans une réaction de 25 µl) et par l'incubation à 65°C pendant 25 min avant de garder sur glace.

### **4.2 Fragments utilisés**

#### **4.2.1 Dans les tests de complémentation dans *E. coli xerCD*<sup>-</sup>**

Le pCS210 a été digéré avec les enzymes *EcoRV* et *NruI*. Le fragment de 787 pb incluant une portion du gène *tet* fut purifiée et marquée ensuite aléatoirement à la DIG, par la Klenow.

## 4.2.2 Dans les réactions *in vitro*

### 4.2.2.1 Tests topoisomérases

Pour les tests de réactions topoisomérases, le plasmide pMIN33 (contenant un site *dif*) a été marqué aléatoirement à la DIG par la Klenow en suivant le protocole recommandé par Roche Diagnostics, Inc.

### 4.2.2.2 Détection de la formation d'une jonction de Holliday

Pour les réactions *in vitro* avec le site *cer* afin de détecter la formation de jonction de Holliday par les recombinaisons, le site *cer* a été marqué à la DIG par une réaction PCR sur le plasmide pKS492 en utilisant 100 pmol des amorces standards M13-F40 et M13-R48, et le fragment obtenu fut purifié sur gel d'agarose.

### 4.2.2.3 Dans les études de retard sur gel

Le marquage du site *cer* est indiqué au point précédent.

Pour les retards avec le site *difEc*, un fragment de 102 pb contenant le site *dif* d'*E. coli* a été marqué à la DIG par une réaction de PCR sur l'ADN de DH<sub>5</sub>α en utilisant 100 pmol des amorces *difF22* et *difR4*.

## 5 Recherche et séquençage de *xerCPm*

Afin de cloner le gène *xerC* de *P. mirabilis* (*xerCPm*), les mêmes conditions que celles précédemment employées dans la recherche de *xerDPm* ont été utilisées (voir Villion et Szatmari, 1998).

Tout d'abord, des amorces dégénérées nous ont permis d'amplifier le fragment C-terminal conservé chez toutes les recombinaisons de type Xer. Ces amorces ont été fabriquées suite à l'alignement des protéines Xer jusqu'alors séquencées (Amorces A1 et C, voir Villion et Szatmari, 1998). Ces amorces nous avaient donné un premier fragment (qui avait permis l'obtention du gène *xerDPm*), puis un second qui s'est révélé différent du premier. Le programme employé alors fut : 94°C/30 s, 37°C/60 s et 72°C/2 min pendant 30 cycles, suivis d'une élongation finale de 5 min à 72°C, avec l'usage de la polymérase Taq et 250 pmol d'amorces, dans un thermocycleur Hybaid. Des amorces spécifiques ont ensuite été synthétisées afin de réaliser un PCR inverse (en suivant un protocole fourni par

E. Campeau, communication personnelle ; voir Villion et Szatmari, 1998) sur l'ADN génomique de *P. mirabilis* (4,5 µg dans un volume final de 55 µl) digéré avec l'enzyme *Pst*I (45 unités). Avant de procéder à la réaction PCR, l'ADN génomique ainsi digéré pendant 16 h à 37°C a été purifié au chloroforme-alcool isoamyle (concentration 24 :1), précipité avec de l'éthanol 70% + 250 mM NaCl à -70°C pendant 0,5 h, avant d'être centrifugé, puis dilué dans de l'eau déionisée. Différentes dilutions de cette suspension ont été employées pour la ligation de nuit à 14°C. Une fois l'ADN prêt, la réaction de PCR inverse a été réalisée avec un thermocycleur Biometra avec 100 pmol des amorces PrC2 et mirC2 (voir le tableau V pour la séquence des amorces). Les conditions de réaction furent : 94°C/30 s, 52°C/1 min, 72°C/2 min pendant 30 cycles, suivi d'une extension finale à 72°C/5 min, en utilisant la polymérase Taq.

Afin de confirmer la séquence *xerC* de *P. mirabilis*, une autre réaction PCR a été réalisée sur l'ADN génomique de l'espèce en utilisant cette fois les amorces PmCF et PmCR. Les conditions employées furent alors : 95°C/15 s, 52°C/30 s, 72°C/5 min pendant 30 cycles, suivi d'une extension finale à 72°C/5 min dans le thermocycleur Hybaid, en utilisant cette fois la polymérase Vent<sup>TM</sup> (NEB).

Le séquençage a ensuite été réalisé sur un produit PCR amplifié par les amorces PmxercF et PmxercR, à l'aide d'un autoséquenceur 373 ABI et réalisé par Patty Singer de l'Université du Maine (voir son site web pour le séquençage facilité : <http://www.ume.maine.edu/dnaseq/index.htm>).

La séquence de *xerCPm* ainsi obtenue a été soumise à GenBank (numéro d'accèsion : AFO33498).

Un autre produit amplifié par PCR avec les amorces PmCF et PmCR (en employant la polymérase Taq) a été cloné dans le vecteur pTZ-PC (voir Villion et Szatmari, 1998). Ce clone, pMVC6 a été employé pour la construction d'autres plasmides, lesquels ont été alors employés dans les études de complémentation.

## 6 Construction de certains plasmides

### 6.1 pMVC<sub>mbp</sub>

Le gène *xerCPm* fut amplifié avec les amorces PrC5' et PmCR en suivant les conditions de réaction employées pour la confirmation de la séquence lors de la recherche

de *xerC*. Une fois purifié, le produit a été phosphorylé avec la T4 polynucléotide kinase et inséré dans le vecteur pMal-c2 digéré avec *XmnI* et déphosphorylé avec la SAP (phosphatase alcaline de crevette). La ligation s'est faite de nuit avec la ligase d'ADN T4 à 14°C. Le produit de ligation obtenu a été ensuite transformé dans la souche DH5 $\alpha$ , et le tout fut étalé sur gélose LA + Ap pour la croissance de nuit à 37°C.

Une fois le clone (pMVC<sub>mbp</sub>) obtenu et la séquence clonée vérifiée, il a été transformé dans la souche DS9029 pour une future expression et purification de XerCPm fusionné à la protéine MBP (protéine se liant au maltose). Les détails de la purification sont indiqués plus loin dans le matériel et méthodes.

## 6.2 pCS210Km<sup>R</sup>

Compte tenu de la résistance naturelle à la tétracycline chez *P. mirabilis*, le plasmide pCS210 servant aux tests de complémentation dans *E. coli* a dû être modifié. Le gène *tet* a été enlevé par une digestion *HindIII* + *NcoI* et fut remplacé par le fragment *HindIII* de pHP45 $\Omega$ Km. Les deux fragments linéaires ainsi obtenus ont été traités à la Klenow, puis purifiés sur gel d'agarose, avant d'être liés par la T4 DNA ligase.

## 6.3 pK $\Delta$ SacKm<sup>R</sup>

Le vecteur pK $\Delta$ SacKm<sup>R</sup> a servi à la construction de la souche de *P. mirabilis* inactivée pour les gènes *xerC* et *D*. Pour le fabriquer, la zone incluant le gène *sacB* et une portion de *sacR* fut retirée de pKNG101 par la digestion avec les enzymes *FspI* et *XmnI*, puis le reste du vecteur fut relié avec la T4 ligase (NEB). Le gène de résistance à la kanamycine, provenant du vecteur pUC119K digéré avec *PstI*, fut ensuite ajouté dans le site *SbfI* (un schéma de ce nouveau vecteur est montré à la Figure 15B).

# 7 Transfert de type Southern

Avant de procéder à une hybridation, les échantillons d'ADN d'intérêt, migrés sur gel d'agarose, ont dû être transférés sur une membrane de nylon chargée positivement. Pour ce faire, le gel a tout d'abord été trempé dans trois solutions différentes afin d'en dénaturer l'ADN. Le gel fut immergé 15 min dans une solution de dépurination (250 mM HCl), puis deux fois dans une solution de dénaturation (0,5 M NaOH, 1,5 M NaCl) pendant 15 min, et enfin deux fois dans une solution de neutralisation (0,5 M Tris-Cl, pH 7,5, 1,5 M

NaCl). Ensuite, l'ADN du gel fut transféré de façon « classique » sur la membrane de nylon chargée positivement (Hybond N<sup>+</sup> de Amersham) pendant la nuit (voir Ausebel *et al.*, 1987-1997). Le tampon employé pour ce transfert fut du 10X SSC (1,5 M NaCl, 150 mM citrate de sodium; pH 7,0).

### **7.1 Préhybridation et hybridation d'une membrane récupérée à la suite d'un transfert Southern :**

La solution employée pour hybrider une membrane sur laquelle de l'ADN a été transféré fut la suivante : 25 ml de 20X SSC, 10 ml Réactif de blocage (Roche Diagnostics, Inc.) à 10% (préparé dans du tampon 1, voir les solutions employées pour la détection et la révélation au point 7.2 de ce chapitre), 1 ml de N-lauryl-sarcosine à 10%, 200 µl de SDS à 10% et 63,8 ml de ddH<sub>2</sub>O stérile. Dans tous les cas, la préhybridation a duré environ 2 à 3 h à la température optimale pour la future sonde employée. Elle fut suivie d'une hybridation faite avec une solution de préhybridation neuve contenant la sonde d'ADN préalablement dénaturée à 100°C/10 min. L'hybridation a duré pendant la nuit à la même température que celle employée pour la préhybridation. Une fois l'hybridation effectuée, la sonde fut mise à part et conservée à -20°C. La membrane récupérée et mise dans un plat fut alors lavée deux fois dans une solution à faible « stringence » (« low stringency buffer » : 2X SSC, 0,1% SDS) à température de la pièce pendant 5 min, puis deux fois dans une solution à plus forte stringence (« high stringency buffer » : 1X SSC, 0,1% SDS) à 65°C pendant 15 min. Ces lavages ainsi que l'hybridation furent faits avec agitation. Par la suite, la membrane fut traitée dans une succession de tampons avant de pouvoir la révéler sur un film Super-Rx (Fujifilm). Ce procédé est décrit dans la section suivante.

### **7.2 Détection et révélation de la digoxygénine**

Que ce soit pour révéler une sonde marquée à la DIG employée dans une hybridation, ou pour révéler un site marqué à la DIG employé dans une réaction de retard sur gel de polyacrylamide, la démarche à suivre pour la détection et la révélation de la DIG fut la même. La membrane fut tout d'abord rapidement rincée dans du tampon 1 (tampon d'acide maléique : 0,1 M acide maléique, 0,15 M NaCl, pH ajusté à 7,5). Ensuite, un blocage de la membrane fut effectué en remplaçant le tampon 1 par 30 ml de tampon 2 [30 ml par membrane de 11 cm X 14 cm ; le tampon 2 est fait de 5 g de réactif de blocage (Blocking Reagent, Roche) dans 500 ml de tampon 1, puis autoclavé], et laissé avec

agitation pendant 30 min. Après 30 min d'agitation, la solution fut remplacée par du tampon 2 neuf contenant 6  $\mu$ l d'un anticorps anti-DIG couplé à la phosphatase alcaline (Anti-Digoxygénine-AP, fragments Fab, 0,75U/  $\mu$ l ; Roche) et laissée encore pendant 30 min avec agitation. La membrane fut ensuite rincée deux fois avec 50 ml de tampon 1 pendant 15 min, puis trempée 2 min dans le tampon de détection (tampon 3 : 100 mM Tris-Cl, pH 9,5, 100 mM NaCl). La membrane fut alors insérée entre deux acétates et, délicatement, la solution de révélation fut ajoutée en couvrant toute la membrane. Pour une membrane de 11 cm X 14 cm, il faut 2,5 ml de solution de révélation [100  $\mu$ l diluant 25X CDP-Star (NEB), 5  $\mu$ l réactif CDP-Star (CDP-Star Reagent ; NEB) et 2,4 ml dH<sub>2</sub>O]. Après 5 min de contact, l'excédent de solution fut enlevé, et les acétates scellées. La membrane fut ensuite exposée sur un film Super-Rx (Medical X-ray Film Super RX ; Fujifilm).

## 8 Purifications protéiques

### 8.1 MBP-XerCPm et MBP-XerCEc :

Les souches DS9029 contenant pMVC<sub>mbp</sub> (pour XerCPm) et DS9029 contenant pGB500 (pour XerCEc), ont été mises en culture dans du LB + antibiotiques et induites à une DO<sub>600</sub> de 0,6 avec 0,1 mM d'IPTG pendant 2h30 à 37°C avec agitation. Une fois les culots récupérés par centrifugation et conservés à -70°C pendant une nuit, ils ont été resuspendus dans du tampon contenant 20 mM Tris-Cl, pH 7,4, 200 mM NaCl, 1 mM EDTA, 0,5 mM PMSF et 1 mg/ml de lysozyme, tel qu'indiqué dans Ausebel *et al.* (1987-1997). Après plusieurs cycles de gels-dégels (-70°C et 37°C), les cellules ont été sonifiées, puis centrifugées. Les surnageants récupérés furent ensuite mis en contact avec 1 ml de résine d'amylose (NEB ; 1 ml permet de récupérer 3 mg de protéine purifiée), pendant plus d'une heure à 4°C. Ce mélange résine-surnageant de protéines fut ensuite transféré sur une colonne de polypropylène pour permettre la purification en tant que telle. Un premier lavage fut fait avec le même tampon employé pour la resuspension des culots, sans le lysozyme, puis un second fut fait avec 20 mM Tris-Cl, pH 7,4, 1 M NaCl, 1 mM EDTA et 0,5 mM PMSF. Les éluions (faites avec le tampon de resuspension contenant 10 mM de maltose) furent récupérées à coup de 500  $\mu$ l et conservées à -70°C. Une partie de ces éluions fut conservée dans 40% glycérol à -20°C pour emploi immédiat.

## 8.2 His-XerDPm et XerDEc :

Les souches HMS174(DE3) contenant le plasmide pMVD72 et BL21 contenant les plasmides pREP4 et pRM130, ont été mises en culture dans du LB + antibiotiques et induites à une  $DO_{600}$  de 0,6 à 0,7 avec 0,1 mM d'IPTG pendant 3 h à 37°C avec agitation. Une fois les culots récupérés par centrifugation et conservés à -70°C pendant une nuit, ils ont été resuspendus dans du tampon contenant du tampon phosphate à pH 8,0 (50 mM  $NaH_2PO_4$ , 20 mM imidazole, 300 mM NaCl, 0,5 mM PMSF, 1 mg/ml de lysozyme). Comme ce fut le cas pour la purification des protéines MBP-XerC ci-haut, plusieurs cycles de gels et dégels (-70°C et 37°C) ont été réalisés avant que les cellules ne soient sonifiées puis centrifugées. Un contact d'une heure à 4°C fut fait avec le surnageant récupéré et 700  $\mu$ l de résine Ni-NTA (NEB ; 8 mg attendus de protéine purifiée par ml de résine), avant que le mélange résine-surnageant ne fût passé sur une colonne de polypropylène. Les échantillons furent lavés avec le tampon de resuspension sans lysozyme, puis avec le même tampon contenant 100 mM d'imidazole. Ensuite, les fractions protéiques furent éluées avec le même tampon contenant cette fois 250 mM d'imidazole. Toutes les fractions, récupérées à coup de 500  $\mu$ l, furent conservées à -70°C, et une portion, conservée dans 40% glycérol, fut gardée à -20°C pour usage immédiat.

## 8.3 His-XerCPm

La purification de His-XerCPm a été faite à partir du plasmide pMV17C transformé dans la souche DS9029. 2 ml d'une culture de nuit ont été ajoutés à 100 ml de LB + Ap + Km, puis incubés à 37°C avec agitation jusqu'à l'atteinte d'une  $DO_{600}$  de 0,6. L'incubation s'est poursuivie pendant 2h à 37°C après l'ajout de 0,1 mM d'IPTG pour l'induction. La purification protéique a été faite comme pour les protéines XerD de *P. mirabilis* et d'*E. coli*.

## 8.4 ArgR

La protéine ArgR d'*E. coli* a été exprimée et purifiée à partir du plasmide pDB169 transformé dans la souche DS955 *recA*<sup>-</sup> (voir les tableaux III et IV), en suivant le protocole décrit par Lim *et al.* (1987).

## 8.5 PepA

La protéine PepA d'*E. coli* a été exprimée et purifiée à partir du plasmide pCS126 transformé dans la souche DS955 *recA*<sup>-</sup>, selon le protocole indiqué par Stirling *et al.* (1989). Cependant, l'étape finale de dialyse n'a pas été réalisée. Ainsi, après une induction de 2 h à 37°C avec 0,5 mM IPTG, le culot récupéré après centrifugation a été lavé dans du tampon TM (10 mM Tris, 10 mM MgCl<sub>2</sub>, pH 8,0), puis gardé dans ce tampon pendant une nuit à -70°C. La lyse a ensuite été réalisée par une série de gels-dégels (-70°C et 37°C) suivie d'une sonification. Les culots ensuite récupérés ont été solubilisés dans du tampon à forte concentration en sels (1M NaCl, 10 mM MgCl<sub>2</sub>, 10 mM Tris-Cl pH 8,0) et la protéine purifiée fut conservée dans ce même tampon contenant 30% glycérol.

Tous les échantillons protéiques furent ensuite vérifiés sur gels SDS-PAGE à 12,5% et colorés au bleu de Coomassie (Dreyfuss *et al.*, 1984).

## 9 Gels SDS-PAGE

L'électrophorèse des échantillons protéiques s'est fait sur un gel de polyacrylamide dénaturant à 12,5% (SDS-PAGE), suivant le protocole de Laemmli (1970) (voir également Ausebel *et al.*, 1997). Un mini-gel (8,3 cm x 7,3 cm) avec des espaceurs de 1,5 mm d'épaisseur fut composé d'un gel de séparation fait avec :

3,7 ml d'une solution d'acrylamide 33,5%- bisacrylamide 0,3%  
 3,8 ml de Tris -Cl 1M, pH 8,8 ou 9,1  
 2,1 ml d'eau distillée  
 100 µl de SDS 10%  
 80 µl d'APS 10%  
 20 µl de TEMED

et d'un « stacking » gel fait avec :

325 µl d'une solution d'acrylamide 30%- bisacrylamide 0,44%  
 312,5 µl de Tris-Cl 1M pH 6,8  
 1,9 ml d'eau distillée  
 25 µl de SDS 10%  
 25 µl d'APS 10%  
 10 µl TEMED

Avant la mise sur gel pour la migration, les échantillons protéiques devaient être mis dans un tampon de chargement dénaturant [50 mM Tris-Cl (pH 6,8), 100 mM DTT, 2

% SDS, 0,1 % Bleu de bromophénol et 10 % glycérol] afin d'être dénaturés à 100°C, pendant quelques minutes.

La migration fut réalisée dans une cuve à électrophorèse contenant du tampon tris-glycine 1X (25 mM Tris, 250 mM glycine pH 8,3 et 0,1% SDS).

### 9.1 Coloration des gels

Le gel fut ensuite coloré au bleu de Coomassie brillant (0,25% bleu Coomassie brillant R250, 50% méthanol, 10% acide acétique) pendant 30 minutes et décoloré dans une solution de méthanol 30%- 10% acide acétique en changeant fréquemment la solution.

Lorsque les bandes furent trop faibles pour être observées au bleu de Coomassie, le même gel a pu être coloré au nitrate d'argent. Le gel fut alors trempé toute la nuit dans du méthanol 50% avec une légère agitation. Le lendemain, après changement de méthanol pour 1 h supplémentaire, le gel fut rincé à l'eau distillée, avant d'être coloré pendant 15 min dans la solution au nitrate d'argent préparée de la façon suivante :

4 ml de nitrate d'argent ( $\text{AgNO}_3$ ) 20% furent ajoutés goutte à goutte à une solution contenant 21 ml de NaOH 0,36% et 1,4 ml de  $\text{NH}_4\text{OH}$  14,8 M (ou 30%). Cette solution fut complétée à 100 ml avec de l'eau distillée.

Le gel fut ensuite lavé pendant 5 min à l'eau distillée, puis révélé avec une solution contenant 1,25 ml d'acide acétique 1% et 125  $\mu\text{l}$  de formaldéhyde ( $\text{HCHO}$ ) 37% dans 250 ml complétés avec de l'eau distillée. Lorsque la révélation adéquate était atteinte, le gel était rincé brièvement avec de l'eau distillée et la réaction, stoppée avec du méthanol 10%.

## 10 Test de Bradford

L'évaluation de la concentration protéique a été faite en suivant le protocole indiqué avec le réactif de Bradford utilisé (Bio-Rad) :

Tout d'abord, une courbe tendance a été réalisée avec du BSA, de 1  $\mu\text{g/ml}$  à 50  $\mu\text{g/ml}$ . Ainsi, à 800  $\mu\text{l}$  contenant le BSA, 200  $\mu\text{l}$  de réactif de Bradford ont été ajoutés. Après 20 min d'attente, la densité optique à 595 nm a été prise et la courbe de la densité optique en fonction de la concentration protéique, réalisée. À partir de cette courbe, la concentration des échantillons fut déduite : 25  $\mu\text{l}$  et 40  $\mu\text{l}$  des diverses éluions furent

mesurés dans 775 et 760  $\mu\text{l}$  d'eau déionisée respectivement, avec l'ajout de 200  $\mu\text{l}$  de réactif de Bradford.

## 11 Retards sur gel

Les retards sur gel furent réalisés en suivant le protocole employé par Blakely et Sherratt (1996). Ainsi, les protéines Xer furent mises en contact avec l'ADN marqué à la digoxygénine à tester dans du tampon TENG (20 mM Tris-Cl, pH 7,5, 1 mM EDTA, 50 mM NaCl et 5% glycérol) contenant 1  $\mu\text{g}$  de polyIdC-polyIdC (Amersham). Les quantités employées furent de 600 nM de protéine pour environ 3 nM de site marqué.

Une fois la réaction réalisée à 37°C pendant 10 à 15 min, les échantillons furent migrés à 80V pendant environ 1h20 sur un gel de polyacrylamide non dénaturant à 6% (préalablement chauffé à 200V pendant 30 min) dans du tampon TBE ou TAE 1X. Les gels ont été ensuite transférés par capillarité sur une membrane N<sup>+</sup> (Amersham) pendant toute la nuit. Une fois la membrane fixée aux UV, la révélation et la détection des sites marqués furent réalisées en suivant le protocole suggéré par Roche, puis avec le système CDP-Star (NEB) tel qu'indiqué pour les réactions de Southern ou de Northern sur un film Super RX Medical X-ray (Fujifilm).

## 12 Réactions *in vitro*

### 12.1 Avec le site *cer*

Les réactions réalisées avec le pCS210 afin de vérifier la formation de jonctions de Holliday ont été réalisées tel que décrit par Colloms *et al.* (1996), mais avec quelques modifications. Toutes les réactions ont été faites pendant 1 h à 37°C dans un volume total de 45  $\mu\text{l}$  dans le tampon suivant : 1 mM L-arginine, 1,36 mM EDTA, 50 mM Tris pH 8,0, 50 mM NaCl, 5 mM spermidine, 25  $\mu\text{g/ml}$  BSA et 20% glycérol. Les protéines Xer et les protéines accessoires ArgR et PepA d'*E. coli* ont été employées aux mêmes concentrations qu'indiquées par Colloms *et al.* (1996), 300 nM de XerC et XerD (monomères), 30 nM d'ArgR et 300 nM de PepA (hexamères). Avant d'être précipitées à l'éthanol, les réactions furent traitées une fois au phénol-chloroforme-alcool isoamyle (dans un ratio de 25 :24 :1 respectivement), puis une fois au chloroforme-alcool isoamyle (dans un ratio de 24 :1). L'ADN récolté fut ensuite remis en solution dans de l'eau déionisée, puis digéré avec

*EcoRV* pendant 1h à 37°C avant d'être analysé sur gel d'agarose 1% dans du tampon TAE 1X.

## 12.2 Avec le site *difEc*

Les réactions avec le pMIN33 (contenant le site *difEc*), afin de vérifier l'activité topoisomérase des recombinaisons, ont été réalisées en suivant le protocole indiqué par Cornet *et al.* (1997), en employant 500 nM de protéines Xer, au lieu de 1,1 µM pour XerC et 125 nM pour XerD, dans le même tampon, mais sans spermidine. Les réactions ont duré 14h30 à 37°C, dans 50 mM Tris, pH 8,0, 50 mM KCl, 1,25 mM EDTA, 0,25 µg/ml BSA et 40% glycérol (v/v), puis l'ADN a été précipité à l'éthanol et mis sur gel d'agarose 1% dans du tampon TAE 1X à 22 V/19 h.

Pour les deux types de réactions *in vitro*, les gels d'agarose ont été transférés de la même façon que le gel résultant des tests de complémentation sur des membranes Gene Screen Plus, avant que les membranes ne soient hybridées avec les sondes adéquates (site *cer* marqué pour les réactions *in vitro* au site *cer*, et site *dif* marqué pour les réactions avec le pMIN33 ; voir le point 4.2.2 de ce chapitre).

## 13 Tests de complémentation dans *E. coli xer<sup>-</sup>*

Des monomères de pCS210 ont été purifiés et transformés dans la souche DS9029 à l'aide du kit TransformAid (Fermentas), en suivant les indications du manufacturier. Les plasmides pEcCD, pPmCD, pPmCEcD, pPmDEcC et pUC19 (voir le Tableau IV) ont ensuite été transformés dans cette nouvelle souche. Certaines des colonies obtenues ont été mises en culture dans du LB+Tc+Ap et l'ADN plasmidique a été extrait à l'aide du kit de QIAGEN (voir la section Minipreps en point 2.2 de ce chapitre). 400 ng de chacun des extraits ont été mis sur gel d'agarose 1% pour une électrophorèse de 12 h à 33 V. Le gel a ensuite été transféré sur une membrane chargée positivement (Hybond<sup>TM</sup>-N<sup>+</sup> d'Amersham Pharmacia Biotech), après la dépurination, dénaturation et neutralisation du gel, comme il est fait dans le protocole standard de transfert de type Southern (point 7 de ce chapitre). Cependant, le transfert a été fait en séchant le gel directement sur la membrane dans un sécheur de gel (75 min à 65°C). La membrane a ensuite été bombardée aux rayons ultraviolets pour la réticulation de l'ADN (UV cross-linking), puis préhybridée à 70°C/4 h dans la solution de préhybridation [5X SSC, 1% réactif de blocage (Roche Diagnostics,

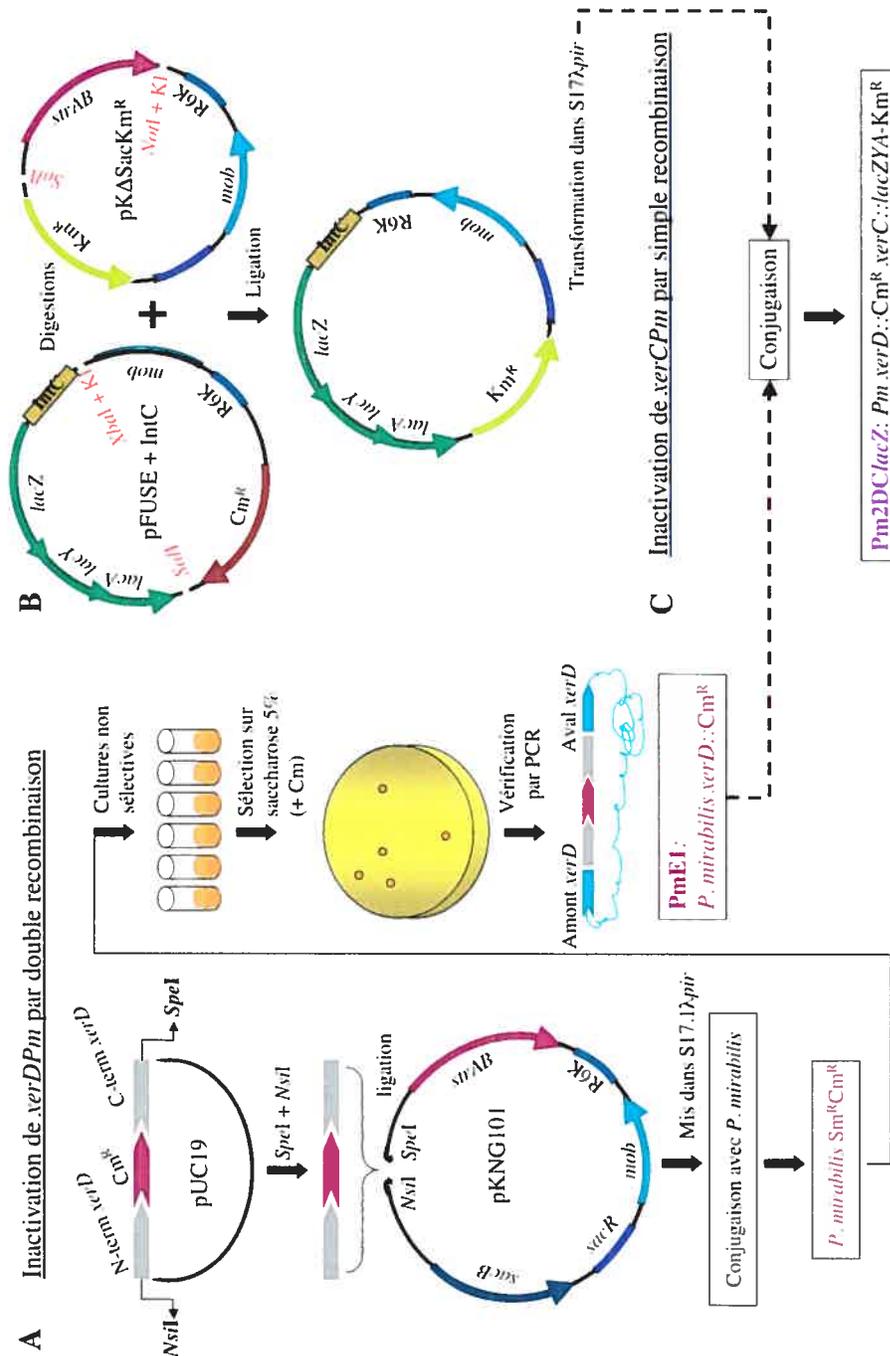
Inc.), 0,1% N-Lauryl-sarcosine, 0,02% SDS]. L'hybridation a suivi dans une solution de préhybridation fraîche et contenant le gène *tet* de pCS210 marqué à la digoxygénine employé comme sonde (à 70°C pendant 16 h). La membrane fut ensuite lavée deux fois dans une solution 2X SSC-0,2% SDS à la température de la pièce, puis deux fois dans une solution 0,1X SSC-0,1% SDS à 65°C avant d'être traitée pour la détection par chemiluminescence (voir la section de détection en point 7.2 pour les transferts par Southern : protocole de détection de la DIG de Roche Diagnostics, Inc.). La détection finale fut faite avec la solution CDP-Star (NEB), suivie par une exposition sur film « Super RX Medical X-ray » (Fujifilm).

### 13.1 Tests de stabilité plasmidique

Afin de vérifier la stabilité du pCS210 dans la souche DS9029 contenant les divers plasmides transformés dans la section précédente, ces souches furent mises en culture dans du LB ne contenant que de l'ampicilline (pour le maintien des clones *xer*), mais pas de tétracycline (servant à la sélection du pCS210), et ce, pour une période de 5 jours. L'ADN plasmidique fut extrait chaque jour, et des dilutions de la culture de nuit furent étalées sur des géloses LA-Tc-Ap et sur des géloses LA-Ap afin de déterminer le nombre de UFCs (unités formatrices de colonies). De plus, 50 µl de la dilution  $10^{-4}$  furent mis dans 5 ml de LB-Ap pour une croissance de nuit à 37°C, afin d'obtenir une croissance correspondant à une vingtaine de générations. Afin de confirmer ces résultats, 50 colonies obtenues sur les géloses LA-Ap furent également étalées par « patch » sur les géloses LA-Tc-Ap et LA-Ap. Le ratio entre le nombre d'UFCs obtenues sur un type de gélose versus l'autre fut établi et mis sur un graphique.

#### 13.1.1 Dilutions

Les dilutions des cultures ont été réalisées dans du PBS [« phosphate buffered saline »; 20 mM phosphate de sodium, pH7,5 (10 mM  $\text{NaH}_2\text{PO}_4$  + 10 mM  $\text{Na}_2\text{HPO}_4$ ) et 100 mM NaCl].



**Figure 15: Schémas de l'inactivation de *xerDPm* par double recombinaison et de la création de *P. mirabilis xerC-xerD*<sup>-</sup>.** A) Inactivation du gène *xerDPm* à l'aide du vecteur pKNG101. La souche ainsi inactivée contient une cassette Cm<sup>R</sup> à la place de la portion centrale de *xerD*. Cette inactivation a été réalisée par événement de double recombinaison. B) Construction d'un vecteur-suicide contenant le fragment interne de *xerCPm*. C) La souche de *P. mirabilis* inactivée pour les gènes *xerC* et *xerD* a pu ensuite être réalisée par la conjugaison de la souche S17λpir contenant le vecteur construit en B et la souche PmE1 construite en A. Le gène *xerC* a été inactivé suite à l'intégration du vecteur dans le chromosome par un seul événement recombinaison, en créant une fusion d'opéron avec *lacZYA*. Tous les détails sont donnés dans le texte. (Kl : Traitement à la Klenow).

## 14 Inactivation des gènes *xer* dans *P. mirabilis*

### 14.1 Inactivation du gène *xerD<sub>Pm</sub>* :

Divers clonages ont été nécessaires avant l'obtention de la souche de *P. mirabilis* inactivée pour *xerD*. La démarche est illustrée en Figure 15A.

Tout d'abord les portions N- et C-terminales du gène *xerD<sub>Pm</sub>* ont été clonées de part et d'autre du site de polyclonage du vecteur pUC19 : la portion N-terminale, qui inclut une portion du gène en amont de *xerD*, a été amplifiée avec les amorces PmxDN5' et PmxDN3' (voir le Tableau V) qui contiennent respectivement les sites *Nsi*I et *Eco*RI, puis digérée avec *Eco*RI. Cette portion a été intégrée entre les sites *Ssp*I et *Eco*RI de pUC19 (pour donner le clone 4D). La portion C-terminale, qui inclut une portion du gène en aval de *xerD<sub>Pm</sub>*, a été amplifiée par PCR avec les amorces [PmxDC5'a] et [PmxDC3'], lesquelles contiennent respectivement les sites *Pst*I et *Spe*I. Ce produit, digéré avec *Pst*I, a été intégré dans le nouveau clone formé précédemment (pour donner le clone 21D). Le clone 4D avait été préalablement digéré *Hind*III, traité à la Klenow, puis digéré avec *Pst*I.

Ensuite, la cassette de résistance au chloramphénicol a été insérée entre les deux parties du gène. Pour ce faire, la cassette a été extraite du vecteur pCAT19 par digestion avec *Xba*I, et intégré dans le site *Xba*I du clone 21D formé à l'étape précédente (pour donner le clone 2<sup>f</sup>).

L'ensemble du fragment contenant la cassette Cm<sup>R</sup> flanquée des portions N- et C-terminales a ensuite été extrait de pUC19 par digestion avec les enzymes *Spe*I et *Nsi*I pour être transféré dans le vecteur suicide pKNG101 préalablement digéré avec les mêmes enzymes. Puisque la réplication du vecteur suicide nécessite la protéine Pir, la sélection d'un clone potentiel a été faite dans la souche CC118λ*pir* (les transformations précédentes avaient toutes été réalisées dans DH5α). Par la suite, le nouveau plasmide (nommé 1D) a été transformé dans la souche S17.1λ*pir* pour lui permettre son transfert par conjugaison dans *P. mirabilis* 240-82.

#### 14.1.1 Conjugaison bactérienne pour l'inactivation du gène *xerD<sub>Pm</sub>*:

L'inactivation complète du gène *xerD<sub>Pm</sub>* s'est faite en 2 étapes. Premièrement, la conjugaison simple entre la souche donatrice (dans ce cas-ci, la souche S17.1λ*pir* contenant

le clone 1D) et la souche réceptrice (*P. mirabilis* 240-82). Cette étape permet la sélection de clones ayant intégré le plasmide contenant deux portions du gène *xerD<sup>Pm</sup>* par recombinaison homologue. L'échange peut avoir eu lieu alors, soit dans la portion N-terminale, soit dans la portion C-terminale. Deuxièmement, une seconde étape de sélection est nécessaire pour débarrasser le génome de *P. mirabilis* du vecteur ayant permis le premier échange. Dans cette étape, on cherche donc à sélectionner des clones ayant subi une deuxième réaction d'échange dans l'autre portion de *xerD* n'ayant pas servi lors de la 1<sup>re</sup> réaction d'échange.

#### 1<sup>re</sup> étape : la conjugaison :

Les souches donatrices et réceptrices ont été mises en culture dans du LB avec les antibiotiques adéquats (Sm et Cm pour la souche S17.1 $\lambda$ *pir* contenant le vecteur 1D et Tc pour *P. mirabilis*) et laissées sans agitation à 28°C pendant une nuit. Le lendemain, 200  $\mu$ l de chacune des cultures furent lavés avec 1 ml de PBS avant d'être resuspendus dans 50  $\mu$ l de LB. Les deux cultures ainsi préparées furent alors passées sur le même filtre à 0,45 $\mu$ , et le filtre fut déposé (côté cultures bactériennes vers le haut) sur une gélose LA ne contenant aucun antibiotique. Cette boîte fut alors laissée, non renversée, à 28°C pour une nuit.

Le lendemain, le filtre fut récolté, mis dans 1 ml de PBS et bien vortexé. Tout le mélange bactérien fut alors étalé à coup de 100  $\mu$ l ou 200  $\mu$ l sur des géloses LSW contenant de la Sm et de la Tc, puis incubé à 28°C pendant 48 h. Les colonies obtenues (très peu en réalité, moins d'une dizaine en tout) furent étalées sur 3 milieux différents : 1) LSW + Sm + Tc, 2) LSW + Cm + Tc et 3) sur LSW sans antibiotique. Les colonies conservées alors sont celles ayant poussé à la fois sur la gélose contenant la Sm et celle contenant le Cm. Une réaction de PCR sur ces clones, avec les amorces PmxDN5' et PmxDC3', nous a permis de confirmer que le plasmide était bel et bien dans la souche bactérienne. La réaction de PCR montrait alors une bande correspondant au gène *xerD<sup>Pm</sup>* intégral et une autre bande correspondant au gène *xerD<sup>Pm</sup>* ayant l'insertion de Cm<sup>R</sup>.

#### 2<sup>e</sup> étape : sélection sur saccharose

Pour nous débarrasser du reste du pKNG101 maintenant intégré dans le génome de *P. mirabilis*, une seconde sélection, sur milieu contenant du saccharose, est nécessaire. En effet, le pKNG101 contient le gène *sacB* codant pour la lévansaccharase de *Bacillus subtilis* (Gay *et al.*, 1983). Cette enzyme non seulement hydrolyse le saccharose, mais aussi

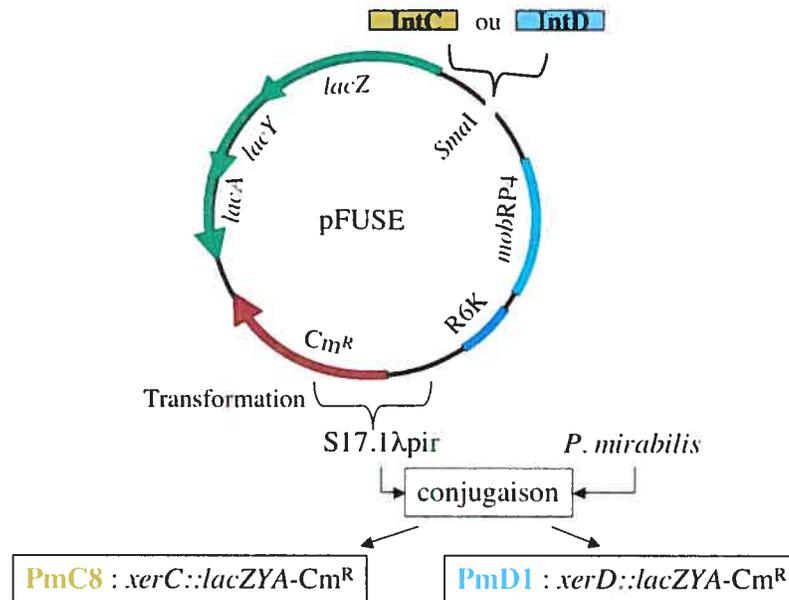
synthétise la « lévane » (polymère de résidus fructose de haut poids moléculaire) dont l'accumulation dans la cellule bactérienne serait toxique (Reyrat *et al.*, 1998). Ainsi, seuls les clones s'étant débarrassés du reste du pKNG101 peuvent survivre. Ils pourront être comme la souche sauvage (2<sup>e</sup> réaction d'échange ayant eu lieu dans la même portion du gène *xerD* que lors du 1<sup>er</sup> échange) ou comme une nouvelle souche ayant acquise la résistance au Cm par son insertion à l'intérieur de *xerD* (2<sup>e</sup> réaction d'échange ayant eu lieu dans la 2<sup>e</sup> portion du gène *xerD*).

Pour ce faire, les colonies étalées obtenues de la 1<sup>re</sup> étape ont été mises en culture dans 5 ml de LB (sans antibiotique), et laissées à 37°C pour 6h avec agitation. 100 µl de ces cultures ont été ensuite étalées sur milieu LSW + saccharose 5% + Cm + Tc, puis incubées à 28°C/48 h. Les colonies obtenues ont ensuite été vérifiées sur LSW + Sm, LSW + Cm et enfin sur LSW sans antibiotique. Seules les colonies ayant poussé sur Cm sans croître sur Sm ont été sélectionnées et vérifiées par PCR avec les amorces PmxDN5' et PmxDC3'.

De toutes les colonies obtenues, une seule a passé tous les critères ; la nouvelle souche de *P. mirabilis* ainsi obtenue fut nommée PmE1.

## 14.2 Fusions d'opéron avec *lacZ* :

Les gènes *xerC* et *xerD* de *P. mirabilis* ont été inactivés par l'insertion de l'opéron *lacZ* en fusion avec le début des gènes (voir la Figure 16). Le vecteur suicide pFUSE a alors été employé (Bäumler *et al.*, 1996).



**Figure 16 : Construction des souches de *P. mirabilis* *xerC*<sup>-</sup> ou *xerD*<sup>-</sup> par événement de simple recombinaison.** Les constructions furent réalisées à l'aide du vecteur pFUSE afin de permettre une fusion d'opéron avec l'opéron *lacZ*. Les fragments internes des gènes *xerCPm* ou *xerDPm* avaient été préalablement amplifiés par PCR, et les clones dont les fragments avaient été insérés dans la bonne orientation ont été transformés dans la souche S17.1λpir afin d'effectuer une conjugaison avec la souche *P. mirabilis* 240-82. (Lire le texte pour les autres détails).

#### 14.2.1 *xerCPm* :

Un fragment interne de *xerCPm* a été amplifié avec les amorces PmC-IntF et PmC-IntR sur l'ADN génomique de *P. mirabilis*, à une température d'hybridation de 49°C et une élongation de 40 s. Ce fragment, une fois purifié, a été inséré dans le vecteur pFUSE digéré avec *SmaI*, purifié également. Le produit de ligation a ensuite été transformé dans la souche CC118λpir. Les colonies obtenues contenant le pFUSE avec insert, ont été vérifiées par PCR avec les amorces PmCIntF et XDL-R (amorce s'hybridant à la partie N-terminale de *lacZ* en direction 3' vers 5') pour sélectionner les clones dont l'insert était dans la bonne orientation. Le clone retenu (21C) a ensuite été transformé dans la souche S17.1λpir avant de pouvoir effectuer la conjugaison avec *P. mirabilis*. La conjugaison a été effectuée sur filtre, essentiellement comme celle réalisée pour l'obtention de PmE1 ci-haut, à l'exception que deux fois plus du receveur a été mis en contact avec le donneur. Cette fois-ci, une simple réaction d'échange était nécessaire et l'ensemble du vecteur fut conservé dans la souche. Le clone final, nommé PmC8, a été vérifié par PCR avec les

amorces PmxCN5' et XDL-R. De plus, ce clone apparaît maintenant bleuté sur gélose contenant de X-gal (ne pas oublier que *P. mirabilis* est une espèce bactérienne *lac*<sup>-</sup>).

#### 14.2.2 *xerDPm* :

Comme pour *xerCPm*, un fragment interne de *xerDPm* a été amplifié avec les amorces IntD-F et IntD-R à une température d'hybridation de 48°C et une élongation de 40 s. Ce fragment, une fois purifié, fut inséré dans le vecteur pFUSE traité de la même façon que précédemment, avant d'être transformé dans la souche CC118λ*pir*. Encore une fois, l'orientation de l'insert fut vérifiée dans les clones obtenus par une réaction PCR avec les amorces PmDIntF et XDL-R. Le clone (nommé 14IntDPm) dans la bonne orientation fut ensuite transformé dans la souche S17λ*pir*. La conjugaison entre ce clone et la souche de *P. mirabilis* a ensuite été effectuée exactement de la même manière que pour l'obtention de PmC8. La vérification de l'insertion de pFUSE dans le chromosome de *P. mirabilis* a été faite par PCR avec les amorces PmxDN5' et XDL-R à une température d'hybridation de 48°C pour une élongation de 1 min 30 s. De plus, le clone obtenu apparaissait bleuâtre sur une gélose contenant du X-gal. La nouvelle souche ainsi obtenue fut nommée PmD1.

#### 14.2.3 *xerCPm* et *xerDPm* :

*P. mirabilis* a également été inactivée pour les deux gènes *xer* (Figure 15C). Tout d'abord, la construction d'un vecteur similaire à pFUSE, mais contenant une cassette de résistance différente a été nécessaire (voir la Figure 15B). Ainsi, le clone 21C vu plus haut (au point 14.2.1) a été digéré avec *Xba*I, traité à la Klenow, puis digéré de nouveau avec *Sal*I pour extraire le fragment contenant la portion IntC (portion interne de *xerCPm* obtenu de la réaction de PCR avec les amorces IntCPmF et IntCPmR) fusionné à l'opéron *lacZYA* de pFUSE. Ce fragment a été inséré dans le vecteur pKΔSackM<sup>R</sup> digéré avec *Not*I, traité à la Klenow, puis redigéré avec *Sal*I. Le résultat de cette ligation fut transformé dans la souche CC118λ*pir* et a donné un clone qui a ensuite été transformé dans la souche S17.1λ*pir* afin de pouvoir effectuer une conjugaison avec la souche PmE1 (*Pm xerD*<sup>-</sup>; voir le point 14.1). La conjugaison s'est faite sur filtre, comme lors de l'inactivation des gènes *xerC* et *xerDPm* par fusion d'opéron avec *lacZYA*. Afin de confirmer l'intégration du vecteur à l'intérieur de *xerCPm*, les clones bleus (sur milieu contenant du X-gal) obtenus ont été vérifiés par PCR avec les amorces PmxCN5' et XDL-R à une température

d'hybridation de 49°C et une élongation de 2 min.,. La nouvelle souche ainsi créée fut nommée Pm2DClacZ.

## 15 Études de l'essaimage (swarming)

Les diverses souches de *P. mirabilis* ont été vérifiées quant à leur capacité de migration sur une gélose ordinaire LA contenant du TTC (2,3,5-triphenyltetrazolium : 1 ml d'une solution à 1% préparée dans de l'eau distillée par tranche de 100 ml. 5 µl d'une culture de nuit ont été déposés au centre de la gélose préalablement asséchée pendant au moins 1 à 2 h à 37°C. Après l'adsorption de la goutte dans l'agar, les géloses ont été incubées à 28°C/ 24 h, puis mises à 4°C pour une autre journée.

Les photos ont ensuite été prises avec un appareil numérique Canon Powershot S40, par Daniel Morency (responsable de laboratoire, département de microbiologie et immunologie, Université de Montréal).

D'autres photos ont été prises des diverses souches de *P. mirabilis* cultivées sur des géloses LA sans TTC. L'incubation a alors duré 48 h à 28°C.

## 16 Microscopie

La préparation des lames de *P. mirabilis* a été réalisée par Julie Proulx (stagiaire pour l'été 2004 au laboratoire de Dr Szatmari). La coloration au DAPI a été faite en suivant le protocole de Chen *et al.* (1999). Tout d'abord, les lamelles ont été préalablement traitées à la poly-L-lysine : une goutte de ce produit a été déposée sur une lamelle propre, et laissée de 10 s à 2 min. Après avoir retiré le reste de poly-L-lysine, la lamelle fut séchée puis lavée deux fois avec de l'eau distillée.

Les bactéries à colorer furent mises en culture jusqu'à une  $DO_{600}$  de 0,6. La fixation des cellules fut faite en ajoutant 100 µl de [formaldéhyde 16% + 0,4 µl glutaraldéhyde 20%] + 200 µl de 0,1 M  $NaPO_4$  à 500 µl de culture. Après une incubation de 15 min à la température de la pièce (RT), les cellules furent immédiatement mises et gardées 15 min sur glace. Les cellules furent ensuite lavées deux fois dans du PBS 1X, puis remises en suspension dans du PBS 1X contenant 0,2 µg/ml de colorant DAPI. Après une incubation de 10 min à la noirceur, les cellules furent lavées deux fois dans du PBS 1X, toujours à la noirceur, avant d'être remises en suspension dans 100 µl de PBS 1X. 10 µl furent ensuite

déposés sur une lamelle prétraitée à la poly-L-lysine et laissés sécher. Une goutte de PBS 1X a ensuite été ajoutée avant de sceller la lamelle sur une lame à l'aide de vernis à ongle, et le tout fut conservé au froid à l'abri de la lumière, jusqu'à l'observation au microscope.

Les photos furent prises par Monique Vasseur (technicienne de laboratoire, Microscopie et Imagerie, Département de biochimie, Université de Montréal) à l'aide d'un microscope inversé à fluorescence Nikon TE2000U, sous un objectif Nikon 100x Plan apo ; N.A. 1,40, et ayant une caméra noir et blanc CCD 12 bits CoolSnap HQ de Photometrics. Les photos des coloration DAPI furent prises sous un filtre DAPI (cube Chroma #31000 DAPI/Hoeschst/AMCA dont les longueurs d'onde étaient 360/40-400-460/50 de qualité D (standard), tandis qu'en lumière transmise, le mode d'observation était en DIC (contraste interférentiel différentiel ; également nommé Nomarski). Enfin, le logiciel employé pour l'acquisition des images était Metamorph de UIC (Universal Imaging Corporation), version 5,2.

## 17 Dosages $\beta$ -galactosidase

Les dosages  $\beta$ -galactosidase ont été effectués essentiellement selon le protocole indiqué par Miller (1972).

Les diverses souches bactériennes à tester furent diluées 1/50 dans 80 ml LB à partir d'une culture de nuit et laissées à 37°C avec agitation jusqu'à la fin de l'expérience. Toutes les demi-heures, 1 ml de la culture en croissance fut prélevé et le dosage effectué. Lorsque la  $DO_{600}$  dépassait 0,7, une dilution de ce millilitre était nécessaire. Par exemple, si la  $DO_{600}$  était rendue à 0,7-0,8 pour un échantillon, le prélèvement suivant était dilué 1 :2 dans du LB gardé sur glace.

Avant d'effectuer le dosage, plusieurs tubes Eppendorf avaient été préparés avec 500  $\mu$ l de solution de lyse (tampon Z : 0,06 M  $Na_2HPO_4$ , 0,04 M  $NaH_2PO_4 \cdot H_2O$ , 0,01 M KCl, 0,001 M  $MgSO_4$ , 0,05 M  $\beta$ -mercaptoéthanol), afin de réaliser le dosage le plus rapidement possible après le prélèvement à la  $DO_{600}$  voulue de l'échantillon. Au moment de devoir prélever l'échantillon, 1 goutte de SDS 0,1% et 2 gouttes de chloroforme ont été ajoutés à ces tubes contenant le tampon Z. L'échantillon prélevé fut mis immédiatement sur glace : 1 ml a servi à une mesure de la  $DO_{600}$ , et 500  $\mu$ l ou 200  $\mu$ l furent ajoutés aux tubes préparés auparavant. Le volume a alors été complété à 1 ml avec du LB. Les tubes

furent ensuite agités vigoureusement au vortex avant d'être mis à 28°C. Après quelques minutes, 200 µl d'ONPG (4 mg/ml, préparé le matin même) furent ajoutés, et les tubes remis à 28°C. Aussitôt qu'une couleur jaunâtre suffisamment intense est apparue, 450 µl de Na<sub>2</sub>CO<sub>3</sub> 1 M ont été ajoutés afin de stopper la réaction. Le délai entre cet ajout et celui de l'ONPG fut noté avec précision. Ensuite, la densité optique fut prise à 420 et 550 nm, afin de calculer le nombre d'unités Miller auquel correspondait l'activité β-galactosidase de l'échantillon mesuré. Le nombre d'unités Miller fut calculé selon la formule suivante :

$$\frac{1000 \times (DO_{420} - 1,75 \times DO_{550})}{(\text{Temps} \times V \times DO_{600})}$$

dans laquelle V est le volume utilisé (en ml) de la culture lysée pour le dosage (0,2 ou 0,5 ml dans les cas mentionnés) et Temps (en min) est la durée de la réaction avant l'apparition de la couleur jaunâtre.

Ensuite des courbes des unités Miller en fonction de la DO<sub>600</sub> ont été dressées afin de comparer les deux souches de *P. mirabilis* testées (*xerC* et *xerD*).

## 18 Extraction d'ARN :

Le protocole suivi pour extraire l'ARN de *P. mirabilis* provient du guide venant avec l'achat de TRIzol (Invitrogen) et adapté pour l'ARN bactérien par Karen Wassarman (Wassarman et Storz, 2000).

Comme pour toute manipulation de l'ARN, les précautions d'usage furent respectées afin d'éviter la contamination des échantillons avec des ribonucléases (RNases). Par exemple, toute verrerie employée fut lavée au savon Sparkleen et passée au Four Pasteur, les plastiques employés étaient certifiés « sans trace de RNases », l'eau distillée utilisée pour la préparation des solutions fut autoclavée pendant 1h, l'eau MilliQ utilisée pour la resuspension des échantillons fut traitée au DEPC, etc.

Tout d'abord, les cellules bactériennes furent récoltées par centrifugation dans un tube Eppendorf de 2 ml (10 ml furent prélevés en tout pour les cellules en phase précoce, et 5 ml pour les cellules en phase exponentielle-tardive). Une fois le culot resuspendu dans [75 µl TE (pH7,5) + 5 µl lysozyme (10 mg/ml)], trois séries de gels (dans un mélange carboglace-alcool) et dégels (dans un bain d'eau à température de la pièce) furent réalisées. 1,75 ml de TRIzol furent ensuite ajoutés à la suspension décongelée et le tout bien passé au

vortex. Après une centrifugation à 12000 g pendant 10 min à 4°C, le surnageant fut transféré dans un tube contenant 320 µl de chloroforme. La phase aqueuse obtenue après centrifugation fut ensuite précipitée avec de l'isopropanol, puis le culot d'ARN obtenu, lavé avec de l'éthanol 75%. Une fois séché partiellement, le culot fut resuspendu avec de l'eau traitée au DEPC. Les échantillons d'ARN ont ensuite été conservés à -70°C.

L'ARN ainsi récupéré a été quantifié par une prise de DO à 260 nm et le degré de pureté de l'échantillon a été évalué par la mesure du ratio  $DO_{260}/DO_{280}$ . Une unité de  $DO_{260}$  correspond à 40 µg/ml, tandis qu'un échantillon resuspendu dans 10 mM Tris-Cl, pH 7,5 est considéré pur lorsque le ratio  $DO_{260}/DO_{280}$  est compris entre 1,8 et 2,1.

### **18.1 Électrophorèse d'ARN sur gel d'agarose avec formaldéhyde**

Le protocole indiqué dans le cahier de Roche Diagnostics (Roche Molecular Biochemicals, D1G Application Manual for Filter Hybridization, ©2000) a été suivi pour la migration des échantillons d'ARN, suivi de leur transfert sur membrane avant de révéler l'ARN s'hybridant à une sonde *xer*.

#### **Matériel nécessaire :**

##### 10X MOPS (acide morpholinéopropanesulfonique), pH 7,0:

- 200 mM MOPS (PM : 209,27)
- 50 mM Acétate de sodium (PM : 136,08)
- 20 mM EDTA (PM : 372,24)

Le pH est ajusté à 7,0 avec du NaOH et la solution est autoclavée pendant 1 h. Il est à noter que, tel qu'indiqué dans le livret de Roche, la solution devenue jaunâtre suite au passage à l'autoclave reste fonctionnelle.

##### Tampon de chargement (à préparer en dernière minute) :

- 250 µl formamide déionisée 100%
- 83 µl formaldéhyde 37%
- 50 µl MOPS 10X
- 57 µl dH<sub>2</sub>O autoclavée 1 h

La quantité préparée ici serait pour 55 échantillons de 4 µl d'ARN. On ajuste le volume à la quantité d'échantillons à préparer, selon le volume nécessaire.

##### Mixture de bleu :

Pour 55 échantillons de 4 µl d'ARN, on aurait une mixture contenant 22 µl de bleu de bromophénol (2,5% ; préparé dans de l'eau-DEPC), 22 µl de bromure d'éthidium (10 mg/ml) et 176 µl d'eau-DEPC.

**Protocole :**

Avant d'être mis sur gel, les échantillons d'ARN ont été dilués à une concentration de 5 µg/ml dans de l'eau-DEPC. Ainsi, pour faire migrer 15 µg d'ARN, 3 µl d'échantillon furent ajoutés à 6 µl de tampon de chargement fraîchement préparé. Les échantillons furent alors dénaturés à 65°C pendant 10 min avant d'ajouter 3 µl de la mixture de bleu.

Les échantillons d'ARN préparés ont alors été migrés sur un gel contenant 1,2% d'agarose avec formaldéhyde. Pour préparer ce type de gel, 0,96 g d'agarose ont été fondus dans 67,7 ml d'eau distillée passée à l'autoclave 1 h. Une fois la température de 65°C atteinte, un mélange contenant 8 ml de MOPS 10X + 4,3 ml de formaldéhyde 37% a été ajouté et l'agarose a été versée dans un plateau de taille moyenne (14 cm X 11 cm) préalablement nettoyée de toute trace de RNases. Une fois l'agarose figé, le plateau fut mis dans une cuve contenant 700 ml de tampon MOPS 1X (avec 0,4 µg/ml de bromure d'éthidium). Cette quantité a permis de remplir la cuve jusqu'à la bordure du gel, sans couvrir les puits. Les échantillons furent alors chargés dans chacun des puits et le courant mis à 80V pour 10 min. Une fois le front de migration sorti des puits, le gel fut recouvert de tampon MOPS 1X (toujours avec 0,4 µg/ml de bromure d'éthidium) et la migration fut de 16 h/22V.

## **19 Transferts de type Northern**

Suite à la migration d'échantillons d'ARN sur gel d'agarose contenant de la formaldéhyde, le gel fut transféré sur une membrane de nylon chargée positivement (N<sup>+</sup>, Amersham). Le transfert, qualifié de type Northern, a été fait exactement tel que suggéré par le catalogue de Roche Diagnostic, Inc., (Roche Molecular Biochemicals, DIG Application Manual for Filter Hybridization, ©2000). Avant le transfert, le gel a été trempé deux fois dans du 20X SSC (300 mM citrate de sodium, 3 M NaCl, pH 7,0 ; autoclavé 1 h), pendant 15 min, afin d'enlever toute trace de formaldéhyde. Le transfert a ensuite été réalisé tel qu'indiqué pour les transferts de Southern, à la différence que deux feuilles de papier Whatman ont été ajoutées sur la membrane, au lieu d'une seule. De plus, la membrane n'a pas été préalablement trempée dans de l'eau déionisée, mais simplement posée sur le gel. Enfin, le tampon employé pour le transfert par capillarité est le 20X SSC au lieu du 10X SSC. Le montage préparé a ensuite été recouvert d'une pellicule plastique et laissé ainsi pendant une nuit.

### 19.1 Hybridation de membrane récupérée d'un transfert d'ARN

Après le transfert de l'ARN sur la membrane de nylon, elle fut récupérée et passé aux UV (UV Stratalinker) afin d'en fixer l'ARN. La membrane fut ensuite installée dans un sac à préhybridation et 20 ml de la solution de préhybridation préchauffée furent ajoutés (il est recommandé d'avoir 20 ml/100 cm<sup>2</sup> de membrane).

Le tampon de préhybridation employé pour une membrane contenant de l'ARN, mais hybridé avec une sonde d'ADN fut un tampon à forte teneur en SDS. Pour le préparer, il a fallu mettre 7 g de SDS dans un erlenmeyer et y ajouter une solution contenant, dans l'ordre :

- 50 ml formamide déionisée ou ultrapure
- 25 ml 20X SSC
- 5 ml Sodium phosphate, pH 7, 0, 1M
- 20 ml Réactif de blocage 10%
- 1 ml N-lauroylsarcosine 10%.

Le réactif de blocage fut préalablement préparé avec 10 g de « Blocking Reagent » (Roche) dans 100 ml de tampon 1 (tampon d'acide maléique, voir le point 7.2 de ce chapitre), et autoclavé. La formamide, si ultrapure, a pu être employée telle quelle. Sinon, il a fallu la préparer en faisant agiter 5 g de résine Amberlite dans 50 ml de formamide pendant 30 à 60 min, puis la filtrer avec du papier Whatman 1 M.

La solution de préhybridation ainsi faite a ensuite été chauffée (à 65°C) en l'agitant doucement, afin d'aider à la dissolution du SDS.

Le sac contenant la membrane et la solution de préhybridation fut ensuite laissé à 50°C pendant 2 à 3 h, dans un four à hybridation muni d'un plateau agitateur. La solution de préhybridation fut ensuite remplacée par 20 ml de solution d'hybridation et le sac fut remis pour la nuit à 50°C. La solution d'hybridation fut une solution de préhybridation préchauffée à 50°C contenant la sonde d'ADN voulue (10 à 20 ng par ml de solution) préalablement dénaturée à 100°C/10 min. Lorsqu'une solution d'hybridation a été réemployée, la dénaturation a été effectuée à 65°C/10 min.

Suite à l'hybridation, la membrane récupérée a été lavée deux fois dans une solution « non-stringente » (2X SSC, 0,1% SDS) pendant 5 min à la température de la pièce, puis deux fois dans une solution plus « stringente » (0,1X SSC, 0,1% SDS), pendant 15 min à 68°C. Par la suite, la révélation et la détection de la DIG furent effectuées comme indiqué

pour les membranes récupérées suite à un transfert de Southern (voir le point 7.2 de ce chapitre).

## 20 Délétion du site *dif* dans *E. coli* :

Pour réaliser la délétion du site *dif* d'*E. coli*, le système de Wanner (Datsenko et Wanner, 2000) a été employé. Tout d'abord, le gène de résistance au Cm du vecteur pKD3 a été amplifié avec les amorces difF1-P1 et difR3-P2. En plus de contenir une région d'homologie avec la cassette de résistance au Cm permettant son amplification, ces amorces contiennent en 5' une région d'homologie avec le site *difEc*. Cette réaction de PCR a été réalisée avec un mélange de 10 : 1 de Taq vs Vent polymérase, dans un appareil à gradient pour déterminer la température d'hybridation adéquate. Le gradient s'échelonnait de 40°C à 70°C ( $\bar{x} = 55^\circ\text{C}/30 \text{ s}$  ; gradient de 30°C) et l'élongation était de 1 min 45 s. Le produit obtenu (de 1,1 kb) à la température de 53,2°C fut purifié. Ce produit a ensuite été digéré avec *DpnI*, et repurifié. Environ 100 ng de ce produit ont été transformés dans la souche BW25113 contenant le plasmide pKD46. La souche BW25113 avait été rendue électrocompétente en ajoutant 50 mM d'arabinose au milieu LB et en faisant croître les cellules à 30°C. La suite du protocole fut tel qu'indiqué dans les autres cas (voir le point 3.2 de ce chapitre). L'arabinose mise dans les milieux sert à induire le système  $\lambda$  Red du pKD46, ce qui facilite la recombinaison homologue d'un fragment linéaire au chromosome bactérien.

Après le choc électrique, les cellules transformées ont été incubées à 28°C/2h30 dans du SOC contenant 25 mM d'arabinose. Elles ont ensuite été étalées sur trois géloses LA + Cm et incubées à 37°C. Après 48 h, une seule colonie fut obtenue. Une réaction de PCR avec les amorces difKpnF et difEcoR a permis de confirmer l'obtention de la souche BW25113 $\Delta$ dif::Cm<sup>R</sup>. De plus, une incubation supplémentaire à 37°C sur milieu contenant de l'ampicilline, a permis de nous assurer que le plasmide pKD46 avait bel et bien été perdu.

Ensuite, pour enlever la résistance au Cm de la nouvelle souche, des cellules compétentes en ont été faites et transformées avec le plasmide pCP20 (contient le gène codant pour la recombinase Flp). Les cellules transformées ont été incubées à 28°C sur milieu LA + Ap, et plusieurs des colonies obtenues ont ensuite subi 4 passages à 37°C pour éliminer le plasmide pCP20. Après ces 4 passages, les colonies ne poussant ni sur LA + Ap

ni sur LA + Cm ont été sélectionnées et vérifiées par PCR avec les amorces difKpnF et difEcoR. Le principe était que, comme le gène de résistance au Cm était flanqué de part et d'autre de deux sites *FRT*, la recombinaison entre ces deux sites par la recombinaison Flp permettait d'enlever le gène Cm<sup>R</sup>. Il ne restait alors qu'un seul site *FRT* dans le génome bactérien. La nouvelle souche ainsi obtenue fut nommée : *BWΔdif*. La Figure 17 montre la zone enlevée dans le génome d'*E. coli*.

CGGAAGTGCAGATGCATGAAATCCTGAGTGACTTTGCCAGAATGATTCCA	50
GCAGCACTGGATAACGTGAAGACTTCATTACCGACAGATTTTCCGGAGAA	100
CGTGGTGACGGCAGTTGAAAGCAATGTGTTGAGGTTGCATGGACGGTTAA	150
GCCGAGAATACGGTAGTAAGTGAGATATGGGGTTTGGTCATTGTTAATGA	200
GCATGACAATCATGACCGCCAACGACTGGATTCTTCCTTGGTTTATATTG	250
GTGCGCATAATGTATATTATGTTAAATCATGTCTGAGTGATGCGAAGTGC	300
TTTTCTGGTAGTCGTTATTCGTTCAAGGTTCAAGAACTTAAGCCTCTCCT	350
CTTTCATAATTGGCCTGTGCATGGAGAAATAGCCTGGTTATTCCCTGCCC	400
CATGCTAAGAGCGCAGTGTCTGTCGCATAGACACTGCATACGAAAACACC	450
AGAGGGTGCTGTTACTTATTTAAGGACGGATCAAGCGGGAATGAGTCCAC	500
TCCGGTATCAGAAATGGCTACGTCTCAATGAAGTCAGGCGACAGATGCTG	550

**Figure 17 : Région entourant le site *dif* chez *E. coli*.** En rouge : site *dif* de 31 pb, de la position 247 à 278. Les autres séquences surlignées indiquent la position des amorces employées : amorces difKpnF, difF22, difR4 et difEcoR. Les amorces difF22 et difR4 ont servi à la fabrication de la sonde *difEc* et au site employé pour les tests de retards sur gel (voir les sections appropriées). La zone délétée de la souche *BWΔdif* est la zone rose encadrée.

## 21 Détection d'un site *dif* potentiel chez *P. mirabilis*.

Une banque génomique de *P. mirabilis* a été réalisée en digérant la totalité de l'ADN génomique extrait par *EcoRV*. Cet ADN, purifié, a ensuite été cloné dans le vecteur pUC19 digéré avec *SmaI*, traité à phosphatase alcaline et purifié. Le produit de cette ligation a ensuite été transformé dans la souche *BWΔdif* (la construction en a été décrite ci-haut en point 20). L'ADN plasmidique de l'ensemble des colonies obtenues fut extrait, fait migré sur un gel d'agarose 1% pendant la nuit à 22V, puis transféré sur une membrane de nylon chargée positivement, avant de procéder à une hybridation avec le site *difEc* marqué à la DIG. Suite au résultat de la révélation, la portion du gel incluant des plasmides variant

de 3,5 à 4 kb fut purifiée et retransformée dans la même souche d'*E. coli*. La dilution  $10^{-1}$  de cette transformation fut étalée à coup de 150  $\mu$ l sur cinq géloses, lesquelles furent ensuite répliquées sur velours.

Afin de vérifier chacune des colonies de la gélose ayant donné un signal positif à une autre hybridation avec le site *difEc*, un méga-mix contenant les amorces, dNTPs, tampon, etc., a été préparé puis réparti dans plusieurs tubes à PCR. L'enzyme employée ici fut la polymérase HotStarTaq de QIAGEN. Une portion de l'étalement des diverses colonies à tester a été mise directement dans l'un des tubes à l'aide d'un cure-dents, puis la réaction de PCR a débuté avec un cycle à 95°C pendant 15 min. Ce préchauffage non seulement est nécessaire à l'activation de la polymérase HotStarTaq de QIAGEN, mais en plus permet l'éclatement des cellules et une bonne dénaturation des brins d'ADN chromosomique. La suite du programme était « classique » avec une hybridation à 53°C et une élongation de 1 min 40 s. Quelques réactions furent vérifiées sur gel d'agarose, puis 10  $\mu$ l de chacune des réactions furent déposées sur une membrane de nylon (N<sup>+</sup>, Amersham) d'environ 15 X 20 cm, après leur dénaturation à 95°C/15 min. Suite à l'hybridation de cette membrane, toujours avec le site *difEc* marqué à la DIG, trois forts signaux et un autre plus faible sont apparus. Ces quatre clones ont été envoyés à séquencer (service de séquençage du CHUM, Montréal).

# RÉSULTATS

## Chapitre III : Résultats

Les résultats obtenus au cours de cette thèse sont divisés en deux parties. La première partie concerne l'étude qui a suivi la découverte et le séquençage du gène *xerD* de *P. mirabilis* et qui concernait alors le sujet de mon mémoire de maîtrise (Villion, 1998). Cette première partie, traitant de la protéine XerC de la même espèce, a d'ailleurs donné lieu à une publication dans FEMS Microbiology Letters en septembre 2003 (Villion et Szatmari, 2003).

La seconde partie concerne la suite de ces résultats, en portant l'accent sur l'espèce *P. mirabilis* en tant que telle. Ainsi, une série d'expériences a mené à la découverte du site *dif* de cette espèce, le phénotype de souches mutantes pour les gènes *xer* a été analysé brièvement et des premières études d'expression ont été réalisées. La discussion reprendra chacun des éléments dans son ensemble.

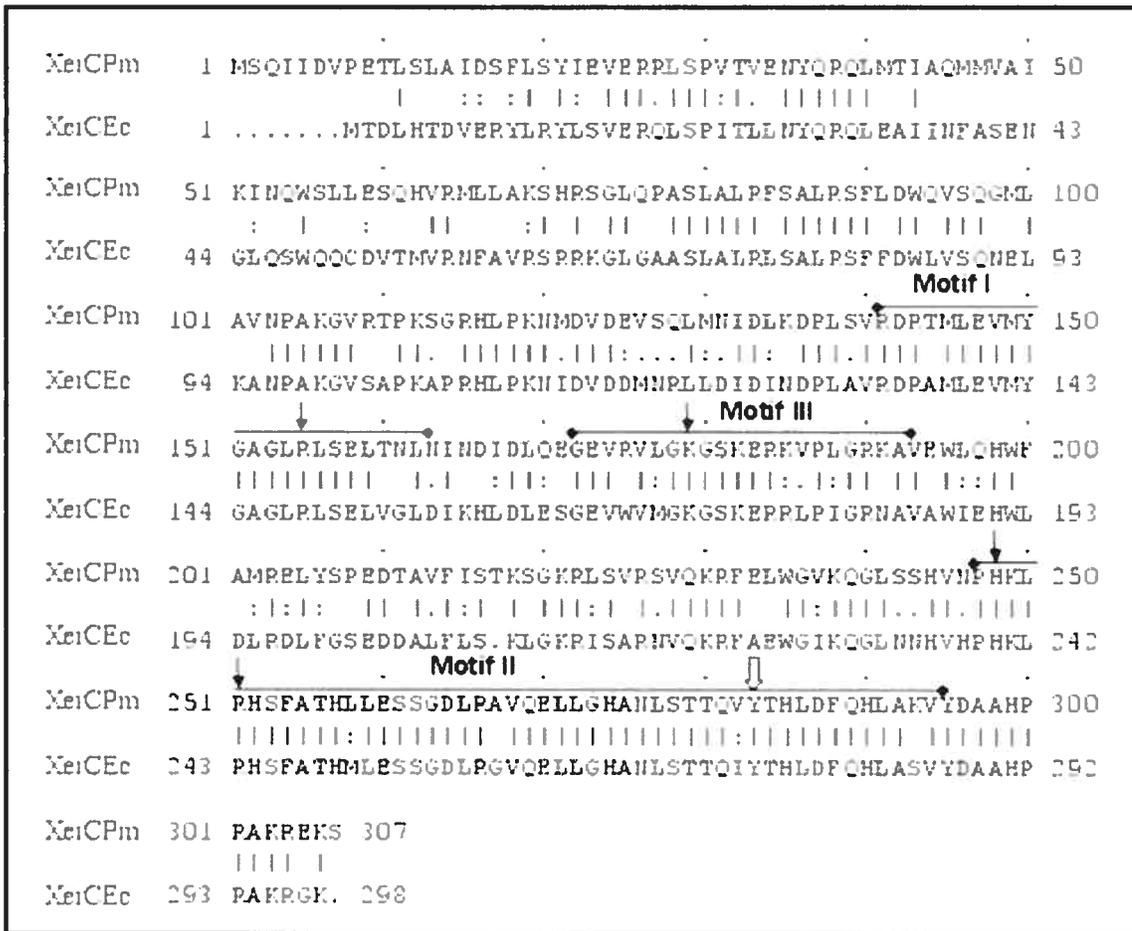
### 1 Première partie

#### 1.1 Obtention et analyse de la séquence XerCPm

Une première étude employant un vecteur surenroulé contenant des sites de recombinaison spécifiques avait laissé supposer que *P. mirabilis* devait posséder un système homologue au système Xer. En effet, un produit de résolution était obtenu suite à la transformation du vecteur dans cette espèce (Sirois et Szatmari, 1995). Notre étude subséquente a d'ailleurs permis de confirmer la présence de *xerD* (Villion et Szatmari, 1998).

Afin de rechercher un homologue de *xerC*, nous avons employé la même méthode que celle employée pour la recherche de *xerD*. Ainsi, les mêmes amorces dégénérées nous ont permis d'amplifier un fragment correspondant à la portion C-terminale conservée des gènes *xer*. Nous avons ainsi amplifié un fragment de 350 pb différent du précédent. Une fois la séquence obtenue, des amorces spécifiques furent synthétisées et un PCR inverse a été réalisé afin de compléter la séquence par les portions immédiatement en amont et en aval du fragment. Finalement, la séquence entière a été réamplifiée par PCR direct avec de nouvelles amorces pour inclure ces régions flanquante. Cette séquence entière obtenue a

permis de révéler un ORF (cadre de lecture ouvert) fortement similaire au gène *xerC* d'*E. coli*. Nous l'avons ainsi nommé *xerCPm* et envoyé à GenBank (numéro : AF033498).



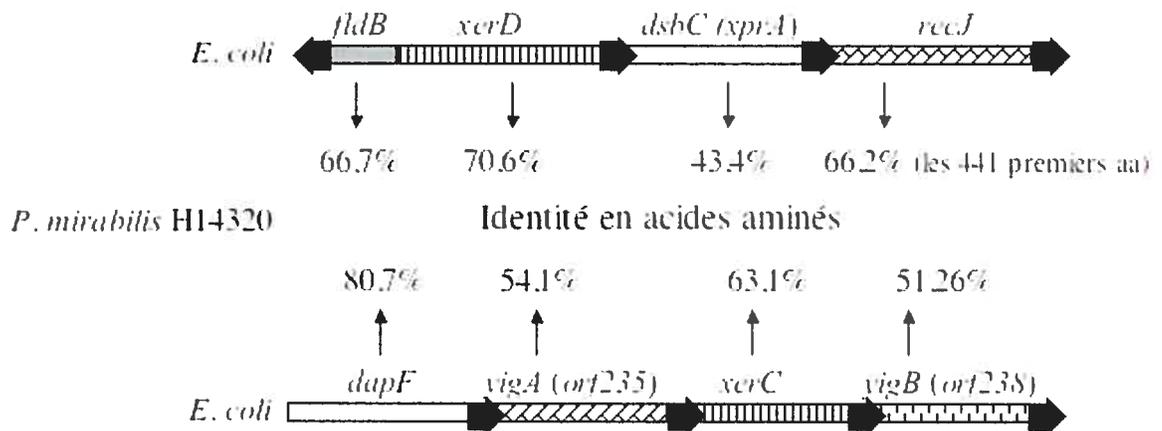
**Figure 18: Comparaison des séquences en acides aminés des XerC d'*E. coli* (XerEc) et de *P. mirabilis* (XerCPm).** L'alignement de type « GAP » a été réalisé avec le programme de « GCG Wisconsin Package ». La protéine XerCPm montre 63% d'identité avec XerCEc. Les flèches noires montrent les résidus conservés de la famille des recombinases de la famille de l'Int $\lambda$  (R-H-R et K), tandis que la flèche ouverte montre la tyrosine catalytique. Les motifs I, II et III conservés sont tels qu'ils ont été présentés par Cao *et al.* (1997). Enfin, les lignes verticales (|) indiquent une identité, les deux points (:), un très fort degré de similarité, et le point (.), un degré plus faible de similarité entre les acides aminés des deux protéines.

La séquence protéique déduite de *xerCPm* est identique à 63,1% avec la protéine XerC d'*E. coli* (voir la Figure 18), et 56,9% avec XerC d'*H. influenzae*. Tout comme c'est le cas pour les protéines Xer d'*E. coli*, XerCPm présente 37% d'identité avec sa protéine partenaire XerDpm. Bien entendu, la région la plus conservée est la région C-terminale qui contient les trois motifs retrouvés chez tous les membres de la famille des intégrases du

phage  $\lambda$  ; cette région présente 71,9% d'identité avec celle retrouvée chez XerCEc, tandis que le domaine N-terminal ne présente que 47,2% d'identité.

### 1.1.1 Organisation des gènes *xerC* et *xerD* chez *P. mirabilis*

Depuis 2004, une souche de *P. mirabilis* est en train d'être séquencée par l'Institut Sanger (voir le site web : [http://www.sanger.ac.uk/Projects/P\\_mirabilis/](http://www.sanger.ac.uk/Projects/P_mirabilis/)) : la souche H14320. Le génome de cette souche ferait environ 3,9 Mb et posséderait un pourcentage en G + C de 30 à 40%. Grâce à ce séquençage, il est maintenant possible de confirmer que l'organisation des gènes *xer* est vraiment similaire à celle retrouvée chez *E. coli*. La Figure 19 montre l'organisation des gènes d'*E. coli* avec les pourcentages d'identité des protéines avec celles de *Proteus mirabilis*. Ainsi, le gène *xerCPm* est flanqué de deux *orf* fortement similaires à *yigA* et *yigB*, tandis que le gène *dsbC* suit le gène *xerDPm*. Tout comme chez *E. coli*, un gène similaire à *fldB* (encodant la flavodoxine II) est situé immédiatement en amont de *xerDPm*, et serait transcrit dans le sens inverse de *xerD*.



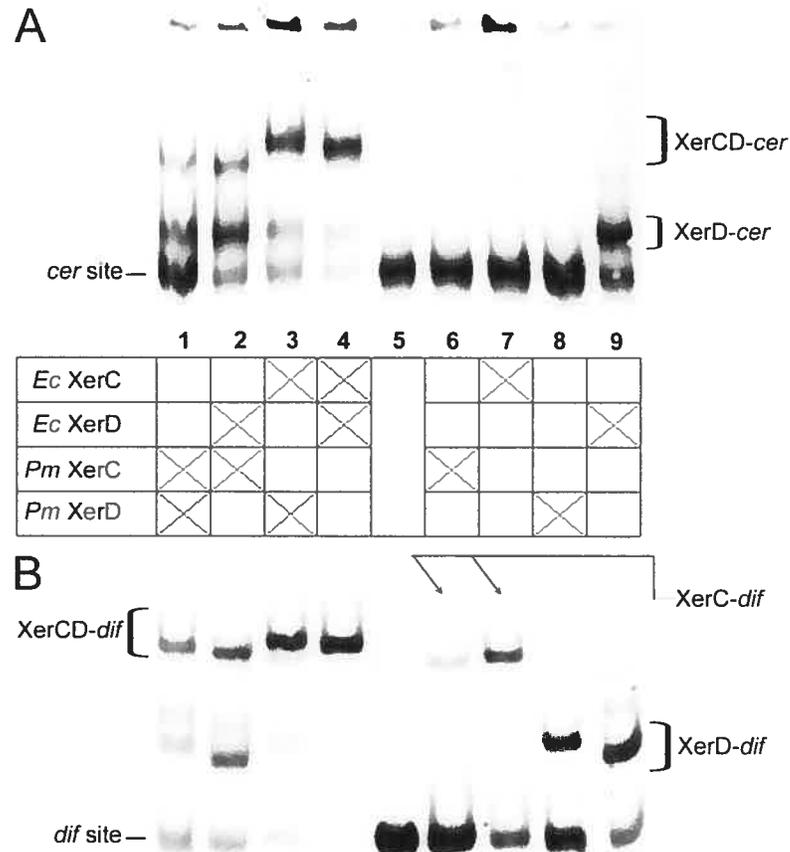
**Figure 19 : Organisation des gènes *xerC* et *xerD* de *P. mirabilis*, souche H14320 comparée à celle retrouvée chez *E. coli*. Le pourcentage d'identité de chacune des protéines des deux espèces est indiqué.**

## 1.2 Expériences de retard sur gel

Afin de vérifier si les protéines de *P. mirabilis* pouvaient bel et bien se lier aux sites de recombinaison spécifique, des tests de retard sur gel ont été réalisés avec les sites *cer* et *dif* marqués à la digoxygénine. En effet, il était présumé que tel devait en être le cas compte tenu de la similarité des séquences des recombinases de *P. mirabilis* et d'*E. coli*, et de la détection de la recombinaison au site *cer* dans l'espèce étudiée (Sirois et Szatmari, 1995).

Les tests de retard ont été réalisés avec la protéine XerCPm fusionnée à la protéine MBP (protéine se liant au maltose ; maltose binding protein). Malgré la taille de cet ajout (42 kDa), des études précédentes avaient montré qu'il n'empêchait pas la liaison spécifique à l'ADN (Arcieszewska et Sherratt, 1995). De plus, la protéine MBP seule n'a pas montré de capacité de liaison aux sites spécifiques employés (données non montrées).

Comme il est montré à la figure Figure 20A, la protéine MBP-XerCPm n'a pas pu se lier au site *cer* seule, tout comme ce fut le cas pour la protéine MBP-XerCEc (puits 6 et 7). Cependant, ce n'est pas le cas pour la protéine XerDEc (puits 9 ; ceci avait été déjà démontré par Blakely *et al.*, 1993). Dans notre cas, la protéine His-XerDPm ne semble pas se lier seule au site *cer* avec les concentrations employées ici, mais nous avons déjà démontré que cette protéine pouvait bel et bien s'y lier (Villion et Szatmari, 1998).



**Figure 20: Tests de liaison aux sites *difEc* et *cer*.** Les différentes recombinases purifiées (600 nM) ont été mises en présence du fragment d'ADN marqué à la DIG, pendant 10 min/37°C, à l'exception du puits 5 qui représente la migration du site testé (*cer* en A ou *dif* en B) en l'absence de recombinase. Un mélange de XerC et XerD provenant de *P. mirabilis* et/ou d'*E. coli* a été mis dans les puits 1 à 4, tandis que XerC ou XerD de chacune des espèces ont été testées seules en présence du site dans les puits 6 à 9.

XerCPm peut se lier au site *cer*, mais seulement en conjonction avec la protéine His-XerDPm ou XerDEc, comme on peut le voir par l'apparition de fragments situés plus haut dans le gel (Figure 20A, puits 1 et 2). On peut remarquer cependant que l'affinité de MBP-XerCPm pour le site *cer* semble moins élevée que celle de MBP-XerCEc lorsqu'elle est en conjonction avec l'une ou l'autre des protéines XerD (Figure 20A, puits 1 et 2 à comparer avec les puits 3 et 4 où il reste moins de site *cer* non lié). De plus, les tests de retard montrent une bande forte au niveau de ce qui représente la liaison de XerD seule au site *cer*, dans les quatre cas de mixtures de protéines XerC avec XerD.

Des résultats similaires ont été obtenus au niveau du site *difEc*. Cependant, à la différence du site *cer*, les protéines XerC ont pu s'y lier seules ; MBP-XerCPm s'y est liée par contre un peu moins efficacement que MBP-XerCEc (Figure 20B, puits 6 et 7). Cette différence est vue aussi lorsque les mixtures XerC-XerD sont employées : alors que les retards faits avec MBP-XerCEc (que ce soit avec His-XerDPm ou XerDEc, Figure 20B, puits 3 et 4) montrent un site *difEc* presque totalement retardé, contenant les deux protéines, une très forte bande indicatrice d'un retard par XerDEc apparaît lorsque la mixture contient MBP-XerCPm et XerDEc (Figure 20B, puits 2). Cette bande est, par contre, moins prononcée lorsque les deux protéines de *P. mirabilis* sont mises ensemble (Figure 20B, puits 1).

Dans les puits 1 et 2, une seconde bande apparaît au-dessus de la bande correspondant à un retard fait par la protéine XerD seule au site *difEc* : ce retard pourrait être dû à la liaison de deux XerD, chacun d'eux situé de part et d'autre de la région d'embranchement du site *difEc*. Il pourrait également correspondre à une interaction réalisée entre deux monomères de XerD liés au site habituel de liaison de la protéine.

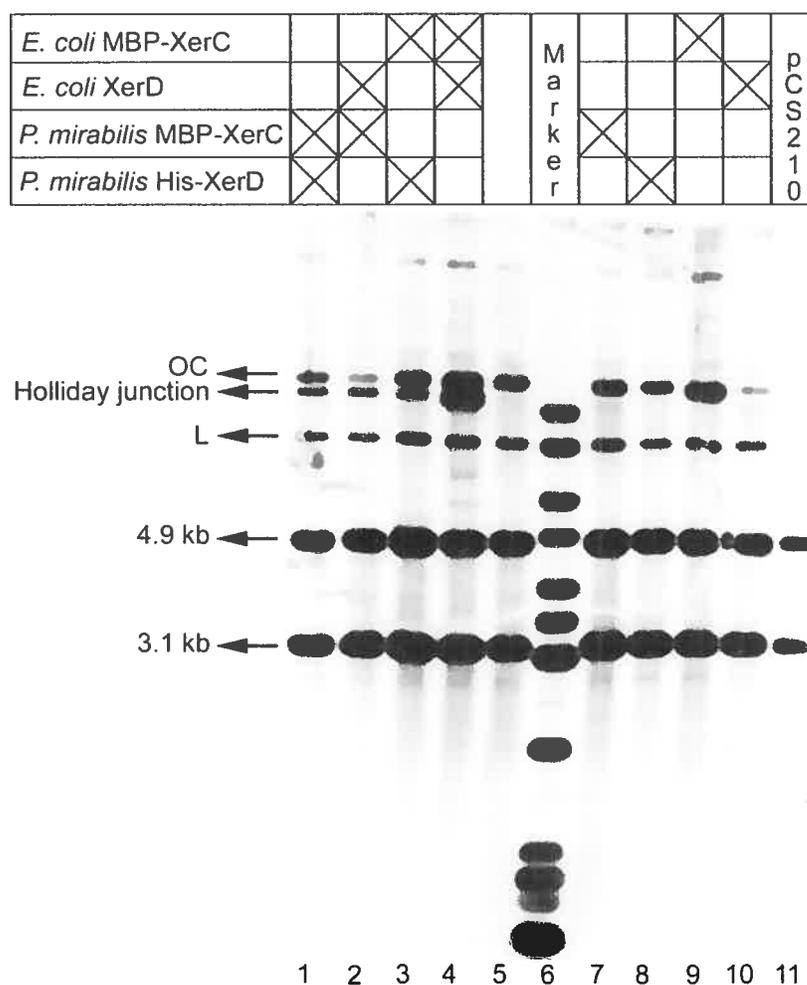
### 1.3 Réactions *in vitro*

#### 1.3.1 Au site *cer* de ColE1

Il a déjà été montré que les recombinaisons d'*E. coli*, en conjonction avec les protéines accessoires ArgR et PepA, donnent lieu à la formation d'une jonction de Holliday suite à la recombinaison entre deux sites *cer* situés sur un plasmide (Colloms *et al.*, 1996). Après la digestion de l'ADN plasmidique ayant subi l'action des recombinaisons, cette

jonction peut être différenciée de la forme linéaire. En effet, cette forme migre plus lentement que la forme linéaire dans un gel d'agarose.

Ainsi, afin de mieux tester l'activité biologique des protéines recombinases de *P. mirabilis*, des réactions de recombinaison *in vitro* ont été réalisées avec le pCS210. Ce vecteur contient deux sites *cer* flanquant le gène *lacZ* (Stirling *et al.*, 1989). Puisque les protéines accessoires ArgR et PepA n'ont pas encore été caractérisées chez *P. mirabilis*, nous avons purifié les protéines accessoires d'*E. coli* afin de compléter la réaction avec les recombinases purifiées MBP-XerCPm et His-XerDPm ou XerDEc. Avant l'analyse sur gel, l'ADN plasmidique a été linéarisé avec l'enzyme *EcoRV* (cette enzyme clive une fois entre les deux sites *cer*, dans le gène *lacZ*, et une fois dans le gène codant pour la résistance à la tétracycline de pCS210) pour donner un fragment de 4,9 kb et un autre de 3,1 kb.



**Figure 21: Réactions *in vitro* au site *cer* de ColE1.** En plus des protéines Xer (environ 300 nM) qui ont été ajoutées dans chacune des réactions (tel qu'indiqué), toutes les réactions ont été réalisées avec les protéines accessoires ArgR (30 nM) et PepA (300 nM) d'*E. coli*. Le puits 5 contient le pCS210 dans le tampon de réaction contenant seulement les protéines accessoires, tandis que le puits 11 contient le pCS210 extrait de la souche DS9029 sans qu'il n'ait subi d'autres traitements. Le puits 6 représente les fragments du marqueur Gene Ruler (Fermentas) qui se sont hybridés à la sonde *cer*. Les puits 1 à 4 montrent les réactions ayant contenu un mélange de XerC et XerD provenant d'*E. coli* et/ou de *P. mirabilis*, tandis que les puits 7 à 10 montrent les réactions avec les protéines XerC ou XerD seules. La jonction de Holliday potentielle est identifiée à gauche de la figure (Holliday junction). Les formes, circulaire ouverte (OC : open circular) et linéaire (L), sont également indiquées.

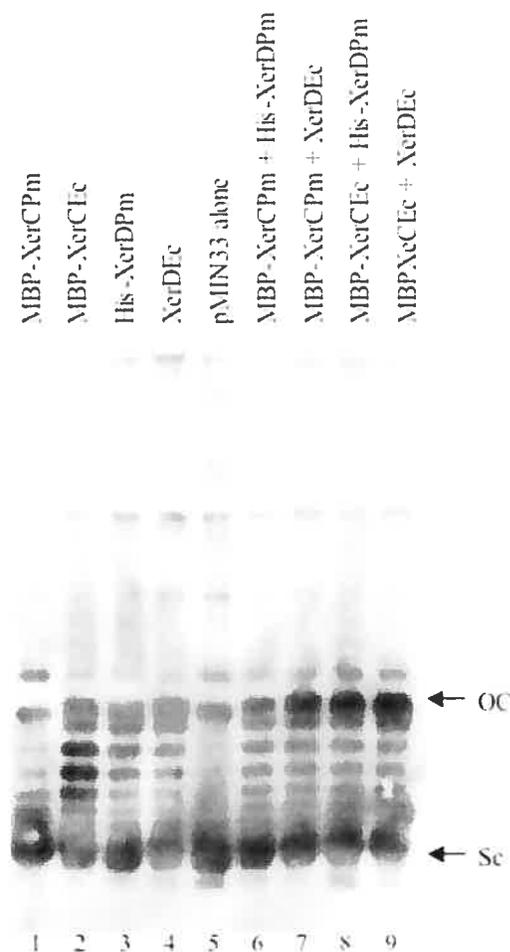
Suite à la réaction *in vitro*, un fragment a été obtenu à la hauteur de migration d'une jonction de Holliday et ce, peu importe la mixture de protéines employée (voir la Figure 21, puits 1 à 4). Ainsi, les recombinases de *P. mirabilis* ont pu exercer leur action de clivage au niveau du site *cer* avec les protéines d'*E. coli*. Comme attendu, les protéines seules (XerC ou XerD) n'ont pas montré la formation de cette jonction (Figure 21, puits 6 à 10), ce qui indique que la première réaction de clivage nécessite bien la présence des deux

protéines (il est à noter que la formation de cette jonction a été observée à au moins deux reprises avec les recombinaisons de *P. mirabilis* dans des expériences séparées).

Au niveau du puits 5 de la Figure 21, il semblerait qu'il y ait eu une légère dégradation du pCS210. Cette dégradation a probablement été causée par des nucléases contaminantes provenant des fractions de ArgR ou PepA, puisqu'elle n'apparaît pas lorsque le pCS210 n'a pas été mis dans le tampon de réaction contenant ces deux protéines (voir alors le puits 11). Cependant, cette dégradation n'a pas empêché de détecter la jonction de Holliday dans tous les cas où les protéines XerC et XerD ont été ajoutées à la réaction. Enfin, on peut voir que la réaction ayant donné le plus haut taux de ces jonctions fut avec les recombinaisons d'*E. coli* (Figure 21, puits 4).

### 1.3.2 Au site *dif* d'*E. coli*

Il a été montré que les recombinaisons XerC et XerD d'*E. coli* possèdent une activité topoisomérase de type I spécifique de site (Cornet *et al.*, 1997). Comme il a été mentionné dans la partie introduction de cette thèse, cette activité topoisomérase a aussi été montrée chez d'autres membres de la famille des intégrases du phage  $\lambda$ , comme Cre (Abremski *et al.*, 1986), et l'intégrase du phage  $\lambda$  (Kikuchi et Nash, 1979). Cette activité est reflétée par l'apparition de topoisomères suite à l'action d'une recombinaison, via leur tyrosine catalytique, sur un vecteur contenant le site *difEc* comme le vecteur pMIN33.



**Figure 22: Réactions *in vitro* au site *difEc*.** Les protéines impliquées dans chacune des réactions sont indiquées au haut de chaque puits. Les réactions ont été réalisées à 37°C pendant 14h30 en présence de pMIN33 (vecteur contenant un site *dif* d'*E. coli*). Les topoisomères se situent entre la forme complètement relaxée du plasmide et la forme surenroulée (Sc). OC : forme circulaire ouverte (open circular) de pMIN33.

Cette activité de clivage a également été montrée avec les recombinases de *P. mirabilis*, comme on peut le voir à la Figure 22 (puits 2 et 4), tout comme ce fut le cas pour les recombinases d'*E. coli* (Cornet *et al.*, 1997, et voir le puits 1 et 3). Cette activité fut non seulement détectée après 14h30, mais également après 3h d'incubation à 37°C avec le plasmide pMIN33 (mais alors, les topoisomères apparaissaient plus faiblement; résultat non montré).

Cependant, contrairement à ce qui a été observé avec les recombinases d'*E. coli* (Cornet *et al.*, 1997), nous pouvons noter que les topoisomères ont été obtenus même en présence des deux recombinases partenaires, et ce, à plus d'une reprise (Figure 22, puits 9, et données non montrées). Ainsi, avec nos purifications, le phénomène d'inhibition

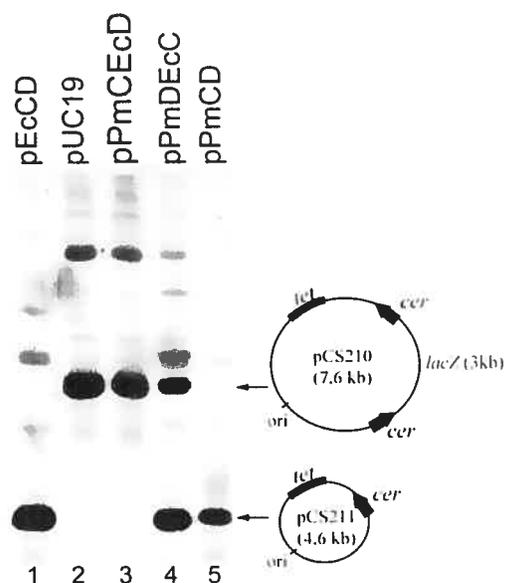
réciproque n'a pas été observé, et ce, peu importe la combinaison des recombinaises employées (Figure 22, puits 6 à 9).

#### 1.4 Tests de complémentation *in vivo* (dans *E. coli xerCD*<sup>-</sup>)

Ainsi, nous avons montré que les recombinaises de *P. mirabilis* étaient actives sur les sites *cer* et *difEc in vitro*. Nous avons ensuite voulu vérifier l'activité biologique de XerCPm par un test de complémentation *in vivo*. Un test similaire avait été réalisé pour tester l'activité de XerDPm; il avait alors été montré qu'un clone contenant le gène *xerDPm* pouvait compléter une souche d'*E. coli xerD*<sup>-</sup>, mais un peu moins efficacement que le clone contenant le gène *xerD* d'*E. coli* (Villion et Szatmari, 1998).

Le test reproduit ici a consisté à mesurer le taux de délétion du gène *lacZ* (flanqué de deux sites *cer*) de pCS210, dans une souche d'*E. coli xerC*<sup>-</sup>*xerD*<sup>-</sup> transformée avec des clones contenant les gènes *xerC* et *xerD* de *P. mirabilis* et/ou d'*E. coli* (voir la section matériel et méthodes pour de plus amples détails). Lorsque la recombinaison survient entre les deux sites *cer* du plasmide, le gène *lacZ* est perdu, et nous pouvons constater *in vitro* la délétion de 3 kb du plasmide.

Malgré le degré de similarité entre les espèces *P. mirabilis* et *E. coli* selon les études phylogénétiques, une souche d'*E. coli* mutante pour le gène *xerC* n'a pas pu être complétée par un clone contenant le gène *xerCPm* afin de résoudre le pCS210 (données non montrées de Villion et Szatmari, 1998). Dans cet essai, le gène *xerDEc* était exprimé sur le chromosome, tandis que le gène *xerC* était exprimé sur un plasmide multicopie. Afin de nous assurer que la recombinaison sur le pCS210 n'était pas influencée par cette différence de niveau d'expression entre les deux protéines partenaires, un autre test de complémentation a été réalisé dans une souche d'*E. coli* mutante pour les deux gènes *xer* (souche DS9029) et contenant le pCS210. Cette souche a ensuite été transformée avec le clone pPmCEcD contenant les gènes *xerCPm* et *xerDEc*. Comme il avait été observé précédemment, seul le pCS210 non résolu fut détecté (voir la Figure 23, puits 3), comme ce fut le cas lorsqu'il n'y avait pas de recombinaises ajoutées dans la souche (Figure 23, puits 2, transformation avec le pUC19).



**Figure 23: Tests de complémentation dans *E. coli xer<sup>-</sup>* : extractions plasmidiques.** Les plasmides ayant les gènes *xer* ont été transformés dans la souche DS9029 (*E. coli xerC<sup>-</sup> xerD<sup>-</sup>*) contenant le vecteur pCS210. Les plasmides ont ensuite été extraits le premier jour après la transformation (jour 0) dans le milieu sélectif, et environ 400 ng de chaque préparation plasmidique ont été séparés sur un gel contenant 1% d'agarose. Après avoir été transféré sur une membrane chargée positivement, le pCS210 a été détecté par le gène *tet* marqué à la DIG. Pour la désignation de chacun des plasmides, voir le **Tableau IV** de la section matériel et méthodes. Le premier puits sert de contrôle positif (les gènes *xerC* et *xerD* sont d'*E. coli*), le puits 2, de contrôle négatif (vecteur seul), tandis que les puits 3 à 5 représentent le pCS210 obtenu des souches contenant des clones avec *xerCPm* et *xerDEc* (puits 3), *xerCEc* et *xerDPm* (puits 4) et *xerC* et *xerD* de *P. mirabilis* (puits 5).

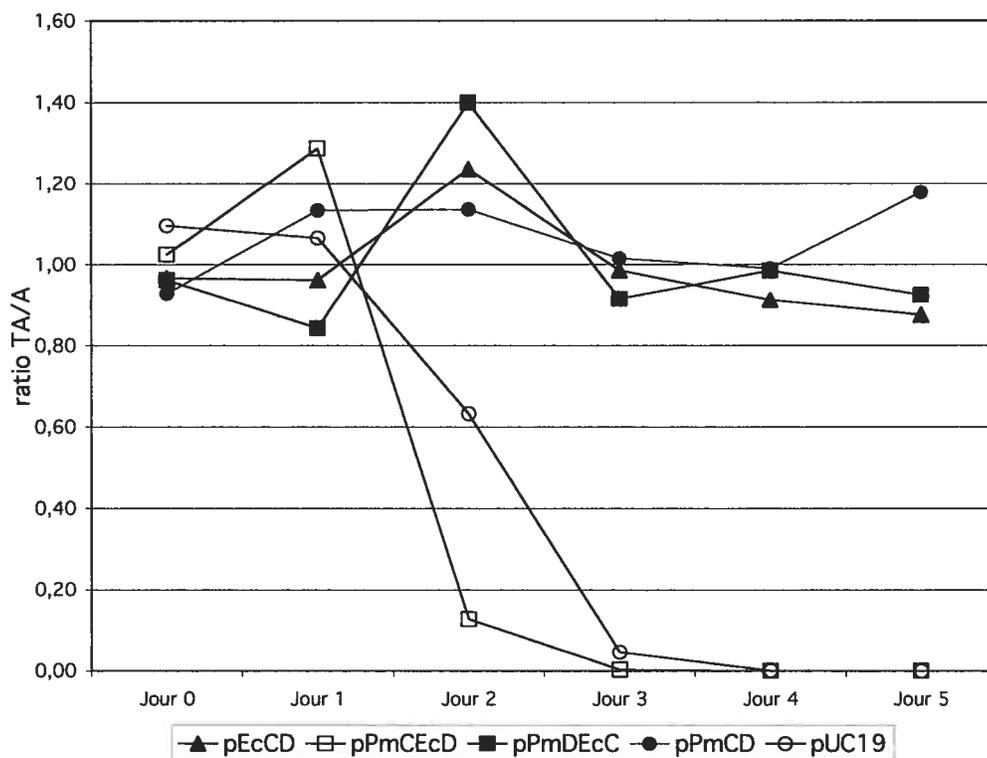
Ce résultat indique que XerCPm n'a pas pu agir en collaboration avec XerDEc afin de résoudre le pCS210. Cependant, un clone contenant les deux gènes provenant de *P. mirabilis* (pPmCD) transformé dans la souche DS9029 contenant le pCS210 a permis la résolution de ce vecteur (Figure 23, puits 5), et ce, de façon aussi efficace que le clone contenant les deux gènes *xer* provenant d'*E. coli* (Figure 23, puits 1, avec le clone pEcCD).

Finalement, un clone contenant les gènes *xerDPm* et *xerCEc* (pPmDEcC) a montré un résultat similaire à ce qui a déjà été publié (Villion et Szatmati, 1998) : le pCS210 a été résolu, mais de façon partielle. Une population mixte de vecteur original et du vecteur résolu a été obtenue (Figure 23, puits 4).

### 1.5 Études de stabilité plasmidique

Afin de déterminer si les recombinases de *P. mirabilis* sont aptes à remplir leur rôle pour le maintien des plasmides ayant un site de recombinaison spécifique, la stabilité du

pCS210 a été mesurée en l'absence de pression de sélection. Les souches contenant le pCS210 et les différents clones (contenant les gènes *xer*) décrits dans la section précédente ont été mises en culture dans un milieu non sélectif pour le pCS210 (sans tétracycline), et laissées ainsi jusqu'à la croissance d'une centaine de générations. Ainsi, la seule façon supposée pour le pCS210 d'être maintenu dans la souche est par l'action du système Xer sur le site *cer*. Tout au long d'une période de 5 jours, le ratio entre le nombre de colonies retrouvées sur le milieu sélectif pour le pCS210 et le nombre total d'unités formatrices de colonies (CFU) retrouvées sur le milieu non sélectif, a été calculé et mis sur un graphique (voir la Figure 24).



**Figure 24: Tests de stabilité plasmidique.** Les plasmides ayant les gènes *xer* ont été transformés dans la souche DS9029 (*E. coli xerC<sup>-</sup> xerD<sup>-</sup>*) contenant le vecteur pCS210, puis les cultures se sont faites dans un milieu non sélectif pour le pCS210. Le graphique représente la stabilité du pCS210 selon le plasmide transformé dans cette souche, pendant une période allant jusqu'à 5 jours (100 générations). Le ratio TA/A est le nombre d'unités viables (UFCs) obtenues sur milieu sélectif (contenant Tc) versus le nombre d'UFCs obtenues sur un milieu non sélectif pour le pCS210 (Ap seulement). Ce graphique constitue la moyenne de deux expériences différentes. Le plasmide pEcCD contient les gènes *xerC* et *xerD* d'*E. coli*; pPmCEcD, *xerCPm* et *xerDEc*; pPmDEcC, *xerCEc* et *xerDpm*; et pPmCD, *xerC* et *xerD* de *P. mirabilis*. De plus amples détails sont retrouvés dans le texte de cette section.

En parallèle, les colonies obtenues sur le milieu sélectif ont été étalées sur les milieux sélectifs et non sélectifs, et la différence de croissance entre les deux milieux a été évaluée. Cependant, comme les deux méthodes ont donné des résultats similaires, les données de cette deuxième méthode ne sont pas montrées ici.

Comme on peut le voir sur le graphique présenté en Figure 24, la souche d'*E. coli* dans laquelle les deux gènes *xer* ont été mutés et non remplacés n'a pas pu maintenir le pCS210 sur plus de 60 générations (ligne contrôle faite avec le pUC19). Ce résultat a été confirmé par l'analyse d'un transfert de Southern (non montré), où le pCS210 n'était plus détectable après 3 jours de croissance. D'un autre côté, la souche DS9029 (mutante pour les deux gènes *xer*), a pu conserver le pCS210 dans sa forme résolue (plasmide nommé pCS211) pour plus de 100 générations lorsqu'un vecteur contenant les deux gènes *xer* d'une même espèce y a été ajouté (voir la Figure 24, données avec le pEcCD et le pPmCD).

La souche contenant le vecteur pPmDEcC (contenant les gènes *xerDPm* et *xerCEc*) a également pu maintenir le pCS211 durant la croissance d'une centaine de générations. Cependant, le clone contenant les gènes *xerCPm* et *xerDEc* (pPmCEcD), non seulement n'a pas pu résoudre le pCS210 (vu par transfert Southern, non montré, et aussi en Figure 23 au jour 0), mais n'a pas pu le maintenir. En effet, il a été perdu presque entièrement après deux jours de croissance.

## 2 Deuxième partie

Nous avons, dans la deuxième partie de cette étude, construit des souches de *P. mirabilis* inactives pour les gènes *xerC* et/ou *xerD*. Les détails de ces constructions sont schématisés aux Figures 15 et 16 de la section matériel et méthodes. Brièvement, nous avons inactivé ces gènes par l'insertion d'un vecteur permettant une fusion d'opéron avec *lacZYA* à l'intérieur même des deux gènes : le vecteur pFUSE. Ce vecteur a pu être transféré dans *P. mirabilis* par conjugaison avec la souche d'*E. coli* donatrice S17.1 $\lambda$ pir. Les souches ainsi obtenues, ont été nommées PmD1 (pour la souche *xerD* ::*lacZYA*-Cm<sup>R</sup>), PmC8 (pour la souche *xerC* ::*lacZYA*-Cm<sup>R</sup>) et Pm2DClacZ (pour la souche préalablement inactivée pour *xerD* et ensuite inactivée par l'insertion d'un vecteur permettant la fusion d'opéron avec *lacZYA* dans *xerC*).

## 2.1 Tests de stabilité plasmidique dans *P. mirabilis*

Outre la vérification des constructions par réactions PCR avec des amorces indiquant l'insertion du gène *lacZ* à la suite de la portion N-terminale des gènes *xer*, nous avons confirmé l'inactivité des gènes *xer* des souches PmD1 et PmC8 à l'aide du vecteur pCS210Km<sup>R</sup>. Ce vecteur est dérivé du pCS210 déjà employé dans les études de stabilité plasmidique (voir plus haut et Villion et Szatmari, 2003). Cependant, l'espèce *P. mirabilis* étant naturellement résistante à la tétracycline, nous avons dû insérer une autre cassette de résistance dans le vecteur, la kanamycine. Comme déjà mentionné, ce vecteur contient deux sites *cer* flanquant le gène *lacZ*. Bien entendu, les souches de *P. mirabilis* vérifiées ici étant déjà *lac*<sup>+</sup>, nous ne pouvions pas employer le phénotype bleu-blanc sur géloses contenant du X-gal, pour une confirmation rapide de la perte ou non de ce *lacZ*. Par contre, une extraction plasmidique nous a permis de constater que le pCS210Km<sup>R</sup> est demeuré intact dans les deux souches construites, c'est-à-dire qu'aucun événement de recombinaison n'a eu lieu entre les deux sites *cer* (non montré). Cependant, nous avons observé que le vecteur a eu tendance à multimériser, contrairement à ce que l'on avait pu observer chez *E. coli*.

Des tests de stabilité plasmidique ont par la suite été tentés dans les souches de *P. mirabilis* inactivées pour *xerC* ou *xerD* (PmC8 et PmD1) et contenant le vecteur pCS210Km<sup>R</sup>.

Ainsi, nous avons transformé les vecteurs pMVC6 (*xerCPm/pTZ-PC*), pSDC105 (*xerCEc/pBAD*) et pTZ-PC (contrôle négatif) dans la souche PmC8 pCS210Km<sup>R</sup>, et les vecteurs pMVD37 (*xerDPm/pTZ-PC*), pRM130 (*xerDEc/pUC18*) et pTZ-PC (contrôle négatif) dans la souche PmD1 pCS210Km<sup>R</sup>. Nous voulions ainsi étudier la complémentation de la mutation *xer* dans *P. mirabilis*, puis vérifier la stabilité du pCS210Km<sup>R</sup> dans l'environnement de ces souches, tout comme il avait été fait dans *E. coli* (voir la première partie des résultats, points 1.4 et 1.5).

Malheureusement, après une nuit à 37°C, aucune croissance n'a été obtenue sur les milieux contenant la Km (antibiotique pour la sélection du vecteur étudié), comme si le pCS210Km<sup>R</sup> avait été perdu après une seule nuit en bouillon sans kanamycine, et ce, peu importe la souche. Après deux jours, quelques colonies finirent par apparaître sur ces géloses, mais cela laissait plutôt suggérer des mutations spontanées. Un autre test, fait de la

même façon, a donné exactement le même résultat. Nous n'avons donc pas pu poursuivre l'expérience.

Ainsi, le pCS210Km<sup>R</sup> n'a pu être maintenu, et ce, même dans la souche sauvage. Il se pourrait qu'un autre système impliqué dans la stabilité plasmidique d'un vecteur contenant un site reconnu par Xer soit absent ou non fonctionnel dans *P. mirabilis*.

## 2.2 Phénotype des souches de *P. mirabilis xerC* et/ou *xerD*

### 2.2.1 Essaimage (swarming)

*P. mirabilis* est connu pour son comportement migratoire caractéristique sur une gélose d'agar. Cette espèce, extrêmement mobile, a tendance à effectuer ce que l'on appelle de l'essaimage (plus connu sous le nom anglais de « swarming »). Ce processus multicellulaire se caractérise par divers cycles de différenciation, migration puis consolidation sur une gélose d'agar, lesquels donneront lieu à des « vagues » visibles au fur et à mesure que l'espèce envahit la gélose entière (Rauprich *et al.*, 1996 et voir le point 7.2.2 du chapitre d'introduction). Suivant la mutation de certains gènes, ce phénomène peut-être modifié et alors, le patron de migration s'en voit changé. Les anneaux de consolidation peuvent être dispersés ou plus rapprochés, ou encore présenter un caractère dendritique. Bien entendu, certaines mutations peuvent tout simplement empêcher la différenciation des cellules en cellules d'essaimage (Fraser *et al.*, 2000).

Ainsi, nous avons voulu savoir si les mutations dans les gènes *xer* pouvaient engendrer une modification du patron de migration chez *P. mirabilis*. Nous avons alors déposé une goutte de 5 µl d'une culture de nuit sur une gélose contenant 1,5% d'agar et du TTC (2,3,5-triphenyltetrazolium; dilution finale de 1 :10 000)<sup>5</sup> et nous l'avons laissé croître à 28°C. Après 24 h, les couleurs obtenues sur chacune des géloses étaient plutôt pâles, mais après avoir laissé ces géloses à 4°C pour une autre période de 24 h, le contraste en a été augmenté, et nous avons pu prendre les photos (voir la Figure 25).

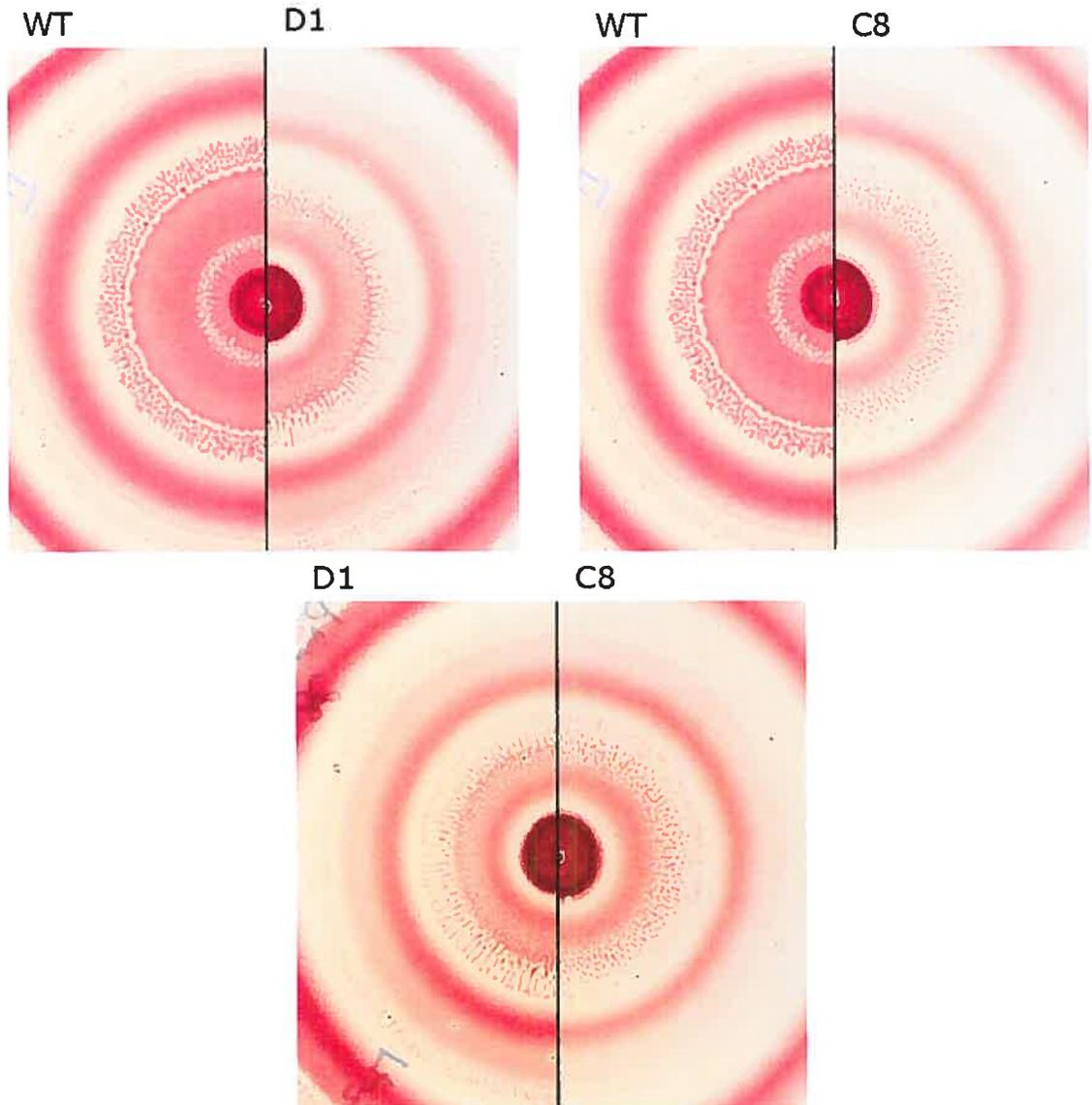
---

<sup>5</sup> Le TTC, également nommé « tetrazolium red », est un colorant non sélectif utilisé surtout dans le milieu m-colibblue24 (Millipore) permettant de détecter spécifiquement les cellules d'*E. coli* (alors en bleu) parmi les autres coliformes colorés en rouge. Pour devenir rouge, le TTC incolore doit être réduit par une déshydrogénase afin de donner du formazan de couleur rouge (voir son emploi sur le site : <http://www.2020seedlabs.ca>, consulté le 24-11-04).

L'emploi de TTC dans les géloses nous a permis de bien visualiser les différentes phases de migration sur la gélose<sup>6</sup>. L'on peut remarquer des zones plus blanchâtres correspondant aux zones d'intenses migrations sur la gélose. Le TTC ne formant sa coloration rouge que lors de l'activité d'une déshydrogénase, cela nous permet de voir que ces phases présentaient une activité enzymatique réduite, puisque l'on ne peut supposer de mort cellulaire à ces endroits. D'ailleurs, il avait déjà été montré que les cellules d'essaimage présentaient une diminution d'activité enzymatique et de la prise d'oxygène. Comme suggéré par Fraser *et al.* (2000), cette diminution d'activité viendrait en quelque sorte compenser le fait de la surexpression des flagelles et de la production de facteurs de virulence pendant l'essaimage.

---

<sup>6</sup> Lorsque nous regardons les photos du mouvement migratoire prises des géloses sans TTC (voir annexe 1), la différence « qualitative » du patron de l'intérieur du premier anneau de consolidation chez les deux mutants PmC8 et PmD1 est encore visible, mais le contraste est beaucoup moins prononcé. (À au moins trois reprises, ce type d'étalement fut réalisé et les mêmes observations ont pu être constatées quant aux profils de migration des différentes souches de *P. mirabilis* sur géloses LA, à 37°C et à 28°C). Quant au patron de Pm2D*ClacZ*, il est fortement similaire à celui remarqué pour PmD1. Une note cependant: toutes les phases de consolidation de Pm2D*ClacZ* présentent le caractère « dentelé », alors qu'elles deviennent plus lisses au fur et à mesure de l'envahissement de la gélose, chez PmD1.



**Figure 25 : Croissance de *P. mirabilis* sur milieu LA contenant du TTC.** Les trois souches ont été cultivées à 28°C pendant 24 h à partir du dépôt de 5 µl d'une culture de nuit sur une gélose préalablement « asséchée » 1-2 h/37°C. Les géloses furent ensuite conservées à 4°C pendant 24 h, afin d'augmenter le contraste des couleurs obtenues grâce au TTC. WT : souche sauvage (Pm 240-82), C8 : *xerC* ::*lacZYA*-Cm<sup>R</sup>, D1 : *xerD* :: *lacZYA*-Cm<sup>R</sup>.

Dans tous les cas le profil de l'essaimage a été décomposé de la manière suivante :

- Une phase de latence plus ou moins longue qui dépend de l'inoculum de départ, requise avant le démarrage de la migration sur gélose (Fraser *et al.*, 2000). Cette phase est représentée par le point central rouge foncé (séparation en deux ici pour la présentation des différentes souches).
- Une 1<sup>re</sup> phase de migration représentée par un halo blanchâtre.
- Une 1<sup>re</sup> phase de consolidation (de couleur rouge) pouvant être séparées en deux : la première partie est plutôt uniforme et la seconde constitue une frange (sur la Figure 25, cette frange présente des motifs différents selon les souches). Durant cette phase, il existe un mélange de cellules « stationnaires » de la première vague de migration en train de se différencier et d'autres cellules (déjà différenciées) sont en train de se multiplier. Cette vague de consolidation continue de s'étendre jusqu'à la fin de la première vague d'essaimage, au moment où la deuxième phase de migration est entamée (Fraser *et al.*, 2000), ce qui serait représenté par la frange.
- La suite est une succession de phases de migration et de consolidations menant à l'invasion de la gélose.

Comme le montre la Figure 25, l'essaimage des diverses souches présente de légères variations. De manière générale, les mutants présentent un premier anneau de consolidation de rayon réduit par rapport à celui de la souche sauvage. Cependant, même si l'emplacement des phases de migration est différent, l'épaisseur (de la zone blanchâtre) semble inchangée.

Pour ce que nous considérons la première « vraie » phase de consolidation, les mutants présentent une diminution en largeur de la première partie et un allongement de la frange. Comparativement à la souche sauvage, la frange de la souche PmD1, forme de courts filaments pointés vers l'extérieur. Cela ne semble pas être un caractère dendritique (voir Allison et Hughes, 1991), car l'ensemble des cercles de consolidation est très uniforme. D'autre part, la frange de la souche PmC8 est plus grande et est plutôt en « pointillé ».

Pour les anneaux suivants, mis à part leur rayon réduit, les principales différences résident en la couleur rouge moins intense par rapport à la souche sauvage et à leur épaisseur réduite.

Les vagues de consolidation sont donc plus « petites » chez les deux mutants (elles font environ 78% de la taille de cette même zone de la souche sauvage). Comme déjà mentionné, l'induction de ce mouvement migratoire dépend de plusieurs facteurs avec entre autres la densité cellulaire et l'âge de la cellule. Ce dernier paramètre serait aussi impliqué dans le signal de fin de migration (lire R.M. Harshey, 2003). Ceci pourrait être dû à un taux de génération différent car les mutants *xer* présentent un ralentissement de leur croissance et donc une diminution de la densité cellulaire à un moment donné. (voir la section des temps de génération, plus loin dans le texte et la partie discussion). D'un autre côté, la phase de migration apparaît similaire dans les trois cas (cette phase serait représentée par les zones plus blanchâtres de la Figure 25).

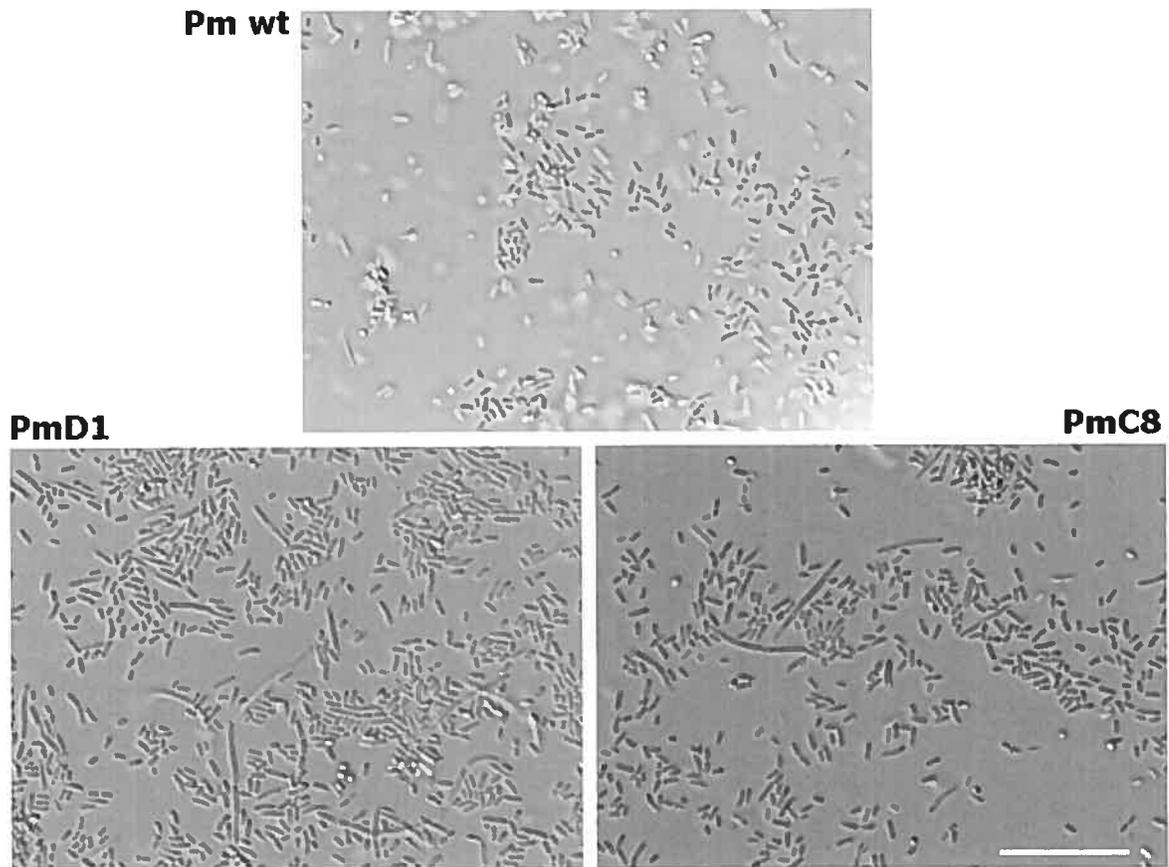
Ainsi, la migration ne semble pas avoir été affectée par les mutations *xer*. Par contre, la « longueur » (l'étalement de la zone) requise pour entamer une deuxième phase de migration a été touchée.

Les diverses souches de *P. mirabilis* ont également été mises à 37°C. La croissance étant beaucoup plus rapide, nous n'avons pas pu bien distinguer les phénotypes de migration. La souche sauvage avait alors envahi la gélose sans laisser de traces bien définies de consolidation. Les « vagues » des autres souches étaient un peu mieux définies, mais malgré tout difficiles à observer.

### 2.2.2 Études microscopiques

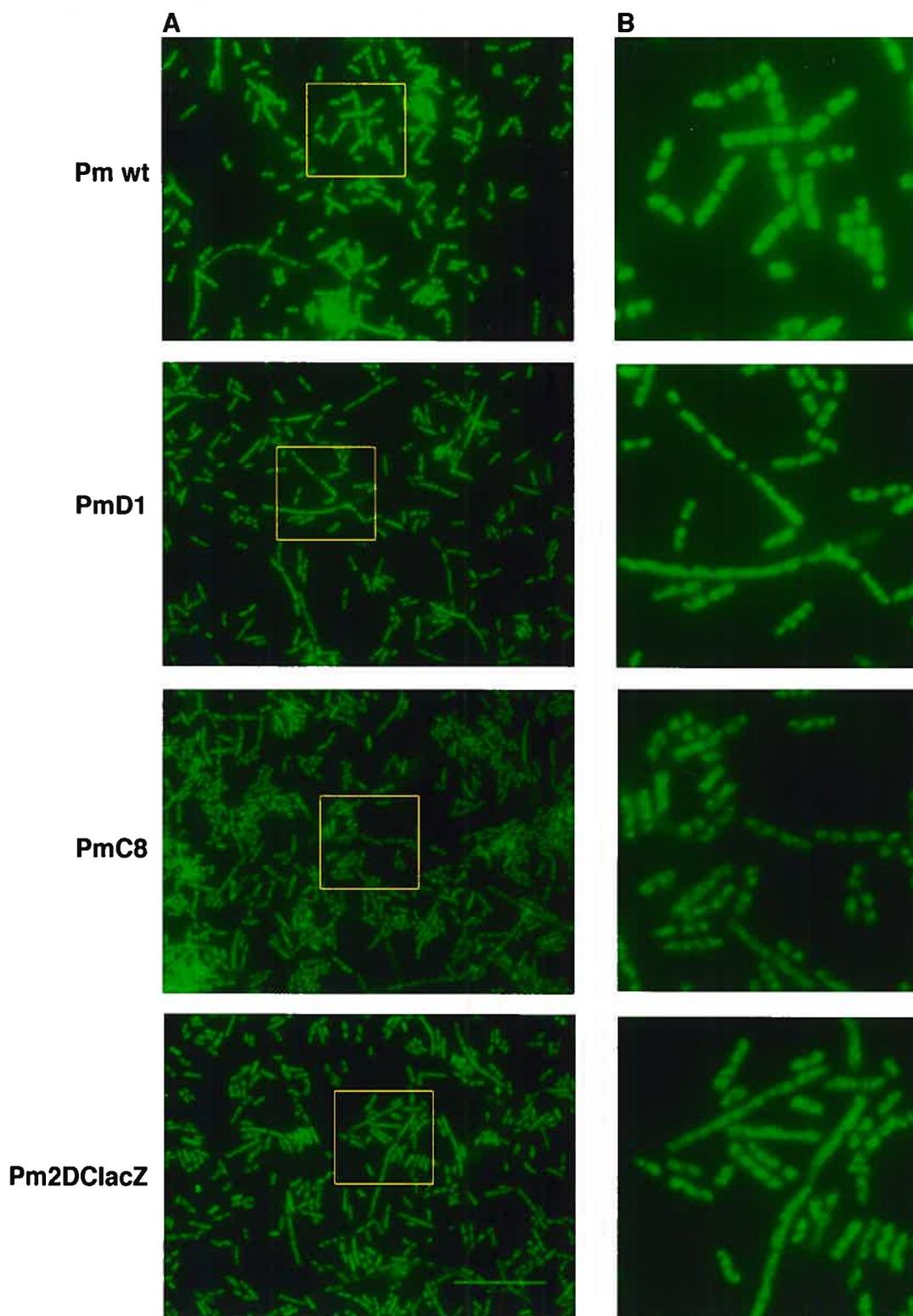
Les diverses souches de *P. mirabilis* ont également été observées sous microscope, en contraste différentiel et intégral. Les cellules furent prises à partir de colonies ayant poussé sur des géloses LSW. Ce milieu permettant le ralentissement de l'essaimage et l'obtention de colonies définies nous assurait à la base de ne pas obtenir de cellules filamenteuses.

Tel qu'attendu, les souches mutantes pour les gènes *xer* ont présenté un phénotype de filamentation comme il avait été observé chez *E. coli* (voir la Figure 26). Les filaments peuvent aller jusqu'à plus de 20  $\mu\text{m}$ , tandis que les cellules de la souche sauvage s'échelonnent de 2  $\mu\text{m}$  à 4,5  $\mu\text{m}$  environ.



**Figure 26 : Étude microscopique de diverses souches de *P. mirabilis*.** Les cellules ont été prélevées sur des géloses LSW avant d’être fixées au formaldéhyde sur lames par le protocole indiqué dans la section Matériel et Méthodes. La barre de calibration représente 20  $\mu\text{m}$ .

En observant les cellules en coloration DAPI, nous pouvons de plus remarquer que le nucléoïde a été affecté par les mutations *xer* (Figure 27). Les cellules sauvages présentent un nucléoïde condensé, tandis que les nucléoïdes des cellules *xer*<sup>-</sup> (PmC8, PmD1 ou Pm2D*ClacZ*) semblent être plus diffus. En fait, les nucléoïdes apparaissent même fragmentés en plusieurs petits morceaux d’ADN tout le long des cellules filamenteuses.

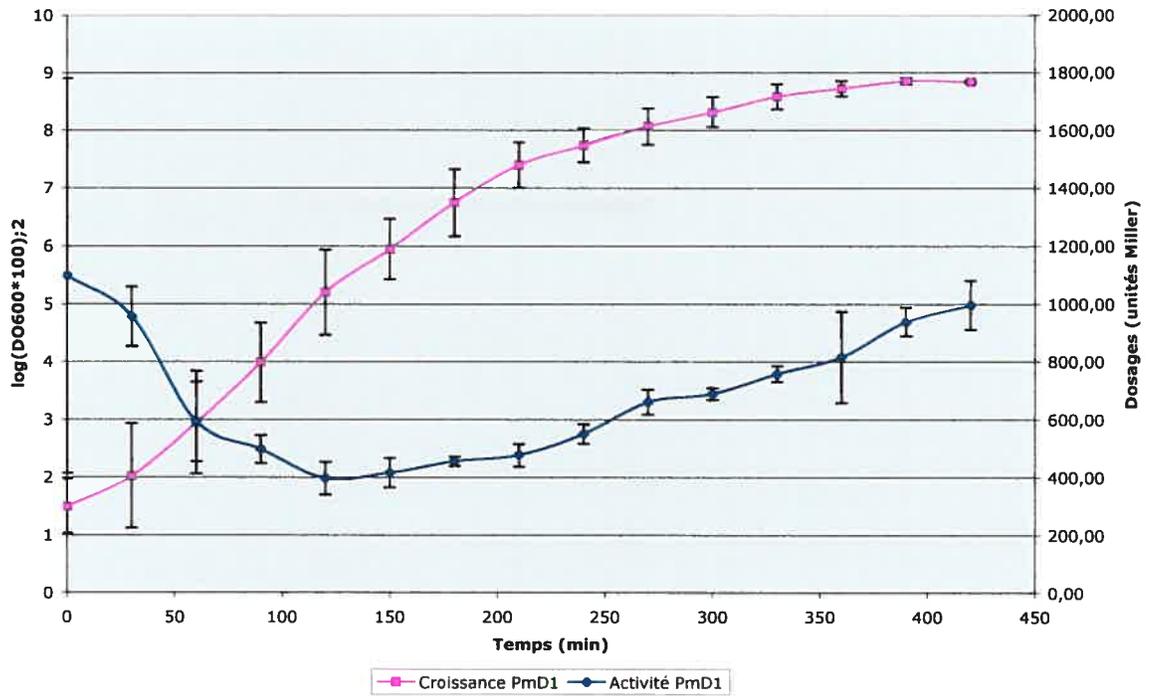


**Figure 27: Étude microscopique de diverses souches de *P. mirabilis* par coloration DAPI.** Les cellules ont été mises en culture dans du milieu LA à 37°C jusqu'à une  $DO_{600}$  de 0,6, avant d'être fixées, puis traitées pour la coloration DAPI. Les photos en B représentent un agrandissement (x 3) du carré jaune sur les photos en A. Pour de plus amples détails, voir le chapitre du matériel et méthodes.

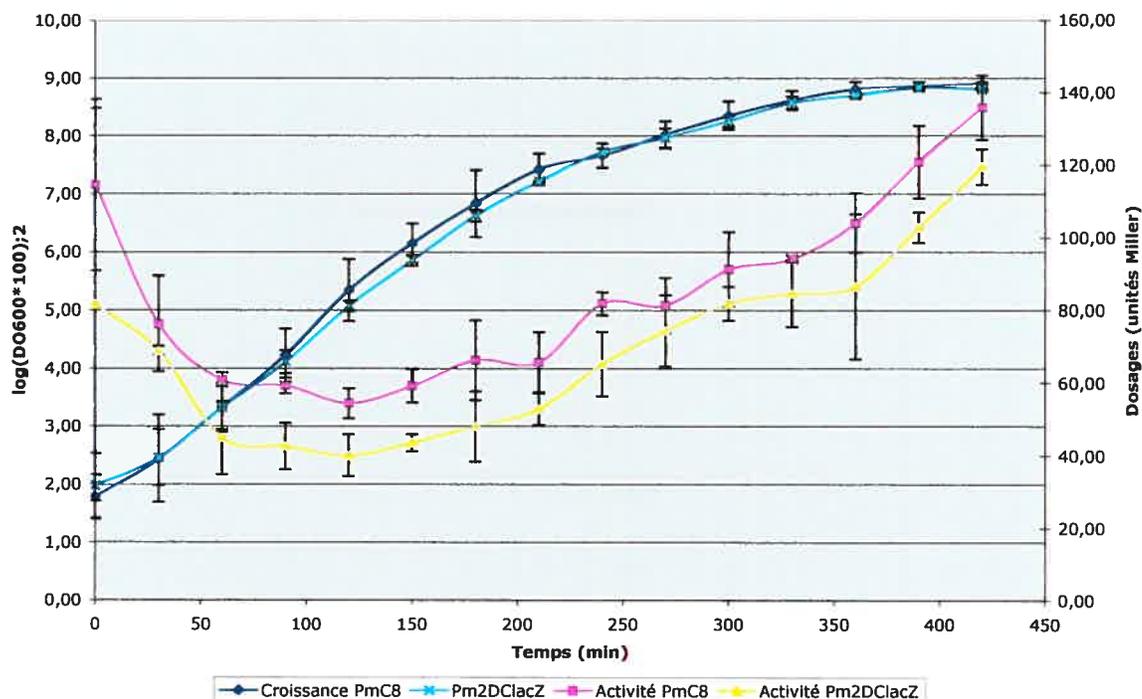
## 2.3 Expression des gènes *xer*

### 2.3.1 Dosages B-gal

Les trois souches de *P. mirabilis* pour lesquelles le gène *xerC* ou *xerD* a été inactivé (par fusion transcriptionnelle) avec l'opéron *lacZYA* à l'aide du vecteur pFUSE, ont été testées pour leur activité  $\beta$ -galactosidase. Ce type de fusion nous permettait d'avoir une idée de l'activité d'un promoteur potentiel de *xerC* ou *xerD*. Ainsi, trois séries de dosages différents furent réalisés et ont permis l'obtention des courbes indiquées en Figure 28 (pour les dosages faits avec la souche PmD1) et en Figure 29 (pour les dosages faits avec les souches *xerC*<sup>-</sup>, PmC8 et Pm2DC*lacZ*). Il est à noter que nous avons observé une certaine variabilité de la croissance en fonction du temps, variabilité qui est révélée par les barres d'erreurs. La phase de latence a en effet été plus ou moins longue selon les essais, surtout ceux de PmD1. Il apparaît que la culture de nuit n'était pas toujours à la même densité avant de débiter l'expérience. Malgré tout, nous avons quand même pu tirer des conclusions générales des divers essais réalisés et d'autres, non présentés, ont révélé la même tendance.



**Figure 28 : Dosages  $\beta$ -galactosidase de *P. mirabilis xerD*<sup>-</sup>.** Le graphique représente la moyenne de trois essais de dosages avec les souches PmD1 (*xerD* ::*lacZYA*-Cm<sup>R</sup>; courbe en bleu foncé), et la moyenne de la courbe de croissance (en rose) faite en même temps (voir le matériel et méthodes pour de plus amples détails). Chacune des valeurs des trois essais de dosages représentaient elles-mêmes la moyenne de deux échantillons. Les densités optiques ont été converties en  $\log_2$  de la ( $DO_{600} \times 100$ ) pour tenir compte de l'aspect logarithmique de la croissance (et la multiplication fut pour éviter les données négatives). La phase exponentielle a été déterminée ici approximativement entre les temps 90 à 200 min. Les barres d'erreurs sont représentées sur chacune des courbes. Comme on le remarque sur les deux courbes, l'erreur a été plus prononcée en début de croissance que vers la fin (voir le texte pour les détails).



**Figure 29: Dosages  $\beta$ -galactosidase des deux souches de *P. mirabilis xerC*.** Comme en Figure 26, le graphique représente la moyenne de trois essais de dosages, mais cette fois avec les souches PmC8 (*xerC* ::*lacZYA*-Cm<sup>R</sup>; courbe en bleu foncé) et Pm2DClacZ (*xerD* ::Cm<sup>R</sup> *xerC* ::*lacZYA*-Km<sup>R</sup>; courbe bleu pâle), ainsi que la moyenne de la courbe de croissance (PmC8, en rose et Pm2DClacZ, en jaune). Ces expériences ont été faites en même temps que celles de PmD1 (voir le matériel et méthodes pour de plus amples détails). Chacune des valeurs des trois essais représentaient elles-mêmes la moyenne de deux échantillons. Pour le reste des détails concernant les données du graphique, voir la Figure 26.

Comme on peut le voir par les deux graphiques, l'activité du promoteur supposé de *xerD* est beaucoup plus fort que celui de *xerC* (voir Figure 28 et Figure 29 les courbes concernant PmD1 et PmC8). En effet, il y aurait une différence d'environ 7 fois entre les deux promoteurs. Par exemple, à une  $DO_{600}$  de 0,4 à 0,5, [correspondant à 5,3 et 5,6 en  $\log_2(DO_{600} \times 100)$ ] l'activité du promoteur de *xerDPm* est d'environ 374 unités Miller, tandis qu'elle est d'environ 54 unités Miller pour le promoteur de *xerCPm* dans la souche PmC8 (voir Figure 28 et Figure 29, et les valeurs retransmises dans le Tableau VI). En début de phase stationnaire, les activités grimpent pour les deux souches, mais l'écart reste le même. Ainsi, à une  $DO_{600}$  d'environ 2 [correspondant à 7,6 en  $\log_2(DO_{600} \times 100)$ ], l'activité du promoteur *xerD* dans la souche PmD1 est de 529,6 en moyenne, et d'environ 73 dans la souche PmC8. Les données inscrites dans le Tableau VI ont été calculées à la main à partir des données brutes afin d'éviter le biais créé par les phases de latence plus ou

moins variables (voir plus haut); les données étaient alors évaluées en fonction de la densité optique plutôt qu'en fonction du temps de la croissance.

**Tableau VI: Comparaison de l'activité des promoteurs *xerC* et *xerD* à partir des dosages  $\beta$ -gal.**

Souche	DO <sub>600</sub>	Activité (unités Miller)
PmD1	0,47±0,05	373,86±24,58
	2,02±0,06	529,6±93,4
PmC8	0,52±0,05	53,59±3,86
	2,11±0,07	72,92±11,28
Pm2DClacZ	0,58±0,04	43,43±2,31
	2,11±0,10	65,20±8,91

Pour toutes les cellules, les relevés en fonction du temps des activités enzymatiques ont permis de mettre en évidence une courbe avec 3 parties. Une première partie qui montre une diminution des activités se déroule pendant la phase de latence. La seconde est un plateau et se produit au moment de la phase exponentielle de croissance. Enfin les dosages des activités augmentent par la suite durant la phase stationnaire.

Le plateau nous laisse penser que durant la phase exponentielle, la production de  $\beta$ -gal est plutôt constante et ce, pour les trois souches testées (voir les Figure 28 et Figure 29 comparez les courbes de croissance avec les courbes des activités). L'augmentation observée en phase stationnaire est quant à elle paradoxale et ne pourrait être due qu'à la stabilité de la  $\beta$ -gal. De plus amples explications seront données dans la discussion.

Pour vérifier rapidement si l'absence de la protéine XerD pouvait influencer l'expression de XerC, le dosage  $\beta$ -galactosidase a également été effectué sur la souche Pm2DClacZ. Comme déjà mentionné, cette souche est inactive pour le gène *xerD* et le gène *xerC* est inactivé par l'insertion de *lacZYA* en fusion d'opéron à l'intérieur de ce gène. Selon le graphique tiré en Figure 29, il semblerait que l'activité révélée par les dosages  $\beta$ -gal est toujours légèrement plus faible dans la souche Pm2DClacZ que dans la souche PmC8. Les valeurs prises individuellement semblent révéler la même tendance (voir les exemples données au Tableau VI). Cependant, les valeurs obtenues étant très faibles (voir l'échelle d'unités Miller à la droite du graphique) et compte tenu des écart-types obtenus

des divers essais, de plus amples essais seraient nécessaires pour en tirer des valeurs statistiquement valables.

L'activité de ces promoteurs a également été vérifiée en présence de *xerC* ou de *xerD* en multcopie. Les souches PmC8 et PmD1 ont été transformées avec les vecteurs pMVC6 (*xerCPm/pTZ-PC*), pMVD37 (*xerDPm/pTZ-PC*) et pTZ-PC (contrôle négatif). Les dosages ont alors été effectués en diluant la culture de nuit 1 :200 (au lieu de 1 :80) avant d'entamer la croissance des diverses souches. Tout d'abord, la présence d'un vecteur dans la souche n'a pas modifié l'activité produite en termes d'unités Miller (contrôle négatif). Nous pouvions donc tester directement l'influence de la protéine sauvage sur son propre promoteur, ou encore, l'influence de la protéine Xer complémentaire en surplus. Cependant, nous n'avons pas remarqué de différences notables avec les souches, non transformées (résultats non montrés). Ainsi, à première vue, il ne semble pas qu'il y ait une quelconque influence des protéines Xer sur leur propre promoteur, ou sur le promoteur de la protéine complémentaire.

#### 2.3.1.1 Temps de génération

À partir des données recueillies pour le calcul de l'activité  $\beta$ -gal (voir plus haut), nous avons pu établir une moyenne des courbes de croissance pour chacune des souches, à 37°C. À partir de ces courbes, les temps de génération ont été établis et les valeurs sont indiquées au Tableau VII (la courbe obtenue de la souche sauvage n'est pas montrée. Elle a été réalisée à partir de deux essais différents faits à part). Globalement, les temps de génération des souches mutées pour les gènes *xer* ont été légèrement augmentés : alors que les simples mutants ont eu des temps de génération assez similaire, le double mutant a vu son temps augmenter d'environ 9 min.

**Tableau VII: Moyenne des temps de générations des diverses souches de *P. mirabilis*.**

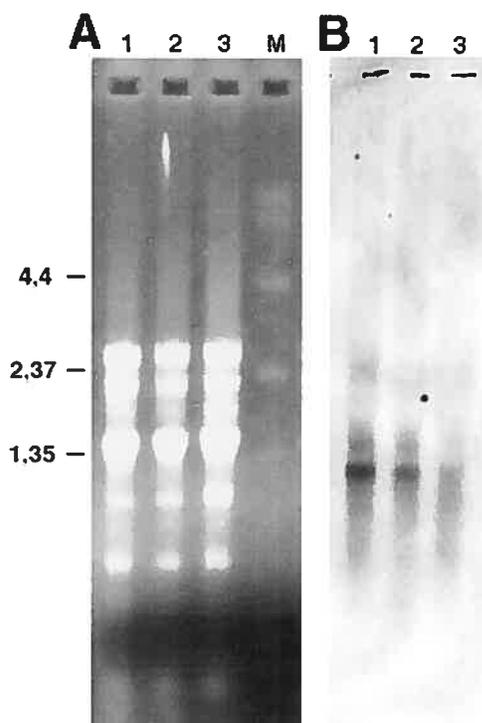
Souche	Moyenne (min)	Écart-type
Pmwt	29,90	0,19
PmD1	32,81	2,15
PmC8	33,57	0,82
Pm2DClacZ	38,71	1,61

### 2.3.2 Hybridations Northern

Pour analyser plus avant l'expression de *xerCPm* et de *xerDPm*, nous avons ensuite extrait l'ARN total de *P. mirabilis* afin de détecter l'ARNm des deux gènes grâce à des sondes d'ADN marquées à la DIG. Ces sondes pouvaient être un fragment interne des gènes respectifs, les gènes entiers ou encore une portion en amont des deux gènes. Dans tous les cas, les mêmes résultats ont été observés.

En premier lieu, nous avons voulu extraire l'ARN à différentes densités optiques à 600 nm, afin de voir si nous pouvions corréler les résultats avec ceux obtenus au cours des dosages  $\beta$ -galactosidases. La Figure 30B montre l'extraction d'ARN hybridé à un fragment interne de *xerDPm*. Nous observons un fragment correspondant à moins de 1,35 kb selon le marqueur d'ARN employé, et une autre bande, plus faible cependant légèrement au-dessus de la 1<sup>re</sup> bande. Selon ce gel, il semblerait que l'expression de l'ARNm de *xerDPm* soit plus forte en début de phase exponentielle et qu'elle diminue légèrement en phase exponentielle tardive (puits 1 vs puits 3). (Cependant, de plus amples études seront nécessaires afin de confirmer les niveaux d'expression d'ARNm de *xerDPm*, puisque les extractions dans le temps n'ont été faites qu'une seule fois).

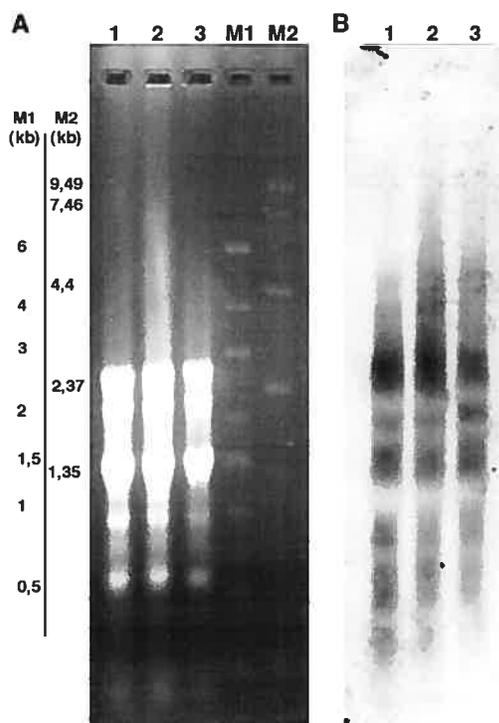
D'un autre côté, nous observons une bande forte (avec une autre beaucoup plus faible au-dessus), ce qui pourrait indiquer que la terminaison de la transcription s'est produite peu après le codon d'arrêt de *xerDPm*. En effet, la bande obtenue correspond à moins de 1,35 kb, ce qui équivaut presque à la taille de *xerDPm*. Une deuxième hybridation de type Northern, réalisée cette fois avec une sonde faite d'un fragment N-terminal de *xerDPm*, a montré le même résultat : un deuxième marqueur a permis de préciser la taille aux alentours de 1 kb (hybridation non montrée). Ainsi, contrairement à ce que nous pouvions penser au départ, l'organisation génomique de *xerDPm* ne serait peut-être pas en opéron avec le gène *dsbC* qui suit (de plus amples détails seront donnés en discussion).



**Figure 30 : Hybridation de type Northern avec la sonde d'ADN *xerDPm* contre l'ARN total de *P. mirabilis*.** A) Migration sur agarose 1,2% de l'ARN total de *P. mirabilis*. 300  $\mu$ l d'une culture de nuit ont été ensemencés dans 50 ml de LB neuf. L'ARN a été extrait au TRIzol à différentes DO<sub>600</sub> : 0,335 (puits 1), 0,676 (puits 2), 1,112 (puits 3), et 15  $\mu$ g ont été mis sur gel d'agarose avant d'être transférés sur membrane de nylon chargée positivement. Le marqueur (puits M) employé est le 0,24 – 9,5 kb d'Invitrogen. B) Hybridation de type Northern. La membrane récupérée du gel indiqué en A a été hybridée à un fragment d'ADN interne du gène *xerDPm* (environ 600 pb) marqué à la DIG.

Nous avons également hybridé l'ARN total de *P. mirabilis* avec une sonde faite du gène *xerCPm* marqué à la DIG (voir la Figure 31). Cependant, l'ARN a été extrait à une DO correspondant à la phase exponentielle. Les résultats obtenus de l'extraction faite à différents temps n'ont pas été suffisamment clairs pour être présentés. Ainsi, nous ne pouvons rien dire ici quant à l'expression de l'ARNm produit pour *xerCPm*.

Contrairement à ce qui a été observé pour *xerD*, plusieurs bandes sont apparues après l'hybridation. Ainsi, une portion correspondant à *xerCPm* s'est retrouvée sur plusieurs ARN de différentes tailles. Les bandes s'échelonnent d'environ 1 kb à entre 4,4 et 6 kb et la bande la plus foncée est à environ 3 kb (Figure 31). Le même profil a été obtenu avec une sonde faite avec seulement un fragment interne de *xerC* ou encore d'ADN situé en amont de *xerCPm* (non montrés). Ce profil de bandes pourrait indiquer que l'ARN de *xerCPm* serait polycistronique.

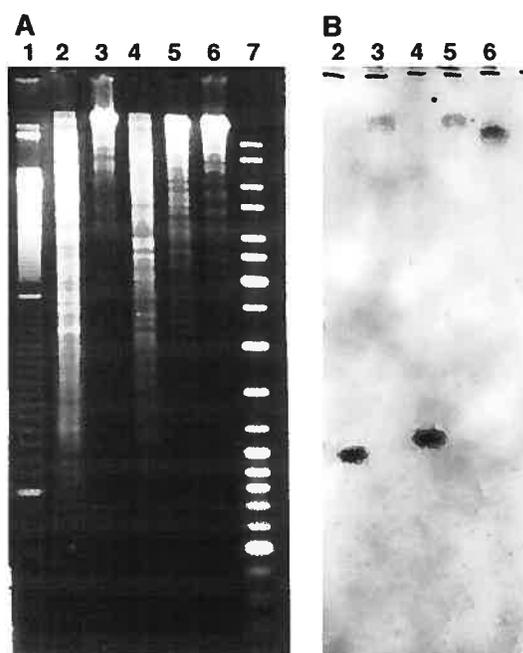


**Figure 31: Hybridation de type Northern avec la sonde *xerCPm* sur l'ARN total de *P. mirabilis*.** A) Migration sur agarose 1,2% de l'ARN total de *P. mirabilis*. 400  $\mu$ l d'une culture de nuit ont été mis en culture dans 80 ml et l'ARN total a été extrait à une  $DO_{600}$  de 0,540, avant d'être mis sur gel. Les puits 1 à 3 représentent 28  $\mu$ g d'ARN de trois échantillons différents extraits à la même DO. Les marqueurs employés sont : High Range de Fermentas (M1) et 0,24 – 9,5 kb de Invitrogen (M2). B) Transfert de type Northern. Le gel obtenu en A a été transféré sur une membrane de nylon chargée positivement, laquelle a par la suite été hybridée au gène *xerC* entier marqué à la DIG.

## 2.4 Recherche d'un site *dif* potentiel chez *Proteus mirabilis*

### 2.4.1 Plusieurs approches essayées

Dans un autre ordre d'idée, nous avons tenté de déterminer un site *dif* potentiel chez *P. mirabilis*. En effet, lorsque nous avons hybridé une sonde faite du site *difEc* à l'ADN génomique de *P. mirabilis* digéré avec plusieurs enzymes de restriction, nous avons été en mesure de détecter une bande indicatrice d'un site similaire au site *dif* d'*E. coli* (voir la Figure 32). Compte tenu de l'obtention de cette bande, nous avons essayé divers protocoles pour tenter de détecter ce site correspondant potentiellement à *dif*.



**Figure 32: Recherche d'un site *dif* chez *P. mirabilis*.** A) Digestion de l'ADN génomique de *P. mirabilis* avec différentes enzymes de restriction. B) Hybridation du gel représenté en A et transféré sur membrane, avec la sonde *difEc* marquée à la DIG. Les puits 1 et 7 représentent deux marqueurs de poids moléculaire : le 100 pb d'Amersham et le 2-log de NEB, respectivement. Les puits 2 à 6 représentent respectivement la digestion avec les enzymes : *EcoRV* (2), *BamHI* (3), *HindIII* (4), *PstI* (5) et *Asp718* (6).

La première approche abordée fut basée sur l'action de recombinaison en tant que telle. Effectivement, une certaine activité de clivage des recombinases XerC et XerD de *P. mirabilis* avait été montrée sur le site *difEc* (observation de topoïsomères sur le vecteur pMIN33, Figure 22). Ainsi, nous espérions observer une intégration d'un vecteur suicide contenant le site *difEc* dans le génome de *P. mirabilis* par l'intermédiaire de la recombinaison spécifique de site. Nous avons ainsi réalisé plusieurs conjugaisons bactériennes entre la souche d'*E. coli* SM10 $\lambda$ *pir* contenant le vecteur pVi11 (pKNG101 portant le site *dif* d'*E. coli* de 33 pb + 70 pb adjacentes) ou le vecteur pDEcK (pKNG101 portant le site *dif* central d'*E. coli* de 33 pb) et la souche de *P. mirabilis* 240-82. Malheureusement, des diverses conjugaisons réalisées, aucune n'a permis l'obtention de clones.

En second lieu, des réactions de PCR ont été tentées avec les amorces employées pour amplifier le site *dif* d'*E. coli* (ces amorces permettaient l'amplification de 102 pb incluant le site *difEc*). Compte tenu des similitudes probables entre les sites *dif* de plusieurs espèces, nous espérions pouvoir amplifier un produit à des températures moins élevées, et

ce, malgré l'emploi d'amorces qui ne se seraient pas parfaitement hybridées à la matrice d'ADN. Un produit a bel et bien été obtenu, de longueur plus courte que celui obtenu avec l'ADN d'*E. coli*. Par contre, le séquençage de ce produit n'a montré aucune similitude avec un site pouvant être reconnu par les recombinases de type Xer. Et, effectivement, comme nous l'avons constaté plus tard, même la région flanquante « rapprochée » du site *dif* était divergente (voir la suite de cette section).

La troisième approche fut basée sur l'activité de liaison des protéines Xer au site *dif*. Comme montré en Figure 20 de la première partie des résultats (voir les puits 6 et 8), XerCPm et XerDPm se sont liés individuellement au site *dif* d'*E. coli*. Ainsi, nous souhaitions récupérer un site de l'ADN génomique de *P. mirabilis* qui aurait été lié par les recombinases. Nous avons alors réalisé une « mini-banque » génomique de l'ADN de *P. mirabilis* digéré avec *EcoRV*. La digestion par cette enzyme avait permis de révéler un fragment d'environ 1-1,2 kb lors des hybridations faites avec la sonde *difEc*; aussi, nous avons cloné dans pUC19 un ensemble de fragments variant d'environ 900 pb à 1,5 kb.

Pour cette approche, la protéine XerDPm, fusionnée à une queue histidine, fut accrochée à des billes magnétiques de nickel (billes magnétiques Ni-NTA de QIAGEN). En passant l'ADN génomique cloné sur ces billes, nous espérions intercepter un fragment reconnu spécifiquement par la recombinase. Malheureusement, après maints ajustements, le vecteur pUC19 sans insert a été récupéré par les billes, fort probablement à cause de la charge positive du nickel. Ainsi, même si nous avons tenté d'augmenter la spécificité de liaison en augmentant, par exemple, la concentration en sels lors des lavages, nous n'avons pas pu détecter de clones potentiels qui auraient contenu un site reconnu par les protéines Xer.

Compte tenu du problème de cette charge, nous aurions pu penser à employer une colonne d'amylose sur laquelle serait fixée la protéine MBP-XerCPm. Cependant, il avait été montré que la constante de dissociation estimée pour la liaison de XerCEc au site *difEc* était de 40 à 160 fois plus forte que celle de XerDEc (Blakely *et al.*, 1993). Même si nous n'avons pas calculé l'affinité des diverses protéines au site *difEc*, il semblerait que la liaison de XerCPm à ce site soit également plus faible que celle de XerDPm (revoir la Figure 20). Il aurait alors fallu construire une fusion MBP avec XerDPm afin de réaliser ce même genre de test sur colonne d'amylose.

En parallèle, à l'aide de cette même banque *EcoRV*, faite dans une souche d'*E. coli* *recA*<sup>-</sup>, nous avons tenté de détecter un clone qui aurait multimérisé. En effet, un vecteur contenant le site *dif* d'*E. coli* a tendance à présenter des formes multimériques (voir le point 4.3.2.1 du chapitre d'introduction), alors que ce n'est pas le cas du vecteur sans insert (Figure 33). Nous pouvions donc supposer qu'il en serait de même avec le site *dif* de *P. mirabilis*. Malheureusement, cette méthode ne nous a pas permis de détecter de clone multimérisant suffisamment pour être sélectionné. De plus, il nous aurait fallu avoir une très bonne représentativité dudit clone supposé contenir le fragment incluant un site *dif* potentiel, chose que nous n'avons pas vérifiée.



**Figure 33 : Multimérisation d'un vecteur contenant le site *dif* d'*E. coli*.** Le 1<sup>er</sup> puits contient un pUC19 sans insert, le 2<sup>e</sup> contient le pMIN33 (contient la région centrale de 33 pb du site *dif* d'*E. coli*). Les deux vecteurs ont été extraits de la souche DH5 $\alpha$ .

Enfin, la quatrième et dernière approche a consisté à détecter un clone positif parmi plusieurs par hybridation sur colonies avec une sonde *difEc* marquée à la DIG. En premier lieu, nous avons tenté de faire une autre banque génomique dans pUC19, mais cette fois avec l'enzyme *Sau3A*-I. La digestion à complétion de l'ADN génomique de *P. mirabilis* nous avait en effet permis de détecter un signal à environ 700 pb suite à une hybridation avec le site *difEc*. Malheureusement, le bruit de fond était trop fort pour que nous puissions distinguer un quelconque regroupement de clones positifs, puisque la banque avait été faite dans DH5 $\alpha$ . De façon illusoire, nous espérions que le fait de retrouver un site potentiel en multicopie suffirait à contourner le fait que la souche était déjà *dif*<sup>+</sup>. Nous avons alors vérifié rapidement quelques groupes de clones par des réactions de PCR avec les amorces M13-F40 et M13-R48 universelles, mais rien n'a été obtenu.

Trois nouvelles banques ont ensuite été préparées avec l'ADN de *P. mirabilis* digéré avec *EcoRV*, *MspI* ou *Sau3AI*, mais cette fois, elles furent réalisées dans une souche d'*E. coli* dont le site *dif* a été retiré : la souche BW $\Delta$ *dif* (voir le point 20 de la section matériel et méthodes pour les détails de la construction de la souche).

Plusieurs nouveaux tests d'hybridation sur colonies ont été faits avec une sonde *difEc* marquée à la DIG. Avant d'entamer ces expériences, un test préalable a été fait en transformant un mélange avec un rapport 1 :10 de pMIN33 (contenant le site *difEc*) et de pUC19, afin d'être certains de détecter efficacement et uniquement les clones contenant le pMIN33. Malgré un bruit de fond assez élevé, nous avons pu mettre en évidence les clones contenant le pMIN33 et nous avons donc poursuivi l'expérience avec les banques de *P. mirabilis*.

Nous avons testé en tout une quarantaine de membranes faites à partir de géloses contenant une centaine de colonies sans parvenir à obtenir de colonies réellement positives. À chaque fois, les colonies ayant donné un signal un peu plus fort que le bruit de fond, ne se sont pas révélées positives par la suite (des hybridations avaient été réalisées avec les produits amplifiés par PCR de la région comprise entre les amorces M13-F40 et M13-R48 des différents clones potentiels).

Finalement, nous avons extrait l'ADN plasmidique de l'ensemble des colonies de quatre des géloses de la banque de *P. mirabilis* digérée par *EcoRV*, pour ensuite séparer partiellement tous les plasmides sur gel d'agarose. Une hybridation subséquente fut réalisée avec la sonde *difEc* marquée à la DIG, ce qui a permis de détecter une très faible bande. Celle-ci correspondait à un peu moins de 4 kb selon le marqueur de poids moléculaire, ce qui représentait bien une insertion d'environ 1,2 kb dans pUC19. Quatre séries de bandes plasmidiques comprises entre 3,5 kb et 4 kb furent alors découpées et purifiées. L'une d'entre elles ayant donné un signal positif à un test rapide d'hybridation (toujours avec la sonde *difEc*) a pu être de nouveau transformée dans la souche BW $\Delta$ *dif*. En procédant de la sorte, nous avons enrichi la banque d'ADN de la séquence souhaitée. De nouveau, les plasmides furent extraits à partir de l'ensemble des colonies obtenues sur une seule gélose, et cette fois encore, un signal a été obtenu avec l'une des extractions. La séquence clonée de chacune des 81 colonies de la gélose ayant servi à la préparation plasmidique (une copie avait, bien entendu, été sauvegardée) a pu alors être amplifiée par

réaction PCR (voir la section matériel et méthodes) afin d'être vérifiées par une autre hybridation avec la sonde *difEc*.

C'est ainsi que nous avons obtenu trois clones fortement positifs et un autre qui a donné un signal plus faible. Deux de ces quatre clones ont révélé un site similaire au site *dif* d'*E. coli*. Le clone sur lequel la sonde *difEc* s'est le moins bien hybridée a révélé une séquence assez similaire à l'ensemble de la sonde employée, mais n'a pas indiqué un site central correspondant à un site *dif*.

```

          10          20          30          40          50
Pm#15. CCTACTCATATTTTAATTATCCCAATAAACTTATTCCAACGGTAAACGA
          60          70          80          90         100
Pm#15. TGTTACTGCACAAGATGAACAAGTTTTAGGTCATCTATTTGTTGTGGCAG
          110         120         130         140
Pm#15. CTAAAATTGCTCAGCAAGAAGGTATTGCAGAAGAGGGATATC-TGAAAAG
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
difEc.          cggaagtgcagatg-----catgaaatc
          10          20
          150         160         170         180         190
Pm#15. CTGCGGTAAATGA--TTACCTCTACAAAAAATTATTTGTGGCGGGGGATT
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
difEc. ctg-----agtgactttgcc-----agaat-----gatt
          30          40
          200         210         220         230         240
Pm#15. TATGCAGTATTGAAGTGATAAACGATCACAAATATATTGGC---GATTT
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
difEc. ccagcagcactg---gat-aacg-tgaagacttcattaccgacagattt
          50          60          70          80          90
          250         260         270         280
Pm#15. TACGGTTAGATTTA--TAGTTACAGCA----AAGAGCAATAGCAATAGCA
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
difEc. tccgg--aga---acgtggtgacggcagttgaa-----
          100         110
          290         300         310         320         330
Pm#15. ATAGCAATATCAACGGTAGAAGTTTGAAATAATAAATAAAAATATTAAGG
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
difEc. --agcaatgt-----gttgaggtt-----
          120         130
          340         350         360         370         380
Pm#15. ATTATTATTGGAGAGATGTCTTTGATTCTTATGATAGCGTCACGGGCATC
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
difEc. -----gcat-
          140
          390         400         410         420         430
Pm#15. AATTTAACTGACGGTATAAACATCTAAAAAGGCTTTACTAAAATCCTTTC
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
difEc. -----ggacggt-taagc-----cgagaa-----
          150
          440         450         460         470         480
Pm#15. TAAAGTAGTATCTGATAAAAAGACAAAAAACATTGGGTC-----CAACC

```

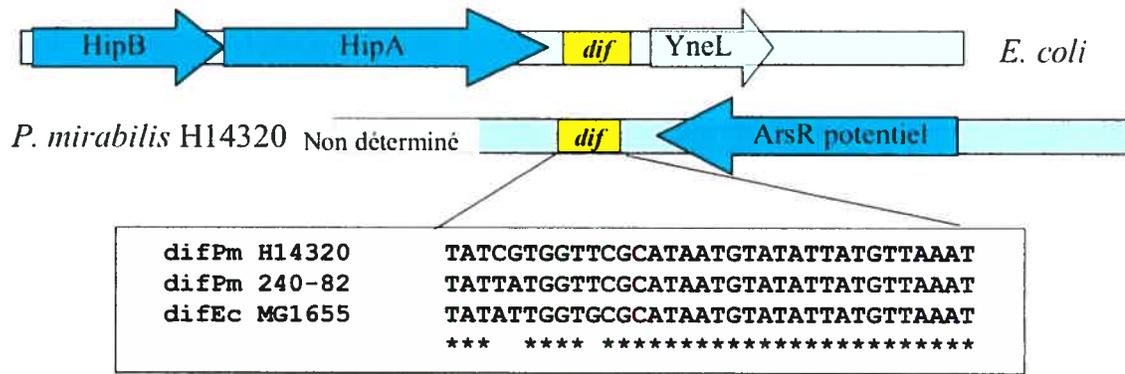


#### 2.4.2 Le site *dif* de *P. mirabilis*

La comparaison de la séquence obtenue précédemment indique que seules 3 bases divergent dans le 33 pb central du site *dif* lorsqu'on le compare avec le site *dif* d'*E. coli* (voir le bas de la Figure 35). Cependant, comme nous pouvons le voir par l'alignement de la séquence obtenue avec celle de la zone de 550 pb entourant le site *difEc*, les séquences entourant ces 33 pb sont très différentes (voir la Figure 34). Nous avons alors tenté de déterminer ce à quoi pouvait ressembler cette séquence par comparaison de séquences (blasts de séquences nucléotidiques et de séquences traduites).

Il est à noter ici qu'au moment d'entamer ces recherches, le génome de *P. mirabilis* n'était pas séquencé, ou du moins, n'était pas disponible. En effet, une souche de cette espèce semblait avoir été séquencée par une compagnie privée, mais le brevet déposé alors nous a empêché d'en avoir l'accès pour une vérification de séquence. Au moment où le site *dif* de l'isolat clinique employé tout au long de cette étude fut enfin découvert, l'Institut Sanger, en collaboration avec le laboratoire de Dr Mobley (Université du Maryland à Baltimore, U.S.A.), entamait le séquençage d'une souche nommée H14320 à des fins non commerciales. Nous avons ainsi pu effectuer certaines comparaisons de séquence avec le clone ayant révélé le site *dif*, et surtout, découvrir une portion plus longue de la région entourant ce fameux site.

Ainsi, alors que le gène *hipA* se retrouve en amont du site *dif* chez *E. coli*, la séquence traduite en aval de la zone *dif* de *P. mirabilis* serait un régulateur potentiel de transcription de la famille de ArsR. Du moins, c'est ce qu'a révélé une analyse génomique faite par comparaison de séquence de la portion *dif* retrouvée dans la souche H14320 avec l'ensemble des génomes séquencés dans le MBGD (<http://mbgd.genome.as.jp/>; consulté le 29 novembre 2004). La famille de protéines ArsR sert à la répression de divers opérons impliqués dans la résistance à des métaux lourds tels que l'arsenic et l'antimoine (Busenlehner *et al.*, 2003).



**Figure 35: Comparaisons entre les régions chromosomiques de *P. mirabilis* et d'*E. coli* où se trouve le site *dif*.** Chez *E. coli* le site *dif* se trouve à la suite du gène encodant *HipA* et juste avant un gène encodant la protéine de fonction inconnue *YneL*. Chez *P. mirabilis* H14320, le site *dif* est situé en amont d'un gène qui présente des similitudes au gène *arsR*, lequel encode un régulateur de transcription. Les séquences des sites *dif* sont indiquées sous le schéma. Tel qu'indiqué sur le site Colibri (<http://www.genolist.pasteur.fr/Colibri>), les gènes *hipA*, *hipB* et *yneL* sont transcrits sur le brin (–) du chromosome d'*E. coli* MG1655, et donc, l'orientation complète du schéma du haut n'est faite que pour les besoins de l'alignement montré ici.

D'un autre côté, l'alignement de la portion en amont du site *dif* retrouvé chez la souche Pm240-82 a montré une forte divergence de la séquence en amont du site *dif* retrouvé chez Pm H14320. Les annexes 3 et 4 illustrent les deux alignements ayant présentés des homologies dans la région du site *dif* de ces deux souches. Alors que l'alignement de l'annexe 3 montre l'homologie au niveau de la région entourant le site *dif*, l'alignement de l'annexe 4 révèle une homologie dans la portion plus en amont de la région de *dif*. Ces divergences seront plus amplement traitées dans la section Discussion.

## 2.5 Liaison des recombinases Xer de *P. mirabilis* au site *dif*<sub>Pm</sub> :

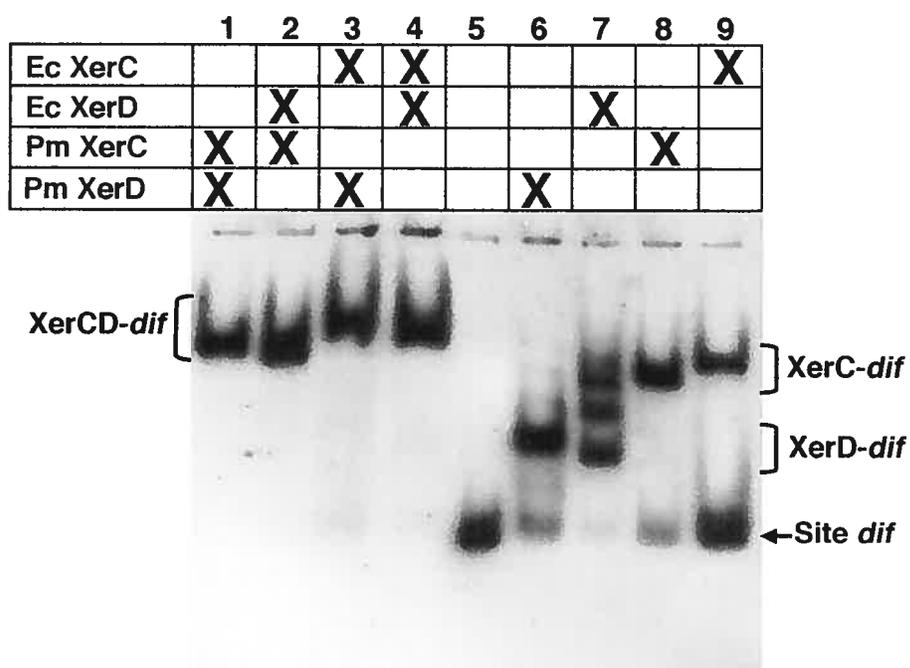
Afin de vérifier si le site découvert était bel et bien reconnu par les recombinases Xer de *P. mirabilis*, nous avons effectué un test de retard sur gel avec ce site marqué à la DIG (les amorces *dif*<sub>Pm</sub>5' et *dif*<sub>Pm</sub>3' ont alors été employées pour amplifier 120 pb incluant *dif*<sub>Pm</sub>), en suivant le même protocole que pour les retards avec *dif*<sub>Ec</sub>.

Comme on peut le voir à la Figure 36, le site *dif*<sub>Pm</sub> a été lié, non seulement par les recombinases XerC et XerD de *P. mirabilis*, mais aussi par celles d'*E. coli* (voir les puits 6 à 9). On peut également remarquer que le XerDEc a tendance à former plusieurs complexes, tout comme il avait été le cas sur le site *dif*<sub>Ec</sub> [le même phénomène a été observé avec deux fois moins de protéine employée (non montré)]. Nous pouvons supposer que des monomères de XerDEc ont eu la capacité de se lier de part et d'autre de la région

centrale du site *dif* et ainsi reconnaître le site de liaison de XerC. La bande la plus retardée pourrait indiquer la liaison d'un dimère de XerD en plus d'un monomère.

À la quantité de protéines employées (600 nM pour le retard montré), le site *difPm* a été entièrement retardé lorsqu'il y a eu combinaison de XerC et de XerD, peu importe l'espèce de laquelle elle provenait.

Les retards avec la protéine MBP-XerCEc ont été plus élevés dans le gel, ce qui serait probablement dû à une courbure différente du site *dif* induite par la protéine. (Cette protéine est de poids moléculaire plus faible que celle de *P. mirabilis* et donc, aurait dû théoriquement migrer plus rapidement dans le gel). En effet, nous pouvons supposer que la courbure induite par XerCEc serait plus grande, puisque « plus la courbure est grande à un point particulier d'un site, plus grand est le retard relatif durant une électrophorèse » (Blake *et al.*, 1997). Cependant, sachant que la même quantité de protéines a été employée dans chacun des puits, il semblerait que la liaison ait été plus faible que pour celle de MBP-XerCPm (la même chose a été observée avec deux fois moins de protéines). Comme suggéré par Ferreira *et al.* (2001) pour des chimères des protéines XerEc, il se pourrait que l'énergie de liaison ait été employée à modifier la conformation du site.



**Figure 36: Retards sur *difPm*.** Les retards sur gel ont été faits sur le site *dif* de *P. mirabilis* marqué à la DIG. 600 nM de protéines ont été employés dans chaque puits pour cet essai. Le puits 5 indique le site *dif* seul, les puits 1 à 4 correspondent à un mélange de XerC et de XerD, tandis que les puits 6 à 9 sont les retards faits avec chacune des recombinaisons, individuellement. Ec : *E. coli*, Pm : *P. mirabilis*.

## **DISCUSSION**

## Chapitre IV : Discussion

### 1 Organisation génomique des recombinases de *P. mirabilis*

Comme le montre l'ensemble des résultats présentés ici, *P. mirabilis* possède bel et bien les gènes *xer*, lesquels encodent les protéines qui peuvent se lier aux sites spécifiques d'*E. coli* et qui ont montré une activité catalytique.

Le gène *xerC* de *P. mirabilis* a été obtenu de la même manière que *xerD* (Villion et Szatmari, 1998). Les mêmes amorces dégénérées ont permis l'amplification de deux fragments différents correspondant à la région C-terminale des gènes *xer*. L'un de ces fragments a montré plus de similarités aux portions C-terminales des gènes *xerD* jusqu'alors séquencés, tandis que le second fragment, a montré des similarités avec les gènes *xerC*. Un PCR inverse fait avec des amorces cette fois spécifiques nous a permis d'obtenir le reste de la séquence de *xerC*.

La structure primaire, déduite de la séquence ainsi obtenue, a montré un plus haut degré de similarité avec les protéines XerC qu'avec les protéines XerD ; elle a montré 63,1% d'identité à la protéine XerC d'*E. coli* et 44,9% d'identité à la protéine XerD de la même espèce. De plus, la séquence en acides aminés contient l'élément ESS indicateur d'une protéine XerC (Hallet *et al.* 1999). Comme mentionné dans la section des résultats, le plus haut niveau de similarité se situe dans le domaine C-terminal où le résidu catalytique tyrosine est retrouvé.

Lors de la recherche des gènes *xerC* et *xerD* de *P. mirabilis* (voir Villion, 1998), nous avons été en mesure de prévoir les séquences adjacentes de ces gènes. Cependant, nous n'en avons alors pas confirmé la séquence. Nous supposons que ces gènes étaient organisés de la même façon que chez *E. coli*. Cette supposition tenait au fait que la même organisation avait été retrouvée pour le gène *sss* (homologue de *xerC*) chez *P. aeruginosa* (Höfte *et al.*, 1994), et par la suite chez *P. fluorescens* (Dekkers *et al.*, 1998) et chez *S. marcescens* (Guynn *et al.*, 1998). Maintenant, grâce au séquençage de la souche PmH14320 par l'institut Sanger, nous pouvons affirmer que l'organisation des gènes *xer* de *P. mirabilis* est similaire à celle retrouvée chez *E. coli* (voir la Figure 19 des résultats).

Il a été supposé que le gène *xerC* d'*E. coli* formait un opéron avec *dapF*. Cela pourrait être aussi le cas pour *xerC*, mais il est aussi possible que le transcrit soit *yigA*-

*xerC-yigB*. En effet, nos premières études de Northern ont révélé, non pas un ARNm d'une seule longueur, mais plutôt un profil d'ARN de différentes longueurs, ce qui indiquait la possibilité d'un ARN subissant différents traitements (processing) lors de la transcription (voir la Figure 31 dans les Résultats).

Par contre, pour ce qui est de l'ARNm de *xerDPm* d'après nos résultats de Northern (Figure 30), il ne serait pas en opéron, mais serait plutôt transcrit seul. Cependant, nous avons pu observer une très faible bande un peu diffuse au-dessus de la bande correspondant à l'ARNm de *xerDPm*. Cette bande pourrait correspondre à une prolongation de la transcription de *xerD* bien après son codon stop, sans que cela n'indique un opéron avec le gène *dsbC*. À moins d'être le reflet d'une dégradation rapide d'un ARNm plus long, la bande obtenue aurait dû faire au moins 2 kb pour révéler un opéron avec *dsbC*. Bien que ces études ne soient que préliminaires, deux hybridations réalisées avec des sondes différentes (deux régions différentes de *xerD* marquées par la DIG), ont montré le même profil d'ARN.

## 2 Études de complémentation

La protéine XerCPm a pu agir de concert avec la recombinase de la même espèce pour résoudre le pCS210, et ce de façon aussi efficace que lorsque les deux recombinases provenaient d'*E. coli* (Figure 23, puits 1 et 5). De plus, ce plasmide ainsi résolu (pCS211) a pu être maintenu sans pression de sélection pour au moins 100 générations (Figure 24). Ceci indique que les recombinases de *P. mirabilis* peuvent agir de concert au site *cer* comme le font les recombinases provenant d'*E. coli*. Il est aussi à noter que la protéine XerDPm peut interagir avec XerCEc de façon suffisamment efficace pour le maintien du pCS210 pour au moins 5 jours, malgré le fait que la recombinaison entre les deux sites *cer* n'est à ce moment pas aussi efficace que lorsque les protéines proviennent de la même espèce (Figure 23, puits 4, et Figure 24).

Cependant, malgré le haut niveau de similarité entre les protéines XerC des deux espèces, la protéine XerCPm n'a pas été capable d'agir de concert avec la protéine XerDEc au niveau du site *cer* afin de compléter une souche d'*E. coli* mutante pour les deux gènes *xer*. La souche d'*E. coli* *xerC<sup>-</sup>xerD<sup>-</sup>* transformée avec le plasmide pPmCEcD a non seulement été incapable de résoudre le pCS210 en pCS211, mais n'a pas pu maintenir le pCS210 en l'absence de pression de sélection (Figure 23, puits 3 et Figure 24). En fait,

dans les conditions employées, le plasmide pCS210 a même été perdu à un taux légèrement plus rapide que le témoin ne contenant pas de gène *xer* (Figure 24, comparer les souches avec pPmCEcD vs pUC19). Cette différence, même si faible, a été reproductible au travers de notre expérimentation.

L'incapacité du clone contenant le gène *xerCPm* de compléter la souche *xerC* d'*E. coli* pour la recombinaison entre deux sites *cer* est surprenante, compte tenu de plusieurs faits, entre autres :

- XerCPm a pu se lier au site *cer* en collaboration avec XerDEc (Figure 20A, puits 2) ;
- Une jonction de Holliday a été détectée dans les réactions *in vitro* avec le pCS210, en utilisant ces mêmes protéines (Figure 21, puits 2). La réaction la plus évidente dans ce cas fut avec les protéines d'*E. coli*, mais malgré tout, un fragment migrant à la hauteur correspondant à une jonction de Holliday fut obtenu dans tous les cas, que ce soit en présence des recombinases de la même espèce, ou en présence des protéines des deux espèces ensemble (Figure 21, puits 1 à 4) ;
- La complémentation fonctionnelle a également été décrite dans *Haemophilus influenzae*, mais pour le site *dif* (Neilson *et al.* 1999).

Ce problème de complémentation inefficace peut simplement être dû au fait que la protéine XerCPm est une protéine plus large que celle d'*E. coli* (311 aa au lieu de 298 aa) et qu'il y a un bon degré de divergence dans la séquence du domaine N-terminal (47,2% d'identité entre les deux protéines). Cette dissemblance pourrait affecter le niveau de coopération entre les protéines Xer.

Il a été suggéré que deux domaines internes dans les protéines Xer servent au contrôle réciproque de la réaction de recombinaison. Par un échange des résidus ESS (retrouvé normalement chez XerC) et NHG (chez XerD), ces protéines mutantes avaient montré une activation d'elles-mêmes alors qu'elles inhibaient l'action de la protéine partenaire (Hallet *et al.*, 1999). Bien que ces résidus soient retrouvés chez les recombinases de *P. mirabilis*, serait-il possible qu'ils ne soient pas disponibles pour exercer leur action? Cependant, il ne semble pas y avoir de séquences pouvant inhiber cette interaction autour du motif, et le domaine ESS retrouvé chez XerCPm, tout comme c'est le cas chez toutes les

recombinases XerC séquencées jusqu'à maintenant, est entouré d'une région très bien conservée.

Ces différences peuvent également affecter la réaction de clivage du site *cer* et ainsi limiter la formation de la jonction de Holliday, ou encore réduire la reconnaissance du site nécessaire à une liaison efficace. De plus, il ne faut pas oublier que le site *cer* est spécifique pour les recombinaisons d'*E. coli*, et ainsi, ne pas l'être forcément pour celles de *P. mirabilis*. Le même cas a d'ailleurs été présenté avec le site *dif* de *Bacillus subtilis* : alors que XerDEc peut reconnaître et se lier à ce site, XerCEc ne le peut pas (Sciochetti *et al.* 2001).

Cependant, puisque les réactions *in vitro* ont montré une activité catalytique de XerCPm au site *cer* en collaboration avec XerDPm et les protéines accessoires d'*E. coli*, il est possible que le problème provienne de la formation de la jonction de Holliday par XerCPm. La conformation de cette jonction peut ne pas être un bon substrat pour la protéine responsable de la résolution de cette jonction. En effet, il a été montré que XerDEc, même si essentielle pour l'activité catalytique de XerC dans la recombinaison au site *cer*, n'était pas la protéine qui permettait la résolution de la jonction (Colloms *et al.* 1996).

Ainsi, il se pourrait que la protéine XerCPm en collaboration avec XerDEc ait formé une jonction Holliday n'ayant pas pu être résolue ou encore être bloquée à ce stade. Une seconde supposition pourrait être que la structure produite par ces protéines ait empêché tout simplement la réplication du plasmide. En effet, des itérons ont montré qu'ils contrôlaient la réplication du plasmide P1 en reliant deux plasmides ensemble dans une conformation en 8. Ce modèle fut nommé « Handcuffing » puisque les plasmides sont en quelque sorte menottés l'un à l'autre sans pouvoir poursuivre leur réplication (Park *et al.*, 2001 et Das et Chattoraj, 2004). Cependant, il n'y a aucune évidence ici pour un tel type de réaction.

De plus amples études sur la nature et la résolution des jonctions de Holliday formées par XerCPm via le complexe avec XerDEc, des études sur les jonctions purifiées, ainsi que leur formation dans les mutants *xer* de *P. mirabilis*, devraient pouvoir aider à comprendre davantage les caractéristiques de cette structure.

Afin de pouvoir tester l'activité des recombinases Xer au site *cer* dans un environnement plus « naturel », nous avons ensuite construit des mutants de *P. mirabilis* pour les gènes *xerC* et *xerD*. Ainsi, nous avons essayé des tests de complémentation dans ces nouvelles souches, tout comme ce que l'on avait fait précédemment dans la souche d'*E. coli* DS9029 (*xerCD*<sup>-</sup>). Malheureusement, nous n'avons pas été en mesure de voir si une différence existait quant à l'action des différentes recombinases au sein de cette espèce ; après une seule journée d'incubation dans un bouillon non sélectif, l'espèce n'a pas pu croître sur le milieu contenant de la Km. Ceci laissait à penser que le vecteur pCS210Km<sup>R</sup> avait été perdu très rapidement, et ce, même si le système Xer de *P. mirabilis* avait été réintroduit en *trans*. Ce phénomène n'a cependant pas été observé lors de nos tests de complémentation chez *E. coli* lorsque les protéines Xer provenaient de *P. mirabilis* (revoir la Figure 24). Ainsi, le phénomène de multimérisation de ColE1 dans *P. mirabilis*, observé par Goebel et Helinski (1968), serait sûrement dû à d'autres protéines retrouvées dans cette espèce, protéines qui ne seraient pas nécessairement liées au système Xer.

### 3 Études d'expression

Au moment d'entamer les recherches présentées ici, aucune étude n'avait montré les niveaux d'expression des protéines Xer au sein de leur cellule-hôte. Très récemment par contre, des fusions transcriptionnelles ont été réalisées avec les protéines Sss (homologue de XerC) et XerD de *Pseudomonas fluorescens*; les niveaux d'expression révélés pour les deux promoteurs étaient alors très faibles, de l'ordre d'une quarantaine d'unités Miller en phase stationnaire (Martínez-Granero *et al.*, 2005). Les résultats obtenus ici pour le promoteur de la protéine XerC de *P. mirabilis* montrent également un très faible taux d'expression, mais de l'ordre de la centaine d'unités Miller.

Nous aurions pu nous attendre à ce que les niveaux d'expression de XerD soient similaires à ceux retrouvés pour XerC, compte tenu de leur rôle conjoint dans la recombinaison spécifique de site. Nous avons été cependant surpris de constater que le promoteur de XerDp<sub>m</sub> permettait une expression 7 fois plus forte de l'opéron *lacZ* que celui de sa protéine partenaire, et ce, peu importe la phase de croissance de *P. mirabilis* (voir Figure 28 et Figure 29). Les études préliminaires n'ont pas montré d'influence des protéines Xer sur leur propre promoteur.

Cette différence d'expression pourrait peut-être s'expliquer par le rôle qu'a XerD au niveau du site *dif*, *in vivo*. En effet, chez *E. coli*, il a été révélé que XerDEc réalisait la première réaction d'échange au niveau du site *dif* suite à une « activation » par FtsK (Aussel *et al.*, 2002). De son côté, XerCEc ayant montré une capacité à effectuer la première réaction d'échange au niveau du site *dif* en l'absence de FtsK, la quantité réduite permettrait de s'assurer qu'elle ne vienne exercer son action avant que XerD ne soit activée. Cette quantité réduite est compensée par une plus grande efficacité de clivage (montrée *in vitro* ; Blakely *et al.*, 1997).

Le taux d'expression des ARNm des gènes *xer* révélé par hybridations Northern semble entrer en contradiction avec ce que nous avons obtenu lors des dosages  $\beta$ -gal. Les dosages  $\beta$ -gal montreraient une certaine stabilité dans l'expression de XerD, tandis que l'ARNm de *xerD<sup>Pm</sup>* a paru plutôt décliner à l'arrivée de la phase stationnaire (comparer la courbe de la Figure 28 avec la Figure 30B).

D'une part, lorsque l'on étudie l'expression en fonction du temps, il apparaît un plateau dans l'expression au moment de la phase exponentielle de croissance. À l'approche de la phase stationnaire, la quantité de  $\beta$ -gal augmente. Cependant, à ce niveau, nous ne pouvons pas savoir s'il y a réellement stabilité dans l'expression protéique ou induction de la production d'ARNm lors du déclenchement de la phase stationnaire ; dans les deux cas, la pente ne peut qu'être positive, compte tenu du ralentissement de la croissance.

D'autre part, au niveau de l'expression de l'ARNm, on remarque plutôt une diminution de cet ARNm au fur et à mesure de l'augmentation de la densité optique. Cette observation ne reste par contre que purement qualitative puisque nous n'avons pas mesuré avec exactitude ce taux d'expression.

Cette contradiction pourrait s'expliquer par la conjonction de plusieurs données. Premièrement, la protéine  $\beta$ -gal est une protéine relativement stable (Bergquist et Truman, 1978). Il est donc fort probable qu'elle s'accumule au cours du temps et qu'elle ne subit pas la dégradation qu'aurait peut-être subi XerD. Deuxièmement, *P. mirabilis* est une espèce naturellement *lacZ* et n'a donc pas le système impliqué dans le contrôle des niveaux de cette protéine dans la cellule. Et enfin, il se pourrait que le contrôle de l'expression de XerD soit réalisé au niveau de l'expression de l'ARNm, par un traitement

de l'ARNm après sa production, ou bien entendu, par l'état de stabilité de la protéine produite.

Lorsque les cellules approchent de la phase stationnaire, le temps de génération s'allonge ; l'ensemble des protéines présentes dans les cellules se partage donc moins rapidement. Dans le même temps, la cellule peut contrôler les niveaux d'ARNm (réduction de la transcription ou traitement de l'ARNm produit). Ainsi, malgré la diminution de l'ARNm, la traduction résiduelle peut s'ajouter à la stabilité de la protéine et au ralentissement de la croissance et produire un effet d'augmentation de l'expression.

Pour ce qui est de l'expression de *xerCPm*, l'extraction de l'ARN total de *P. mirabilis* à différents temps de croissance suivie de l'hybridation à une sonde *xerCPm* (marquée à la DIG) n'a pas permis de déterminer s'il y avait réellement une différence dans les niveaux d'expression (résultat non montré). Le traitement élevé (« processing ») de l'ARNm produit et révélé par un profil d'ARN s'échelonnant de 1 kb à plus de 4 kb, a rendu la première analyse difficile. Peut-être y avait-il des niveaux de clivage de l'ARNm différents selon la protéine à produire ? Nous avons donc simplement présenté les résultats de l'extraction à un temps défini en phase exponentielle pour montrer les différents ARNm produits, indication d'un opéron quant à la transcription de *xerCPm* (voir plus haut dans la discussion).

Bien entendu, ces études d'expression restent préliminaires. Les résultats montrés ici devront être confirmés et il faudrait pouvoir mesurer avec plus d'exactitude les niveaux de transcription des gènes *xer*. Il serait de plus important d'effectuer des études d'immunobuvardages de type Western afin de connaître avec plus de justesse l'expression protéique selon les phases de croissance de la cellule. En effet, il ne faut pas oublier que les dosages  $\beta$ -gal effectués ici ont permis de mesurer le niveau d'activité d'un promoteur (fusions d'opéron), mais ont pu biaiser les résultats par une trop grande modification de l'ARNm et ne nous permettent pas de connaître la stabilité de la protéine une fois exprimée ni son niveau « réel ». À ce niveau donc, nous ne pouvons que conclure qu'il existe une différence dans les niveaux d'expression entre *xerC* et *xerD*.

Des anticorps polyclonaux auraient déjà été fabriqués contre XerC d'*E. coli*; dans des immunobuvardages de type Western, ils auraient montré une spécificité suffisamment forte pour ne pas interférer avec XerDEc (Blakely *et al.*, 1993 dans les données non

montrées). Cependant, advenant le cas où la production de ces anticorps soit ardue pour une quelconque raison, il serait possible d'effectuer des marquages des protéines Xer dans le chromosome lui-même, de type fusion His-tag, MBP ou HA, molécules contre lesquelles des anticorps ont déjà été purifiés. Ces résultats pourraient alors refléter un niveau plus juste du taux de protéines Xer dans la cellule bactérienne, puisqu'elles seraient exprimées sous leur propre promoteur et n'auraient alors subi que de faibles modifications.

#### 4 Promoteurs potentiels ?

Comme déjà mentionné, avec l'avancée du séquençage de la souche H14320 de *P. mirabilis*, il devient maintenant possible d'étudier plus aisément l'organisation génomique des séquences elles-mêmes et les séquences entourant les gènes *xer*.

Comme le profil d'ARN obtenus suite à l'hybridation de *xerCPm* laisse présager un opéron, nous aurions pu supposer que XerCPm soit exprimée à partir du promoteur *dapF* comme il en a été supposé pour *xerCEc*. À partir du programme FgenesB de Softberry (<http://www.softberry.com/berry.phtml>), le gène *xerCPm* formerait un transcrite constitué de *dapF-yigA-xerC-yigB*. Un promoteur est alors retrouvé en amont (164 pb) du codon de départ de *dapF*. Cependant, un autre promoteur serait situé entre *dapF* et *yigA* (à 17 pb seulement de *yigA*), ce qui permettrait la formation d'un transcrite *yigA-xerC-yigB*. Le profil d'ARN obtenu indique d'ailleurs un fragment à 3 kb, taille qui pourrait correspondre à un tel transcrite.

Le programme FgenesB de Softberry a prédit également un opéron formé par les gènes *xerD* et *dsbC*, mais, comme discuté plus haut, cela ne semble pas être le cas d'après le profil d'ARN obtenu suite à l'hybridation avec une sonde *xerDPm*. Un terminateur Rho indépendant serait d'ailleurs retrouvé dans la séquence *dsbC* selon le programme Findterm de Softberry. Deux promoteurs potentiels sont situés directement en amont de *xerD* : l'un est à 691 pb, (mais serait alors situé dans le gène *fldB*; comme il est transcrite dans le sens inverse de *xerD*, cela créerait un ARN anti-sens qui viendrait inhiber la synthèse de flavodoxine II), tandis que le second est placé à 234 pb en amont du codon « start » supposé de *xerD*. Le second nous semble donc le plus probable, compte tenu de la taille de l'ARNm obtenu.

À partir des données obtenues des hybridations Northern et de ces brèves recherches de promoteurs potentiels par le biais d'un logiciel retrouvé sur le web, des expériences

d'extension d'amorces (« Primer Extension ») ou encore de cartographies par la nucléase S1 devraient être envisagées afin de déterminer plus précisément ces promoteurs.

## 5 Phénotype des souches de *P. mirabilis xerC<sup>-</sup>* et/ou *xerD<sup>-</sup>*

Chez *E. coli*, les mutants *xer* sont incapables de résoudre les dimères chromosomiques pouvant être formés suite à un nombre impair d'événements de recombinaison homologe. Ainsi, un nombre accru de filaments apparaît, filaments qui contiennent des nucléoïdes aberrants qui sont incorrectement partitionnés (les masses d'ADN sont mal ségréguées) (Blakely *et al.*, 1991, 1993). Cette filamentation serait causée par le système SOS, lui-même probablement induit par le déchirement de l'ADN par le septum de division cellulaire (phénomène de « guillotine » ; Hendricks *et al.*, 2000 et revoir le point 6.9 de l'introduction).

Chez *Proteus mirabilis*, la filamentation est un processus normal du cycle cellulaire de *Proteus* lorsque les cellules entrent en contact avec une surface solide (différenciation en cellules d'essaimage). Cependant, les cellules se présentent sous forme de courts bâtonnets lorsqu'elles sont « nageuses » (en bouillon) ou dites végétatives lorsque la croissance se fait sur des géloses réduisant ou inhibant le mouvement pas essaimage (Fraser *et al.*, 2000). Ainsi, les cellules sauvages mises sur gélose LSW étaient de courts bâtonnets (environ 1-2 µm de longueur; Figure 26). Cependant, tout comme cela avait été le cas pour les cellules d'*E. coli*, les mutants *xer* de *P. mirabilis* ont présenté des filaments.

D'un autre côté, afin d'observer la « forme » de l'ADN à l'intérieur des diverses souches, des colorations DAPI ont été tentées, cette fois sur des cellules cultivées en bouillon. Elles ont alors révélé une certaine différence au niveau des nucléoïdes, entre la souche « sauvage » et les souches mutantes (Figure 27) : ce phénotype correspondait aussi à ce qui avait déjà été observé chez les mutants *xer* d'*E. coli* (Blakely *et al.*, 1993).

Cependant, contrairement à ce que nous aurions dû observer, la souche sauvage elle-même a présenté un certain niveau de filamentation lorsque prélevée d'un bouillon. Ce résultat nous apparaissait bizarre puisque les cellules en bouillon, soit les cellules dites « nageuses » sont normalement individuelles (voir plus haut). Par contre, le nucléoïde des filaments de la souche sauvage est apparu « normal », tandis qu'il semblait « granuleux » dans certains des filaments des souches mutantes.

Nous ne savons pas pourquoi ces cellules sauvages se sont présentées allongées en bouillon, alors qu'elles apparaissent normales sur une gélose inhibant l'essaimage. Nous en avons simplement supposé que cette filamentation pouvait être une caractéristique de la souche employée. En effet, comme déjà mentionné, cette souche était un isolat clinique. Bien qu'elle présente toutes les caractéristiques de l'espèce *mirabilis* du genre *Proteus*<sup>7</sup>, il se pourrait que des « mutations » spontanées existaient déjà dans la souche 240-82, mutations qui auraient pu expliquer ce phénotype particulier. D'autres souches du département, par la suite vérifiées par microscopie, n'ont pas présenté ce niveau de filamentation (non montré). Le phénotype engendré par les mutations *xer* chez ces souches (et chez la souche H14320 en train d'être séquencée par l'Institut Sanger) serait donc à vérifier.

La propriété qu'a l'espèce de migrer de façon particulière sur une gélose LA, n'a pas été inhibée par les mutations dans les gènes *xer* de *P. mirabilis*. Cependant, le phénotype en a été légèrement modifié et il était différent selon le gène *xer* touché et le mutant *xerCD<sup>-</sup>* a montré un profil de migration très similaire à celui de la souche *xerD<sup>-</sup>* (voir l'annexe 1).

Il est évident que le phénomène d'essaimage est un phénomène complexe impliquant une panoplie de gènes. Ainsi, le mouvement migratoire ne dépend pas uniquement de l'augmentation de la synthèse flagellaire, mais aussi de la densité et du contact avec une surface solide (revoir le point 7.2.2 de l'introduction). De plus, les dernières études de ce phénomène chez *P. mirabilis* montrent l'implication d'autres gènes, même si leur fonction première ne semble pas liée à la migration en tant que telle. Par exemple, la putrescine, qui est un constituant de la membrane externe de *P. mirabilis* et ayant diverses fonctions, servirait de signal extracellulaire pour la différenciation en cellules d'essaimage (Sturgill et Rather, 2004). D'un autre côté, les gènes impliqués dans la synthèse des LPS (lipopolysaccharides) sont impliqués dans l'élongation des cellules d'essaimage (Belas *et al.*, 1995).

Nous avons de la difficulté à comprendre pourquoi les mutations dans les gènes *xer* ont montré un profil d'essaimage différent de celui de l'isolat clinique non muté. Chez

---

<sup>7</sup> La vérification de l'isolat clinique avait été réalisée préalablement par Micheline Desmeule au département, à l'aide de tests classiques employés pour l'identification des entérobactéries (Bergey *et al.*, 1984).

*S. typhimurium*, un mutant *xerC*<sup>-</sup> a perdu sa capacité d'essaimage (Rasika M. Harshey, comm. Pers.). Chez *P. mirabilis*, il n'est que modifié, très légèrement. La propriété de migration n'a donc pas été touchée. Nous ne pouvons donc pas vraiment tirer de conclusion quant à ce phénomène : il y a eu formation des vagues caractéristiques, elles semblent avoir été formées adéquatement, dans les mêmes temps (bien entendu, en tenant compte des différences des taux de croissance)...

Se pourrait-il que la mutation dans *xerD* ait engendré un effet polaire, et ainsi affecter le gène *dsbC* qui suit ? [Ce gène, chez *E. coli*, code pour l'isomérase disulfide et a un rôle dans la formation des ponts disulfures de plusieurs protéines périplasmiques (Missiakas *et al.*, 1994)]. Il serait possible alors qu'elle soit indirectement liée au mouvement migratoire de *P. mirabilis* de la même façon que *galU* (glucose-1-phosphate uridylyltransférase), impliquée dans la synthèse de la paroi cellulaire, affecte l'élongation des cellules d'essaimage (Belas *et al.*, 1995).

Afin de vérifier cet éventuel effet polaire, nous pourrions d'abord examiner l'essaimage à la suite de la complémentation par un vecteur en faible nombre de copies exprimant le gène *xerD*<sup>Pm</sup> sauvage et, si la modification du phénotype est toujours maintenue, par la complémentation par un vecteur exprimant le gène *dsbC*. Par la suite, l'inactivation de *dsbC* en tant que telle devrait être testée, sans que le gène *xerD* ne soit touché.

Le ralentissement de la croissance causé par les mutations *xer*, pourrait expliquer en partie les phénotypes observés sur ces géloses. Le taux de génération de la souche *xerD*<sup>-</sup> ne montre pas une très grande différence avec le taux de génération de la souche sauvage (Tableau VII). En effet, il semblerait que la souche PmD1 ait un taux de génération de près de 33±2,15 min, ce qui est légèrement plus grand que celui de la souche sauvage (29,2±0,2 min). La différence est donc faible, mais tout de même présente. De la même façon, le taux de génération obtenue suite à 3 expériences pour la souche *xerC*<sup>-</sup> est d'environ 33,6±0,8 min. Dans tous les cas, le ralentissement de croissance n'a pas empêché l'obtention d'une densité cellulaire adéquate pour l'enclenchement du mouvement de migration sur gélose LA, mais il pourrait expliquer la différence observée dans l'épaisseur des phases de consolidation. Il ne faut pas oublier, de plus, que les taux de génération ont été mesurés à 37°C et non pas à 28°C, température d'incubation employée afin d'observer plus adéquatement les divers phénotypes des mouvements par essaimage.

## 6 Le site *dif* de *P. mirabilis*

Dans un tout autre ordre d'idée, nous avons déterminé le site *dif* de *P. mirabilis* (Figure 35), site qui fut reconnu par les protéines Xer (Figure 36). Ce site s'est révélé fortement identique au site *dif* d'*E. coli*, avec seulement 3 pb de différence, toutes situées dans le site de liaison à XerC.

Tel qu'attendu, les protéines XerEc ont également pu s'y lier de façon individuelle. Cependant, on a pu noter que la protéine XerCPm s'y est liée de façon un peu plus efficace que la protéine XerCEc et a probablement induit une courbure plus grande du site (en effet, malgré le poids moléculaire plus grand de MBP-XerCPm, cette protéine a permis la migration un peu plus rapide du site *dif* que n'a pu le faire MBP-XerCEc). Ceci pourrait être expliqué par cette différence de 3 pb dans le site de liaison de XerC. Du côté des protéines XerD, la liaison fut aussi efficace que ce qui fut observé au niveau du site *difEc* (comparez la Figure 20B à la Figure 36).

Tout comme il a été vu avec les protéines d'*H. influenzae* pouvant interagir avec les recombinases d'*E. coli* (Neilson *et al.* 1999), il en a été de même avec les protéines de *P. mirabilis*. Cela a démontré encore une fois que les recombinases Xer ont des domaines communs permettant de former une structure définie pouvant mener à bien la recombinaison au site *dif*.

Au départ, cette recherche du site *dif* de *P. mirabilis* nous apparaissait intéressante afin de rechercher les potentiels éléments RAG (voir l'introduction de cette thèse). En effet, les régions flanquant le site *dif* de *Salmonella enterica* serovar Typhi ou Typhimurium, espèces très proches d'*E. coli* d'un point de vue phylogénétique, ont très peu de similarités avec la zone retrouvée chez *E. coli*, et pourtant, le même biais a été retrouvé pour les motifs RAG dans cette région (Corre et Louarn, 2002). Il aurait été ainsi intéressant de pouvoir comparer ces régions avec la région terminale de *P. mirabilis* afin de renforcer le modèle de l'équipe Louarn voulant que ces éléments aient un rôle important à jouer dans le contrôle de la recombinaison au site *dif*.

Comme mentionné dans la section des résultats, au moment de réaliser cette thèse, le séquençage de *Proteus* était inaccessible. À l'heure actuelle, l'assemblage et les vérifications du séquençage de la souche PmH14320 sont pratiquement terminés; ainsi, la recherche d'un tel octamère dans la région entourant le site *difPm* est maintenant possible.

Ce que nous pouvons remarquer cependant, c'est la divergence de l'organisation de la séquence de la région entourant le site *dif* entre les deux souches de *P. mirabilis*, alors que le site *dif* de 28 pb est parfaitement conservé (voir l'annexe 3). En effet, la région adjacente au site *dif* de l'isolat clinique Pm240-82 UM correspond à un *orf* similaire à *asrA* ou à un régulateur potentiel de transcription. Par contre, lorsque nous avons cherché à savoir ce à quoi pouvait correspondre la séquence plus en amont de cette région chez Pm240-82 UM (alignement montré en annexe 4), une recherche dans la souche PmH14320 nous a indiqué une séquence très fortement identique à *ycfH* d'*E. coli*, une désoxyribonucléase potentielle. Cette séquence n'est pas à la proximité de *dif* dans la souche PmH14320. Il apparaît donc intéressant de constater qu'une forme de réarrangement chromosomique s'est produite dans ce que nous supposons la région terminale, au sein d'une même espèce bactérienne.

# **CONCLUSION**

## Conclusion

Les divers travaux présentés ici ont permis de montrer que les séquences retrouvées chez *P. mirabilis* correspondant à *xer* pouvaient bel et bien servir de recombinases chez *E. coli*. Ces recombinases, non seulement ont pu se lier à des sites spécifiques de recombinaison (*cer* et *dif*), mais ont également montré une activité de clivage au niveau de ces mêmes sites. Cependant, la complémentation d'une souche d'*E. coli xer<sup>-</sup>* n'a pas pu se faire avec un plasmide contenant les gènes *xerCPm* et *xerDEc*. Plusieurs hypothèses ont été émises dans la discussion afin d'essayer de comprendre pourquoi la coopération entre ces deux protéines a été si inefficace. D'un autre côté, le vecteur-test contenant le site *cer* (pCS210Km<sup>R</sup>) s'est révélé très instable dans la souche *P. mirabilis xerCD<sup>-</sup>*. Ainsi, même si le système Xer de *P. mirabilis* est efficace en tant que système de résolution au niveau du site *cer*, il ne l'est pas ou est insuffisant pour le maintien d'un vecteur contenant ce site.

Au niveau cellulaire, les mutations des gènes *xer* chez *P. mirabilis* n'ont pas induit de phénotype réellement différent de celui rencontré chez *E. coli*. Par contre, nous avons constaté que le profil de migration par essaimage sur gélose, caractéristique importante de l'espèce, avait légèrement été modifié, mais cela pourrait être simplement une conséquence du ralentissement de la croissance engendré par les mutations *xer*.

D'un autre côté, nous avons pu montrer que le promoteur de la recombinase XerD de *P. mirabilis* était sept fois plus actif chez cette espèce que sa protéine partenaire XerC (par des dosages de  $\beta$ -galactosidase), et que la taille de l'ARNm de *xerDPm* obtenu après croissance en bouillon LA à 37°C, pouvait indiquer un ARNm monocistronique. Enfin, nous avons été en mesure de déterminer le site *dif* de *P. mirabilis*.

De plus amples études seront nécessaires afin de déterminer avec plus d'exactitude les niveaux d'expression des recombinases au sein de la cellule bactérienne. Les données relevées ici laissent supposer que les niveaux d'expression protéique ont été plutôt stables au cours de la croissance de *P. mirabilis*, et qu'une accumulation de la  $\beta$ -gal serait à l'origine de l'augmentation des niveaux d'unités Miller en phase stationnaire. Au contraire, la quantité d'ARNm de *xerD* n'a pas semblé constante d'après une première étude d'hybridation de type Northern. D'autres hybridations de type Northern devront être réalisées afin de lever toute ambiguïté quant aux résultats ainsi obtenus. Des conditions de croissance variées pourraient également être testées afin de connaître les changements

d'expression de ces recombinases dans d'autres situations. Il faudrait de plus évaluer la stabilité de l'ARNm en parallèle avec la stabilité protéique.

Et, qu'en est-il de la localisation de ces protéines au sein de la cellule ? Sont-elles diffuses au sein de la cellule, ou bien logées précisément près de leur site d'action ? Il semblerait qu'aucune localisation n'ait été observée chez *E. coli* (F.X. Barre, comm. Pers.). Au sein du laboratoire, des études préliminaires ont été aussi réalisées (chez *P. mirabilis* cette fois) avec un marquage par un anticorps anti-His, mais aucun résultat n'a été obtenu pour l'instant.

Et au sujet de la collaboration avec la protéine FtsK ? Les études concernant le système Xer tendent à se concentrer de plus en plus sur cette molécule FtsK qui relie la résolution des dimères chromosomiques et la division cellulaire. Une brève recherche dans le génome de PmH14320 montre que cette protéine existe également dans cette espèce (identité d'environ 76%). Y a-t-il un véritable contact entre ces différentes protéines afin de stabiliser le complexe pour la recombinaison au site *dif* ?

*E. coli* demeure un modèle de choix (simplicité pour la croissance, d'utilisation) pour une première compréhension des divers systèmes régissant la vie bactérienne dans son ensemble. Par contre, on se rend compte que le système Xer peut servir à d'autres usages que la monomérisation des chromosomes par le biais de la recombinaison au site *dif*. Que son rôle peut s'avérer vital chez d'autres espèces, et même, qu'il peut être impliqué dans la virulence chez certaines d'entre-elles. Pourrait-il en être de même chez *P. mirabilis* ? Comme déjà mentionné, il existe un lien entre la différenciation cellulaire de cette espèce et la production de facteurs de virulence. Bien que le mouvement migratoire par essaimage n'ait été modifié que légèrement par l'inactivation des gènes *xer*, est-ce que le système Xer pourrait être impliqué indirectement dans l'efficacité de la colonisation du tractus urinaire ?

Beaucoup reste à définir sur les rôles et sur les mécanismes exacts du système Xer au sein de l'ensemble des espèces bactériennes. Par l'étude de ce système chez *P. mirabilis* nous voulions vérifier si des généralités pouvaient encore être établies.

## Bibliographie

- Abremski, K. E., et Hoess, R. H. (1992). Evidence for a second conserved arginine residue in the integrase family of recombination proteins. *Protein Eng* 5, 87-91.
- Abremski, K., Wierzbicki, A., Frommer, B., et Hoess, R. H. (1986). Bacteriophage P1 Cre-*loxP* site-specific recombination. Site-specific DNA topoisomerase activity of the Cre recombination protein. *J Biol Chem* 261, 391-396.
- Alén, C., Sherratt, D. J., et Colloms, S. D. (1997). Direct interaction of aminopeptidase A with recombination site DNA in Xer site-specific recombination. *EMBO J* 16, 5188-5197.
- Allison, C. et Hughes, C. (1991). Closely linked genetic loci required for swarm cell differentiation and multicellular migration by *Proteus mirabilis*. *Mol Microbiol* 5, 1975-1982.
- Arciszewska, L. K., Baker, R. A., Hallet, B., et Sherratt, D. J. (2000). Coordinated control of XerC and XerD catalytic activities during Holliday junction resolution. *J Mol Biol* 299, 391-403.
- Arciszewska, L. K., et Sherratt, D. J. (1995). Xer site-specific recombination *in vitro*. *EMBO J* 14, 2112-2120.
- Arciszewska, L. K., Grainge, I., et Sherratt, D. J. (1997). Action of site-specific recombinases XerC and XerD on tethered Holliday junctions. *EMBO J* 16, 3731-3743.
- Arciszewska, L., Grainge, I., et Sherratt, D. (1995). Effects of Holliday junction position on Xer-mediated recombination *in vitro*. *EMBO J* 14, 2651-2660.
- Argos, P., Landy, A., Abremski, K., Egan, J. B., Haggard-Ljungquist, E., Hoess, R. H., Kahn, M. L., Kalionis, B., Narayana, S. V., Pierson, L. S., 3rd, et *et al.* (1986). The integrase family of site-specific recombinases: regional similarities and global diversity. *EMBO J* 5, 433-440.
- Ausebel, F. M., Bent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. et Stuhl, K. (1987-1997). *Current protocols in molecular biology*. Wiley, New York.
- Aussel, L., Barre, F. X., Aroyo, M., Stasiak, A., Stasiak, A. Z., et 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.
- Azaro, M. A., et Landy, A. (1997). The isomeric preference of Holliday junctions influences resolution bias by lambda integrase. *EMBO J* 16, 3744-55.
- Barre, F. X., Aroyo, M., Colloms, S. D., Helfrich, A., Cornet, F., et Sherratt, D. J. (2000). FtsK functions in the processing of a Holliday junction intermediate during bacterial chromosome segregation. *Genes Dev* 14, 2976-2988.
- Bath, J., Sherratt, D. J., et Colloms, S. D. (1999). Topology of Xer recombination on catenanes produced by lambda integrase. *J Mol Biol* 289, 873-883.

- Bath, J., Wu, L. J., Errington, J., et Wang, J. C. (2000). Role of *Bacillus subtilis* SpoIIIE in DNA transport across the mother cell-prespore division septum. *Science* 290, 995-997.
- Bäumler, A. J., Tsolis, R. M., van der Velden, A. W., Stojiljkovic, I., Anic, S., et Heffron, F. (1996). Identification of a new iron regulated locus of *Salmonella typhi*. *Gene* 183, 207-213.
- Becker, J., et Brendel, M. (1996). Molecular characterization of the *xerC* gene of *Lactobacillus leichmannii* encoding a site-specific recombinase and two adjacent heat shock genes. *Curr Microbiol* 32, 232-236.
- Begg, K. J., Dewar, S. J., et Donachie, W. D. (1995). A new *Escherichia coli* cell division gene, *ftsK*. *J Bacteriol* 177, 6211-6222.
- Belas, R., Erskine, D., et Flaherty, D. (1991). Transposon mutagenesis in *Proteus mirabilis*. *J Bacteriol* 173, 6289-6293.
- Belas, R., Goldman, M., et Ashliman, K. (1995). Genetic analysis of *Proteus mirabilis* mutants defective in swarmer cell elongation. *J Bacteriol* 177, 823-8.
- Bergey, D. H., Holt, J. G., et Murray, R. G. E. (1984). *Bergey's manual of systematic bacteriology*. Editeur en chef : John G. Holt. Direction éditoriale : R.G.E. Murray...[*et al.*] ; avec la contribution de 124 collaborateurs. Williams & Wilkins. Baltimore.
- Bergquist, P. L., et Truman, P. (1978). Degradation of missense mutant  $\beta$ -galactosidase proteins in *Escherichia coli* K-12. *Mol Gen Genet* 164, 105-108.
- Bigot, S., Corre, J., Louarn, J. M., Cornet, F., et 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, 876-886.
- Black, D. S., Kelly, A. J., Mardis, M. J., et Moyed, H. S. (1991). Structure and organization of *hip*, an operon that affects lethality due to inhibition of peptidoglycan or DNA synthesis. *J Bacteriol* 173, 5732-5739.
- Blake, J. A., Ganguly, N., et Sherratt, D. J. (1997). DNA sequence of recombinase-binding sites can determine Xer site-specific recombination outcome. *Mol Microbiol* 23, 387-398.
- Blakely, G. W., Davidson, A. O., et Sherratt, D. J. (1997). Binding and cleavage of nicked substrates by site-specific recombinases XerC and XerD. *J Mol Biol* 265, 30-39.
- Blakely, G. W., et Sherratt, D. J. (1994). Interactions of the site-specific recombinases XerC and XerD with the recombination site dif. *Nucleic Acids Res* 22, 5613-5620.
- Blakely, G., Colloms, S., May, G., Burke, M., et Sherratt, D. (1991). *Escherichia coli* XerC recombinase is required for chromosomal segregation at cell division. *New Biol* 3, 789-798.
- Blakely, G., et Sherratt, D. (1996). Determinants of selectivity in Xer site-specific recombination. *Genes Dev* 10, 762-773.

- Blakely, G., May, G., McCulloch, R., Arciszewska, L. K., Burke, M., Lovett, S. T., et Sherratt, D. J. (1993). Two related recombinases are required for site-specific recombination at *dif* and *cer* in *E. coli* K12. *Cell* 75, 351-361.
- Blattner, F. R., Plunkett, G., 3rd, Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B., et Shao, Y. (1997). The complete genome sequence of *Escherichia coli* K-12. *Science* 277, 1453-1474.
- Boyle, D. S., Grant, D., Draper, G. C., et Donachie, W. D. (2000). All major regions of FtsK are required for resolution of chromosome dimers. *J Bacteriol* 182, 4124-4127.
- Bregu, M., Sherratt, D. J., et 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., et 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.
- Busenlehner, L. S., Pennella, M. A., Giedroc, D. P. (2003). The SmtB/ArsR family of metalloregulatory transcriptional repressors: Structural insights into prokaryotic metal resistance. *FEMS Microbiol Rev* 27, 131-43.
- Campbell, J. L., Richardson, C. C. and Studier, F. W. (1978). Genetic recombination and complementation between bacteriophage T7 and cloned fragments of T7 DNA. *Proc Natl Acad Sci U S A* 75, 2276-2280.
- Cannon, P. M., et Strike, P. (1992). Complete nucleotide sequence and gene organization of plasmid NTP16. *Plasmid* 27, 220-230.
- Cao, Y., et Hayes, F. (1999). A newly identified, essential catalytic residue in a critical secondary structure element in the integrase family of site-specific recombinases is conserved in a similar element in eucaryotic type IB topoisomerases. *J Mol Biol* 289, 517-527.
- Cao, Y., Hallet, B., Sherratt, D. J., et Hayes, F. (1997). Structure-function correlations in the XerD site-specific recombinase revealed by pentapeptide scanning mutagenesis. *J Mol Biol* 274, 39-53.
- Capiaux, H., Cornet, F., Corre, J., Guijo, M. I., Perals, K., Rebollo, J. E., et Louarn, J. M. (2001). Polarization of the *Escherichia coli* chromosome. A view from the terminus. *Biochimie* 83, 161-170.
- Capiaux, H., Lesterlin, C., Perals, K., Louarn, J. M., et Cornet, F. (2002). A dual role for the FtsK protein in *Escherichia coli* chromosome segregation. *EMBO Rep* 3, 532-536.
- Chalker, A. F., Lupas, A., Ingraham, K., So, C. Y., Lunsford, R. D., Li, T., Bryant, A., Holmes, D. J., Marra, A., Pearson, S. C., Ray, J., Burnham, M. K., Palmer, L. M., Biswas, S., et Zalacain, M. (2000). Genetic characterization of gram-positive homologs of the XerCD site-specific recombinases. *J Mol Microbiol Biotechnol* 2, 225-233.

- Charlier, D., Gigot, D., Huysveld, N., Roovers, M., Pierard, A., et 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., et 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 ColEI multimers. *J Mol Biol* 250, 392-406.
- Charlier, D., Kholti, A., Huysveld, N., Gigot, D., Maes, D., Thia-Toong, T. L., et Glansdorff, N. (2000). Mutational analysis of *Escherichia coli* PepA, a multifunctional DNA-binding aminopeptidase. *J Mol Biol* 302, 411-426.
- Chatwin, H. M., et Summers, D. K. (2001). Monomer-dimer control of the ColE1 P(*cer*) promoter. *Microbiology* 147, 3071-3081.
- Chen, J. C., Weiss, D. S., Ghigo, J. M., et Beckwith, J. (1999). Septal localization of FtsQ, an essential cell division protein in *Escherichia coli*. *J Bacteriol* 181, 521-530.
- Chen, S. H., Merican, A. F., et Sherratt, D. J. (1997). DNA binding of *Escherichia coli* arginine repressor mutants altered in oligomeric state. *Mol Microbiol* 24, 1143-1156.
- Chen, Y., et Rice, P. A. (2003). New insight into site-specific recombination from Flp recombinase-DNA structures. *Annu Rev Biophys Biomol Struct* 32, 135-159.
- Chen, Y., Narendra, U., Iype, L. E., Cox, M. M., et Rice, P. A. (2000). Crystal structure of a Flp recombinase-Holliday junction complex: assembly of an active oligomer by helix swapping. *Mol Cell* 6, 885-897.
- Cherepanov, P. P., et Wackernagel, W. (1995). Gene disruption in *Escherichia coli*: Tc<sup>R</sup> and Km<sup>R</sup> cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. *Gene* 158, 9-14.
- Clerget, M. (1991). Site-specific recombination promoted by a short DNA segment of plasmid R1 and by a homologous segment in the terminus region of the *Escherichia coli* chromosome. *New Biol* 3, 780-8.
- Colloms, S. D., Alén, C., et 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.
- Colloms, S. D., Bath, J., et Sherratt, D. J. (1997). Topological selectivity in Xer site-specific recombination. *Cell* 88, 855-864.
- Colloms, S. D., McCulloch, R., Grant, K., Neilson, L., et Sherratt, D. J. (1996). Xer-mediated site-specific recombination *in vitro*. *EMBO J* 15, 1172-1181.

- Colloms, S. D., Sykora, P., Szatmari, G., et Sherratt, D. J. (1990). Recombination at ColE1 *cer* requires the *Escherichia coli xerC* gene product, a member of the lambda integrase family of site-specific recombinases. *J Bacteriol* 172, 6973-6980.
- Cornet, F., Hallet, B., et 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.
- Cornet, F., Louarn, J., Patte, J., et 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., Mortier, I., Patte, J., et 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 site dif*. *J Bacteriol* 176, 3188-3195.
- Corre, J., Cornet, F., Patte, J., et Louarn, J. M. (1997). Unraveling a region-specific hyper-recombination phenomenon: genetic control and modalities of terminal recombination in *Escherichia coli*. *Genetics* 147, 979-989.
- Corre, J., et 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.
- Corre, J., Patte, J., et Louarn, J. M. (2000). Prophage lambda induces terminal recombination in *Escherichia coli* by inhibiting chromosome dimer resolution. An orientation-dependent cis-effect lending support to bipolarization of the terminus. *Genetics* 154, 39-48.
- Craig, N. L. (2002). Mobile DNA : an introduction. *In* Mobile DNA II. Edited by Craig, N. L., Craigie, R., Geller, M., et Lambowitz, A. M. ASM Press. Washington D.C., 3-11.
- Czaplewski, L. G., North, A. K., Smith, M. C., Baumberg, S., et Stockley, P. G. (1992). Purification and initial characterization of AhrC: the regulator of arginine metabolism genes in *Bacillus subtilis*. *Mol Microbiol* 6, 267-275.
- Czirok, A., Matsushita, M., et Vicsek, T. (2001). Theory of periodic swarming of bacteria: application to *Proteus mirabilis*. *Phys Rev E Stat Nonlin Soft Matter Phys* 63, 031915.
- Das, N., et Chattoraj, D. K. (2004). Origin pairing ('handcuffing') and unpairing in the control of P1 plasmid replication. *Mol Microbiol* 54, 836-849.
- Datsenko, K. A., et Wanner, B. L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* 97, 6640-6645.
- Dekkers, L. C., Phoelich, C. C., van der Fits, L., Lugtenberg, B. J. (1998). A site-specific recombinase is required for competitive root colonization by *Pseudomonas fluorescens* WCS365. *Proc Natl Acad Sci U S A* 95, 7051-7056.
- Donachie, W. D. (2002). FtsK: Maxwell's Demon? *Mol Cell* 9, 206-207.

- Draper, G. C., McLennan, N., Begg, K., Masters, M., et Donachie, W. D. (1998). Only the N-terminal domain of FtsK functions in cell division. *J Bacteriol* 180, 4621-4627.
- Dreyfuss, G., Adam, S. A., et Choi, Y. D. (1984). Physical change in cytoplasmic messenger ribonucleoproteins in cells treated with inhibitors of mRNA transcription. *Mol Cell Biol* 4, 415-423.
- Dröge, P., et Cozzarelli, N. R. (1992). Topological structure of DNA knots and catenanes. *Methods Enzymol* 212, 120-130.
- Ennifar, E., Meyer, J. E., Buchholz, F., Stewart, A. F., et Suck, D. (2003). Crystal structure of a wild-type Cre recombinase-*loxP* synapse reveals a novel spacer conformation suggesting an alternative mechanism for DNA cleavage activation. *Nucleic Acids Res* 31, 5449-5460.
- Errington, J., Bath, J., et Wu, L. J. (2001). DNA transport in bacteria. *Nat Rev Mol Cell Biol* 2, 538-545.
- Espeli, O., Lee, C., et Marians, K. J. (2003). A physical and functional interaction between *Escherichia coli* FtsK and topoisomerase IV. *J Biol Chem* 278, 44639-44644.
- Esposito, D., et Scocca, J. J. (1997). The integrase family of tyrosine recombinases: evolution of a conserved active site domain. *Nucleic Acids Res* 25, 3605-3614.
- Fellay, R., Frey, J., et Krisch, H. (1987). Interposon mutagenesis of soil and water bacteria: a family of DNA fragments designed for in vitro insertional mutagenesis of gram-negative bacteria. *Gene* 52, 147-154.
- Fernandes, B. L. et da Costa, S. O. (1996). High efficiency of transformation of *Proteus mirabilis* with a pUC19 derivative vector directs the expression and secretion of *Bacillus subtilis*  $\alpha$ -amylase gene. *J Microbiol Methods* 26, 147-150.
- Ferreira, H., Butler-Cole, B., Burgin, A., Baker, R., Sherratt, D. J., et 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.
- Ferreira, H., Sherratt, D., et Arciszewska, L. (2001). Switching catalytic activity in the XerCD site-specific recombination machine. *J Mol Biol* 312, 45-57.
- Fraser, G. M., Furness, R. B., et Hughes, C. (2000). Swarming migration by *Proteus* and related bacteria. *In* Prokaryotic development. Edited by Brun, Y. V., et Shimkets, L. J. American Society for Microbiology, Washington, DC., 381-401.
- Fuqua, W. C. (1992). An improved chloramphenicol resistance gene cassette for site-directed marker replacement mutagenesis. *Biotechniques* 12, 223-225.
- Gay, P., Le Coq, D., Steinmetz, M., Ferrari, et E., Hoch, J. A. (1983). Cloning structural gene *sacB*, which codes for exoenzyme levansucrase of *Bacillus subtilis*: expression of the gene in *Escherichia coli*. *J Bacteriol* 153 1424-31.

- Glew, M. D., Marena, M., Rosengarten, R., et Citti, C. (2002). Surface diversity in *Mycoplasma agalactiae* is driven by site-specific DNA inversions within the *vpma* multigene locus. *J Bacteriol* *184*, 5987-5998.
- Goebel, W., et Helinski, D. R. (1968). Generation of higher multiple circular DNA forms in bacteria. *Proc Natl Acad Sci U S A* *61*, 1406-1413.
- Goodman, S. D., et Scocca, J. J. (1989). Nucleotide sequence and expression of the gene for the site-specific integration protein from bacteriophage HP1 of *Haemophilus influenzae*. *J Bacteriol* *171*, 4232-4240.
- Goodwin, T. J., Butler, M. I., et Poulter, R. T. (2003). Cryptons: a group of tyrosine-recombinase-encoding DNA transposons from pathogenic fungi. *Microbiology* *149*, 3099-3109.
- Gopaul, D. N., Guo, F., et Van Duyne, G. D. (1998). Structure of the Holliday junction intermediate in Cre-*loxP* site-specific recombination. *EMBO J* *17*, 4175-4187.
- Gourlay, S. C., et Colloms, S. D. (2004). Control of Cre recombination by regulatory elements from Xer recombination systems. *Mol Microbiol* *52*, 53-65.
- Grainge, I., et Jayaram, M. (1999). The integrase family of recombinase: organization and function of the active site. *Mol Microbiol* *33*, 449-456.
- Grainge, I., et Sherratt, D. J. (1999). Xer site-specific recombination. DNA strand rejoining by recombinase XerC. *J Biol Chem* *274*, 6763-6769.
- Grant, S. G. N., Jesse, J., Bloom, F. R. et Hanahan, D. (1990) Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. *Proc Natl Acad. Sci USA* *87*, 4645-4649.
- Guhathakurta, A., et Summers, D. (1995). Involvement of ArgR and PepA in the pairing of ColE1 dimer resolution sites. *Microbiology* *141 (Pt 5)*, 1163-1171.
- Guhathakurta, A., Viney, I., et Summers, D. (1996). Accessory proteins impose site selectivity during ColE1 dimer resolution. *Mol Microbiol* *20*, 613-620.
- Guijo, M. I., Patte, J., del Mar Campos, M., Louarn, J. M., et Rebollo, J. E. (2001). Localized remodeling of the *Escherichia coli* chromosome: the patchwork of segments refractory and tolerant to inversion near the replication terminus. *Genetics* *157*, 1413-1423.
- Guo, F., Gopaul, D. N., et van Duyne, G. D. (1997). Structure of Cre recombinase complexed with DNA in a site-specific recombination synapse. *Nature* *389*, 40-46.
- Guo, F., Gopaul, D. N., et Van Duyne, G. D. (1999). Asymmetric DNA bending in the Cre-*loxP* site-specific recombination synapse. *Proc Natl Acad Sci U S A* *96*, 7143-7148.
- Guynn, L. J., Dai, W., et Benedik, M. J. (1998). Nuclease overexpression mutants of *Serratia marcescens*. *J Bacteriol* *180*, 2262-2264.

- Hakenbeck, R., Balmelle, N., Weber, B., Gardes, C., Keck, W., et de Saizieu, A. (2001). Mosaic genes and mosaic chromosomes: intra- and interspecies genomic variation of *Streptococcus pneumoniae*. *Infect Immun* 69, 2477-86.
- Hakkaart, M. J., van den Elzen, P. J., Veltkamp, E., et Nijkamp, H. J. (1984). Maintenance of multicopy plasmid Clo DF13 in *E. coli* cells: evidence for site-specific recombination at *parB*. *Cell* 36, 203-209.
- Hallet, B., Arciszewska, L. K., et 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.
- Hallet, B., et 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.
- Haren, L., Ton-Hoang, B., et Chandler, M. (1999). Integrating DNA: transposases and retroviral integrases. *Annu Rev Microbiol* 53, 245-281.
- Harshey, R. M. (2003). Bacterial motility on a surface: many ways to a common goal. *Annu Rev Microbiol* 57, 249-273.
- Hayes, F., et Sherratt, D. J. (1997). Recombinase binding specificity at the chromosome dimer resolution site *dif* of *Escherichia coli*. *J Mol Biol* 266, 525-537.
- Hayes, F., Lubetzki, S. A., et Sherratt, D. J. (1997). *Salmonella typhimurium* specifies a circular chromosome dimer resolution system which is homologous to the Xer site-specific recombination system of *Escherichia coli*. *Gene* 198, 105-110.
- Hendricks, E. C., Szerlong, H., Hill, T., et Kuempel, P. (2000). Cell division, guillotining of dimer chromosomes and SOS induction in resolution mutants (*dif*, *xerC* and *xerD*) of *Escherichia coli*. *Mol Microbiol* 36, 973-81.
- Henrichsen, J. (1972). Bacterial surface translocation: a survey and a classification. *Bacteriol Rev* 36, 478-503.
- Herrero, M., de Lorenzo, V., et Timmis, K. N. (1990). Transposon vectors containing non-antibiotic resistance selection markers for cloning and stable chromosomal insertion of foreign genes in gram-negative bacteria. *J Bacteriol* 172, 6557-6567.
- Hickman, A. B., Waninger, S., Scocca, J. J., et Dyda, F. (1997). Molecular organization in site-specific recombination: the catalytic domain of bacteriophage HP1 integrase at 2.7 Å resolution. *Cell* 89, 227-237.
- Hill, T. M., et Mariani, K. J. (1990). *Escherichia coli* Tus protein acts to arrest the progression of DNA replication forks *in vitro*. *Proc Natl Acad Sci U S A* 87, 2481-2485.
- Himmelreich, R., Plagens, H., Hilbert, H., Reiner, B., et Herrmann, R. (1997). Comparative analysis of the genomes of the bacteria *Mycoplasma pneumoniae* and *Mycoplasma genitalium*. *Nucleic Acids Res* 25, 701-712.

- Hiraga, S., Ichinose, C., Onogi, T., Niki, H., et Yamazoe, M. (2000). Bidirectional migration of SeqA-bound hemimethylated DNA clusters and pairing of *oriC* copies in *Escherichia coli*. *Genes Cells* 5, 327-341.
- Hodgman, T. C., Griffiths, H., et Summers, D. K. (1998). Nucleoprotein architecture and ColE1 dimer resolution: a hypothesis. *Mol Microbiol* 29, 545-558.
- Höfte, M., Dong, Q., Kourambas, S., Krishnapillai, V., Sherratt, D., et Mergeay, M. (1994). The *sss* gene product, which affects pyoverdinin production in *Pseudomonas aeruginosa* 7NSK2, is a site-specific recombinase. *Mol Microbiol* 14, 1011-1020.
- Hojgaard, A., Szerlong, H., Tabor, C., et Kuempel, P. (1999). Norfloxacin-induced DNA cleavage occurs at the *dif* resolvase locus in *Escherichia coli* and is the result of interaction with topoisomerase IV. *Mol Microbiol* 33, 1027-1036.
- Huber, K. E., et Waldor, M. K. (2002). Filamentous phage integration requires the host recombinases XerC and XerD. *Nature* 417, 656-659.
- Iuchi, S., et Weiner, L. (1996). Cellular and molecular physiology of *Escherichia coli* in the adaptation to aerobic environments. *J Biochem (Tokyo)* 120, 1055-1063.
- James, R., Alfano, J., Ham, H. et Collmer, A. (1995). Use of Tn5tac1 to clone a *pel* gene encoding a highly alkaline, asparagine-rich pectate lyase isosyme from an *Erwinia chrysanthemi* EC16 mutant with deletions affecting the major pectate lysase isosymes. *J Bacteriol* 177, 4553-4556.
- Jayaram, M. (1997). The *cis-trans* paradox of integrase. *Science* 276, 49-51.
- Jouan, L., et Szatmari, G. (2003). Interactions of the *Caulobacter crescentus* XerC and XerD recombinases with the *E. coli dif* site. *FEMS Microbiol Lett* 222, 257-262.
- Kamadurai, H. B., Subramaniam, S., Jones, R. B., Green-Church, K. B., et Foster, M. P. (2003). Protein folding coupled to DNA binding in the catalytic domain of bacteriophage lambda integrase detected by mass spectrometry. *Protein Sci* 12, 620-626.
- Kaniga, K., Delor, I., et Cornelis, G. R. (1991). A wide-host-range suicide vector for improving reverse genetics in gram-negative bacteria: inactivation of the *blaA* gene of *Yersinia enterocolitica*. *Gene* 109, 137-141.
- Kikuchi, Y., et Nash, H. A. (1979). Nicking-closing activity associated with bacteriophage lambda *int* gene product. *Proc Natl Acad Sci U S A* 76, 3760-3764.
- Klemm, P. (1986). Two regulatory *fim* genes, *fimB* and *fimE*, control the phase variation of type 1 fimbriae in *Escherichia coli*. *EMBO J* 5, 1389-93.
- Kolot, M. N. (1990). [Structural-functional organization of the *par* region of the ColN plasmid]. *Mol Biol (Mosk)* 24, 495-500.
- Krogh, B. O., et Shuman, S. (2002). A poxvirus-like type IB topoisomerase family in bacteria. *Proc Natl Acad Sci U S A* 99, 1853-1858.

- Kuempel, P. L., Henson, J. M., Dircks, L., Tecklenburg, M., et 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.
- Kuempel, P., Hogaard, A., Nielsen, M., Nagappan, O., et Tecklenburg, M. (1996). Use of a transposon (Tndif) to obtain suppressing and nonsuppressing insertions of the *dif* resolvase site of *Escherichia coli*. *Genes Dev* 10, 1162-1171.
- Kwon, H. J., Tirumalai, R., Landy, A., et Ellenberger, T. (1997). Flexibility in DNA recombination: structure of the lambda integrase catalytic core. *Science* 276, 126-131.
- Kwon, O., Georgellis, D., et Lin, E. C. (2000). Phosphorelay as the sole physiological route of signal transmission by the arc two-component system of *Escherichia coli*. *J Bacteriol* 182, 3858-3862.
- Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680-5.
- Lau, I. F., Filipe, S. R., Soballe, B., Okstad, O. A., Barre, F. X., et Sherratt, D. J. (2003). Spatial and temporal organization of replicating *Escherichia coli* chromosomes. *Mol Microbiol* 49, 731-743.
- Lee, Y. S., Han, J. S., Jeon, Y., et Hwang, D. S. (2001). The *arc* two-component signal transduction system inhibits *in vitro* *Escherichia coli* chromosomal initiation. *J Biol Chem* 276, 9917-9923.
- Lemon, K. P., et Grossman, A. D. (2001). The extrusion-capture model for chromosome partitioning in bacteria. *Genes Dev* 15, 2031-2041.
- Leslie, N. R., et Sherratt, D. J. (1995). Site-specific recombination in the replication terminus region of *Escherichia coli*: functional replacement of *dif*. *EMBO J* 14, 1561-1570.
- Lesterlin, C., Barre, F. X., et Cornet, F. (2004). Genetic recombination and the cell cycle: what we have learned from chromosome dimers. *Mol Microbiol* 54, 1151-1160.
- Li, Y., Youngren, B., Sergueev, K., et Austin, S. (2003). Segregation of the *Escherichia coli* chromosome terminus. *Mol Microbiol* 50, 825-834.
- Liaw, S.-J., Lai, H.-C., Ho, S.-W., Luh, K.-T., Wang, W. B. (2003). Role of RsmA in the regulation of swarming motility and virulence factor expression in *Proteus mirabilis*. *J Med Microbiol* 52, 19-28.
- Liaw, S.-J., Lai, H.-C., Wang, W. B. (2004). Modulation of swarming and virulence by fatty acids through the RsbA protein in *Proteus mirabilis*. *Infect Immun* 72, 6836-6845.
- Lim, D. B., Oppenheim, J. D., Eckhardt, T., et 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 U S A* 84, 6697-6701.

- Liu, G., Draper, G. C., et Donachie, W. D. (1998). FtsK is a bifunctional protein involved in cell division and chromosome localization in *Escherichia coli*. *Mol Microbiol* 29, 893-903.
- Liu, X., et De Wulf, P. (2004). Probing the ArcA-P modulon of *Escherichia coli* by whole genome transcriptional analysis and sequence recognition profiling. *J Biol Chem* 279, 12588-12597.
- Lobry, J. R. (1996). Asymmetric substitution patterns in the two DNA strands of bacteria. *Mol Biol Evol* 13, 660-665.
- Lobry, J. R., et Louarn, J. M. (2003). Polarisation of prokaryotic chromosomes. *Curr Opin Microbiol* 6, 101-108.
- Louarn, J. M., Louarn, J., Francois, V., et Patte, J. (1991). Analysis and possible role of hyperrecombination in the termination region of the *Escherichia coli* chromosome. *J Bacteriol* 173, 5097-5104.
- Louarn, J. M., Kuempel, P., et Cornet, F. (2004). The terminus region of the *Escherichia coli* chromosome, or all's well that ends well. *In* The bacterial chromosome. Edited par Higgins, P. N., ASM Press, Section II, chapitre 13.
- Louarn, J., Cornet, F., Francois, V., Patte, J., et Louarn, J. M. (1994). Hyperrecombination in the terminus region of the *Escherichia coli* chromosome: possible relation to nucleoid organization. *J Bacteriol* 176, 7524-7531.
- Lovett, S. T., et Clark, A. J. (1985). Cloning of the *Escherichia coli* *recJ* chromosomal region and identification of its encoded proteins. *J Bacteriol* 162, 280-285.
- Lovett, S. T., et Kolodner, R. D. (1991). Nucleotide sequence of the *Escherichia coli* *recJ* chromosomal region and construction of *recJ*-overexpression plasmids. *J Bacteriol* 173, 353-364.
- Maas, W. K. (1994). The arginine repressor of *Escherichia coli*. *Microbiol Rev* 58, 631-640.
- Martínez-Granero, F., Capdevila, S., Sánchez-Contreras, Martín, M., Rivilla, R. (2005). Two site-specific recombinases are implicated in phenotypic variation and competitive rhizosphere colonization in *Pseudomonas fluorescens*. *Microbiol* 151, 975-983.
- Massey, T. H., Aussel, L., Barre, F. X., et Sherratt, D. J. (2004). Asymmetric activation of Xer site-specific recombination by FtsK. *EMBO Rep* 5, 399-404.
- McCulloch, R., Burke, M. E., et Sherratt, D. J. (1994a). Peptidase activity of *Escherichia coli* aminopeptidase A is not required for its role in Xer site-specific recombination. *Mol Microbiol* 12, 241-251.
- McCulloch, R., Coggins, L. W., Colloms, S. D., et Sherratt, D. J. (1994b). Xer-mediated site-specific recombination at *cer* generates Holliday junctions *in vivo*. *EMBO J* 13, 1844-1855.

- Messer, W. (2002). The bacterial replication initiator DnaA. DnaA and oriC, the bacterial mode to initiate DNA replication. *FEMS Microbiol Rev* 26, 355-374.
- Miller, J. H. (1972). *Experiments in molecular genetics*. Cold Spring Harbor Laboratory. Cold Spring Harbor, N.Y. 466 p.
- Miller, V. L., et Mekalanos, J. J. (1988). A novel suicide vector and its use in construction of insertion mutations: osmoregulation of outer membrane proteins and virulence determinants in *Vibrio cholerae* requires *toxR*. *J Bacteriol* 170, 2575-2583.
- Missiakas, D., Georgopoulos, C., et Raina, S. (1994). The *Escherichia coli dsbC (xprA)* gene encodes a periplasmic protein involved in disulfide bond formation. *EMBO J* 13, 2013-2020.
- Mizuuchi, K., et Baker, T. A. (2002). Chemical mechanisms for mobilizing DNA. *In* *Mobile DNA II*. Edited by Craig, N. L., Craigie, R., Geller, M., et Lambowitz, A. M. ASM Press. Washington D.C., 12-23.
- Molina, F., et Skarstad, K. (2004). Replication fork and SeqA focus distributions in *Escherichia coli* suggest a replication hyperstructure dependent on nucleotide metabolism. *Mol Microbiol* 52, 1597-1612.
- Morlon, J., Chartier, M., Bidaud, M., et Lazdunski, C. (1988). The complete nucleotide sequence of the colicinogenic plasmid ColA. High extent of homology with ColE1. *Mol Gen Genet* 211, 231-243.
- Mukherjee, K. J., Rowe, D. C., Watkins, N. A., et Summers, D. K. (2004). Studies of single-chain antibody expression in quiescent *Escherichia coli*. *Appl Environ Microbiol* 70, 3005-3012.
- Nash, H. A. (1996). Site-specific recombination : integration, excision, resolution and inversion of defined DNA segments. *In* *Escherichia coli and Salmonella typhimurium Cellular and molecular biology*. Second Edition. Neihardt F. C., ASM Press. Washington. Vol. II, 2363-2376.
- Neilson, L., Blakely, G., et Sherratt, D. J. (1999). Site-specific recombination at *dif* by *Haemophilus influenzae* XerC. *Mol Microbiol* 31, 915-926.
- Ni, J., Sakanyan, V., Charlier, D., Glansdorff, N., et Van Duyne, G. D. (1999). Structure of the arginine repressor from *Bacillus stearothermophilus*. *Nat Struct Biol* 6, 427-432.
- Niki, H., Jaffe, A., Imamura, R., Ogura, T., et Hiraga, S. (1991). The new gene *mukB* codes for a 177 kd protein with coiled-coil domains involved in chromosome partitioning of *E. coli*. *EMBO J* 10, 183-193.
- Niki, H., Yamaichi, Y., et Hiraga, S. (2000). Dynamic organization of chromosomal DNA in *Escherichia coli*. *Genes Dev* 14, 212-223.

- Nunes-Düby, S. E., Azaro, M. A., et Landy, A. (1995). Swapping DNA strands and sensing homology without branch migration in lambda site-specific recombination. *Curr Biol* 5, 139-148.
- Nunes-Düby, S. E., Kwon, H. J., Tirumalai, R. S., Ellenberger, T., et Landy, A. (1998). Similarities and differences among 105 members of the Int family of site-specific recombinases. *Nucleic Acids Res* 26, 391-406.
- Nunes-Düby, S. E., Matsumoto, L., et Landy, A. (1987). Site-specific recombination intermediates trapped with suicide substrates. *Cell* 50, 779-788.
- Pargellis, C. A., Nunes-Düby, S. E., de Vargas, L. M., et Landy, A. (1988). Suicide recombination substrates yield covalent lambda integrase-DNA complexes and lead to identification of the active site tyrosine. *J Biol Chem* 263, 7678-7685.
- Park, K., Han, E., Paulsson, J., et Chatteraj, D. K. (2001). Origin pairing ('handcuffing') as a mode of negative control of P1 plasmid copy number. *EMBO J* 20, 7323-7332.
- Patient, M. E., et Summers, D. K. (1993). ColE1 multimer formation triggers inhibition of *Escherichia coli* cell division. *Mol Microbiol* 9, 1089-1095.
- Paul, S., et Summers, D. (2004). ArgR and PepA, accessory proteins for XerCD-mediated resolution of ColE1 dimers, are also required for stable maintenance of the P1 prophage. *Plasmid* 52, 63-68.
- Perals, K., Capiiaux, H., Vincourt, J. B., Louarn, J. M., Sherratt, D. J., et Cornet, F. (2001). Interplay between recombination, cell division and chromosome structure during chromosome dimer resolution in *Escherichia coli*. *Mol Microbiol* 39, 904-913.
- Perals, K., Cornet, F., Merlet, Y., Delon, I., et 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.
- Prikryl, J., Hendricks, E. C., et Kuempel, P. L. (2001). DNA degradation in the terminus region of resolvase mutants of *Escherichia coli*, and suppression of this degradation and the Dif phenotype by *recD*. *Biochimie* 83, 171-176.
- Rather, P. N. (2005). Swarmer cell differentiation in *Proteus mirabilis*. *Envir Microbiol* 0, doi :10.1111/j.1462-2920.00806.x.
- Rauprich, O., Matsushita, M., Weijer, C. J., Siegert, F., Esipov, S. E., et Shapiro, J. A. (1996). Periodic phenomena in *Proteus mirabilis* swarm colony development. *J Bacteriol* 178, 6525-6538.
- Recchia, G. D., Aroyo, M., Wolf, D., Blakely, G., et Sherratt, D. J. (1999). FtsK-dependent and -independent pathways of Xer site-specific recombination. *EMBO J* 18, 5724-5734.
- Recchia, G. D., et Sherratt, D. J. (1999). Conservation of xer site-specific recombination genes in bacteria. *Mol Microbiol* 34, 1146-1148.

- Recchia, G. D., et Sherratt, D. J. (2002). Gene acquisition in bacteria by Integron-mediated site-specific recombination. *In* Mobile DNA II. Edited by Craig, N. L., Craigie, R., Geller, M., et Lambowitz, A. M. ASM Press. Washington D.C., 162-176.
- Rehel, N., et Szatmari, G. (1996). Characterization of the stable maintenance of the *Shigella flexneri* plasmid pHS-2. *Plasmid* 36, 183-190.
- Reyrat, J.M., Pelicic, V., Gicquel, B., et Rappuoli, R. (1998). Counterselectable markers: untapped tools for bacterial genetics and pathogenesis. *Infect Immun* 66 4011-7.
- Reichmann, P., et Hakenbeck, R. (2002). A XerD recombinase with unusual active site motifs in *Streptococcus pneumoniae*. *J Mol Microbiol Biotechnol* 4, 101-110.
- Richaud, C., et Printz, C. (1988). Nucleotide sequence of the *dapF* gene and flanking regions from *Escherichia coli* K12. *Nucleic Acids Res* 16, 10367.
- Richaud, C., Higgins, W., Mengin-Lecreulx, D., et Stragier, P. (1987). Molecular cloning, characterization, and chromosomal localization of *dapF*, the *Escherichia coli* gene for diaminopimelate epimerase. *J Bacteriol* 169, 1454-1459.
- Rowe, D. C., et Summers, D. K. (1999). The quiescent-cell expression system for protein synthesis in *Escherichia coli*. *Appl Environ Microbiol* 65, 2710-2715.
- Rózalski, A., Sidorczyk, Z., et Kotelko, K. (1997). Potential virulence factors of *Proteus* bacilli. *Microbiol. Mol Biol Rev* 61, 65-89.
- Sadowski, P. D. (1995). The Flp recombinase of the 2-microns plasmid of *Saccharomyces cerevisiae*. *Prog Nucleic Acid Res Mol Biol* 51, 53-91.
- Saleh, O. A., Perals, C., Barre, F. X., et Allemand, J. F. (2004). Fast, DNA-sequence independent translocation by FtsK in a single-molecule experiment. *EMBO J* 23, 2430-2439.
- Sambrook, J., Fritsch, E. F., et Maniatis, T. (1989). *Molecular cloning: a laboratory manual*, 2<sup>nd</sup> Ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Salzberg, S. L., Salzberg, A. J., Kerlavage, A. R., et Tomb, J. F. (1998). Skewed oligomers and origins of replication. *Gene* 217, 57-67.
- Sciochetti, S. A., et Piggot, P. J. (2000). A tale of two genomes: resolution of dimeric chromosomes in *Escherichia coli* and *Bacillus subtilis*. *Res Microbiol* 151, 503-511.
- Sciochetti, S. A., Piggot, P. J., et Blakely, G. W. (2001). Identification and characterization of the *dif* site from *Bacillus subtilis*. *J Bacteriol* 183, 1058-1068.
- Sciochetti, S. A., Piggot, P. J., Sherratt, D. J., et Blakely, G. (1999). The *ripX* locus of *Bacillus subtilis* encodes a site-specific recombinase involved in proper chromosome partitioning. *J Bacteriol* 181, 6053-6062.
- Shaikh, A. C., et Sadowski, P. D. (1997). The Cre recombinase cleaves the *lox* site in *trans*. *J Biol Chem* 272, 5695-5702.

- Sharpe, M. E., Chatwin, H. M., Macpherson, C., Withers, H. L., et Summers, D. K. (1999). Analysis of the ColE1 stability determinant Rcd. *Microbiology* 145 ( Pt 8), 2135-2144.
- Sherratt, D. J. (2003). Bacterial chromosome dynamics. *Science* 301, 780-785.
- Sherratt, D. J., Arciszewska, L. K., Blakely, G., Colloms, S., Grant, K., Leslie, N., et McCulloch, R. (1995). Site-specific recombination and circular chromosome segregation. *Philos Trans R Soc Lond B Biol Sci* 347, 37-42.
- Sherratt, D. J., et Wigley, D. B. (1998). Conserved themes but novel activities in recombinases and topoisomerases. *Cell* 93, 149-152.
- Sherratt, D. J., Lau, I. F., et Barre, F. X. (2001). Chromosome segregation. *Curr Opin Microbiol* 4, 653-659.
- Sherratt, D. J., Soballe, B., Barre, F. X., Filipe, S., Lau, I., Massey, T., et Yates, J. (2004). Recombination and chromosome segregation. *Philos Trans R Soc Lond B Biol Sci* 359, 61-69.
- Simon, R., Prifer, U., et Pühler, A. (1983). A broad host range mobilization system for *in vivo* genetic engineering: transposon mutagenesis in gram negative bacteria. *Biotechnology*. 784-791.
- Sirois, S., et Szatmari, G. (1995). Detection of XerC and XerD recombinases in gram-negative bacteria of the family *Enterobacteriaceae*. *J Bacteriol* 177, 4183-4186.
- Slack, F. J., Serror, P., Joyce, E., et Sonenshein, A. L. (1995). A gene required for nutritional repression of the *Bacillus subtilis* dipeptide permease operon. *Mol Microbiol* 15, 689-702.
- Smith, M. C., et Thorpe, H. M. (2002). Diversity in the serine recombinases. *Mol Microbiol* 44, 299-307.
- Spiers, A. J., et Sherratt, D. J. (1997). Relating primary structure to function in the *Escherichia coli* XerD site-specific recombinase. *Mol Microbiol* 24, 1071-1082.
- Spiers, A. J., et Sherratt, D. J. (1999). C-terminal interactions between the XerC and XerD site-specific recombinases. *Mol Microbiol* 32, 1031-1042.
- Steiner, W. W., et Kuempel, P. L. (1998a). Cell division is required for resolution of dimer chromosomes at the *dif* locus of *Escherichia coli*. *Mol Microbiol* 27, 257-268.
- Steiner, W. W., et Kuempel, P. L. (1998b). Sister chromatid exchange frequencies in *Escherichia coli* analyzed by recombination at the *dif* resolvase site. *J Bacteriol* 180, 6269-6275.
- Steiner, W., Liu, G., Donachie, W. D., et 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., et 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., et 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., et 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.

Sträter, N., et Lipscomb, W. N. (1995). Transition state analogue L-leucinephosphonic acid bound to bovine lens leucine aminopeptidase: X-ray structure at 1.65 Å resolution in a new crystal form. *Biochemistry* 34, 9200-9210.

Sträter, N., Sherratt, D. J., et 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.

Strohmaier, H., Noiges, R., Kotschan, S., Sawers, G., Hogenauer, G., Zechner, E. L., et Koraimann, G. (1998). Signal transduction and bacterial conjugation: characterization of the role of ArcA in regulating conjugative transfer of the resistance plasmid R1. *J Mol Biol* 277, 309-316.

Studier, F. W., et Moffatt, B. A. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol* 189, 113-130.

Sturgill, G., et Rather, P. N. (2004). Evidence that putrescine acts as an extracellular signal required for swarming in *Proteus mirabilis*. *Mol Microbiol* 51, 437-46.

Subramanya, H. S., Arciszewska, L. K., Baker, R. A., Bird, L. E., Sherratt, D. J., et Wigley, D. B. (1997). Crystal structure of the site-specific recombinase, XerD. *EMBO J* 16, 5178-5187.

Summers, D. (1998). Timing, self-control and a sense of direction are the secrets of multicopy plasmid stability. *Mol Microbiol* 29, 1137-1145.

Summers, D. K. (1989). Derivatives of *ColE1 cer* show altered topological specificity in site-specific recombination. *EMBO J* 8, 309-315.

Summers, D. K., Beton, C. W., et Withers, H. L. (1993). Multicopy plasmid instability: the dimer catastrophe hypothesis. *Mol Microbiol* 8, 1031-1038.

Summers, D. K., et 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.

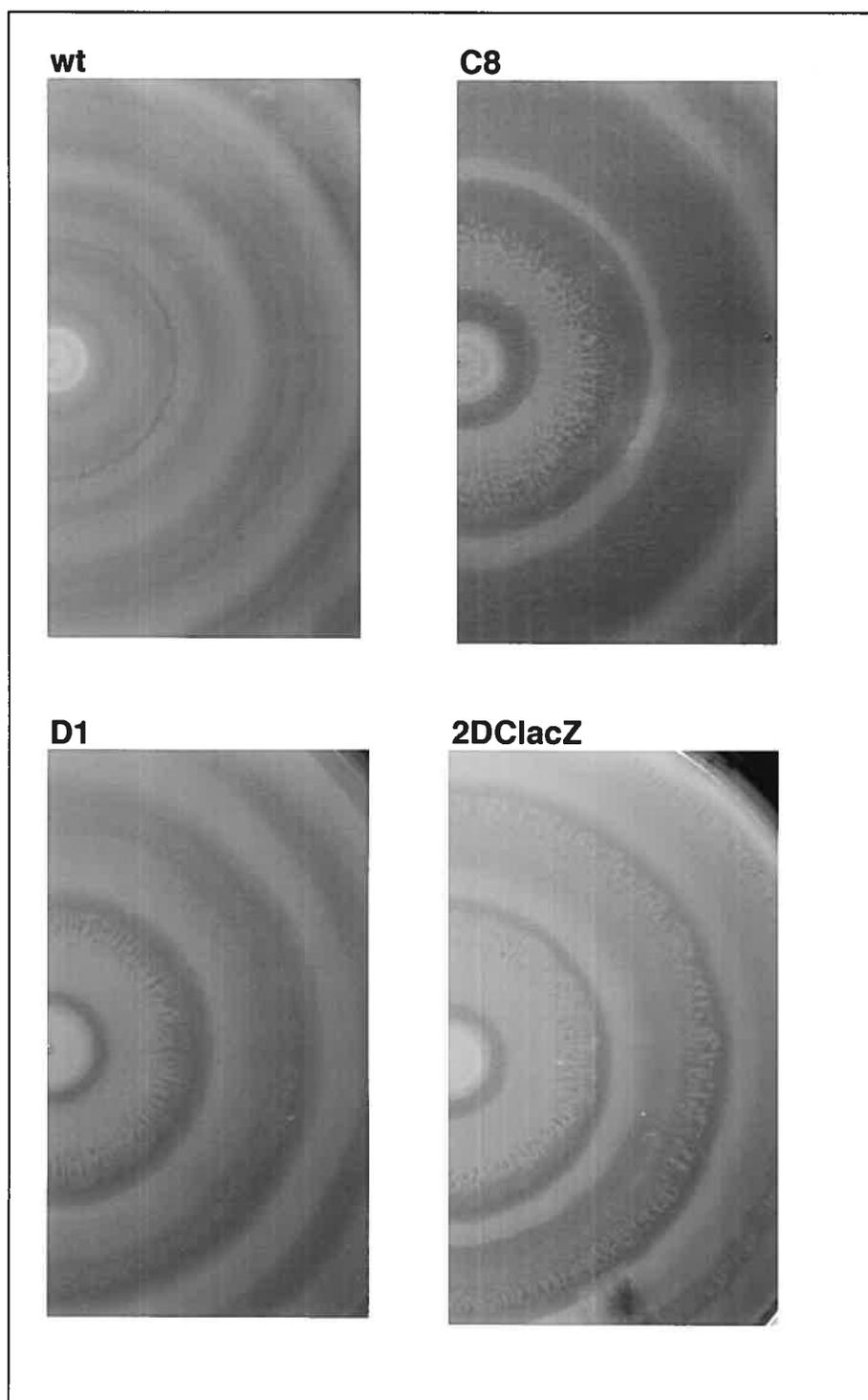
Summers, D. K., et Sherratt, D. J. (1988). Resolution of *ColE1* dimers requires a DNA sequence implicated in the three-dimensional organization of the *cer* site. *EMBO J* 7, 851-858.

- Summers, D., Yaish, S., Archer, J., et Sherratt, D. (1985). Multimer resolution systems of ColE1 and ColK: localisation of the crossover site. *Mol Gen Genet* 201, 334-338.
- Sunnerhagen, M., Nilges, M., Otting, G., et 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.
- Swalla, B. M., Gumport, R. I., et Gardner, J. F. (2003). Conservation of structure and function among tyrosine recombinases: homology-based modeling of the lambda integrase core-binding domain. *Nucleic Acids Res* 31, 805-818.
- Taylor, A. (1993). Aminopeptidases : towards a mechanism of action. *Trends Bioch Sci* 18, 167-172.
- Tecklenburg, M., Naumer, A., Nagappan, O., et 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 U S A* 92, 1352-1356.
- Tian, G., Lim, D., Carey, J., et Maas, W. K. (1992). Binding of the arginine repressor of *Escherichia coli* K12 to its operator sites. *J Mol Biol* 226, 387-397.
- Tribble, G., Ahn, Y. T., Lee, J., Dandekar, T., et Jayaram, M. (2000). DNA recognition, strand selectivity, and cleavage mode during integrase family site-specific recombination. *J Biol Chem* 275, 22255-22267.
- Valens, M., Penaud, S., Rossignol, M., Cornet, F., et Boccard, F. (2004). Macrodomain organization of the *Escherichia coli* chromosome. *EMBO J* 23, 4330-4341.
- Van Duyne, G. D., Ghosh, G., Maas, W. K., et Sigler, P. B. (1996) Structure of the oligomerization and L-arginine binding domain of the arginine repressor of *Escherichia coli*. *J Mol Biol* 256, 377-91.
- Van Duyne, G. D. (2001). A structural view of Cre-*loxP* site-specific recombination. *Annu Rev Biophys Biomol Struct* 30, 87-104.
- Van Duyne, G. D., (2002). A structural view of tyrosine-recombinase site-specific recombination. *In Mobile DNA II*. Edited by Craig, N. L., Craigie, R., Geller, M., et Lambowitz, A. M. ASM Press. Washington D.C., 93-117.
- Villion, M. (1998). Clonage et caractérisation du gène *xerD* de *Proteus mirabilis*. Mémoire présenté en vue de l'obtention du grade de Maître ès sciences (M. Sc.) en microbiologie. Université de Montréal. 78 p.
- Villion, M., et Szatmari, G. (1998). Cloning and characterisation of the *Proteus mirabilis* *xerD* gene. *FEMS Microbiol Lett* 164, 83-90.
- Villion, M., et Szatmari, G. (2003). The XerC recombinase of *Proteus mirabilis*: characterization and interaction with other tyrosine recombinases. *FEMS Microbiol Lett* 226, 65-71.

- Waldor, M. K., et Mekalanos, J. J. (1996). Lysogenic conversion by a filamentous phage encoding cholera toxin. *Science* 272, 1910-1914.
- Wang, L., et Lutkenhaus, J. (1998). FtsK is an essential cell division protein that is localized to the septum and induced as part of the SOS response. *Mol Microbiol* 29, 731-740.
- Wassarman, K. M., Storz, G. (2000). 6S RNA regulates *E. coli* RNA polymerase activity. *Cell* 101, 613-623.
- Wojciak, J. M., Sarkar, D., Landy, A., et Clubb, R. T. (2002). Arm-site binding by lambda-integrase: solution structure and functional characterization of its amino-terminal domain. *Proc Natl Acad Sci U S A* 99, 3434-3439.
- Yanisch-Perron, C., Vieira, J., et Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* 33, 103-119.
- Yasueda, H., Horii, T., et Itoh, T. (1989). Structural and functional organization of ColE2 and ColE3 replicons. *Mol Gen Genet* 215, 209-216.
- Yates, J., Aroyo, M., Sherratt, D. J., et Barre, F. X. (2003). Species specificity in the activation of Xer recombination at *dif* by FtsK. *Mol Microbiol* 49, 241-249.
- Yu, X. C., Tran, A. H., Sun, Q., et Margolin, W. (1998a). Localization of cell division protein FtsK to the *Escherichia coli* septum and identification of a potential N-terminal targeting domain. *J Bacteriol* 180, 1296-1304.
- Yu, X. C., Weihe, E. K., et Margolin, W. (1998b). Role of the C terminus of FtsK in *Escherichia coli* chromosome segregation. *J Bacteriol* 180, 6424-6428.

# **ANNEXES**

## Annexe 1



**Croissance de *P. mirabilis* sur milieu LA.** Les quatre souches ont été cultivées à 28°C pendant 48 h. WT : souche sauvage (Pm 240-82), C8 : *xerC* ::*lacZYA*-Cm<sup>R</sup>, D1 : *xerD*::*lacZYA*-Cm<sup>R</sup>, 2DClacZ : *xerD* ::Cm<sup>R</sup> *xerC* ::*lacZYA*-Km<sup>R</sup>.

## Annexe 2



FEMS Microbiology Letters 226 (2003) 65–71

**FEMS**  
 MICROBIOLOGY  
 Letters

www.fems-microbiology.org

## The XerC recombinase of *Proteus mirabilis*: characterization and interaction with other tyrosine recombinases

Manuela Villion, George Szatmari\*

Département de microbiologie et immunologie, Université de Montréal, CP 6128, Succ. Centre-Ville, Montréal, QC, Canada H3C 3J7

Received 10 March 2003; received in revised form 21 May 2003; accepted 7 July 2003

First published online 16 August 2003

### Abstract

XerC and XerD are two site-specific recombinases, which act on different sites to maintain replicons in a monomeric state. This system, which was first discovered and studied in *Escherichia coli*, is present in several species including *Proteus mirabilis*, where the XerD recombinase was previously characterized by our laboratory. In this paper, we report the presence of the *xerC* gene in *P. mirabilis*. Using in vitro reactions, we show that the two *P. mirabilis* recombinases display binding and cleavage activity on the *E. coli* *dir* site and the ColE1 *cer* site, together or in collaboration with *E. coli* recombinases. In vivo, *P. mirabilis* XerC and XerD are able to resolve and monomerize a plasmid containing two *cer* sites, increasing its stability. However, *P. mirabilis* XerC, in combination with *E. coli* XerD, is unable to perform these functions.

© 2003 Published by Elsevier B.V. on behalf of the Federation of European Microbiological Societies.

**Keywords:** *Proteus mirabilis*; Site-specific recombination; Tyrosine recombinase; XerC; *cer*

### 1. Introduction

The *Escherichia coli* Xer system reacts on specific sites such as the *dir* site found in the terminal replication region of the chromosome, the *cer* site found in ColE1, or the *psi* site from pSC101 [1,2], to maintain these replicons in a monomeric state (reviewed in [3]).

This system comprises two enzymes, XerC and XerD, which belong to the integrase/tyrosine recombinase family [1,4]. Sequences similar to the *xerCD* genes are found in many bacterial genera (reviewed in [5]); and see as example the tyrosine recombinase website: <http://mywebpages.comcast.net/donnespot/home.html>. They have been found in Gram-positives, such as *Bacillus subtilis* [6,7], *Lactobacillus leichmannii* [8], *Lactobacillus casei* (Flandin and Szatmari, in preparation), *Staphylococcus aureus* or *Streptococcus pneumoniae* [9], as well as Gram-negative bacteria,

such as *E. coli* and other Enterobacteriaceae [10], *Pseudomonas aeruginosa* [11], *Haemophilus influenzae* [12], *Vibrio cholerae* [13], and many others.

In 1968, Goebel and Helinski demonstrated the multimerization of the ColE1 plasmid in *Proteus mirabilis* strains [14]. Considering this result, we could expect either that this species does not contain an efficient multimer resolution system, namely the Xer system, or that this system is not functional on the *cer* site. Such a result would be surprising considering the importance of this system in *E. coli* and the phylogenetic link between this species and *P. mirabilis*. It was shown that Xer sites may be recombined in *P. mirabilis*. Despite this observation, these plasmids displayed a high degree of multimerization and were rapidly lost from this host [10]. We have already cloned and sequenced the *P. mirabilis xerD* gene. Although highly similar to the *E. coli xerD* gene, it has been shown to poorly complement an *E. coli xerD* mutant. We demonstrated this was not due to inefficient binding to the *cer* site used in the test, as was shown by a gel retardation assay [15].

In this paper, we report the cloning of the *P. mirabilis xerC* gene and the characterization of its product by DNA binding assays, complementation tests and in vitro recombination reactions.

\* Corresponding author. Tel.: +1 (514) 343-5767.  
 Fax: +1 (514) 343-5701.

## 2. Materials and methods

### 2.1. Bacterial strains, plasmids and growth conditions

The *E. coli* strains and plasmids used in this study are given in Table 1. The *P. mirabilis* strain was a clinical isolate (UM 240-82, see [10]). Bacteria were grown in LB broth or plated on LB agar containing antibiotics, 5-bromo-4-chloro-3-indolyl- $\beta$ -D-galactopyranoside (X-gal) and/or isopropyl- $\beta$ -D-thiogalactoside (IPTG) as required. Ampicillin (Ap) was used at 100  $\mu\text{g ml}^{-1}$ , tetracycline (Tc) at 6.25  $\mu\text{g ml}^{-1}$ , kanamycin at 50  $\mu\text{g ml}^{-1}$ , X-gal at 40  $\mu\text{g ml}^{-1}$ .

### 2.2. DNA manipulations

All routine DNA manipulations were performed as described in [16] and [17]. DNA purifications of extracted fragments from agarose gels were done with Sephaglas BandPrep kit (Amersham Pharmacia Biotech) or with QIAquick Gel Extraction kit (Qiagen). Southern hybridizations were done as described in [10].

Restriction enzymes, Klenow fragment of *E. coli*, T4 polynucleotide kinase and T4 DNA ligase were obtained from New England Biolabs (NEB). Shrimp alkaline phosphatase (SAP) was from Roche Diagnostics.

### 2.3. DNA sequencing

To clone the *P. mirabilis xerC* gene (*xerCPm*), the same polymerase chain reaction (PCR) conditions were used as for *P. mirabilis xerD* (*xerDpm*): using degenerate primers. PCR was performed to amplify a 450-bp fragment, which was shown to be different from that obtained for *xerDpm* [15]. Inverse PCR was then carried out as previously described [15], except that genomic DNA was digested with *Pst*I before proceeding to the ligation further used for inverse PCR. This PCR was done using a Biometra thermocycler with 100 pmol of the following primers: PrC2 (CTTTACCTAATACTCGGACTTC) and mirC2 (CTCTTTTGCCACTCATTTAC) and cycling conditions: 94°C/30 s, 52°C/1 min, 72°C/2 min for 30 cycles, with a final extension at 72°C/5 min, using Taq polymerase.

To confirm the sequence of *xerCPm*, PCR was performed on genomic DNA using 100 pmol of the following primers: PmcF (CAAGAAGGTATGGGAACG) and PmcR (CGGATAAACGCTAAGGTC). Cycling conditions were: 95°C/15 s, 52°C/30 s, 72°C/45 s for 30 cycles, with a final extension at 72°C/5 min, using Vent polymerase with a Hybaid thermocycler. The *xerCPm* sequence was subsequently submitted to GenBank (accession number AF033498) after sequencing the PCR product with an ABI 373 autosequencer using the following primers: PmxerF (GGATAAACTCGCGCAG) and PmxerR (CCATCACAGGAACATTGTC).

Another amplified PCR product was used using Taq

polymerase with primers PmcF and PmcR) to obtain clones with *xerCPm* in plasmid pIZ-PC (see [15]). These clones were used to construct other plasmids for complementation tests.

### 2.4. Protein purifications

*xerCPm* was amplified by PCR with 100 pmol of the following primers: PrCS' (GATCGCAAGCATATGAGCCA) and PmcR. Cycling conditions were exactly the same as for sequence analysis. Once purified, the product obtained was phosphorylated with T4 polynucleotide kinase and inserted into pMal-c2 digested with *Xba*I, and dephosphorylated with SAP followed by an overnight ligation with T4 DNA ligase. Once the clone (pPmC<sub>mbp</sub>) was obtained and its sequence verified, it was transformed in DS9029. 5 ml of an overnight culture was used to inoculate 200 ml LB Ap medium and cells were incubated at 37°C to reach an optical density at 600 nm (OD<sub>600</sub>) of 0.6, then induced with 0.2 mM IPTG for 2 h. Once harvested, pellets were resuspended in column buffer (20 mM Tris-HCl pH 7.4, 200 mM NaCl, 1 mM EDTA), freeze-thawed and sonicated. Supernatants were passed through an amylose column prepared according to [17]. Elutions were done using the same protocol, except that one more step of washing in 20 mM Tris-HCl pH 7.4, 1 M NaCl, 1 mM EDTA was performed before washing with column buffer and eluting with 10 mM maltose in column buffer. Most of the protein eluted in the first fraction and was more than 80% pure, as visualized on a 12.5% sodium dodecyl sulfate-polyacrylamide gel stained with Coomassie blue [18].

His-XerDPm (*P. mirabilis* XerD fused to an amino-terminal His tag) was induced from pMVD72 in HMS174(DE3) as described in [15]. MBP-XerCEc (*E. coli* XerC fused to the maltose binding protein (MBP)) was induced from pGB500 in DS9029 and purified as for MBP-XerCPm. XerDEc (*E. coli* XerD) was induced from pRM130 in BL21 pREP4 with 0.2 or 0.5 mM IPTG for 2 h at 37°C. It was then purified as described by [19] on a His-trap column from Amersham Pharmacia Biotech, omitting the MonoS column step. *E. coli* ArgR was expressed and purified from pDB169 in DS955 *recA*7, essentially as described in [20]. *E. coli* PepA was expressed and purified from pCS126 in DS955 *recA*7 according to [21], omitting the final dialysis step. Pellets were resuspended in high salt buffer (1 M NaCl, 10 mM MgCl<sub>2</sub>, 10 mM Tris-HCl pH 8.0), and stored in high salt buffer containing 30% glycerol. Protein concentrations were determined by a Bradford assay.

### 2.5. DNA binding assays

Gel shift assays were done as described by Blakely and Sherratt [22] using TENg buffer (20 mM Tris-HCl pH 7.5, 1 mM EDTA, 50 mM NaCl and 5% glycerol) with 1  $\mu\text{g}$

polydIdC-polydIdC (Amersham Pharmacia Biotech) and *dit*-dig or *cer*-dig. The *cer* site was dig-labelled by a PCR reaction with pKS492 using 100 pmol of the M13F-40 and M13R-48 primers, and a 102-bp fragment containing the *E. coli dit* site was dig-labelled from a PCR reaction using 100 pmol of the following primers: *diff*22 (CA-GAAAAGCACTTCGCATCAC) and *dit*R4 (CAATCAT-GACCGCCAACGAC). Both substrates were gel purified.

Reactions were performed for 10 min at 37°C and electrophoresed at 4°C on a 6% polyacrylamide gel in 1×TBE buffer at 8 V cm<sup>-1</sup>. Gels were then transferred onto Gene Screen Plus (NEN<sup>®</sup> Life Science Products) positively charged nylon membranes and UV cross-linked. The transferred DNAs were detected using standard dig detection methods (Roche Diagnostica), with CDP-Star (NEB), followed by exposure to Super RN Medical X-ray Film (FujiFilm).

### 2.6. *In vitro* recombination reactions

Reactions were performed as described by Colloms et al. [23] in 45 µl of reaction buffer for 1 h at 37°C. Xer proteins and *E. coli* ArgR and PepA were used at the same concentrations as in [23]. DNA was further digested with *EvoRV* for 1 h at 37°C before analysis on a 1% agarose gel in TAE buffer. Gels were then transferred onto Gene Screen Plus membranes, which were hybridized with a dig-labelled *cer* site and detected in the same way as in binding reactions.

### 2.7. Complementation and plasmid stability tests

Monomers of pCS210 were purified and transformed in

DS9029 with the TransformAid kit (Fermentas). Plasmids pEcCD, pPmCEcD, pPmDEcC, pPmCD and pUC19 (see Table 1) were then transformed in this strain. Plasmid DNA was isolated by a standard alkaline lysis method and about 400 ng of each extraction was electrophoresed on a 1% agarose gel at 2.36 V cm<sup>-1</sup> for 12 h. The gel was transferred onto Hybond<sup>™</sup>-N+ (Amersham Pharmacia Biotech) positively charged nylon membranes and hybridized under stringent conditions using the dig-labelled Tc<sup>R</sup> gene of pCS210. Final detection was done as for binding reactions.

To verify the stability of pCS210 in the different strains according to the *xer* genes involved in *cer* resolution, the strains were grown in LB medium containing only ampicillin to maintain the *xer* clones for a period of 5 days (about 100 generations). Plasmid DNA was extracted each day, and dilutions of the overnight culture were plated onto LB-Tc-Ap agar and LB-Ap agar media for the determination of colony forming units (CFU). In addition, 50 µl of 10<sup>-4</sup> dilution was cultured in LB-Ap and grown overnight at 37°C in order to obtain 20 generations of growth.

## 3. Results

### 3.1. Sequence analysis

It has been shown by plasmid recombination assays that *P. mirabilis* should have genes encoding homologues of recombinases XerC and XerD [10]. Indeed, our previous study demonstrated the presence of the *xerD* gene in this species [15]. The search for a *xerC* homologue in *P. mira-*

Table 1  
Strains and plasmids used in this study

	Description	References
<i>Strains</i>		
DS941	AB1157 <i>recF lacZ lacZΔM15</i>	[4]
DS9029	DS941 <i>xerD</i> <sup>+</sup> <i>Tp</i> <sup>R</sup> <i>xerC</i> <sup>+</sup> <i>mmiMu</i> PR13	[23]
DS955 <i>recA</i>	DS941 <i>argR</i> <sup>+</sup> <i>Tp</i> <sup>R</sup> <i>pepA</i> <sup>+</sup> <i>Tn5Δ<sub>int</sub>-recA-309</i> <i>Tn10</i>	G. Szatmari
BL21 pREP4	F <i>amp</i> <sup>r</sup> <i>hsdS</i> <sub>0</sub> ( <i>r</i> <sub>0</sub> <i>m</i> <sub>0</sub> ) <i>gal dem</i> with plasmid pREP4 containing <i>lacF</i>	[27]
HMS174(DE3)	F <i>recA hsdR</i> <sub>0</sub> ( <i>r</i> <sub>K12</sub> <i>m</i> <sub>K12</sub> ) <i>Ril</i> <sup>R</sup> <i>ΔDE3</i>	[28]
DI15α	F <i>Φ80ΔlacZΔM15 Δ<sub>hucZ</sub>Δ-argP</i> U169 <i>endA1 recA1 hsdR17</i> ( <i>r</i> <sub>K12</sub> <i>m</i> <sub>K12</sub> ) <i>deoR thi-I smpL44</i> <i>gyrA96 relA1</i>	[29]
<i>Plasmids</i>		
pCS210	pACYC184 derivative with a functional <i>lacZ</i> between two <i>cer</i> sites	[7]
pCS211	Resolved pCS210 with one <i>cer</i> site	
pMVD72	<i>P. mirabilis xerD</i> in pET-14b (Novagen)	[15]
pPmC <sub>mbp</sub>	<i>P. mirabilis xerC</i> in pMal-e2 (NEB)	This study
pCS126	Plasmid in which <i>pepA</i> is overexpressed from a <i>lac</i> promoter	[24]
pDB169	<i>argR</i> overexpressing plasmid	[20]
pRM130	<i>E. coli xerD</i> overexpressing Ap <sup>R</sup> pMB1	[1]
pSDC105	<i>E. coli xerC</i> overexpressing plasmid	[4]
pGB500	<i>E. coli xerC</i> in pMal-e2	[30]
pKS492	ColEI <i>cer</i> site in pUC18	[31]
pEcCD	<i>E. coli xerC</i> from pSDC105 in pRM130	This study
pPmCEcD	<i>P. mirabilis xerC</i> in pRM130	This study
pPmDF <sub>2</sub> C	<i>P. mirabilis xerD</i> in a pUC19 containing <i>E. coli xerC</i> from pSDC105	This study
pPmCD	<i>P. mirabilis xerC</i> and <i>xerD</i> in pUC19	This study

*billy* was made in the same way as for *xerD*, using degenerate primers to amplify the C-terminal region, followed by inverse PCR to obtain the remaining sequence. The deduced amino acid sequence of the gene found in this way possesses all of the conserved residues found in the tyrosine recombinase family. Moreover, it reveals the ESS motif indicative of a XerC protein [24], and displays more similarities to XerCEc (63.1% identity; see Fig. 1) than to XerDEc (44.9% identity). The gene was then named *xerCPm* (GenBank number AF033495). The highest similarity between the two XerCs is found in the C-terminal domains (71.9% identity), where the catalytic tyrosine is found.

### 3.2. *In vitro* DNA binding and recombination activity

Considering the similarity between the *P. mirabilis* recombinases and those of *E. coli*, plus the fact that *cer* recombination was detected in *P. mirabilis* [10], it was assumed that *P. mirabilis* proteins should recognize and bind to specific sites normally recombined by the *E. coli* Xer system. To verify this, gel retardation tests were done with *cer* and *dif* using XerCPm fused to MBP. MBP alone was not able to bind to the specific sites (data not shown).

MBP-XerCPm, like *E. coli* XerC, cannot bind to the *cer* site alone (Fig. 2A, lanes 6 and 7). This is not the case for *E. coli* XerD [1] and see Fig. 2A, lane 9). It should be noted that even though we cannot see any retarded fragment with the concentrations used here for His-XerDPm (Fig. 2A, lane 8), we have already shown that this protein can also bind to the *cer* site [15]. Nonetheless, higher retarded fragments were observed with MBP-XerCPm when

XerCPm	1	NSGICGIVPPELGLAIDIFELSTIEVPELGGVTVYENQPELHCAIDQMAI	50
XerCEc	1	.....HTLLHIVDPVPLRYLQVYPLDGLDPIQLDQVPELSEAIHFAGSP	41
XerCPm	51	FIDPMLLEKQHYPIKLAFLKPEKQDFADGALPTFALPFLQMVQDQD	100
XerCEc	44	GLQWQDQEVYDGMFAVPLKRIQLGAAAGALRQALPSPFDWLVQDEL	91
		Motif I	
XerCPm	101	AVMFAKGVYTFKSGPRLKFNQYVREYVQELMIDGKELQSYVDFPIELQHY	150
XerCEc	96	FARDKGVSAFKAQDFKPEFLVYDQMTLQELIDKELFLAVLQDAGEVQK	143
		Motif III	
XerCPm	151	GAHGLSELEDFKRIIDIRLQSGEVVDFQGFPEFVYDGRFAVEMQIMF	200
XerCEc	144	GAHGLSELPVQDRIYQGRDSEEVVQVQKQKPEPRLDIPRHAVMIDKQ	191
XerCPm	201	AMPELVYRPEETAVYIDTFKQPLVYVPEVDFYKSNQVACDLSSEVYRFL	250
XerCEc	194	QLRDLFGSDCALPESLQKQPIKAPINQVQFFAENQYDGLINRVEPHEL	242
		Motif II	
XerCPm	251	PHSFATHELESQDPAVCEELGHARLQDTTNYCHDFDIDAFVYFAARF	300
XerCEc	243	PHSFATHELESQDPAVCEELGHARLQDTTNYCHDFDIDAFVYFAARF	292
XerCPm	301	FAKPPES	307
XerCEc	291	FAKPKL	298

Fig. 1. Amino acid sequence comparison between *E. coli* XerC (XerCEc) and *P. mirabilis* XerC (XerCPm). Gap alignments were done with GCG<sup>®</sup> Wisconsin Package<sup>®</sup>. *P. mirabilis* XerC shares 67% identity with *E. coli* XerC. Motifs I, II and III are as described by Cao et al. [23]. The catalytic residue (Y) is indicated by an open arrow, and the conserved residues R-H-R and K, by straight arrows. Vertical lines (|) correspond to identity, double dots (:) to a high degree of similarity and single dots (.) to a lower degree of similarity between amino acids.

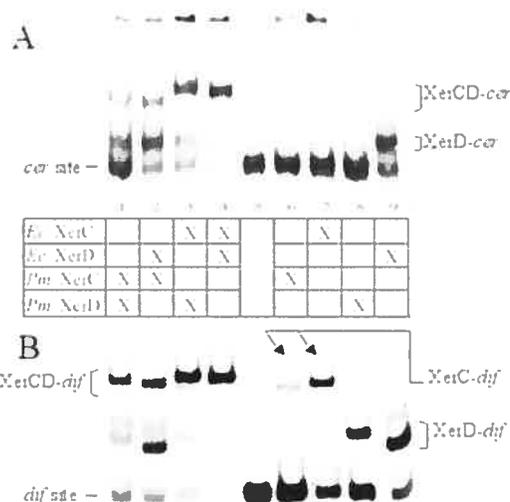


Fig. 2. Binding reactions at *dif* and *cer*. In each lane, different recombinases were put in the presence of dig-labelled fragment for 10 min/37°C, except in lane 5 which represents migration of the specific site (*cer* in A or *dif* in B) without any protein. Lanes 1–4 are with a mix of XerC and XerD from either *P. mirabilis* or *E. coli*, and lanes 5–9 are with XerC or XerD alone.

in collaboration with both XerD proteins (Fig. 2A, lanes 1 and 2). However, it should be noted that the binding of MBP-XerCPm did not appear to be as efficient as that of *E. coli* MBP-XerC used in collaboration with the two XerD proteins (Fig. 2A, compare lanes 1 and 2 with lanes 3 and 4). Moreover, a strong retarded fragment indicative of a XerD binding alone to the site was observed with the mix containing MBP-XerCPm.

Similar results were observed with the *dif* site, except that MBP-XerCPm binds to it alone, like MBP-XerCEc, although less efficiently (Fig. 2B, lanes 6 and 7). When using mixtures of XerC-XerD recombinases, the complexes formed with MBP-XerCEc almost completely retarded the *dif* fragment (no matter which XerD was present); few bands appear indicative of His-XerDPm or XerDEc binding alone to the site. However, when a XerDEc-MBP-XerCPm mixture was used, a strong band indicative of XerDEc binding appears (Fig. 2B, lane 2). This band was more evident than what was observed with His-XerDPm binding (Fig. 2B, lane 1).

In order to further characterize the biological activity of the *P. mirabilis* recombinases, *in vitro* recombination reactions were performed.

*E. coli* Xer recombinases form Holliday junctions on a plasmid containing two *cer* sites in the presence of accessory proteins ArgR and PepA [23]. After digestion of the plasmid DNA having undergone the action of recombinases, this junction can be distinguished from the linear form of the plasmid in an agarose gel; because of its particular structure, it migrates more slowly than the lin-



The test was done with a clone containing *verCPm* and *verDEc* on the same plasmid (named pPmCEcD) in an *E. coli* mutant strain for both *ver* genes (DS9029) having a resident pCS210. Here, only native pCS210 was detected (see Fig. 4A, lane 3), as if there were no recombinases in the strain (Fig. 4A, lane 2 with pUC19). This indicates that XerCPm could not act with XerDEc to resolve pCS210. However, when both *P. mirabilis verC* and *verD* genes are put together in an *E. coli verC<sup>-</sup>verD<sup>-</sup>* double mutant, pCS210 resolution occurs as efficiently as with a clone containing *verCEc* and *verDEc* (Fig. 4A, lanes 1 and 5).

Finally, a similar result to what has been already published ([15] and see above) was observed with a clone containing *verDPm* and *verCEc* (pPmDEcC): a partial resolution of pCS210 was observed (Fig. 4A, lane 4).

### 3.4. Plasmid stability studies

In order to determine if *P. mirabilis* recombinases are able to fulfill their assumed role in plasmid maintenance, the stability of plasmid pCS210 in the absence of selective pressure was tested. Strains containing the different cloned *ver* genes described earlier and pCS210 were grown in media non-selective for pCS210 (without tetracycline), for about 100 generations. The only way for pCS210 to remain stable is by the Xer recombination system acting on *cer*. Over a 5-day period, the ratio between the number of colonies found on a selective medium for pCS210 and the total number of CFU found on the non-selective medium was deduced and plotted on a graph (see Fig. 4B).

In Fig. 4B, a strain in which both *ver* genes are missing was not able to maintain pCS210 beyond 60 generations (control with pUC19). This was confirmed by Southern blot analysis (not shown). On the other hand, *verC<sup>-</sup>verD<sup>-</sup>* strains were able to maintain the resolved form of pCS210 for over 100 generations in the non-selective medium when a vector containing both *ver* genes from the same species was added to these strains (see curves with pFcCD and pPmCD in Fig. 4B). In a similar fashion, clone pPmDEcC (containing *verDPm* and *verCEc*) was able to maintain pCS210 for about 100 generations. However, clone pPmCEcD (containing *verCPm* and *verDEc*) was not able to maintain pCS210 at all, as this plasmid was almost entirely lost after 2 days of growth. This was also confirmed by Southern blot analysis (not shown).

## 4. Discussion

*P. mirabilis* possesses *ver* genes, which encode proteins that can bind to *E. coli* specific sites and show catalytic activity. Despite the high degree of similarity of these proteins with the *E. coli ver* genes, XerCPm cannot act with XerDFc to complement an *E. coli verC<sup>-</sup>verD<sup>-</sup>* strain. In fact, under these conditions, plasmids containing a *cer* site

are lost at a slightly higher rate than the non-*cer* containing control. This difference, although slight, was reproducible.

The observations of Goebel and Helinski [14] on ColE1 multimerization in *P. mirabilis* were not seen in our complementation studies with *E. coli ver* mutants, suggesting that this phenomenon may be due to some other *Proteus* protein(s). Further studies currently under way in our laboratory using *P. mirabilis ver* mutants will help us determine whether this phenomenon is related to *ver* or not.

The inability of a cloned *verC* from *P. mirabilis* to complement an *E. coli verC* mutant for *cer-cer* recombination is surprising, considering the fact that XerCPm binds to the *cer* site in collaboration with XerDFc and that a Holliday junction was detected in *in vitro* reactions using these proteins (Fig. 2A, lane 2 and Fig. 3, lane 2). Moreover, *verC<sup>-</sup>* complementation has been observed with the more distantly related *ssv* gene of *P. aeruginosa* [11]. Functional complementation of *ver* mutants has also been previously described in *Salmonella typhimurium* (for the *cer* site, [26]) and in *H. influenzae* (for the *dif* site, [12]).

This non-complementation problem could simply be due to the fact that the *P. mirabilis* XerC protein is larger (311 aa instead of 298 for XerCFc) and that there is a maximum of divergence in the N-terminal sequence of the protein (47.2% identity between both XerCs). These differences could also affect the cleavage at the *cer* site and so limit the formation of the Holliday junction, or reduce the recognition of the site needed for effective binding. Furthermore, we cannot ignore the fact that the *cer* site is specific for *E. coli* recombinases, and not necessarily for those of *P. mirabilis*.

However, because *in vitro* tests showed catalytic activity of XerCPm in collaboration with XerDFc and with accessory proteins from *E. coli*, it is possible that the problem stems from the Holliday junction formed by XerCPm. The conformation of this junction formed may not be a good substrate for the protein which resolves this junction. In fact, for *E. coli* recombination at *cer*, XerD does not resolve the Holliday junction formed by XerC, even though it is required for XerC's catalytic activity [23].

Studies on the further resolution of the Holliday junctions formed by XerCPm/XerDFc, as well as their formation in *P. mirabilis ver* mutants, should help to shed some light on the unique characteristics of this structure.

### Acknowledgements

We would like to thank Nancy Saint-Cyr, François Alier, Tien Lien Mi and Valérie Hal-Gagné for their help during this work, Patty Singer of the University of Maine for her assistance in automated sequence analysis, and Drs. Marc Drolet and Josée Harel for their comments. This work was supported by an operating grant from NSERC (Natural Sciences and Engineering Research

Council of Canada). M.V. was supported by a NSERC fellowship at the beginning of the study, then by a FCAR (Fonds pour la Formation de Chercheurs et l'Aide à la Recherche du Québec) fellowship.

## References

- [1] Blakely, G., May, G., McCulloch, R., Arciszewska, L.K., Burke, M., Lovett, S.T. and Sherratt, D.J. (1993) Two related recombinases are required for site-specific recombination at *dir* and *cro* in *E. coli* K12. *Cell* 75, 351–361.
- [2] Cornet, F., Mortier, L., 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* site *dif*. *J. Bacteriol.* 176, 3188–3195.
- [3] 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. *Phil. Trans. R. Soc. Lond. B Biol. Sci.* 347, 37–42.
- [4] Colloms, S.D., Sykora, P., Szatmari, G. and Sherratt, D.J. (1990) Recombination at ColE1 *rec* requires the *Escherichia coli* *recC* gene product, a member of the lambda integrase family of site-specific recombinases. *J. Bacteriol.* 172, 6973–6980.
- [5] Recchia, G.D. and Sherratt, D.J. (1999) Conservation of *rec* site-specific recombination genes in bacteria. *Mol. Microbiol.* 34, 1146–1148.
- [6] Slack, F.J., Serron, P., Joyce, F. and Sonnenheim, A.L. (1995) A gene required for nutritional repression of the *Bacillus subtilis* dipeptide permease operon. *Mol. Microbiol.* 15, 689–702.
- [7] Sirochetu, S.A., Piggot, P.J., Sherratt, D.J. and Blakely, G. (1999) The *copA* locus of *Bacillus subtilis* encodes a site-specific recombinase involved in proper chromosome partitioning. *J. Bacteriol.* 181, 6053–6062.
- [8] Becker, J. and Brendel, M. (1998) Molecular characterization of the *recC* gene of *Leu. tobacophilus leichnamni* encoding a site-specific recombinase and two adjacent heat shock genes. *Curr. Microbiol.* 32, 232–236.
- [9] Chalkei, A.F., Lupas, A., Ingraham, K., So, C.Y., Lunsford, R.D., Li, T., Bryant, A., Holmes, D.J., Marra, A., Pearson, S.C., Ray, J., Burnham, M.K.R., Palmer, L.M., Biswas, S. and Zalcman, M. (2000) Genetic characterization of Gram-positive homologs of the XerC/D site-specific recombinases. *J. Mol. Microbiol. Biotechnol.* 2, 225–233.
- [10] Sirois, S. and Szatmari, G. (1995) Detection of XerC and XerD recombinases in Gram-negative bacteria of the family *Enterobacteriaceae*. *J. Bacteriol.* 177, 4183–4186.
- [11] Höfte, M., Dong, Q., Koumbas, S., Krishnapilla, V., Sherratt, D. and Mergoy, M. (1994) The *oxy* product, which affects pyoverdinin production in *Pseudomonas aeruginosa* 78SK2, is a site-specific recombinase. *Mol. Microbiol.* 14, 1011–1020.
- [12] Neilson, L., Blakely, G. and Sherratt, D.J. (1999) Site-specific recombination at *dif* by *Haemophilus influenzae* XerC. *Mol. Microbiol.* 31, 915–926.
- [13] Haber, K.E. and Walker, M.K. (2002) Filamentous phage integration requires the host recombinases XerC and XerD. *Nature* 417, 656–659.
- [14] Couebel, W. and Helinski, D.R. (1968) Generation of higher multiple circular DNA forms in bacteria. *Proc. Natl. Acad. Sci. USA* 61, 1406–1413.
- [15] Villion, M. and Szatmari, G. (1998) Cloning and characterization of the *Bacillus subtilis* *recD* gene. *FEMS Microbiol. Lett.* 164, 83–90.
- [16] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- [17] Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Stahl, K. (1987–1997) *Current Protocols in Molecular Biology*. Wiley, New York.
- [18] Dreyfuss, G., Adam, S.A. and Choi, Y.D. (1984) Physical change in cytoplasmic messenger ribonucleoproteins in cells treated with inhibitors of mRNA transcription. *Mol. Cell. Biol.* 4, 415–423.
- [19] Arciszewska, L.K., Grange, I. and Sherratt, D.J. (1997) Action of site-specific recombinases XerC and XerD on tethered Holliday junctions. *EMBO J.* 16, 3741–3743.
- [20] Lim, D., Oppenheim, J.D., Eckhardt, T. and Meas, 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.
- [21] Stirling, C.J., Colloms, S.D., Collins, J.F., Szatmari, G. and Sherratt, D.J. (1989) *recB*, 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.
- [22] Blakely, G. and Sherratt, D. (1986) Determinants of selectivity in Xer site-specific recombination. *Genes Dev.* 10, 762–773.
- [23] Colloms, S.D., McCulloch, R., Grant, K., Neilson, L. and Sherratt, D.J. (1996) Xer-mediated site-specific recombination in vitro. *EMBO J.* 15, 1172–1181.
- [24] 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.
- [25] Cornet, F., Hallet, B. and Sherratt, D.J. (1997) Xer recombination in *E. coli* site-specific DNA topoisomerase activity of the XerC and XerD recombinases. *J. Biol. Chem.* 272, 21927–21931.
- [26] Hayes, F., Lubetka, S.A. and Sherratt, D.J. (1997) *Salmonella typhimurium* specifies a circular chromosome dimer resolution system which is homologous to the Xer site-specific recombination system of *Escherichia coli*. *Gene* 198, 105–110.
- [27] Studier, F.W. and Moffatt, B.A. (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* 189, 113–130.
- [28] Campbell, J.L., Richardson, C.C. and Studier, F.W. (1978) Genetic recombination and complementation between bacteriophage T7 and cloned fragments of T7 DNA. *Proc. Natl. Acad. Sci. USA* 75, 2276–2280.
- [29] Grant, S.G.N., Jesse, J., Bloom, F.R. and Hanahan, D. (1990) Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. *Proc. Natl. Acad. Sci. USA* 87, 4645–4649.
- [30] Arciszewska, L.K. and Sherratt, D.J. (1995) Xer site-specific recombination in vitro. *EMBO J.* 14, 2112–2120.
- [31] Stirling, C.J., Szatmari, G., Stewart, G., Smith, M.C.M. and Sherratt, D.J. (1988) The arginine repressor is essential for plasmid stabilizing site-specific recombination at the ColE1 *rec* locus. *EMBO J.* 7, 4389–4395.
- [32] Cao, Y., Hallet, B., Sherratt, D.J. and Hayes, F. (1997) Structure-function correlations in the XerD site-specific recombinase revealed by pentapeptide scanning mutagenesis. *J. Mol. Biol.* 274, 39–53.

## Annexe 3

```

Pm#15. ---GNT-----TCGCCAAGCTTGCATGCCTGCAGGTGCGACTCTAG--
      : : : : : : : : : : : : : : : : : : : : : :
H14320 TCAGGTATTTTTTTGTAGCCCATCGGTACTACCACAATAACCCGTCTGGTT
      10 20 30 40 50
      40 50 60
Pm#15. ---AGGATC-----CCCATCGT-----TTTTC-----AAGA--
      : : : : : : : : : : : : : : : : : : : : : :
H14320 GGTAAGGCTCAGAAAACCAAGCCTAGCATAATAATTAGCGCTCGAAGGGG
      60 70 80 90 100
      70 80 90
Pm#15. ---TGATACTGTCACTGCA--TTTCG--TGATATT--TCACCCCAAG----
      : : : : : : : : : : : : : : : : : : : : : :
H14320 CAGTGAAACTTTTAGGAAAGGTTTCAGCTAATTTTGTCACTGGAACAAC
      110 120 130 140 150
      100 110 120
Pm#15. -CGCCTACTCA-----TATTTTAATTATCCCAATA-
      : : : : : : : : : : : : : : : : : : : : : :
H14320 TCGCCTACATAGCCAGTTTCCCTGCCATTATTTTAATT--TCACCATCAG
      160 170 180 190
      130 140 150 160 170
Pm#15. -AACTTATCCAACGGTAAACG-ATGTTACTGCACAAGATGAACAAGTTT
      : : : : : : : : : : : : : : : : : : : : : :
H14320 TAGCATAGCTTAATGTTGTACCTATTGTCATGAGTAATATAAAAAATGGT
      200 210 220 230 240
      180 190 200 210
Pm#15. TAGG---TCATCT---ATTTGTTG--TGGCAG---CTAAAA---TTGC
      : : : : : : : : : : : : : : : : : : : : : :
H14320 GAGTATTTTCAATTTTCATAAATGTTCCCTAGGAATTTACTAAAAATTATTT
      250 260 270 280 290
      220 230
Pm#15. TCAGCAAGA---AGGTATT-----GCAGA---AGAGGG
      : : : : : : : : : : : : : : : : : : : : : :
H14320 TTATTAAGAGACAAGGTTTTTATTAAATTAGCCTAGCATATAAAAAAGGA
      300 310 320 330 340
      240 250 260
Pm#15. ATATC--TGAAAAGCTG-----CGGTAAATGATTAC
      : : : : : : : : : : : : : : : : : : : : : :
H14320 ATATTCATAAAAAATTTTATATTCCAATATGATTATCTTTATGCGACTAT
      350 360 370 380 390
      270 280 290
Pm#15. CTCTACA--AAAAATTATTTGTGGCG-----GGGGATTTA-----
      : : : : : : : : : : : : : : : : : : : : : :
H14320 CTCTATAGGAATAATTTACTTTAACTATAATGGAGCTTTAATTTTATCTA
      400 410 420 430 440
      300 310 320 330
Pm#15. ---TGCAGTATTGAA-----GTGATAA-ACGATCAC--AAATATAT-----
      : : : : : : : : : : : : : : : : : : : : : :
H14320 TTCTGATGTATTAATATAGAAATAAGACGAACAAGGAAATATACAAAA
      450 460 470 480 490
      340 350 360
Pm#15. TGGCGATTTTACGGTTAGAT---TTATAG-TTA-----CAGCAA--
      : : : : : : : : : : : : : : : : : : : : : :
H14320 TCACTCTTATATA-TGAGATGTGTTATATATTATGAACATGTCATCATT
      500 510 520 530 540

```

```

      370          380          390          400
Pm#15. ---AGAGCAA-----TAGCAAT--AGCAATAGCA---ATATCAA---
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
H14320 CCTATAACAAACCTCGGCTAGCAATGTAGCATCAGCATATAAAATTAAGA
      550          560          570          580          590

      410          420
Pm#15. ----CGGTAG-----AAGTT-TGA-----AATAATAAA
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
H14320 AAGTCGGTTGAAATACTAAAGTAGTGACCAAGTCCAATTGAATAATCAA
      600          610          620          630          640

      430          440
Pm#15. T-----AAAAATA-----TTAAGGATTAT---
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
H14320 TCAGCCAAAAACAGAGATTTAGCGCCTATCCGTATTTACCTATCATGGT
      650          660          670          680          690

      450
Pm#15. ----TATTG-----GAGAGATGTC---
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
H14320 TCGCATAATGTGTGACGCTATGTTACAAAAGGGTCTGAGCGATTGCGGT
      700          710          720          730          740

      460          470
Pm#15. -----TTT-----GATTC---TTATG---A
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
H14320 AGAATGGGGTTGGGGTTTACCGCAATCGAGAAGGTTCGTCTTTTACATA
      750          760          770          780          790

      480          490          500
Pm#15. TAGCG--TCACGGGCATCAAT-----TTAACTGACGGTATAAAC
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
H14320 TAGTGGATTATGTGAATCAAGAACCTTCTAACTTAACTGACGGTAGAAAC
      800          810          820          830          840
      510          520          530          540          550
Pm#15. ATCTAAAAAGGCTTTACTAAAATCCTTTCTAAAGTAGTATCTGATAAAAA
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
H14320 ATCTAAAAAGGCTTTATTAATAATCCTTTCTAAAGTAGTATCTGATAAAAA
      850          860          870          880          890
      560          570          580          590          600
Pm#15. GACAAAAAACATTGGGTCCAACCAACATAATCAATCAGCCAAAAACAGA
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
H14320 GACAAAAAA-CATTGGGTCCAACCAACATAATCAATCAGCCAAAAACAGA
      900          910          920          930          940
      610          620          630          640          650
Pm#15. GATTTAGCGCCTATCCGTATTTACCTATTATGGTTCGCATAATGTATAT
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
H14320 GATTTAGCGACTATCCGTATTTACCTATCGTGGTTCGCATAATGTATAT
      950          960          970          980          990
      660          670          680          690          700
Pm#15. TATGTTAAATAGCATGTAATAACCAATAACTTTATGGATGCTTCTTTTA
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
H14320 TATGTTAAATGGCATGTAATAACCAATAACTTTATGGATGCTTCTTTTA
      1000         1010         1020         1030         1040
      710          720          730          740          750
Pm#15. ATGATTTAGCGCTTCAA-GCCCTATTGACTCAGATACGTGTTA-GAATTT
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
H14320 ATGATTTAGCGCTTCAAAGCCCTATTGACTCAGATATGTGTTAAGAATTT
      1050         1060         1070         1080         1090
      760          770          780          790          800
Pm#15. AATA-CTCCCCTTCTAGCCAATCTGAAAATTTA-ATGGGATACTCGATAC
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
H14320 AATAACTCCACTTCTAACCAATCTGAGAATTTTTATGGGATACTCGATAT
      1100         1110         1120         1130         1140

```

```

      810      820      830      840
Pm#15. AA-ATTTAC---GTATAGATAGAAAATATAGACT-GACTTTATTATCANCA
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
H14320 GACATTTACTATGTAGAGATAAAAAATATATAGTAGAACTAAGTCACAAGA
      1150      1160      1170      1180      1190

      850      860      870
Pm#15. A-----CTTTGGAAA---AATATTCT-----GAAG
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
H14320 ATATTACACCATTCAAAATCGTAATAGTTTTTTTATTAGCCTTATAGATG
      1200      1210      1220      1230      1240

      880      890      900
Pm#15. -----TAAACGA-----GGAATCNTTCAGGGGCGTA---TAAT---
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
H14320 CATTGTTATATGATCCTGCTGGAAGCTATACCGGATGTAAAAAATAATCAA
      1250      1260      1270      1280      1290

      910
Pm#15. -----AGA-----ACNCC-----CAATN
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
H14320 TGTGAAGATGGTTCCTTACACATTAAGGGGGTCAGGTGATTTCTTTCAATA
      1300      1310      1320      1330      1340

      920      930
Pm#15. -CNTAAAATCAAAT-----TTAC-----AGGGGCT----
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
H14320 TCCTGATTTCAATTGGGATGATTACCTTCTTACCAGTTATGGGATGGTG
      1350      1360      1370      1380      1390

      940      950
Pm#15. -----TTTTATN-----AGCGGAANG
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
H14320 ATGATATCGTTATTATAGAGTTTGTAATCCGAGAGAACAAGCTGAACAC
      1400      1410      1420      1430      1440

      960      970      980      990
Pm#15. ----CCANNCTTTTTTNAANNTTCCNCGCCNAGGCCTNCCTTN---
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
H14320 ATTACCAAACCTTATTATGAACTTGGTGGTGCTAATGCTTTATTTTGTGC
      1450      1460      1470      1480      1490

      1000      1010
Pm#15. -----AANNA-----TNANNTNAGC-----
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
H14320 TACTAGAATCACTCATGTTTTAAAGTTATCTGGTGGTGTATTTAGTAATC
      1500      1510      1520      1530      1540

Pm#15. -----CC---
      : :
H14320 TTGATACTCCAAT
      1550      1560

```

**1<sup>er</sup> alignement du clone Pm#15 contenant le site *difPm* de la souche Pm240-82UM et d'une portion du chromosome de la souche PmH14320.** Cet alignement a été fait avec le programme Stretcher de EMBOSS (<http://cbrmain.cbr.nrc.ca/EMBOSS/index.html>) qui recherche le meilleur alignement possible entre deux séquences données. La portion du génome de la souche H14320 fut prélevée à partir d'une recherche faite tout d'abord par Blast sur le site de l'institut Sanger ([http://www.sanger.ac.uk/Projects/P\\_mirabilis/](http://www.sanger.ac.uk/Projects/P_mirabilis/)). Il est à noter que les 46 premières bases de Pm#15 correspondent à la séquence du pUC19 dans lequel a été cloné le fragment de *P. mirabilis* et qu'il y a eu des incertitudes dans le séquençage des 215 dernières bases présentées ici. Les 28 pb surlignées en jaune correspondent au site *dif* en tant que tel. Le pourcentage d'identité global est de 49,1% avec 35,3% de « gaps ».

Dans le génome de la souche PmH14320, en aval de la portion indiquée en vert incluant le site *dif*, se situe un gène transcrit en sens inverse de la séquence montrée ; cette séquence encoderait un régulateur de transcription potentiel similaire à ArsR.

Annexe 4 :

Pm#15.	GN-----	10	20			
H14320	CAAACCACCCGTAATTTT	GCAACACTTTT	CAAATACCAAACCTCGCAGA			
	10	20	30	40	50	
Pm#15.	-----	30				
H14320	CTAAGTTGTTCAATAGAAATATTGGCGAGTGATGGCTAGGAATTAATA					
	60	70	80	90	100	
Pm#15.	-----	40				
H14320	AATAATAAGGTGATAAAATGGCAGAAGAGACAATTTT					
	110	120	130	140	150	
Pm#15.	-----	50	60	70		
H14320	TCGTGGTGAATTCCTGCAGATATCGTTTTTCAAGATGATACTGTC					
	160	170	180	190	200	
Pm#15.		80	90	100	110	120
H14320	CATTTTCGTGATATTTACCCCAAGCGCCTACTCATATTTTAATTATCCCC					
	210	220	230	240	250	
Pm#15.		130	140	150	160	170
H14320	AATAAACTTATCCAACGGTAAACGATGTACTGCACAAGATGAACAAGT					
	260	270	280	290	300	
Pm#15.		180	190	200	210	220
H14320	TTAGGTCATCTATTTGTTGTGGCAGCTAAAATTGCTCAGCAAGAAGGTA					
	310	320	330	340	350	
Pm#15.		230	240	250	260	270
H14320	TTGCAGAAGAGGGATATCTGAAAAGCTGCGGTAAATGATTACCTCTACAA					
	360		370	380		
Pm#15.		280	290	300	310	320
H14320	AAATTTATTTGTGGCGGGGATTTATGCAGTATTGAAGTGA----TAAAC					
	390	400	410	420	430	
Pm#15.		330	340	350	360	
H14320	GATCACAATATATTGGCGATT-TTACGGTTAGATT-TATAGTTACAGCA					
	440	450	460	470	480	
Pm#15.		370	380	390	400	410
H14320	GTGAAAACCATTAGGGCCATTATTAAGCAAATAATCTAATGTGATA--A					
	490	500	510	520	530	



	940	950	960	970	980
Pm#15.	GGCTTTT	TATNAGCGGA	ANGGCCANN	CCTTTTT	TNAANNTTCC
	:::	:: :	:	:	:
H14320	GGCCATTC	CAGCCATTA	AATTAGAGAG	ATGTCTGT	CACGGATGATT
	1120	1130	1140	1150	1160
	990	1000	1010		
Pm#15.	NAGGCTNC	CCTTTNA	ANN---	ATNANNT	NAGCCC
	:::	::	:: :	:	:
H14320	CTGGCAA	AATTATT	AGTGGTT	GATAATG	TAAAAA-
	1170	1180	1190	1200	

**2<sup>e</sup> alignement du clone Pm#15 contenant le site *dif*<sub>Pm</sub> de la souche Pm240-82UM et d'une portion du chromosome de la souche PmH14320 montrant une identité en amont de *dif*.** Cet alignement a été fait avec le programme Stretcher de EMBOSS (<http://cbrmain.cbr.nrc.ca/EMBOSS/index.html>) qui recherche le meilleur alignement possible entre deux séquences données. La portion du génome de la souche H14320 fut prélevée à partir d'une recherche faite tout d'abord par Blast sur le site de l'institut Sanger ([http://www.sanger.ac.uk/Projects/P\\_mirabilis/](http://www.sanger.ac.uk/Projects/P_mirabilis/)). Il est à noter que les 46 premières bases de Pm#15 correspondent à la séquence du pUC19 dans lequel a été cloné le fragment de *P. mirabilis* et qu'il y a eu des incertitudes dans le séquençage des 215 dernières bases présentées ici. Les 28 pb surlignées en jaune correspondent au site *dif* en tant que tel. Le pourcentage d'identité global est de 50,2% avec 20,5% de « gaps».

L'amont de la séquence indiquée en rose à partir du génome de PmH14320 a montré une séquence fortement similaire à *ycfH* d'*E. coli*, gène encodant une désoxyribonucléase potentielle.