

Université de Montréal

**Étude des mécanismes moléculaires influençant la variation de phase des
adhésines P, F165₁ et CS31A présentes chez des souches d'*Escherichia coli*
pathogènes.**

par

Richard Graveline

GREMIP et CRIP

Département de pathologie et de microbiologie

Faculté de médecine vétérinaire

Thèse présentée à la Faculté de médecine vétérinaire

en vue de l'obtention du grade de

Philosophiæ Doctor (Ph.D.)

en sciences vétérinaires

option microbiologie

Septembre 2011

© Richard Graveline, 2011

Université de Montréal
Faculté de médecine vétérinaire

Cette thèse intitulée

**Étude des mécanismes moléculaires influençant la variation de phase des
adhésines P, F165₁ et CS31A présentes chez des souches d'*Escherichia coli*
pathogènes.**

présentée par

Richard Graveline

a été évaluée par un jury composé des personnes suivantes :

Président rapporteur : Daniel Dubreuil, Ph.D., GREMIP, CRIP,
Université de Montréal

Directrice de Recherche : Josée Harel, Ph.D., GREMIP, CRIP,
Université de Montréal

Codirectrice : Christine Martin, Dr, HDR, Laboratoire
de Microbiologie, INRA de Theix

Membre du jury : Mario Jacques, Ph.D., GREMIP, CRIP,
Université de Montréal

Examineur externe : Michel Frenette, Ph.D., GREB,
Université de Laval

Représentante du doyen: Marie Archambault, D.M.V., Ph.D.,
Université de Montréal

Résumé

F165₁, les pili Pap et l'antigène CS31A associé aux antigènes de surface K88 sont tout trois des membres de la famille de type P des facteurs d'adhérence jouant un rôle prépondérant lors de l'établissement d'une maladie causée par des souches *Escherichia coli* pathogènes, en particulier des souches d'*E. coli* pathogènes extra-intestinales (ExPEC, *Extra-intestinal pathogenic E. coli*). Leur expression est sous le contrôle d'un mécanisme de régulation transcriptionnel dépendant de l'état de méthylation de l'ADN, résultant dans l'existence de deux populations définies, l'une exprimant l'adhésine (population ON) et l'autre ne l'exprimant pas (population OFF). Malgré de fortes identités de séquences, ces trois systèmes diffèrent l'un de l'autre, principalement par le pourcentage de cellules ON rencontrées. Ainsi, quand CS31A est systématiquement orienté vers un état considéré comme OFF, F165₁ présente une phase ON particulièrement élevée et Pap montre deux états OFF et ON bien distincts, selon le phénotype de départ. La protéine régulatrice sensible à la leucine (Lrp, Leucine-responsive regulatory protein) joue un rôle essentiel dans la réversibilité de ce phénomène épigénétique et il est supposé que les différences de séquences au niveau de la région régulatrice modifient la localisation à ces sites de fixation de Lrp; ce qui résulte, en final, aux différences de phase existant entre CS31A, F165₁ et Pap.

À l'aide de divers techniques parmi lesquelles l'utilisation de gènes rapporteurs, mutagénèses dirigées et d'analyse des interactions ADN-protéines *in vitro*, nous montrons dans ce présent projet que la phase OFF prédominante chez CS31A est principalement due à une faible interaction de Lrp avec la région distale de l'opéron *clp*, et que la présence d'un homologue du régulateur local PapI joue un rôle également clef dans la production de CS31A. Dans le cas de F165₁, nous montrons dans cette étude que le taux élevé de cellules en phase ON est dû à une altération dans le maintien de Lrp sur les sites répresseurs 1-3. Ceci est dû à la présence de deux nucléotides spécifiques, situé de part et d'autre du site répresseur 1, qui défavorisent la fixation de Lrp sur ce site précis. Tout comme dans le cas de CS31A, la formation d'un complexe, activateur ou répresseur de la

phase ON, dépend également de l'action de du régulateur local FooI, qui favorise alors le déplacement de Lrp des sites répresseurs 1-3 vers les sites activateurs 4-6.

Mots clés : adhésine, variation de phase, Lrp, *clp*, *foo*, *pap*

Abstract

F165₁, the pyelonephritis-associated pili (Pap) and the K88-related surface antigen CS31A are three members of the type P family of adhesive factors that play a key role in the establishment of disease caused by Extra-intestinal *Escherichia coli* (ExPEC) strains. They are all under the control of methylation-dependent transcriptional regulation that defines the number of fimbriated (ON) and afimbriated (OFF) cells within a clonal population. Despite a high similarity in DNA sequence, these three adhesive systems nonetheless differ in the ratio of ON cells. While CS31A is always turned toward the OFF state, F165₁ presents a particularly high level of ON cells and Pap shows two distinct OFF and ON states, depending on the starting phenotype. The leucine-responsive regulatory protein (Lrp) plays an essential role in the reversibility of this epigenetic switch and it is believed that the difference in nucleotides within the regulatory region of each operons could modify the binding of Lrp and, in turn, CS31A, Pap and F165₁ phase variation. Using a variety of techniques including gene expression, site-directed mutagenesis, and *in vitro* protein-DNA interaction analysis, we demonstrate that the preferential OFF state observed in CS31A-positive cells is mainly due to a weak interaction of Lrp with the *clp* distal region and that the presence of a PapI homologue within the cell plays a key role in CS31A production. For F165₁, we show in this study that the high level of ON cells found during F165₁ phase variation is due to an altered stability of the DNA complex formed by Lrp at its repressor binding sites 1-3. Again, after each cell cycle, complex formation is modulated by the local regulator FooI (homologue to PapI) which promotes the transit of Lrp toward its activator binding sites 4-6. Furthermore, we identify two nucleotides (T490, G508) surrounding the Lrp-binding site 1 that are critical to maintaining a high OFF to ON switch rate during F165₁ phase variation, as well switching Pap fimbriae toward the OFF state.

Keywords : adhesin, phase variation, Lrp, *clp*, *foo*, *pap*

TABLE DES MATIÈRES

RÉSUMÉ	iii
ABSTRACT	v
LISTE DES TABLEAUX	xii
LISTE DES FIGURES	xiii
LISTE DES SIGLES ET ABRÉVIATIONS	xvi
REMERCIEMENTS	xviii
SECTION I – INTRODUCTION	1
SECTION II – RECENSION DE LA LITTÉRATURE	4
<i>A- ESCHERICHIA COLI</i>	5
1- Les souches ExPEC	6
2- Maladies causées par les souches ExPEC	6
a) Infections du tractus urinaire	7
b) Méningites néonatales	9
c) Septicémies	11
d) Infections chez les animaux	12
-i- Infections chez le porc	15
-ii- Infections chez les bovins	15
-iii- Infections chez les oiseaux	16
3- Phylogénie des souches ExPEC	17
<i>B- RÉGULATION DE L'EXPRESSION DES GÈNES BACTÉRIENS</i>	20
1- Introduction	20
2- Transcription	21
3- Régulation transcriptionnelle	23
4- Mécanismes de répression transcriptionnelle	24
5- Mécanismes d'activation transcriptionnelle	25
6- Contraintes spatiales sur l'architecture de la séquence promotrice	27
7- Mécanismes de coopérativité et contrôle du	28

promoteur	
C- RÉGULATEURS GLOBAUX	32
1- Régulation et organisation de la structure du génome	32
2- Leucine-responsive regulatory protein, Lrp	32
a) Généralités	32
b) Structure de Lrp	33
3- Répression catabolique et le complexe CRP-AMPc	36
4- Histone-Like Nucleoid Structuring protein (H-NS)	36
5- La Dam méthyltransférase	37
D- FACTEURS DE VIRULENCE	40
1- Généralités	40
2- L'adhésion bactérienne et son rôle dans la pathogénèse	42
3- Adhésines bactériennes	42
4- Récepteurs	44
5- Biogénèse des adhésines bactériennes	45
6- Fimbriæ de type 1	48
7- Les fimbriæ de type P	50
a) Généralités	50
b) Structure et organisation génétique de l'opéron <i>pap</i>	50
c) Effet de H-NS et de RimJ sur l'expression de l'opéron <i>pap</i>	54
8- Les fimbriæ F165 ₁	54
a) Généralités	54
b) Structure et organisation génétique de l'opéron <i>foo</i>	55
9- Adhésines afimbriaires	56
a) CS31A	56
b) Structure et organisation génétique de <i>clp</i>	58

10- Association des différents fimbriæ entre eux	58
E- RÉGULATION ÉPIGÉNÉTIQUE	60
1- Introduction	60
2- Phénomènes stochastiques et régulation transcriptionnelle	61
3- Variations antigéniques et variations de phase	64
a) Variation de phase de l'opéron <i>pap</i>	66
b) Influences des régulateurs locaux PapI et PapB	68
-i- PapI	68
-ii- PapB	70
c) Variation de phase de F165 ₁ et CS31A	70
SECTION III – MÉTHODOLOGIE ET RÉSULTATS	73
A- ARTICLE I : LRP-DNA COMPLEX STABILITY DETERMINES THE LEVEL OF ON CELLS IN TYPE P FIMBRIAE PHASE VARIATION	75
1- Abstract	76
2- Introduction	76
3- Results	79
a) Regulatory region of <i>foo</i> operon confers preferential ON state of F165 ₁ -positive cells	79
b) Differential phase variation is associated with altered Lrp stability at <i>foo</i> or <i>pap</i> proximal regions	80
c) Interaction at Lrp binding site 1 differs between <i>foo</i> and <i>pap</i>	81
d) Stability of Lrp binding to proximal region depends on the number of GN(2-3)TTT motifs	82
e) FooI and PapI stabilize Lrp binding to both distal and proximal regions	83

f) Influence of specific sequences on phase variation rate	83
4- Discussion	84
5- Experimental procedures	88
a) Bacterial strains, plasmids and growth conditions	88
b) Phase variation and measurement of switching frequencies	90
c) Surface plasmon resonance	90
d) Footprint analysis	92
6- Acknowledgments	93
7- References	93
8- Legends	99
9- Supporting information	113
a) Supplemental experimental procedures	113
-i- plasmid construction	113
-ii- purification of recombinant proteins	113
-iii- electrophoretic mobility shift assays	114
-iv- multiple alignment and phylogenetic tree	115
b) Supplemental legends	116
c) Supplemental references	132
B- ARTICLE II : Leucine-responsive regulatory protein (Lrp) and PapI homologues control during phase variation of CS31A surface antigen	135
1- Abstract	136
2- Introduction	136
3- Materials and methods	139
a) Bacterial strains, plasmids and growth condition	139

b) Plasmid construction	139
c) Purification of recombinant proteins	140
d) Surface plasmon résonance	141
e) Footprint analysis	143
f) Phase variation and measurement of switching frequencies	143
4- Results	144
a) Differences in stability of Lrp interacting with either <i>clp</i> proximal or distal region	144
b) Lrp binding site occupancies within the <i>clp</i> regulatory region correlates with différence in affinity between the <i>clp</i> distal and proximal regions	145
c) PapI and FooI increase Lrp affinity to both distal and proximal regions	145
5- Discussion	147
6- Acknowledgments	151
7- References	152
8- Legends	156
C- ARTICLE III : Single cell measurements of type P fimbriæ expression in <i>Escherichia coli</i> show a higher level of complexity during phase variation	166
1- Abstract	167
2- Introduction	167
3- Materials and Methods	169
a) Media and growth conditions	169
b) Strains and plasmids	169
c) Measurement of phase variation using the <i>lacZYA</i> reporter system	171
d) Flow cytometry acquisition and analysis	172

4- Results and discussion	172
a) The <i>foo</i> and <i>pap</i> regulatory regions confer different levels of ON colonies	172
b) Real-time measurement of <i>foo</i> and <i>pap</i> phase variation reveals a third sub-population at the single-cell level	174
5- Concluding remarks	177
6- References	178
7- Legends	180
SECTION IV – DISCUSSION - CONCLUSIONS	189
A- DISCUSSION	190
1- Les bases moléculaires de la variation de phase	191
2- L'importance d'être stable	193
3- La différence entre <i>pap</i> et <i>foo</i> ne tient donc qu'à peu de choses	194
4- F165 ₁ et Pap partagent plus de points en commun qu'il ne semble	197
5- Un régulateur pas si local	198
6- Analyse de la variation de phase en temps réel	202
7- Une infection est une chose trop sérieuse pour être laissée à des facteurs de virulence	205
B- CONCLUSION	210
SECTION V – BIBLIOGRAPHIE	211
ANNEXES	
ANNEXE 1 : Influence de la leucine et de l'alanine sur la variation de phase de Pap et de F165 ₁ , ainsi que sur la fixation de Lrp aux sites 1-3 et 4-6 des opérons <i>pap</i> et <i>foo</i>	xix
ANNEXE 2 : Effects of DNA supercoiling and topoisomerases on the expression of genes coding for F165 ₁ , a P-like fimbriæ	xxi

Liste des tableaux

SECTION III

ARTICLE I

Table 1 : Phylogenic and phenotypic characterization of ExPEC strains used in this study 110

Table 2 : Apparent kinetic rate constants for Lrp, FooI, PapI binding to *foo* and *pap* DNA fragments, as determined by SPR 111

Table S1 : Apparent dissociation rate constants for Lrp binding to *foo/pap* DNA fragments in the absence or presence of FooI and PapI, as determined by SPR 125

Table S2 : Phylogenic and phenotypic characterization of strains used in this study 127

Table S3 : Primers used in this study 129

ARTICLE II

Table 1 : Apparent Surface Plasmon Resonance (SPR) rate constants for Lrp, FooI, and PapI binding to *clp* binding sites 160

Table 2 : Effects of PapI homologues on phase variation of *E. coli* *clp-lacZYA* fusion lysogen 164

ARTICLE III

Table 1 : Strains and plasmids used in this study 183

Table 2 : Frequency of phenotypic switch 185

Table 3 : Percentages corresponding to the same populations represented by FACS 187

Liste des figures

SECTION II

Figure 1. Pathogénèse des souches ExPECs chez les animaux	14
Figure 2. Arbre phylogénétique issu de l'analyse de la séquence de 8 gènes de ménage provenant de 234 souches d' <i>E. coli</i> .	19
Figure 3. Interaction de l'ARNpolymérase (ARNpol) avec sa séquence promotrice.	22
Figure 4. Mécanisme moléculaire de régulation transcriptionnelle	26
Figure 5. Localisation des sites de fixation des facteurs transcriptionnels selon leur action (répresseur ou activateur)	27
Figure 6. Logique booléenne appliquée aux mécanismes de régulation transcriptionnelle	31
Figure 7. Représentation tri-dimensionnelle de la structure de Lrp sous forme octamérique (gauche) et monomérique (droite)	35
Figure 8. Représentation schématique de la voie de synthèse des fimbriæ de type « chaperone-usher »	46
Figure 9. Structure des fimbriæ de type 1 avec notamment la structure de l'adhésine FimH liant le mannose	49
Figure 10. Représentation de la structure des gènes sous forme d'un opéron, impliqué dans la synthèse de l'adhésine fimbriaire Pap	51
Figure 11. Régulation de la synthèse des fimbriæ Pap via le système CpxAR	53
Figure 12. Organisation génétique de l'opéron <i>clp</i> codant l'antigène de surface CS31A	58
Figure 13. Représentation schématique du mécanisme moléculaire régissant la variation de phase de l'opéron <i>pap</i>	67
Figure 14. Représentation schématique de la structure tridimensionnelle de PapI	69

SECTION III

ARTICLE I

Figure 1. Schematic representation of <i>foo</i> and <i>pap</i> phase variation	102
Figure 2. F165 ₁ -positive cells present a high percentage of ON cells	103
Figure 3. Association and dissociation kinetics of Lrp binding to wild-type and mutant proximal regions of <i>foo</i> and <i>pap</i>	104
Figure 4. Differences in DNA sequence affect Lrp localization on <i>foo</i> regulatory region	105
Figure 5. FooI and PapI increase both association and stability of Lrp at the <i>foo</i> and <i>pap</i> proximal and distal regions	107
Figure 6. Presence of additional GN(2-3)TTT motifs, combined with the action of FooI, explains the relative expression of <i>foo</i> operon	108
Figure 7. Schematic model illustrating differences in phase variation between <i>foo</i> and <i>pap</i>	109
Figure S1. Comparison of the nucleotide sequences of the regulatory regions of <i>foo</i> and <i>pap</i> operons	119
Figure S2. DNA-binding pattern of Lrp to either <i>foo</i> or <i>pap</i> intergenic region	120
Figure S3. Association and dissociation kinetics of His6-Lrp binding to the <i>foo</i> and <i>pap</i> regions	121
Figure S4. Representative curve fitting on SPR data	122
Figure S5. Association and dissociation kinetics of FooI binding to the distal or proximal regions of <i>foo</i> and <i>pap</i> operons	123
Figure S6. Association and dissociation kinetics of PapI binding to the distal or proximal regions of <i>foo</i> and <i>pap</i> operons	124
Figure S7. Presence of Lrp binding sites within the intergenic region of Pap variants	126
Figure S8. DNA sequence of <i>pap</i> -like intergenic regions among various ExPEC strains correlates with phylogenetic analysis	128
ARTICLE II	
Figure 1. Comparison of the nucleotide sequences of the regulatory regions of the <i>clp</i> and <i>pap</i> operons	158

Figure 2. Association and dissociation kinetics of His ₆ -Lrp binding to the <i>clp</i> proximal regions	159
Figure 3. Lrp localization on <i>clp</i> regulatory region	161
Figure 4. PapI and F _{oo} I increase both affinity and stability of Lrp at the <i>clp</i> proximal and distal regions	162
Figure 5. The presence of a PapI homologue leads to the bistable expression of the <i>clp</i> operon	163
ARTICLE III	
Figure 1. Schematic representation of <i>foo</i> and <i>pap</i> phase variation	182
Figure 2. schematic representation of the reporter system used in this study	184
Figure 3. Monitoring of the phase variation of individualized bacterial cells	186
Figure 4. Hypothetical model for phase variation of fimbrial operons of P fimbriae	188
SECTION IV	
Figure 15. Représentation schématique de l'ensemble des motifs GN(2-3)TTT présents au niveau de la région régulatrice de différents homologues de <i>pap</i>	195
Figure 16. Modèle représentant la configuration des deux phases ON et OFF des opérons <i>foo</i> et <i>pap</i>	196
Figure 17. Modèle de la variation de phase de <i>clp</i>	201
Figure 18. Modèle hypothétique de la variation de phase des opérons fimbriaires codant les adhésines de type P	204
Figure 19. Suivi du devenir de 2 lignées de cellules (une se caractérisant par une fréquence de passage d'un phénotype à l'autre lente, l'autre rapide) mises de façon alternée dans deux conditions de culture distinctes	207
Figure 20. Analyse phylogénétique basée sur la comparaison de séquence de la région intergénique de 22 homologues de <i>pap</i>	209

Liste des sigles et abréviations

A ou Ala	L-alanine
ADN (<i>DNA</i>)	Acide désoxyribonucléique (<i>deoxyribonucleic acid</i>)
Afa	<i>Afimbrial adhesin</i>
Amp	Ampicilline
AMPC (<i>cAMP</i>)	AMP cyclique ou adénosine monophosphate cyclique (<i>cyclic adenosine monophosphate</i>)
Amp ^R	Résistance à l'ampicilline
Antigène O	Antigène oligosaccharidique
ARN (<i>RNA</i>)	Acide ribonucléique (<i>ribonucleic acid</i>)
ARNm	ARN messenger
ATP	Adénosine triphosphate
CAP ou CRP	<i>Catabolite Activator Protein</i> ou <i>cAMP Receptor protein</i> ou protéine activateur catabolique ou protéine récepteur de l'AMPC
<i>clp</i>	<i>Capsule-like</i>
Cam	Chloramphénicol
Cam ^R	Résistance au chloramphénicol
Dam	<i>Déoxyadénosine méthylase, DNA Adénine Méthylase</i> ou Déoxyadénosine méthyltransférase
DO (<i>OD</i>)	Densité optique (<i>Optical density</i>)
<i>foo</i>	<i>f</i> <u>o</u> <i>ne hundred sixty-five</i> <u>o</u> <i>ne</i>
<i>fot</i>	<i>f</i> <u>o</u> <i>ne hundred sixty-five</i> <u>t</u> <i>wo</i>
g	Nombre de génération bactérienne qui ont donné à une colonie
GATC ^{dist}	GATC distal
GATC ^{prox}	GATC proximal
HIS6	<i>His-tag</i> ou queue de poly-histidine
H-NS	<i>Histone-like nucleoid structuring protein</i> ou protéine

	similaire à une histone structurant le nucleoïde
HTH	Hélice-Tour-Hélice
Km	Kanamycine
Km ^R	Résistance à la kanamycine
L ou Leu	L-leucine
Lac	Phénotype lactose
LB	Luria-Bertani
Lrp	<i>Leucine-responsive regulatory protein</i> ou protéine de réponse à la leucine
<i>M</i>	Nombre de cellules ayant effectuées une variation de phase
<i>N</i>	Nombre total de cellules évaluées
OFF	Phase OFF ou phase de non expression de l'opéron
ON	Phase ON ou phase d'expression de l'opéron
<i>pap</i>	<i>Pyelonephritis associated pili</i> ou fimbriæ associé à la pyélonéphrite
PCR	<i>Polymerase chain reaction</i> ou réaction de polymérisation en chaîne
Tet	Tétracycline
Tet ^R	Résistance à la tétracycline
X-gal	5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside
β -gal	β -galactosidase

Remerciements

Je tiens à remercier tous ceux et celles sans qui ce projet de doctorat n'aurait pas été réalisé.

En premier, je remercie infiniment Josée Harel pour m'avoir permis d'intégrer son laboratoire et, surtout, d'avoir été capable de me diriger tout en étant à l'écoute. Cela a été une grande chance pour moi.

Je remercie également fortement Christine Martin qui a été d'une aide précieuse tout au long de ce projet. Sa gentillesse et son sens critique ont grandement contribué à la maturation de ce projet.

Je remercie l'ensemble des chercheurs du GREMIP pour leur disponibilité et leur enthousiasme de tous les jours. En particulier, je tiens à remercier chaleureusement Michaël Mourez pour sa disponibilité. Ce fut un vrai plaisir.

Je tiens également à remercier Daniel Dubreuil et Mario Jacques que j'ai eu la chance de rencontrer durant la préparation des cours de microbiologie ou dans l'organisation des journées de la recherche. Ces expériences ont été extrêmement enrichissantes.

Je remercie les membres du laboratoire de Josée Harel, passés et présents, ainsi que ceux du laboratoire de diagnostique qui ont fait de ce projet un pur plaisir. Je remercie tous les membres du GREMIP et du CRIP. Au plaisir de les revoir!

Je remercie également Diane, Hélène, Manon, Marie-Pier, Nancy et, last but not least, Micheline pour leurs aides multiples.

Enfin, je remercie tous mes amis, les Sherman's, Mike, ainsi que ma famille pour leur soutien. Cette thèse est dédiée à mon père.

SECTION I- Introduction

La capacité à répondre de façon optimale à des environnements changeants est un processus essentiel à tout organisme vivant pour se multiplier. Ce phénomène implique l'action de régulateurs globaux qui modulent à moindre coût énergétique la présence et absence de plusieurs voies métaboliques simultanément. Alors que les bactéries non pathogènes utilisent le milieu environnemental, les souches pathogènes, quant à elles, utilisent l'organisme qu'elles infectent pour subvenir à leurs besoins en nutriments. Ce mécanisme s'accomplit principalement par la synthèse d'une multitude de déterminants de virulence, capables de maintenir la bactérie pathogène à l'intérieur de l'hôte ou encore de lyser les cellules de l'hôte et rendre ainsi accessible une quantité importante de macromolécules entrant dans le métabolisme bactérien. Ainsi, l'attachement des bactéries pathogènes à différentes surfaces dépend de la présence de molécules (adhésines) qui interagissent spécifiquement avec la surface des cellules de l'hôte.

F165₁ et CS31A sont deux adhésines bactériennes apparentées à la famille des fimbriae de type P, et présentes chez des souches d'*E. coli* associées à des cas de septicémies et de diarrhée chez le porcelet et les veaux. Leur expression est sous le contrôle des opérons *foo* et *clp* qui partagent une organisation génétique homologue à celle des autres membres de cette famille, comme *pap*. En particulier, l'expression de ces opérons est régie par un mécanisme de régulation complexe dit variation de phase qui se caractérise par l'apparition d'une population hétérogène de bactéries exprimant (phase ON) ou n'exprimant pas (phase OFF) l'adhésine au sein d'une population clonale. D'un point de vue moléculaire, ce mécanisme repose sur la compétition entre la fixation de Lrp, un régulateur à action global, et la méthylation par la protéine Dam au niveau de la région régulatrice de l'opéron fimbriaire. Si l'ensemble des membres de la famille des fimbriae de type P présente un mécanisme régulationnel commun, les fimbriae F165₁ et CS31A se distinguent néanmoins des autres membres de cette famille par une proportion importante de bactéries en phase ON, dans le cas de F165₁, et par une très faible proportion de bactéries

en phase ON dans le cas de CS31A. Outre leur prévalence de phase, les adhésines F165₁ et CS31A se distinguent aussi par une répression partielle ou totale en présence de leucine et d'alanine dans le milieu, respectivement. Du fait de l'importance de cette famille d'adhésines dans le processus infectieux, il est de ce fait important de définir les relations existant entre le métabolisme des souches *E. coli* pathogènes et cette famille d'adhésine.

Du fait de l'importance de la famille des fimbriae de type P dans le processus infectieux, nous nous sommes intéressés à comprendre au niveau moléculaire les mécanismes de régulation qui distinguent les adhésines F165₁ et CS31A de ceux existant chez Pap.

SECTION II- Recension de la littérature

A- ESCHERICHIA COLI

E. coli est un bacille à gram négatif, non sporulent et motile appartenant à la famille des Entérobactéries. D'un point de vue métabolique, ces bactéries correspondent à des microorganismes aérobies facultatifs, leur permettant d'occuper des niches fortement diversifiées. Ainsi, *E. coli* fait partie intégrante de la microflore du tractus intestinal des mammifères et oiseaux et représente l'espèce aérobie la plus abondante dans l'intestin humain (Ron, 2006).

Dans la grande majorité des cas, les souches d'*E. coli* sont commensales et jouent un rôle bénéfique pour l'hôte en favorisant l'apport de nutriments, en contribuant à la maturation du système immunitaire, et en interférant avec la colonisation du tube digestif par d'autres microorganismes potentiellement pathogènes (Kaper *et al.*, 2004, Salyers *et al.*, 2004, Johnson & Russo, 2002, Yan & Polk, 2004). Certaines souches d'*E. coli* s'avèrent en revanche nuisibles pour l'hôte (Kaper *et al.*, 2004). Ces souches se divisent alors en deux grandes catégories: les souches causant des maladies intestinales (*InPEC* pour *intestinal pathogenic E. coli*) et celles causant des maladies extra-intestinales (*ExPEC* pour *Extraintestinal pathogenic E. coli*)(Kaper *et al.*, 2004, Nataro & Kaper, 1998). Chaque catégorie est également subdivisée selon le pathotype et la présence de gènes de virulence spécifiques. Ainsi, on distingue chez le InPEC les souches d'*E. coli* entérotoxigéniques (*enterotoxigenic E. coli*, *ETEC*), les souches d'*E. coli* entérohémorragiques ou productrices de shigatoxine (*enterohaemorrhagic E. coli/shigatoxin producing E. coli*, *EHEC/STEC*), les souches d'*E. coli* entéroaggrégative (*enteroaggregative E. coli*, *EAEC*), ou encore les souches d'*E. coli* à adhérence diffuse (*diffusely adherent E. coli*, *DAEC*)(Kaper *et al.*, 2004).

1- Les souches ExPEC

Le groupe correspondant aux ExPEC englobe les souches d'*E. coli* présentes dans le tractus intestinal de façon asymptomatique mais pouvant se disséminer et causer la maladie une fois en dehors de l'appareil digestif (Kaper et al., 2004, Russo & Johnson, 2000).

Sauf exception, les souches ExPEC sont incapables de produire des infections intestinales. Ce sont donc des pathogènes opportunistes retrouvés dans les selles des sujets sains avec une fréquence variable selon les individus et les populations humaines étudiées (Smith *et al.*, 2007). Par conséquent, l'acquisition digestive de souches ExPEC par l'hôte est insuffisante à elle seule pour produire une infection et ces souches doivent avoir accès à un site extra-intestinal. La physiopathologie de l'infection par les ExPEC débute par conséquent par la colonisation d'une muqueuse et par l'éviction des systèmes de défense de l'hôte. Elle se poursuit ensuite par la multiplication dans ce site, voire la dissémination vers d'autres (Gunther *et al.*, 2002, Smith et al., 2007).

2- Maladies causées par les souches ExPEC

Chez l'homme, les infections extra-intestinales dues à *E. coli* touchent toutes les tranches d'âges et tous les sites anatomiques. Elles comprennent les infections du tractus urinaire, les méningites (survenant le plus souvent chez le nouveau-né), diverses infections intra-abdominales, les pneumonies, les infections suite à la pose de dispositifs intra-vasculaires, les ostéomyélites ou encore les infections des tissus mous (Smith et al., 2007, Kaper et al., 2004).

On distingue également au sein du groupe des souches ExPEC les souches uropathogènes (*uropathogenic E. coli*, UPEC), les souches causant des septicémies et des méningites (*meningitis-associated E. coli*, MNEC), ou encore les souches causant diverses maladies extra-intestinales chez les oiseaux (*avian pathogenic E. coli*, APEC) (Rodriguez-Siek *et al.*, 2005b).

Aux États-Unis, 40 000 décès par an sont la conséquence directe d'une septicémie à *E. coli*, soit 17 % des sepsis (Hyde *et al.*, 2002). Les bactériémies constituent de ce fait un problème majeur de santé publique de par leur fréquence et leur gravité.

a) Infections du tractus urinaire

Les infections urinaires constituent une des formes d'infection bactérienne les plus fréquentes chez l'homme (Foxman, 2002). Aux États-Unis, les souches UPEC sont responsables de plus de 70% de ce type d'infection acquise en communauté ainsi que de 50% des cas acquis en milieu hospitalier (Ferry *et al.*, 2007, Foxman, 2002, Hooton *et al.*, 1996, Stamm & Hooton, 1993, Kucheria *et al.*, 2005).

Les différentes étapes impliquées dans le développement d'UTIs incluent l'adhérence, la colonisation, la résistance aux mécanismes de défense de l'hôte et la capacité de créer des dommages tissulaires (Johnson *et al.*, 1998, Kennedy *et al.*, 1965, Yamamoto *et al.*, 1997).

Une personne atteinte d'UTI présente des bactéries dans son système urinaire à différents niveaux tels que la vessie (cystite), les reins (pyélonéphrite), l'urine (bactériurie) ou la prostate (prostatite) (Drekonja & Johnson, 2008, Foxman, 2010, Kucheria *et al.*, 2005, Marrs *et al.*, 2005). Généralement, une UTI due à une souche UPEC survient suite à la contamination fécale du tractus génito-urinaire (Johnson & Russo, 2002, Russo *et al.*, 2005). En effet, l'emplacement de ce tractus, surtout chez la femme, favorise l'ascension des bactéries de la flore fécale dans l'urètre jusqu'à atteindre la vessie (Foxman, 2010). Sur le même principe, un reflux d'urine d'une vessie déjà infectée peut occasionner la contamination des reins.

Chez un individu sain, une cystite se guérit généralement de façon naturelle sans développement de séquelles. En revanche, dans certains cas rares, elle peut évoluer vers une pyélonéphrite aiguë (Smith et al., 2007). Ce type d'infection se caractérise par un taux de morbidité élevé et peut occasionner des complications sévères allant jusqu'à la mort du patient (Smith et al., 2007). Certains sujets présentent cependant un nombre élevé de bactéries dans leur urine sans pour autant entraîner le moindre symptôme. On parle alors de bactériurie asymptomatique (Nicolle, 2003, Smith et al., 2007).

La prévalence des UTIs diffère grandement d'un individu à l'autre. Ainsi, les enfants, les femmes, notamment les femmes enceintes, et les personnes âgées présentent un risque plus élevé de développer une UTI durant leur vie. Il en va de même pour toute personne atteinte de diabète, de sclérose en plaques (ou dont l'intégrité de la moelle épinière est altérée), du VIH ou portant une sonde urinaire. Aux Etats-Unis, seul 20% des cas d'UTI répertoriés touche les hommes, alors que 50 à 60% des femmes est atteinte d'UTI au moins une fois dans leur vie (Griehling, 2005, Smith et al., 2007). De plus, une femme ayant été atteinte d'une UTI a un risque de récurrence plus élevé, de l'ordre de 25 à 30% des cas (Bower *et al.*, 2005, Foxman, 2002, Kucheria et al., 2005, Smith et al., 2007). Chez les femmes préménopausées, les souches ExPEC sont également responsables de 85 à 95% des cas de cystites sans complication, représentant environ de 6 à 8 millions de cas par an (Russo & Johnson, 2003). Concernant les pyélonéphrites, 90% sont dues à une souche ExPEC sur les 250,000 cas recensés chez ce même groupe de femmes et environ 100,000 nécessitent une hospitalisation (Russo & Johnson, 2003, Smith et al., 2007). Les UTIs représentent également l'infection bactérienne principale chez les femmes enceintes. En effet, 4 à 10 % de cette tranche de la population présentent une bactériurie asymptomatique, 1 à 4 % d'entre elles sont atteintes de cystite et 1 à 2% développent une pyélonéphrite aiguë. Dans ce dernier cas, le développement d'une pyélonéphrite durant le dernier trimestre de grossesse peut avoir des conséquences sévères pour le fœtus, allant du retard mental, jusqu'à la naissance d'un enfant mort-né

(Foxman, 2002). De plus, les femmes ayant déjà eu des antécédents d'UTI ont plus de chance d'en développer une seconde si elles sont enceintes.

Les personnes âgées représentent une autre tranche de population à risque. En effet, 11 à 25% des personnes âgées non institutionnalisées développent une bactériurie asymptomatique, et 10% d'entre elles développent une UTI avec apparition de symptômes (Foxman, 2002). Cette incidence augmente davantage chez le groupe de personnes âgées institutionnalisées (Tal *et al.*, 2005). La pose de cathéter urinaire est également une des causes majeures d'UTI d'origine nosocomiale, avec plus d'un million de cas par an aux Etats-Unis dont 25 à 35% dus à une souche ExPEC (Foxman, 2002, Russo & Johnson, 2003).

b) Méningites néonatales

La barrière hémato-encéphalique est une barrière physiologique séparant le système nerveux central (SNC) du système vasculaire. Elle est constituée des cellules endothéliales microvasculaires (BMECs pour *brain microvascular endothelial cells*) recouvrant les capillaires sanguins (Huang & Jong, 2001, Huang *et al.*, 2000). Elle a pour fonction majeure de maintenir l'homéostasie des liquides interstitiels cérébraux, en particulier la régulation des échanges entre le sang et le cerveau. Ainsi, cette couche de cellules endothéliales prévient le passage de macromolécules au travers du SNC et constitue une barrière physique contre l'entrée des pathogènes du sang vers le cerveau (Huang & Jong, 2001, Huang *et al.*, 2000).

Aux Etats-Unis, les méningites néonatales causées par *E. coli* s'accompagnent d'un taux de mortalité de 8% et, dans la majorité des cas, les sujets ayant survécu présentent des troubles neurologiques et du développement (Russo & Johnson, 2003, Smith *et al.*, 2007). L'infection peut survenir de différentes manières mais est généralement associée à une

contamination pré- ou postnatale (contamination lors du passage du cordon ombilical ou via le canal vaginal) (Smith et al., 2007).

La méningite néonatale bactérienne est une inflammation des membranes du cerveau ou de la moelle épinière du nouveau-né. Elle se caractérise par l'apparition d'exsudats purulents dans les méninges, d'une inflammation périvasculaire et d'œdèmes au cerveau. L'incidence de ce type de méningite bactérienne est faible, de l'ordre de 2 à 5 cas par tranche de 10,000 habitants dans les pays développés; en revanche, elle est dix fois plus élevée dans les pays en voie de développement (Smith et al., 2007). Aux Etats-Unis, 20 à 40% des cas de méningite néonatale répertoriée chaque année est dû à *E. coli* et plus de la moitié d'entre eux est due aux souches ExPEC (Russo & Johnson, 2003). De plus, près de 80% de ces isolats présentent l'antigène capsulaire K1, suggérant un rôle important de ce type de capsule dans la pathogénie de ces souches (Pong & Bradley, 1999, Smith et al., 2007). Des études ont montré que la présence d'acide polysialique (PSA) dans les capsules de type K1, faiblement immunogène, contribuerait aux mécanismes de résistance aux effets bactéricides du sérum et à l'opsonisation par les cellules phagocytaires de l'hôte (Cieslewicz & Vimr, 1997, Brusés & Rutishauser, 2001, Nakayama *et al.*, 1995). La résistance de ces souches face au système immunitaire de l'hôte serait ainsi due à un mécanisme de mimétisme entre la présence de PSA dans les capsules de type K1 et celle trouvée à la surface des cellules neurales embryonnaires et du nouveau-né (Smith et al., 2007). De plus, il semblerait que la présence de PSA dans les capsules de type K1 interfère avec le développement normal du cerveau (Smith et al., 2007).

La pathogénèse d'une méningite néonatale causée par une souche ExPEC se déroule en plusieurs étapes : les bactéries se trouvent d'abord dans le sang (bactériémie); elles atteignent ensuite les BMECs, puis adhèrent et envahissent ces cellules; elles traversent alors les méninges et se retrouvent finalement dans le SNC (Xie *et al.*, 2004, Xie *et al.*, 2006). Comme dans beaucoup de processus

infectieux, la première étape est déterminante. Il existe en effet une corrélation positive entre la présence d'une bactériémie et le développement de la méningite. Ainsi, Kim *et al.* ont montré que la capacité de développer une méningite est significativement plus élevée quand le compte bactérien d'*E. coli* dans le sang atteint au moins 10^3 bactéries/mL (Kim, 2001). Une fois dans le sang, *E. coli* adhère alors à la surface des BMECs et les bactéries sont internalisées par transcytose où elles demeurent à l'intérieure de vacuoles. La présence de l'antigène K1 est à nouveau requise dans la survie intracellulaire et il est suggéré que sa présence inhibe la fusion des vacuoles d'internalisation avec les protéines lysosomales (Xie et al., 2004). Les bactéries traversent ensuite les BMECs jusqu'à se retrouver dans le SNC où elles se multiplient et induisent la libération de molécules pro-inflammatoires (cytokines, chimiokines, oxyde nitrique), menant à l'augmentation de la perméabilité de la barrière hémato-encéphalique et à une pléiocytose (augmentation du nombre de leucocytes dans le liquide céphalo-rachidien) (Xie et al., 2004, Smith et al., 2007). Tout ceci mène finalement à l'apparition d'œdèmes cérébraux et à l'augmentation de la pression crânienne.

c) Septicémies

La septicémie est une condition médicale due à une infection incontrôlée du système circulatoire sanguin et qui s'accompagne d'une réponse inflammatoire systémique. Dans ses formes sévères, une septicémie mène à un dysfonctionnement d'un ou de plusieurs organes et peut même s'accompagner d'une hypotension menant au choc septique (Annane *et al.*, 2005). La septicémie peut être induite par l'entrée de bactéries à différents sites comme les reins (après une UTI), les viscères (péritonites), la peau (cellulite) ou les poumons (pneumonie).

Du fait de l'absence d'un système immunitaire mature, les premiers jours de la vie d'un nouveau-né sont particulièrement à risque dans le développement

d'une septicémie. Il est communément admis que les microorganismes atteignent le fœtus via l'ascension du tractus génital de la mère, ou en traversant les membranes amniotiques si la mère présente déjà une bactériémie. Quatre à 5 millions d'enfants meurent ainsi chaque année durant les 4 premières semaines de leur vie, et 98% de ces morts surviennent dans les pays non développés (Vergnano *et al.*, 2005, Zupan, 2005). Les souches ExPEC représentent la deuxième cause de septicémie néonatale après les streptocoques du groupe B (Hyde *et al.*, 2002). Lors de l'accouchement, *E. coli* est présent dans la filière génitale chez 13% des femmes et près de 50% des nouveaux-nés sont colonisés à la naissance. La fréquence des infections materno-fœtales causée par *E. coli* est de 0,6 pour 1000 naissances et 40% des septicémies néonatales est due à une souche présentant l'antigène capsulaire K1 (Smith *et al.*, 2007). *E. coli* est également une cause majeure de septicémie chez les personnes âgées et la première cause d'hospitalisation aux Etats-Unis pour les personnes de plus de 65 ans atteintes de ce type d'infection (McBean & Rajamani, 2001, Jackson *et al.*, 2005). Parmi les facteurs aggravants, la présence de diabète ou d'une incontinence urinaire est fréquemment associée à une susceptibilité accrue à une bactériémie causée par *E. coli* (McBean & Rajamani, 2001).

E. coli est également l'agent responsable de septicémies dans ses formes les plus graves dans 17% des cas aux Etats-Unis. Ainsi, Russo et Jonhson ont estimé qu'en 2001, les souches d'*E. coli* ont été responsables de 127,500 cas de septicémies sévères, occasionnant près 40,000 décès, soit un taux de mortalité de 30%, et que l'ensemble a eut un coût estimé entre 1,1 et 2,8 milliards de dollars (Russo & Johnson, 2003).

d) Infections chez les animaux

Au même titre que les humains, les animaux sont contaminés par des souches pathogènes (Smith *et al.*, 2007, Gyles & Fairbrother, 2010). En effet, les bactéries pathogènes présentes dans l'environnement sont ingérées par les

animaux susceptibles et colonisent alors le tractus intestinal (ou le tractus respiratoire dans le cas des oiseaux). Également considérées comme des pathogènes opportunistes, les souches ExPECs font partie de la microflore normale et colonisent différentes surfaces muqueuses, dont celle des tractus intestinal et respiratoire. Dans le cas d'un animal affaibli en revanche, la bactérie peut alors profiter de cette situation avantageuse pour traverser plus facilement la muqueuse et atteindre la circulation sanguine. La bactérie produit alors des toxines qui endommagent les tissus et, dans les cas d'infections localisées, il peut y avoir interaction avec les matrices extracellulaires et entraîner différentes conditions telles qu'une pneumonie, une sérosité, une mammité ou des infections du tractus urinaire (Gyles & Fairbrother, 2010).

Ces bactéries se trouvent alors dans un environnement qui leur permet de résister aux effets létaux du complément et des phagocytes, et ainsi de persister et de se multiplier (Figure 1). Par conséquent, le tube digestif des mammifères et oiseaux est à considérer comme un réservoir des souches ExPEC. De plus, elles peuvent coloniser et se multiplier dans le sang, notamment grâce à leur capacité d'acquérir le fer et de résister à l'effet bactéricide du sérum ou à la phagocytose (Gyles & Fairbrother, 2010).

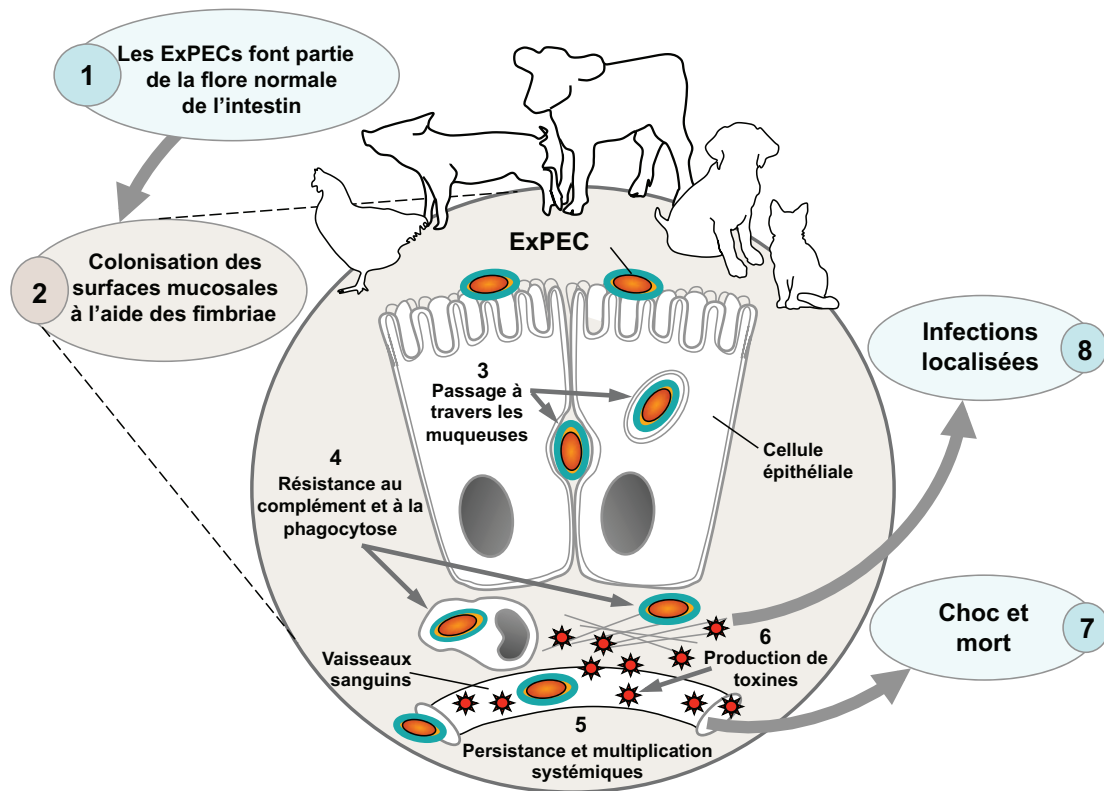


Figure 1. Pathogénèse des souches ExPECs chez les animaux. Les bactéries pathogènes présentes dans l'environnement sont ingérées par les animaux et entrent dans le tractus intestinal (1). Ces pathogènes opportunistes colonisent différentes surfaces muqueuses, dont celle des tractus intestinal et respiratoire (2). Chez l'animal affaibli, ou chez le nouveau-né, la bactérie traverse plus facilement la muqueuse jusqu'à la circulation sanguine (3). Les bactéries internalisées ont la capacité de résister aux effets létaux du complément et des phagocytes (4), et ainsi persistent et se multiplient dans le système (5). La bactérie produit ensuite des toxines qui endommagent les tissus (6). Cette libération d'entérotoxines par la bactérie peut déclencher des réponses cytokinaires pouvant entraîner un choc et la mort de l'animal (7). Dans les cas d'infections localisées, il peut y avoir interaction avec les matrices extracellulaires pouvant entraîner différentes conditions telles qu'une pneumonie, une sérosite ou des infections du tractus urinaire (8). (adapté de Gyles & Fairbrother, 2010).

-i- Infections chez le porc

Chez le porc, les infections causées par les ExPECs correspondent majoritairement à des septicémies, majoritairement chez les nouveaux-nés et, dans une moindre mesure, les jeunes porcelets. Elles se caractérisent par une infection systémique aigüe, accompagnée dans les phases tardives de diarrhée et résultent dans la quasi-totalité des cas dans la mort de l'animal au bout de 3 à 8 heures (Gyles & Fairbrother, 2010). Les infections peuvent être cependant de type localisé comme dans le cas de polyarthrites, de pneumonies, de métrites, ou de méningites (Fairbrother *et al.*, 2005). On observe également fréquemment une congestion de l'intestin, des ganglions lymphatiques mésentériques, ou de certains organes extra-intestinaux (Fairbrother *et al.*, 2005). Dans le cas des formes sévères menant au choc septique et à la mort, les lésions histopathologiques observées correspondent aux cas typiques de septicémies, telle qu'une gastroentérite hémorragique, une hémorragie dans les reins, ou encore la formation de thrombi sur la muqueuse du petit intestin et de l'estomac (Gyles & Fairbrother, 2010).

-ii- Infections chez les bovins

Comme chez le porc, les septicémies causées par *E. coli* apparaissent principalement durant les premiers jours de la vie d'un veau. Il est communément admis que les bactéries entrent dans l'organisme en passant à travers l'épithélium intestinal ou via le cordon ombilical. Les signes cliniques sont également similaires à ceux observés chez le porc. Les bactéries sont alors retrouvées dans la majorité des fluides d'un animal infecté (sécrétion nasale, urine, fèces) (Gyles & Fairbrother, 2010). Lors d'un examen post mortem, on observe fréquemment la présence de pétéchies sur l'épicarde et la membrane séreuse et, occasionnellement, la présence d'œdèmes pulmonaires et d'un gonflement de la rate. Dans les formes chroniques, la méningite et l'arthrite fibrineuse sont également observées (Gyles & Fairbrother, 2010).

-iii- Infections chez les oiseaux

Les infections dues aux souches d'*E. coli* pathogènes aviaires (Avian pathogenic *E. coli*, APEC) se caractérisent par un taux de morbidité et un taux de mortalité élevés (Ewers *et al.*, 2007, Gyles & Fairbrother, 2010, Rodriguez-Siek *et al.*, 2005b, Pourbakhsh *et al.*, 1997). Ce type d'infection est de ce fait un problème majeur en industrie agroalimentaire et cause des pertes économiques considérables. Généralement, les souches pathogènes entrent après inhalation de poussières contaminées (Pourbakhsh *et al.*, 1997, Gyles & Fairbrother, 2010). Les bactéries colonisent alors les voies respiratoires, puis les sacs aériens où elles sont responsables d'aérosacculite. Les oiseaux développent dès lors une septicémie avec notamment l'apparition de pneumonie, de péricardite, de périhépatite, de péritonite jusqu'à la mort de l'oiseau. L'aérosacculite se manifeste généralement par un épaissement de la paroi des sacs aériens ainsi que la présence d'exsudat caséux (Gyles & Fairbrother, 2010). D'un point de vue histopathologique, les phases précoces d'une aérosacculite s'accompagnent d'œdèmes, d'un exsudat séreux ou fibrineux, de l'infiltration d'hétérophiles et de macrophages, résultant en une périhépatite et/ou une péricardite (Gyles & Fairbrother, 2010). Les souches APEC infectent également l'oviducte des poules pondeuses via le sac aérien abdominal, occasionnant alors une salpingite et l'incapacité de pondre. Les souches APEC sont aussi retrouvées dans la cavité péritonéale où elles causent alors une péritonite menant à la mort de l'oiseau. Les souches APEC causent également la maladie dite des grosses têtes (swollen head syndrome) qui se caractérise par la présence d'œdèmes gélatineux dans la tête et les tissus périorbitaux ainsi que d'un exsudat caséux sur les conjonctives et des glandes lacrymales (Pourbakhsh *et al.*, 1997). Parmi l'ensemble des sérotypes causant ce type d'infection, les sérotypes O2 et O78 sont à eux seuls responsables de 80% des cas répertoriés dans le monde (Gyles & Fairbrother, 2010).

3- Phylogénie des souches ExPEC

Avec des génomes allant de 4,5 à 5,5 mégabases et des pourcentages d'identité de l'ordre de 70%, la diversité des souches *E. coli* est grande (Foxman & Riley, 2001). Cependant, des études basées sur l'analyse de PCR ciblant des régions spécifiques du génome ont montré qu'*E. coli* est une espèce clonale pouvant néanmoins être classée en plusieurs groupes phylogénétiques distincts, correspondant aux groupes A, B1, B2, C, D, E et F (Figure 2) (Herzer *et al.*, 1990; Clermont *et al.*, 2011). Les souches ExPEC forment quant à elles un groupe phylogénétique et épidémiologique distinct des autres souches d'*E. coli*. Ainsi, la majorité des souches UPEC appartiennent au groupe B2, et dans une moindre mesure au groupe D (Escobar-Páramo *et al.*, 2004, Ewers *et al.*, 2009, Kanamaru *et al.*, 2006, Moreno *et al.*, 2005, Picard *et al.*, 1999). Elles se caractérisent par l'acquisition de plusieurs facteurs de virulence spécifiques. Au sein même du groupe B2 existe également une très grande diversité (Figure 2). À ce titre, aucun des facteurs de virulence présents au sein des souches du groupe B2 n'est suffisant pour caractériser une souche UPEC. Ceci suggère que les souches UPEC résultent de l'acquisition à maintes reprises de plusieurs facteurs de virulence (Zhang & Foxman, 2003). Il semble en revanche qu'il existe une relation entre l'appartenance à un des groupes phylogénétiques et l'origine des isolats d'ExPEC. En effet, les souches ExPEC appartenant aux groupes phylogénétiques A and B1 sont considérées comme bénignes voir bénéfique pour l'hôte. Généralement, elles se retrouvent dans le tractus intestinal humain et présentent moins de facteurs de virulence spécifiques aux souches pathogènes que les souches appartenant aux groupes phylogénétiques B2 ou D (Nowrouzian *et al.*, 2006, Nowrouzian *et al.*, 2005). En effet, en comparant le profil génétique d'isolats d'origine humaine et animale, Maynard *et al.* ont montré que si la majorité des souches d'origine humaine appartiennent au groupe de souches virulentes B2 et D, celles d'origine animale appartiennent majoritairement aux groupes A et B1 (Maynard *et al.*, 2004). Les souches APEC quant à elles sont fortement apparentées aux souches UPEC humaines (Rodriguez-Siek *et al.*,

2005a, Rodriguez-Siek et al., 2005b). De plus, la plupart des facteurs de virulence présents chez les souches APEC sont également retrouvés chez les souches ExPEC causant des méningites néonatales et des UTIs chez l'homme, suggérant que les souches APEC pourraient être agents de zoonose, voir servir de réservoir pour les gènes de virulence des souches ExPEC (Smith et al., 2007).

Les souches d'appartenance aux groupes A et B1 sont également impliquées dans ces cas d'infection. En effet, bien que dans le cas d'une méningite néonatale les souches ExPEC appartiennent majoritairement au groupe B2, l'appartenance aux groupes A, B1 et D survient généralement lorsque le nouveau-né présente déjà des complications médicales tel qu'un système immunitaire particulièrement affaibli (Dezfulian *et al.*, 2004, Moreno et al., 2005). Ainsi, les méningites néonatales causées par *E. coli* sont plus associées à un système immunitaire immature qu'à une susceptibilité accrue du cerveau du nouveau-né à l'adhésion et l'invasion par *E. coli* (Smith et al., 2007). Il semble donc que le pouvoir pathogène des souches ExPEC dépend également de la capacité de l'hôte à réagir face à une infection.

Sur le même principe, à partir d'un échantillon de dix-huit isolats d'*E. coli*, Dezfulian *et al.* ont montré que toutes les souches prélevées sur des porcelets et veaux atteints de septicémies et appartenant aux groupes phylogénétiques A et B1 partagent plusieurs facteurs de virulence avec des souches ExPEC d'origine humaine (Dezfulian et al., 2004). Ils ont également montré que ces isolats étaient létaux dans un modèle d'infection expérimentale chez le porcelet. De façon similaire, Girardeau *et al.* ont montré qu'il existe une forte similarité en terme de facteurs de virulence entre des isolats de souches ExPEC associés à des cas d'UTI ou de septicémies humaines et des isolats associés à des cas de bactériémies chez le porcelet et le veau (Girardeau *et al.*, 2003). Par conséquent, la présence de facteurs de virulence *sensu stricto* ne suffit pas à expliquer la différence de virulence des souches fortement pathogènes appartenant aux groupes phylogénétiques B2 et D et celle des souches généralement considérées comme opportunistes et appartenant aux groupes phylogénétiques A et B1.

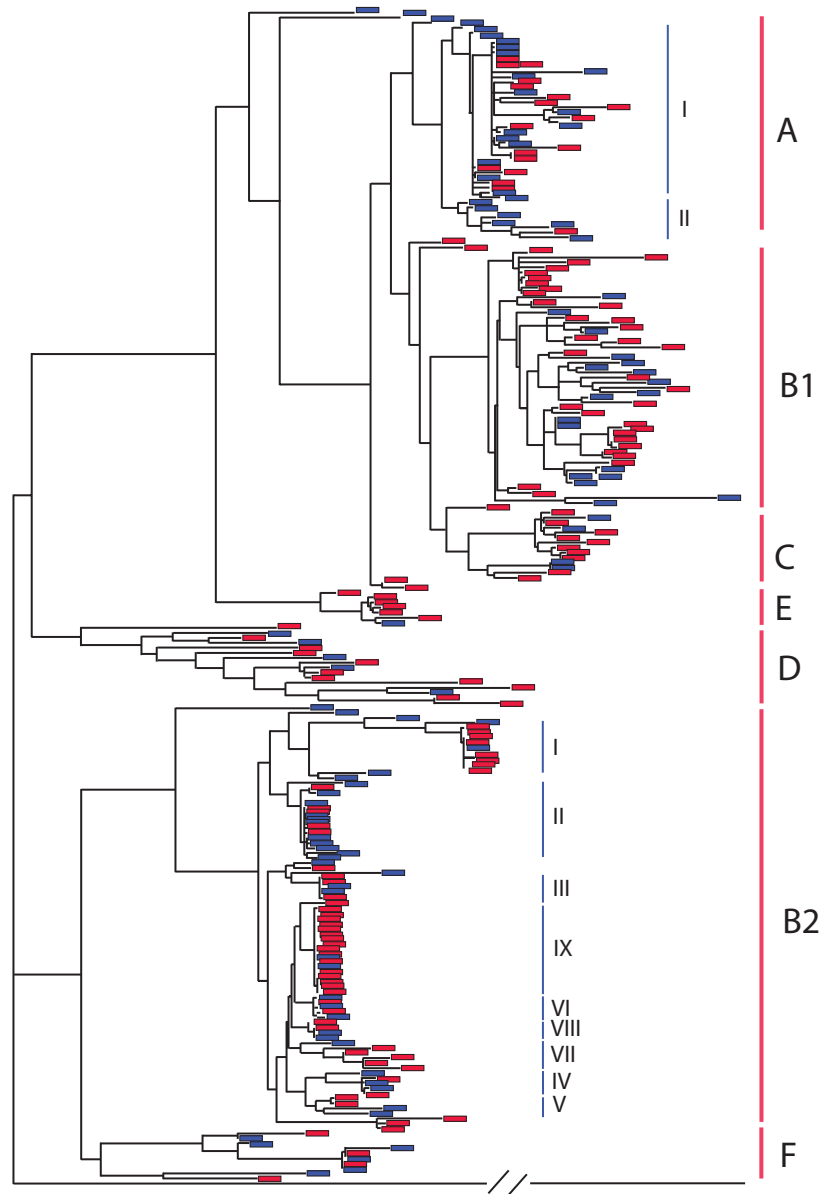


Figure 2. Arbre phylogénétique issu de l'analyse de la séquence de 8 gènes de ménage provenant de 234 souches d'*E. coli*, d'origine humaine ou animale, avec notamment le classement selon l'appartenance à chacun des groupes phylogénétiques. Les souches pathogènes sont représentées en rouges et les souches non pathogènes sont représentées en bleues. Adapté de Clermont et al (2001).

B- RÉGULATION DE L'EXPRESSION DES GÈNES BACTÉRIENS

1- Introduction

Les bactéries réagissent aux conditions environnementales en utilisant différents modes de régulation : régulation métabolique, régulation traductionnelle, régulation transcriptionnelle...etc... Pour ce faire, leurs gènes sont organisés sous forme de réseaux hiérarchisés et interconnectés entre eux, correspondant aux régulons. Ces réseaux n'en restent pas moins flexibles en fonction des conditions environnementales de croissance que les bactéries sont amenées à rencontrer au cours de leur cycle cellulaire.

L'expression des régulons est elle-même sous le contrôle de protéines régulatrices (facteurs transcriptionnels) qui reconnaissent des régions spécifiques dans la séquence nucléotidique sur lesquelles elles se fixent. Dans certains cas, la présence de molécules coeffectrices est nécessaire pour observer l'activité de ces régulateurs. Ainsi, l'activation (ou l'inhibition) d'un régulon, pour une condition de croissance donnée, dépend directement de la concentration en facteur transcriptionnel sous sa forme active dans la cellule bactérienne.

Au niveau de la séquence promotrice, cela se manifeste par l'acquisition d'un ou de plusieurs signaux entrants qui sont intégrés en un signal régulationnel sortant unique. Par conséquent, la force du signal sortant dépend directement de la force de l'ensemble des signaux entrant dans la région promotrice. Les différents paramètres influençant la force de transcription sont plus particulièrement la localisation des sites de fixation par rapport à la position du promoteur, la fidélité du motif de reconnaissance du régulateur par rapport au motif consensus et le type de facteur transcriptionnel impliqué (à savoir s'il s'agit d'un régulateur local ou s'il s'agit au contraire d'un régulateur à action beaucoup plus globale).

2- Transcription

La transcription est le processus biologique qui consiste en la copie des régions dites codantes de l'ADN en une molécule d'ARN (ARNm, ARNt, ARNr, petits ARN). Chez les procaryotes, l'enzyme impliquée dans cette activité est unique et correspond à l'ARN polymérase (Browning & Busby, 2004).

La transcription consiste en 5 étapes distinctes :

- i- la préinitiation (isomérisation),
- ii- l'initiation,
- iii- le retrait du site promoteur,
- iv- l'élongation et,
- v- la terminaison.

Durant le processus de préinitiation, l'ARN polymérase se fixe sur la partie centrale de la région promotrice correspondant aux positions -10 et -35, par rapport au site de début de transcription, en amont du gène. Une fois l'ARN polymérase fixée sur l'ADN, l'ensemble recouvre une zone comprise entre les positions -10 et +2; on parle alors d'isomérisation (Browning & Busby, 2004).

Chez les bactéries, l'initiation de la transcription requiert la présence de protéines appelées facteurs sigma (σ). Ces facteurs sont essentiels dans la reconnaissance de la région promotrice par l'ARN polymérase et varient fortement en nombre. En début d'initiation, le facteur σ s'associe à l'ARN polymérase et permet la reconnaissance des séquences -35 et -10. Le retrait du site promoteur commence à partir du moment où la transcription des premières bases est amorcée (Figure 3).

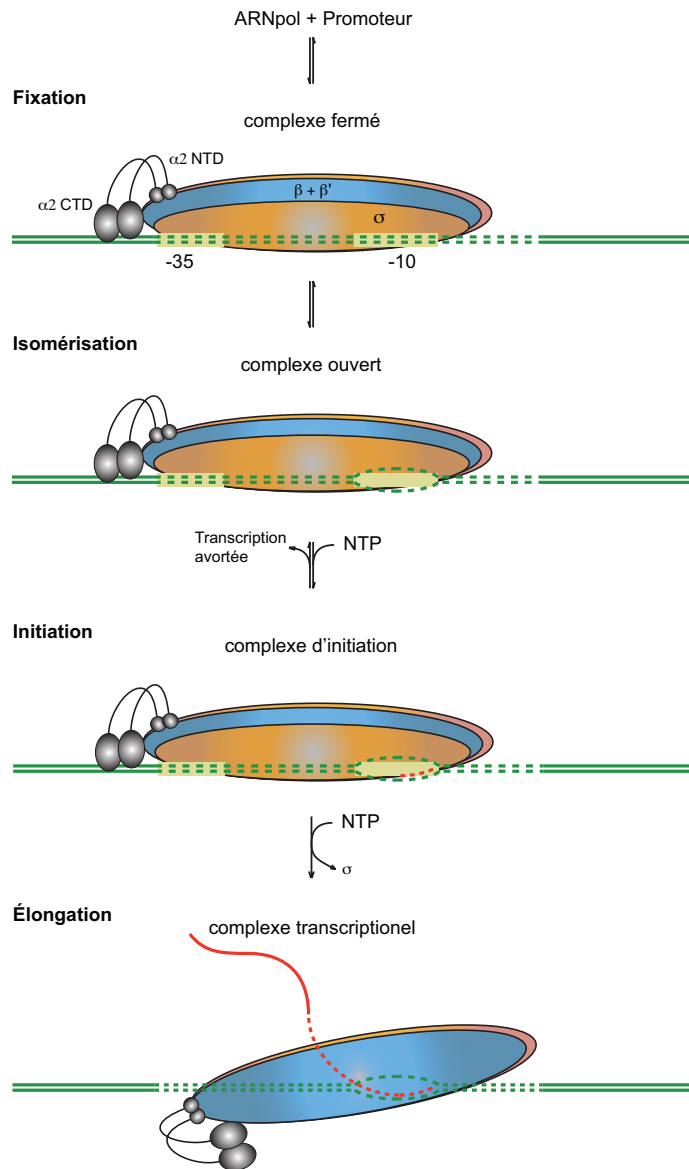


Figure 3. Interaction de l'ARNpolymérase (ARNpol) avec sa séquence promotrice. L'ARNpol est représentée avec ses deux sous-unités α (en gris), ses sous-unités β et β' (en bleu et rose) et le co-facteur σ (en orange). En ce qui concerne les sous-unités α , chacun des domaines carboxyterminaux ($\alpha 2$ CTD) et aminoterminaux ($\alpha 2$ NTD) sont représentés interagissant avec l'ADN et les sous-unités β et β' , respectivement.

Durant les premières étapes, l'ARN polymérase patine régulièrement le long de l'ADN, provoquant ainsi la formation de transcrits incomplets (initiation avortée); ce n'est seulement qu'à partir du moment où environ 20 pb sont transcrites que l'ARN polymérase devient fidèle et qu'elle poursuit alors l'élongation jusqu'à atteindre la fin de la séquence d'ARN à synthétiser.

La terminaison de la transcription correspond alors au détachement de l'ARN polymérase et survient soit suite à la rencontre d'une structure en forme d'épingle à cheveu le long de l'ARN (terminaison rho-indépendante), soit suite à la fixation du cofacteur rho, provoquant alors la dissociation du complexe ADN-ARN-polymérase-ARN (terminaison rho-dépendante) (Ciampi, 2006, Henkin, 2000, Nudler & Gottesman, 2002).

3- Régulation transcriptionnelle

Afin de réguler l'expression génétique, les bactéries utilisent des facteurs transcriptionnels spécialisés qui interagissent avec des régions spécifiques le long de l'ADN. Pour un facteur transcriptionnel donné, les motifs reconnus sur l'ADN peuvent varier en séquence et en longueur mais ils partagent dans la grande majorité des cas une séquence consensus qui représente le motif théorique idéal reconnu par ce facteur transcriptionnel donné (van Hijum *et al.*, 2009). En général, la localisation du site de fixation détermine si le facteur transcriptionnel joue le rôle d'activateur ou de répresseur. De plus, le degré de fidélité du site de fixation par rapport au motif consensus théorique détermine la force de la fixation du facteur transcriptionnel sur sa séquence cible en terme d'activité transcriptionnelle (Lässig, 2007).

La longueur d'un site de fixation est généralement comprise entre 12 et 30 pb; de plus, ces sites se trouvent souvent sous la forme de répétitions directes ou de palindromes, allant de pair avec l'état d'oligomérisation du facteur

transcriptionnel, se présentant généralement sous forme de dimère, comme dans le cas d'AraC chez *E. coli* (Rodionov, 2007).

La fixation et la libération de ces répresseurs ou activateurs sont elles-mêmes dépendantes de la fixation de cofacteurs. Ces-derniers correspondent à des molécules qui varient en taille et en nature chimique (ions, nucléotides, groupements phosphate liés de façon covalente au facteur transcriptionnel, peptides, voir même protéines entières) (Aki *et al.*, 1996, Fabret *et al.*, 1999, Hantke, 2001, Stülke & Hillen, 1999).

4- Mécanismes de répression transcriptionnelle

La présence d'un facteur transcriptionnel présente une action inhibitrice de l'activité transcriptionnelle selon 4 scénarios différents (Browning & Busby, 2004, Lloyd *et al.*, 2001, Rodionov, 2007):

- i- répression par encombrement stérique,
- ii- répression par blocage de l'étape d'élongation (souvent en se fixant sur le codon d'initiation),
- iii- répression par formation d'une boucle d'ADN avec des sites de fixation situés de part et d'autre de la séquence promotrice,
- iv- répression par modulation de l'action positive d'un second facteur transcriptionnel (Figure 4).

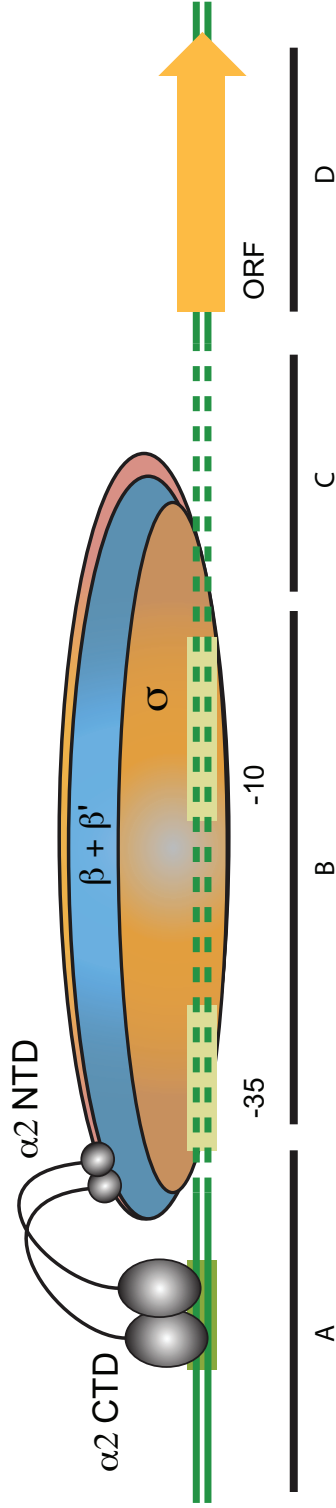
Dans le cas d'une répression par modulation de la présence d'un activateur, la fixation d'un facteur transcriptionnel masque, partiellement ou en totalité, la présence d'un second site de fixation pour un activateur, empêchant ainsi son action positive sur la transcription.

5- Mécanismes d'activation transcriptionnelle

De façon réciproque, quatre modes d'activation sont à envisager (Barnard *et al.*, 2004, Browning & Busby, 2004, Lloyd *et al.*, 2001, Rodionov *et al.*, 2011, Rodionov, 2007, Smits *et al.*, 2007) :

- i- activation par fixation d'un facteur transcriptionnel en amont de la séquence promotrice et interagissant avec les sous-unités α de l'ARN polymérase (activation de classe I);
- ii- activation par fixation d'un facteur transcriptionnel sur une région adjacente à la région promotrice qui recrute l'entrée du facteur σ (activation de classe II);
- iii- activation par changement conformationnel de l'ADN, rendant ainsi la séquence promotrice accessible à l'entrée d'un facteur σ , généralement en tordant la double hélice d'ADN;
- iv- activation par modulation de l'action inhibitrice d'un répresseur (Figure 4).

Cette dernière voix d'activation correspond également au mécanisme d'antirépression décrit dans le cas de l'activateur de la compétence ComK chez *Bacillus subtilis*. En effet, en se fixant sur l'ADN à proximité des répresseurs Rok et CodY sur la séquence promotrice du gène *comK* codant sa propre synthèse, cette protéine provoque une levée de répression sans toutefois déplacer physiquement ces deux répresseurs (Smits *et al.*, 2007).



Type	Mécanisme	Action	Site de fixation du facteur transcriptionnel			
			En amont (A)	Séquence promotrice (B)	En aval (C)	ORF (D)
Répression	gène stérique	absence de l'ARNpol		+		
Répression	barrage	arrêt prématuré de l'élongation			±	+
Répression	boucle d'ADN	absence de l'ARNpol	+		+	±
Répression	modulation de l'action d'un activateur	obstruction de l'entrée d'un activateur	±			
Activation	Classe I	interaction avec la sous-unité σ	+			
Activation	Classe II	favorise la fixation de la sous-unité σ	+			
Activation	conformation de l'ADNII	torsion de l'hélice d'ADN		+		
Activation	modulation de l'action d'un répresseur	obstruction de l'entrée d'un répresseur	±		±	±

Figure 4. Mécanisme moléculaire de régulation transcriptionnelle. Les quatre principaux types de répression et d'activation sont représentés; (+) correspond à la fixation du facteur transcriptionnel sur cette région; (±), plus d'un seul facteur transcriptionnel peut se fixer dans cette région. La phase ouverte de lecture (ORF pour *Open Reading Frame*) correspond au gène dont l'expression est régulée au niveau de la séquence promotrice. Adapté de van Hijum *et al.* (2009).

6- Contraintes spatiales sur l'architecture de la séquence promotrice

La plupart des sites de fixation pour un répresseur se situe entre les positions -60 et +60 par rapport au site d'initiation de la transcription et la force de la répression dépend de la distance séparant le site de fixation par rapport à la séquence promotrice (Collado-Vides *et al.*, 1991, Espinosa *et al.*, 2005, Moreno-Campuzano *et al.*, 2006). Les sites activateurs sont, quant à eux, généralement situés en amont de, ou juste, après la séquence -35 (Rodionov, 2007). Dans le cas des activateurs de classe I, interagissant avec les sous-unités α de l'ARN polymérase, le facteur transcriptionnel se lie majoritairement dans des sites plus éloignés (entre les positions -60 et -95) comparé aux activateurs de classe II pour lesquels le site de fixation va jusqu'à se superposer avec la séquence -35 (Figure 5) (van Hijum *et al.*, 2009).

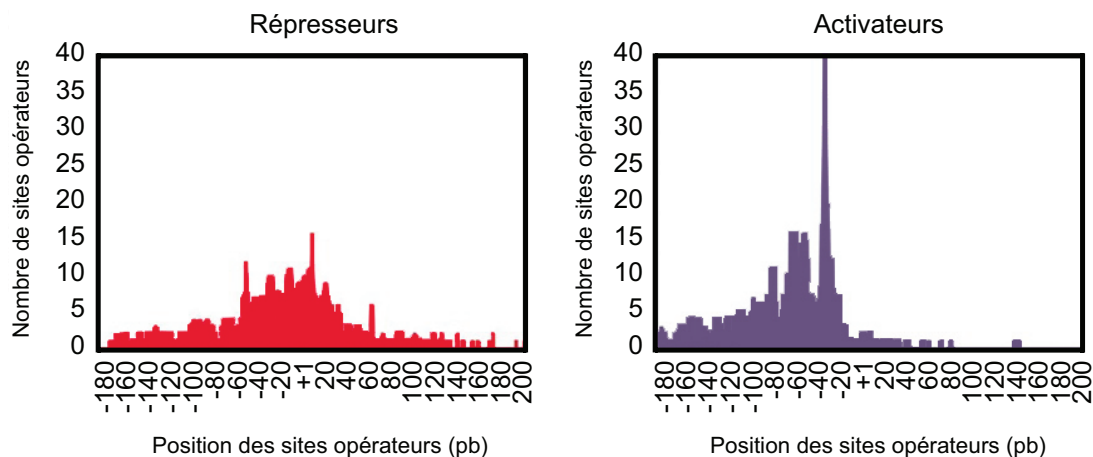


Figure 5. Localisation des sites de fixation des facteurs transcriptionnels selon leur action (répresseur ou activateur) pour plus de 1,100 promoteurs chez *E. coli*. La distribution des sites de fixation des facteurs transcriptionnels est représentée en fonction de leur position par rapport au site d'initiation de la transcription (site +1). Adapté de van Hijum *et al.* (2009).

7- Mécanismes de coopérativité et contrôle du promoteur

Chez *E. coli*, comme chez l'ensemble des bactéries, la majorité des unités transcriptionnelles possède plus d'un site de fixation par facteur transcriptionnel. De plus, ces gènes sont souvent régulés par plus d'un facteur transcriptionnel. Par conséquent, la présence simultanée d'activateurs et de répresseurs résulte en des actions antagonistes sur une séquence promotrice particulière (régulation compétitive) (Buchler *et al.*, 2003, Gertz & Cohen, 2009, Hermsen *et al.*, 2006).

À l'inverse, plusieurs activateurs peuvent agir de concert sur l'expression d'un ou de plusieurs gènes. Chacun de ces facteurs transcriptionnels répondent alors à un signal environnemental donné (régulation coopérative) (Browning & Busby, 2004, Buchler *et al.*, 2003). Sur ce principe, il est possible d'imaginer qu'une unité transcriptionnelle puisse être activée de manière indépendante par des activations de classe I et II. Il est tout aussi envisageable que plusieurs activateurs interagissent avec l'ADN de façon coopérative, ou encore qu'un activateur interfère avec l'action d'un répresseur pendant qu'un second facteur transcriptionnel active directement la transcription.

Deux modes de coopérativité peuvent être envisagés (Hermsen *et al.*, 2006, van Hijum *et al.*, 2009). En premier, l'homo-coopérativité correspond au cas pour lequel plus d'une molécule d'un même facteur transcriptionnel interagissent avec plusieurs sites de fixation situés dans la région promotrice. Dans l'autre cas, l'hétéro-coopérativité correspond au cas de figure pour lequel la fixation d'un facteur transcriptionnel favorise la fixation d'un second type de facteur transcriptionnel. Ainsi, les notions de coopérativité ou de compétitivité ajoutent des degrés de complexité pour les régions régulatrices situées en *cis*.

Des portes d'entrée faisant référence à la logique booléenne peuvent dès lors être appliquées sur les mécanismes de régulation entre une séquence

régulatrice, deux facteurs transcriptionnels distincts et l'ARN polymérase (van Hijum et al., 2009). Bien que représentant des modèles théoriques, la mise en place de ces différents scénarios s'avère être d'une grande utilité dans la compréhension des mécanismes de régulation transcriptionnelle.

À titre d'illustration, voici quelques exemples pouvant être rencontrés (Figure 6) (Buchler et al., 2003, Hermsen et al., 2006, Silva-Rocha & de Lorenzo, 2008, van Hijum et al., 2009):

- Dans le cas d'une porte d'entrée *AND*, l'unité transcriptionnelle est active si et seulement si les deux facteurs transcriptionnels sont présents sous forme active à forte concentration. Ceci peut être dû à la présence de 2 sites de fixation agissant de manière coopérative.
- Dans le cas d'une porte d'entrée *OR*, la présence d'un des 2 facteurs transcriptionnels est suffisante pour activer l'unité transcriptionnelle. Ce cas de figure peut s'observer quand les sites de fixation agissent de façon indépendante sur la même séquence cible.
- Dans le cas des portes d'entrée *NAND* (*not and*), l'unité transcriptionnelle est réprimée si et seulement si les deux facteurs transcriptionnels sont présents sous forme active à forte concentration pour exercer leur fonction de répresseur. Ceci peut correspondre au cas d'un promoteur « fort » pour lequel les sites de fixation des deux facteurs transcriptionnels présentent une dégénérescence de séquence les rendant ainsi peu affins pour chacun des deux facteurs transcriptionnels. Par conséquent, l'action répressive n'est obtenue que si ces deux facteurs transcriptionnels agissent de manière coopérative entre eux, afin d'atteindre un seuil de répression suffisant.
- Dans le cas d'une porte d'entrée *XOR* (*excluded or*), un des 2 facteurs transcriptionnels est suffisant pour activer l'unité transcriptionnelle, ceci étant dû à la présence de sites de fixation affins et agissant de manière indépendante pour chacun des deux facteurs transcriptionnels. De plus, la présence de sites de fixation additionnels de moindre affinité entraîne la répression de l'unité

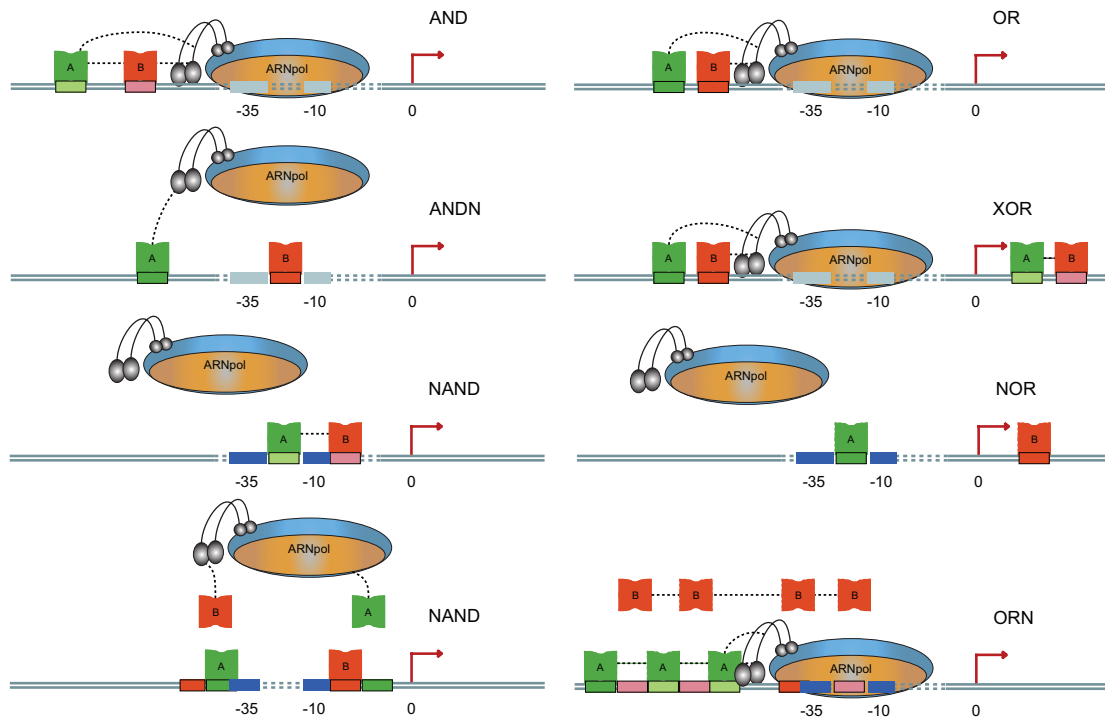
transcriptionnelle suite à l'augmentation de la concentration des 2 facteurs transcriptionnels, agissant alors de façon coopérative.

- Dans le cas d'une porte d'entrée *EQU (equals)*, l'unité transcriptionnelle est activée si et seulement si les 2 facteurs transcriptionnels se trouvent présent sous forme active et à concentration équivalente. Ce modèle s'applique dans le cas où 2 facteurs transcriptionnels possèdent chacun un site de fixation à forte affinité interférant dans l'activité transcriptionnelle d'un promoteur fort. De plus, dans ce modèle, les 2 facteurs transcriptionnels présentent également une activité d'antirépresseur sur l'autre site de fixation, menant ainsi à une levée de répression à fortes concentrations.

- Dans le cas d'une porte d'entrée *ANDN (and not)*, un des facteurs transcriptionnels agit en tant qu'activateur alors que l'autre joue le rôle de répresseur. Ainsi, l'unité transcriptionnelle ne peut être activée que si l'activateur se trouve en forte concentration sous forme active et que le répresseur est faiblement présent.

- Dans le cas d'une porte d'entrée *NOR (not or)*, la présence d'un des 2 facteurs transcriptionnels est suffisante pour réprimer l'unité transcriptionnelle, ceci étant dû à la présence de 2 sites de fixation affins et agissant de manière indépendante.

- Dans le cas d'une porte d'entrée *ORN (or not)*, un facteur transcriptionnel (activateur) entre en compétition avec un second facteur transcriptionnel (inhibiteur) pour leur site de fixation respectif. Ainsi l'unité transcriptionnelle ne peut être activé que si l'activateur se trouve sous forme active à forte concentration ou si l'inhibiteur se trouve en faible concentration.



FTs		Activité de l'unité transcriptionnelle							
C_A	C_B	AND	ANDN	NAND	EQU	OR	XOR	NOR	ORN
Faible	Faible	Off	Off	On	On	Off	Off	On	On
Faible	Élevée	Off	Off	On	Off	On	On	Off	Off
Élevée	Faible	Off	On	On	Off	On	On	Off	On
Élevée	Élevée	On	Off	Off	On	On	Off	On	On

Figure 6. Logique booléenne appliquée aux mécanismes de régulation transcriptionnelle. Les différentes portes d'entrée résultant dans l'action de deux facteurs transcriptionnels (TFs) aux concentrations C_A et C_B combinent plusieurs signaux entrant en un seul et unique signal sortant (transcription, ON; pas de transcription, OFF). La couleur des différentes boîtes le long de la séquence d'ADN indique la force de l'interaction (rose pâle, bleu pâle et vert pâle indiquent des forces d'affinité plus faibles que rouge, bleu et vert foncés).

C- RÉGULATEURS GLOBAUX

1- Régulation et organisation de la structure du génome

La régulation fait également intervenir des mécanismes influençant l'organisation globale du chromosome bactérien. De tels mécanismes sous-entendent en général un mécanisme de régulation beaucoup plus global que ceux décrits ci haut et permettent donc de regrouper simultanément une grande quantité d'unités transcriptionnelles. Ce processus fait intervenir une classe particulière de protéines, les nucléo-protéines (ou protéines de type histone). Parmi ces protéines se trouvent HU, Fis, H-NS, StpA, IHf, Lrp et Dps. En plus de leur rôle de régulateurs dans la topologie de l'ADN, ces protéines peuvent également imposer un profil transcriptionnel bien déterminé.

2- Leucine responsive regulatory protein, Lrp

a) Généralités

Lrp est une protéine basique (pI de 9,2) de 18,8 kDa qui se lie à l'ADN sous forme de dimère (Calvo & Matthews, 1994, Newman & Lin, 1995). Cette protéine fait partie de la famille des protéines régulatrices AsnC et constitue une protéine fortement conservée parmi l'ensemble des eubactéries et des archéobactéries (Yokoyama *et al.*, 2006).

Elle contrôle l'expression d'un ensemble de plusieurs gènes, dit le régulon Lrp, qui sont impliqués dans divers processus biologiques tels que la synthèse et la dégradation des acides aminés, le transport de petites molécules, l'osmolarité ou encore d'autres fonctions biologiques touchant notamment au métabolisme bactérien ou la synthèse de pili (Newman & Lin, 1995). Son action est globale dans la mesure où une délétion de *lrp* affecte, directement ou non, la transcription de 10% des gènes chez *E. coli* (Cho *et al.*, 2008).

Lrp ne présente pas de séquence cible bien identifiée sur de l'ADN. En effet, l'utilisation de la méthode Systematic Evolution of Ligands by EXponential enrichment (SELEX), une méthode de sélection *in vitro* des oligonucléotides capables de fixer une protéine donnée, a identifié la séquence (C/T)AG(A/C/T)A(A/T)ATT(A/T)T(A/G/T)CT(A/G) comme motif consensus reconnu par Lrp. Cette séquence n'est toutefois pas toujours suffisante pour prédire un site de reconnaissance pour Lrp dans la mesure où la présence d'une partie de ce motif GN₂₋₃TTT (pour lequel N correspond à n'importe quel nucléotide) et/ou la présence de région riche en A/T suffirait à la présence de Lrp sur l'ADN (Cui *et al.*, 1995, Yokoyama *et al.*, 2009). Récemment, une étude a également montré qu'un autre motif pourrait également être impliqué dans la fixation de Lrp sur l'ADN (Kawamura *et al.*, 2011).

b) Structure de Lrp

Lrp consiste en deux domaines distincts. Cette protéine possède ainsi un domaine amino-terminal constitué d'un motif Hélice-Tour-Hélice (HTH) impliqué dans l'interaction avec l'ADN et un second domaine effecteur, carboxy-terminal, important dans la régulation de l'activité transcriptionnelle (de los Rios & Perona, 2007).

Lrp forme également des structures oligomériques allant du dimère à très faible concentration, jusqu'à l'hexadécamère à des concentrations de l'ordre du micromolaire. La forme octamérique (constituée de quatre dimères) est la forme préférentielle et celle rencontrées en présence de fortes concentrations en leucine (Figure 7) (Chen & Calvo, 2002, Chen *et al.*, 2001b). Chaque sous-unité présente en effet un site à forte affinité pour la leucine qui favorise la dimérisation et un second de moindre affinité qui favorise le passage d'une forme hexadécamérique à une forme octamérique (Chen & Calvo, 2002).

L'activité de Lrp en tant que facteur transcriptionnel dépend également de la présence d'acides aminés aliphatiques (leucine, alanine, isoleucine et valine),

jouant alors le rôle de cofacteur. Dans certains cas, la présence d'acides aminés dans le milieu environnemental promeut l'action de Lrp; dans d'autres cas, sa présence interfère avec l'action de Lrp ou n'a simplement aucune influence (Chen *et al.*, 2001a).

Les mécanismes de régulation par Lrp interviennent généralement lors de la transition entre une condition environnementale de croissance riche en nutriments et une situation de carence alimentaire, comme dans le cas du passage entre la phase exponentielle et l'entrée en phase stationnaire (Landgraf *et al.*, 1996). De plus, l'induction de l'expression de *lrp* coïncide avec la synthèse de l'alarmone ppGpp, lui-même impliqué dans les processus de réponse au stress (Landgraf *et al.*, 1996).

Du fait de ses différents degrés d'oligomérisation, de sa concentration relativement élevée dans la cellule (de l'ordre de 6000 molécules) et de la grande variété de séquences interagissant avec cette protéine, Lrp est également considéré comme une nucléoprotéine sur le même principe que le rôle joué par H-NS ou par Fis (D'Ari *et al.*, 1993). Lrp est également impliqué dans la régulation de facteurs jouant un rôle important dans l'établissement de la pathogénèse. Ainsi, la transcription de la plupart des opérons fimbriaires est régulée par Lrp (Casadesús & Low, 2006, Low & Casadesús, 2008). De plus, une délétion de *lrp* provoque un défaut important dans l'adhésion aux tissus animaux (Baek *et al.*, 2009).

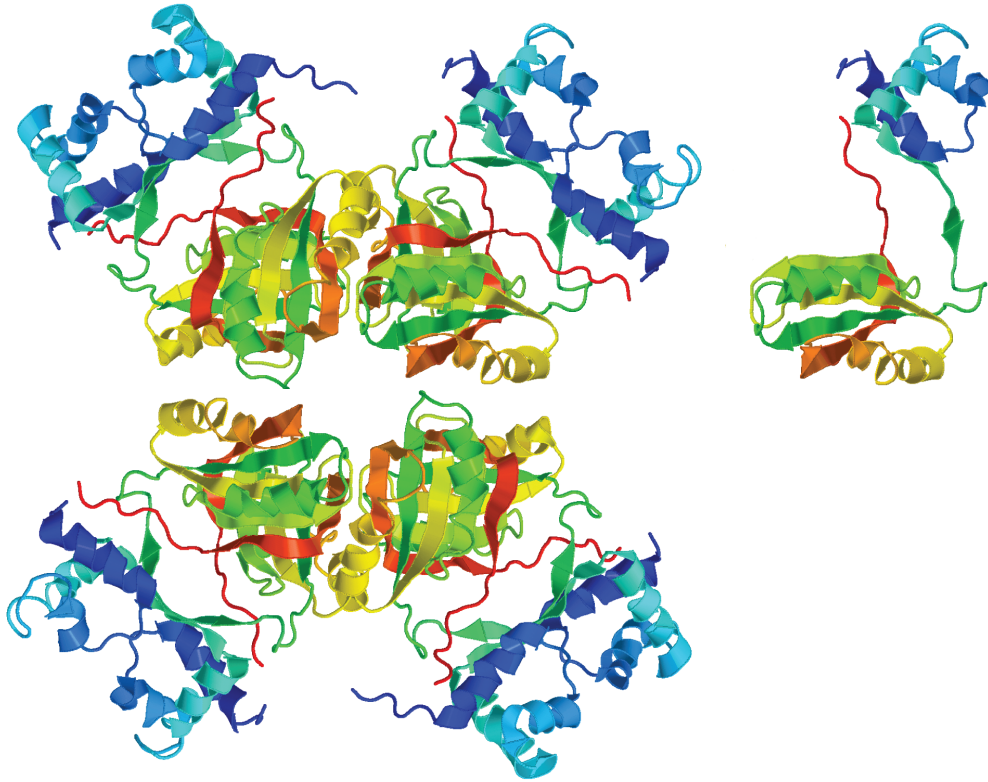


Figure 7. Représentation tridimensionnelle de la structure de Lrp sous forme octamérique (gauche) et monomérique (droite). La double hélice d'ADN s'enroule autour de Lrp sous forme octamérique en interagissant avec chacun des domaines aminoterminaux de Lrp (représentés en dégradé de bleu). Adapté de de los Rios & Perona (2007).

3- Répression catabolique et complexe CRP-AMPc

La protéine réceptrice à l'AMPc (*CRP, cAMP receptor protein*) est un des régulateurs globaux les mieux caractérisés chez *E. coli*. Via CRP, l'AMPc est un signal impliqué dans le contrôle global des gènes chez cette espèce bactérienne. Initialement, ce signal a été décrit comme un cofacteur activant le processus de répression catabolique (Balsalobre *et al.*, 2006, Forsman *et al.*, 1992). Depuis, son rôle a été étudié de façon intensive. En premier, il a été démontré que la présence de glucose inhibe l'expression des gènes codant les enzymes nécessaires dans le métabolisme d'autres sources de carbone (Balsalobre *et al.*, 2006). En effet, l'incorporation de glucose provoque une diminution de la concentration en protéine CRP ainsi que celle en AMPc disponible. Par conséquent, les variations dans le niveau du complexe CRP-AMPc contribue à la régulation précise de plusieurs opérons dont le contrôle est nécessaire pour le métabolisme bactérien. Globalement, le complexe CRP-AMPc contrôle l'expression de plusieurs centaines de gènes, bien que seule une partie d'entre eux est directement impliqué dans le catabolisme bactérien.

4- H-NS (Histone-Like Nucleoid Structuring Protein)

Avec environ 20,000 copies dans le cytosol bactérien, H-NS est une des protéines les plus abondantes (Dorman, 2009a, Dorman, 2009b). Elle agit principalement sur la structure de l'ADN en le compactant et en altérant sa superhélicité par ajout de supertours négatifs. Cette protéine est ainsi impliquée dans l'organisation de la structure du génome bactérien.

À ce jour, aucune séquence consensus à proprement parlé n'a été identifiée pour H-NS. Cette protéine semble plutôt interagir avec les régions courbées de l'ADN ou celles riches en A/T (Dorman & Kane, 2009). En plus de sa fonction dans l'organisation du génome, H-NS est également impliquée dans la régulation de l'expression des gènes en réprimant la transcription à basse

température. De plus, cette protéine joue un rôle important dans la répression des gènes nouvellement acquis par transfert horizontal (Dorman & Kane, 2009).

Directement ou non, H-NS régule l'expression de 5% des gènes d'*E. coli* et module la production d'environ 60 protéines en milieu faible en nutriments. Environ 30% de ces protéines sont impliquées dans la formation de l'enveloppe bactérienne et dans l'adaptation aux conditions environnementales, 20% sont des protéines impliquées dans la transcription et la traduction (Hommais *et al.*, 2001). Les conditions environnementales affectant l'expression de gènes via H-NS sont la température, le pH, la pression osmotique et la présence d'oxygène. H-NS est également impliquée dans la régulation de l'expression génétique de facteurs de virulence comme dans le cas des fimbriæ et des flagelles (thermorégulation des opérons *pap*, *daa*, *sfa*, *fim*, et de *fliC* chez *E. coli*) (Göransson *et al.*, 1990, Harel & Martin, 1999, Landini & Zehnder, 2002, Morschhauser *et al.*, 1993, Olsen *et al.*, 1998, White-Ziegler *et al.*, 2000).

Le mécanisme moléculaire par lequel H-NS régule la transcription n'est pas entièrement connu. En revanche, on considère que H-NS affecte l'expression des gènes en modifiant le degré de surenroulement requis lors de la transcription, ou inhibe directement la transcription en interférant avec l'accessibilité de la séquence promotrice (Dorman, 2009c, Dorman & Corcoran, 2009, Dorman & Kane, 2009). Par conséquent, en modifiant la topologie de l'ADN, cette protéine affecterait l'expression, principalement en altérant la topologie de l'ADN requise pour la transcription.

5- La Dam méthyltransférase

La Dam méthyltransférase est la première méthyltransférase pour laquelle un rôle ayant des répercussions dans la physiologie bactérienne a été mis en évidence. Cette protéine est impliquée dans le transfert d'un groupement méthyle provenant de la S-adénosyl-méthionine à un résidu adénine compris dans un motif nucléotidique 5'-GATC-3' (Wion & Casadesús, 2006). La

méthylation survient assez rapidement après la réplication bactérienne bien que cela ne soit pas immédiat. De ce fait, le passage de la fourche de réplication laisse transitoirement les sites GATC sous forme hémi-méthylée. De plus, la Dam méthyltransférase méthyle indifféremment les sites non méthylés ou hémi-méthylés avec des efficacités similaires.

Chez *E. coli*, l'origine de réplication *oriC* contient 11 sites GATC dans une région de 254 pb, soit une densité 10 fois supérieures à celle due au simple hasard. Ceci va dans le sens d'un rôle de la protéine Dam dans les processus de réplication bactérienne (Wion & Casadesús, 2006). En effet, la réplication est initiée à partir du moment où la protéine initiateur DnaA se fixe sur la région *oriC* et sépare les 2 brins de la double hélice d'ADN, processus qui n'est possible que si les sites GATC sont complètement méthylés (Boye *et al.*, 2000). Après passage de la fourche de réplication, les sites GATC nouvellement formés restent dans un état hémi-méthylée durant environ un tiers de la durée du cycle cellulaire bactérien. En effet, la présence de la protéine SeqA qui présente une forte affinité pour les sites GATC hémi-méthylés réprime l'action de la Dam méthyltransférase pour l'*oriC* en séquestrant ces sites de l'ADN. L'origine de réplication reste donc sous forme hémi-méthylée tant que la protéine SeqA est présente et tout nouveau cycle de réplication est retardé (Campbell & Kleckner, 1990). La Dam méthyltransférase joue également un rôle important dans l'expression du gène codant DnaA. En effet un des promoteurs de *dnaA* (situé à moins de 50 kpb d'*oriC*) se trouve sous forme active si et seulement si trois de ses sites GATC se trouvent sous forme méthylée (Campbell & Kleckner, 1990, Riber & Lobner-Olesen, 2005). De ce fait, l'hémi-méthylation d'*oriC* représente un signal en soit qui provoque la séquestration du chromosome des 2 cellules filles.

Chez les souches pathogènes, la Dam méthyltransférase est impliquée dans les processus de contrôle des facteurs de virulence (Low *et al.*, 2001, Wion & Casadesús, 2006, Hernday *et al.*, 2002). Ainsi, une mutation par délétion du

gène *dam* chez *Salmonella* Enterica entraîne une atténuation de la virulence dans un modèle murin d'infection animale ainsi que plusieurs défauts dans l'expression de gènes de virulence, dans la stabilité de la structure de l'enveloppe, un relargage anormal de protéines et de vésicules membranaires, une sensibilité accrue en présence de sels biliaires, ou encore une expression élevée de l'opéron fimbriaire *stdABC*, résultant dans l'exportation de protéines fortement immunogènes (Heithoff *et al.*, 1999, Garcia-Del Portillo *et al.*, 1999, Pucciarelli *et al.*, 2002, Giacomodonato *et al.*, 2004). De même, chez *Yersinia pseudotuberculosis* et *Haemophilus influenzae*, l'inactivation de la Dam méthyltransférase provoque des effets similaires à ceux observés chez *S. Enterica*, avec notamment la perte de la virulence et une réduction dans la capacité à envahir les cellules épithéliales (Watson *et al.*, 2004, Taylor *et al.*, 2005, Wion & Casadesús, 2006).

D- FACTEURS DE VIRULENCE

1- Généralités

Tout organisme pathogène a en commun la nécessité de réaliser diverses étapes afin d'exercer son pouvoir pathogène. Ces différentes étapes correspondent à :

- (i) la colonisation des surfaces (muqueuses ou cutanées) ;
- (ii) le franchissement des muqueuses (ou de la peau) pour pénétrer dans la circulation sanguine ou dans les tissus ;
- (iii) la résistance aux défenses internes de l'organisme dans le sang et les tissus ;
- (iv) la production d'un effet toxique, responsable des lésions tissulaires et organiques et, *in fine*, de l'apparition des signes cliniques (Kaper et al., 2004).

La virulence se définit comme la capacité d'un organisme à causer la maladie dans un hôte donné. Chez *E. coli*, la virulence résulte en l'effet cumulatif de plusieurs facteurs de virulence. Ainsi, la virulence d'une souche donnée peut être positivement corrélée avec le nombre de gènes de virulence portés (Kaper et al., 2004). Dans le cas des souches pathogènes intestinales, le profile de virulence dépend de la présence de gènes bien définis. En revanche, celui des souches ExPEC découle de la combinaison d'une multitude de facteurs variés (Kaper et al., 2004, Smith et al., 2007). Certains d'entre eux sont néanmoins plus importants que d'autres et des mécanismes communs peuvent être retrouvés d'une souche à l'autre. Par conséquent, la pathogénèse associée aux ExPEC varie grandement d'une souche à l'autre.

Une souche d'*E. coli* est considérée comme ExPEC si elle possède au moins deux des marqueurs de virulence suivants : *papA* (gène codant la sous-

unité structurale des fimbriæ P) et/ou *papC* (gène codant une protéine de membrane impliquée dans l'assemblage des fimbriæ P), *sfa/foc* (gènes codant la sous-unité structurale des fimbriæ S et F1C, respectivement), *afa/dra* (gènes codant l'adhésine afimbriaire Afa), *kpsMT II* (gène impliqué dans la synthèse de la capsule polysaccharidique de groupe 2), et *iutA* (gène codant l'aérobactine) (Johnson & Russo, 2005). D'autres facteurs peuvent cependant être également associés à la virulence des souches ExPEC.

Les souches ExPEC produisent aussi des toxines qui contribuent à la pathogénie en causant des nécroses, des lyses cellulaires ou des dysfonctionnement au site de l'infection. Une des toxines cytotoxiques les mieux caractérisées est l' α -hémolysine (Johnson, 1991, Wiles *et al.*, 2008). Cette toxine appartenant à la famille des toxines RTX (Repeats in ToXins (Hacker *et al.*, 1997)) cause la lyse des érythrocytes, des leucocytes et des tubes rénaux de mammifère. Sa synthèse est sous le contrôle de l'opéron *hlyCABD* également présent chez certaines souches commensales mais significativement moins associé à ces souches qu'il ne l'est pour les souches ExPEC (Nagy *et al.*, 2006).

Du fait de la faible concentration en fer et de son importance dans divers processus biologiques, la présence de mécanismes d'acquisition du fer est également à considérer comme un facteur influençant grandement le pouvoir pathogène des souches ExPEC (Nielubowicz & Mobley, 2010). Un des systèmes les plus efficaces correspond au sidérophore permettant de chélater les ions ferriques Fe^{3+} et de les acheminer vers le cytoplasme bactérien. Chez les souches UPEC, les systèmes d'acquisition du fer et impliqués dans la pathogénèse incluent l'aérobactine, l'entérobactine, la yersiniabactine, la salmochéline (Nielubowicz & Mobley, 2010). Ils ne présentent pas tous la même capacité à chélater le fer. Des études ont montré que l'aérobactine, associée au cas de cystite et de pyélonéphrite, se caractérise par une plus grande efficacité comparée à l'entérobactine présente chez la majorité des souches *E. coli*.

2- L'adhésion bactérienne et son rôle dans la pathogénèse

L'adhérence aux cellules de l'hôte implique la synthèse de molécules d'adhésion qui jouent un rôle important dans les différentes phases de l'infection. En effet, les adhésines contribuent à la virulence en favorisant l'adhérence et la colonisation sur la surface des cellules de l'hôte. Se faisant, les bactéries évitent d'être balayées par les fluides corporels. De plus, le contact intime entre bactéries et cellules de l'hôte permet une libération et une action plus efficace des molécules impliquées dans les étapes subséquentes de l'infection.

On retrouve chez *E. coli* une grande variété d'adhésines qui diffèrent entre elles de part leur structure et de part leur spécificité de récepteur. De plus, une même souche présente la capacité de synthétiser plus d'une adhésine, avec pour chacune une spécificité de récepteur lui permettant de coloniser différentes niches. On distingue en général les adhésines fimbriaires correspondant à des complexes multi-protéiques filamenteux des adhésines afimbriaires qui se distinguent par l'absence de structures visibles autour de la cellule bactérienne.

3- Adhésines fimbriaires

La base de l'architecture des fimbriæ classiques et des fibrilles est une protéine d'une masse moléculaire comprise entre 14 et 22 kDa (Jones *et al.*, 1995, Kuehn *et al.*, 1992, Waksman & Hultgren, 2009). L'assemblage des différentes sous-unités se réalise en une hélice dextrogyre de plusieurs centaines d'exemplaires, formant une structure cylindrique typique aux fimbriæ et aux fibrilles. Cette protéine, la sous-unité majeure (ou fréquemment protéine A), est responsable des propriétés antigéniques des fimbriæ du fait de l'importance en taille de la structure fimbriaire et de sa présence en grande quantité (Kuehn *et al.*, 1992). Dans certains cas, différents membres d'une même

familles de fimbriæ présentent des variations antigéniques, dû à des variations de séquence de la sous-unité majeure (van der Woude & Bäumlér, 2004). Les fimbriæ/fibrilles comprennent également diverses sous-unités mineures, présentes en quelques exemplaires, voire un seul (Waksman & Hultgren, 2009). Ainsi, la base des fimbriæ/fibrilles est attachée à une protéine spécifique incorporée dans la couche externe de la membrane bactérienne, dénommée protéine d'ancrage (ou fréquemment protéine C) (Dodson *et al.*, 1993, Hultgren *et al.*, 1993, Hultgren *et al.*, 1991, Lindberg *et al.*, 1989). Le rôle de certaines autres sous-unités mineures est d'assurer la bonne fixation des fimbriæ/fibrilles sur la protéine d'ancrage (Waksman & Hultgren, 2009). Sur certains fimbriæ, quelques sous-unités mineures peuvent également former une structure courte (10 à 80 nm) de type fibrillaire qui coiffe les fimbriæ (= « tip fibrillum ») (Kuehn *et al.*, 1992). Parmi les sous-unités mineures se trouve la protéine responsable des propriétés spécifiques d'adhésion des fimbriæ/fibrilles : il s'agit de l'adhésine proprement dite (ou fréquemment protéine G) (Waksman & Hultgren, 2009). Cette sous-unité est ainsi responsable de la spécificité de l'interaction avec le récepteur présent sur la cellule eucaryote. L'adhésine G comprend deux domaines : l'extrémité carboxyterminal qui assure l'incorporation de cette sous-unité sur les fimbriæ/fibrilles et l'extrémité aminoterminal qui participe quant à elle à la reconnaissance et à l'interaction avec le récepteur (Hung *et al.*, 2002, Merckel *et al.*, 2003, Dodson *et al.*, 2001). Dans certaines familles de fimbriæ, des variations existent dans la séquence de l'adhésine G, amenant la synthèse de fimbriæ identiques (même sous-unité majeure A), mais montrant des variations dans la spécificité du récepteur (Waksman & Hultgren, 2009). Le plus fréquemment cette adhésine est placée à l'extrémité des fimbriæ/fibrilles en un seul exemplaire. Elle peut cependant être présente en quelques copies intercalées dans les fimbriæ/fibrilles, entre les sous-unités majeures.

4- Récepteurs

Sur la surface de l'hôte, les récepteurs correspondent en général à des entités carbohydratées exposées à la membrane cytoplasmique des cellules eucaryotes. Cette molécule est en général liée de manière covalente à une chaîne polypeptidique. Ces récepteurs représentent ainsi le premier maillon d'une chaîne d'activation métabolique à l'intérieur de la cellule eucaryote, qui fixe des messagers hormonaux de l'hôte. Par conséquent, l'activation de la cascade intracellulaire qui s'ensuit résulte en des modifications biochimiques ou structurales, que la bactérie parasite indirectement.

La première base de la spécificité de récepteur des fimbriæ/fibrilles classiques consiste en la structure carbohydratée la plus simple qui est capable d'inhiber l'adhésion bactérienne : des hexoses, comme le D-mannose pour les fimbriæ de type 1 ; des méthylpentoses ; des acétylhexosamines ; de l'acide sialique. Cette première base de la spécificité de récepteur définit les familles de fimbriæ/fibrilles (Westerlund-Wikstrom & Korhonen, 2005).

À l'intérieur de chaque famille, des variations dans l'identité exacte du récepteur peuvent également exister. Ainsi, la famille des fimbriæ de type P montre comme première base de spécificité de récepteur l'entité α -D-galactopyranosyl-(1-4)- β -D-galactopyranoside, ou Gal (α 1-4) Gal, présente dans des glycolipides faisant partie du groupe sanguin P des érythrocytes humains. Toutefois, trois variants d'adhésines ont été décrits à ce jour: Gal-(α 1-4)-Gal-(β 1-4)-Glucose, ou globotriaose, ou Gb03 pour l'adhésine PapG1 ; GalNAc-(β 1-3)-Gal-(α 1-4)-Gal-(β 1-4)-Glucose, ou globotétraose, ou Gb04 pour l'adhésine PapG2 ; GalNAc-(α 1-3)-GalNAc-(β 1-3)-Gal-(α 1-4)-Gal-(β 1-4)-Glucose, ou globopentaose, ou Gb05 pour l'adhésine PapG3 (Lane & Mobley, 2007).

Cette spécificité de récepteur est souvent reflétée dans la situation *in vivo*. En effet, les récepteurs Gb03 et Gb04 reconnus spécifiquement par les adhésines

PapG1 et PapG2 sont présents surtout sur les cellules épithéliales du tractus urinaire humain et peu sur celles des animaux, tandis que le récepteur GbO5 spécifique pour l'adhésine PapG3 est présente essentiellement sur les cellules des tractus urinaire des animaux, particulièrement les chiens et chats (Smith et al., 2007, Gyles & Fairbrother, 2010). Ainsi, la spécificité de l'interaction adhésines/récepteurs est en grande partie responsable de la spécificité d'hôtes et de tissus des bactéries pathogènes.

5- Biogenèse des adhésines bactériennes

La majorité des fimbriæ/fibrilles classiques sont codés par des ensembles de 4 à 11 gènes regroupés soit sur le chromosome bactérien (parfois un îlot de pathogénicité), soit sur un plasmide. Pour certains fimbriæ, ces gènes peuvent avoir une localisation chromosomique ou plasmidique selon le variant antigénique et/ou la souche qui les produit.

Divers mécanismes de synthèse des fimbriæ/fibrilles ont été identifiés, mais celui appelé « chaperone-usher » est le mieux décrit. Dans ce mécanisme, la biogenèse des fimbriæ/fibrilles implique la traversée de la membrane cytoplasmique des différentes composantes sous-unitaires par un système général de sécrétion des protéines, le système Sec (Figure 8) (Soto *et al.*, 1998).

Dans le périplasme, les différentes sous-unités forment un complexe avec les protéines chaperons, qui montrent une structure générale de type immunoglobuline, afin d'éviter leur auto-assemblage précoce (Waksman & Hultgren, 2009). Les complexes migrent vers la protéine d'ancrage C incorporée dans la membrane externe, qui subit un changement de structure avec ouverture d'un pore (« placier ») au travers duquel les différentes sous-unités passent, après libération de la protéine chaperon (Waksman & Hultgren, 2009). L'assemblage des différentes composantes des fimbriæ suit alors un ordre précis qui repose sur l'affinité des protéines chaperons pour les différentes sous-unités

et sur celle de la protéine d'ancrage pour les complexes protéines chaperon/sous-unité. Cet ordre fait en sorte que la construction des fimbriæ sur la protéine d'ancrage commence par son extrémité distale et se termine par sa base. Les fimbriæ naissants sont alors propulsés progressivement vers le milieu extérieur par l'ajout des sous-unités sous la partie des fimbriæ déjà formée. L'assemblage des sous-unités se fait sans dépense d'énergie suite à des changements conformationnels thermodynamiquement favorables aux interactions entre différentes sous-unités.

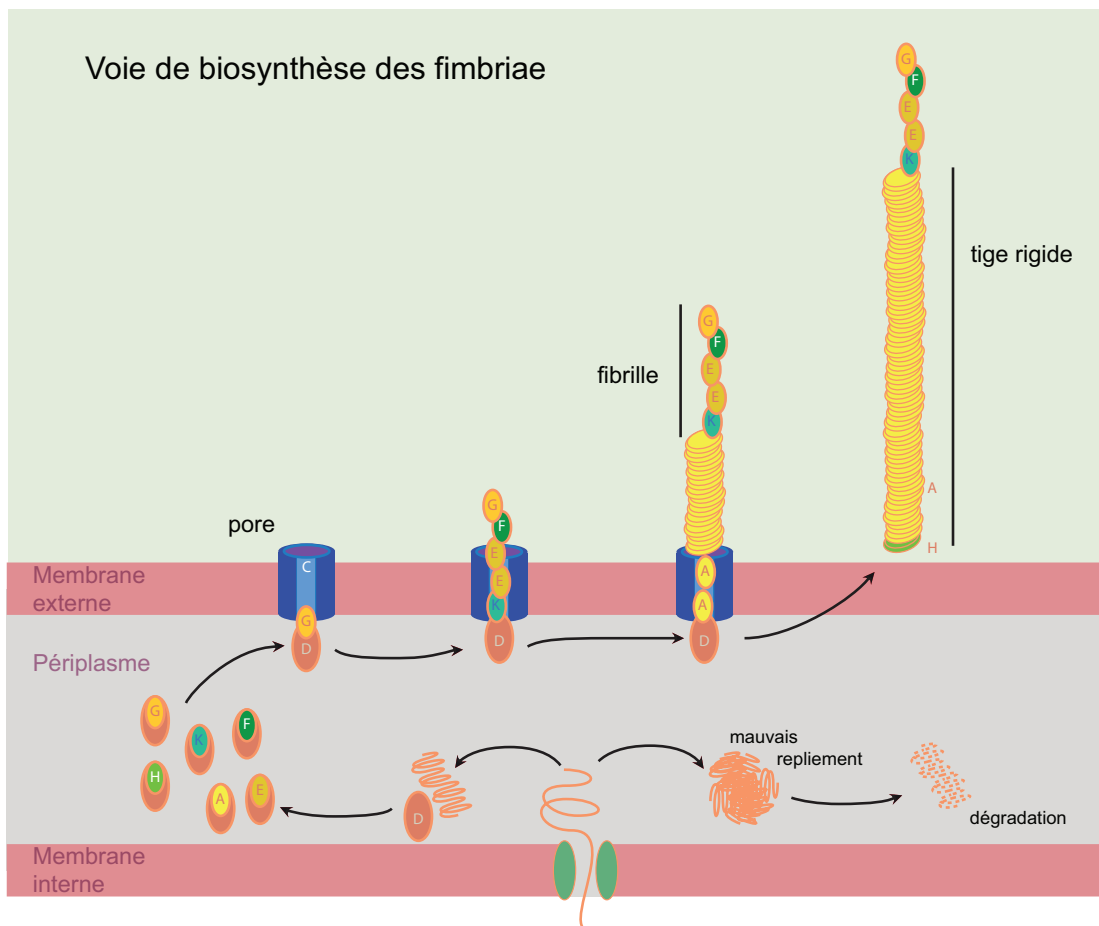


Figure 8. Représentation schématique de la voie de synthèse des fimbriæ de type « chaperone-usher ».

Dans le cas des fimbriæ de type P, la protéine chaperon PapD est impliquée dans l'assemblage des fimbriæ selon une architecture de type tige (Waksman & Hultgren, 2009). Son rôle consiste à former des complexes de pré-assemblage périplasmiques avec chacune des sous-unités fimbriaires avant leur incorporation dans la structure du fimbriæ. La concentration de chacun des types de sous-unités est un facteur important pour la régulation et la définition de la taille de la pointe fibrillaire et de la tige des fimbriæ. La translocation initiale des sous-unités des fimbriæ Pap à travers la membrane cytoplasmique se fait par l'intermédiaire de la machinerie de sécrétion Sec-dépendante. Les sous-unités naissantes sont retenues à l'intérieur de la membrane cytoplasmique par leur portion carboxyterminale, via une interaction hydrophobe (Waksman & Hultgren, 2009). En présence de PapD, les sous-unités sont envoyées dans l'espace périplasmique sous la forme de complexes de préassemblage protéine chaperon/sous-unités. Ces derniers sont alors dirigés vers le placier PapC localisé dans la membrane externe pour l'assemblage.

PapC est constitué d'au moins six sous-unités et forme un pore par lequel les sous-unités des fimbriæ sont transloquées à travers la membrane externe. En théorie, la pointe fibrillaire linéaire d'une largeur de 2 nm passe facilement par le canal de 2-3 nm de diamètre du placier (Li *et al.*, 2004, Forsman *et al.*, 1989, So & Thanassi, 2006). En revanche, la tige des fimbriæ, d'une largeur de 6-8 nm nécessite d'être sous forme déroulée (représentant alors une structure linéaire de 2 nm de diamètre) (Waksman & Hultgren, 2009). La tige des fimbriæ serait ainsi transloquée à travers la membrane externe sous cette forme linéaire pour ensuite adopter sa conformation hélicoïdale finale seulement après avoir atteint la surface externe. Ce phénomène serait impliqué dans le processus à l'origine de la croissance de cette structure à la surface bactérienne. Le placier joue de ce fait un rôle plus actif dans l'assemblage des fimbriæ de type P que la simple fonction de pore. Il est communément admis que PapC reconnaît de façon différentielle les complexes protéines chaperon/sous-unité selon leur position finale dans le fimbriæ. Ainsi, le complexe protéine chaperon/adhésine se lie plus fortement et

plus rapidement au placier (Waksman & Hultgren, 2009). La cinétique de séparation des complexes adhésine/adhésine au placier serait un facteur déterminant dans la localisation de l'adhésine à la pointe. De plus, pour tous les complexes sous-unité/sous-unité, les taux de dissociation du placier sont faibles, suggérant qu'après la formation d'un complexe avec le placier, la sous-unité est destinée à l'assemblage des fimbriæ. La complémentarité de surface sous-unité/sous-unité jouerait également un rôle clé dans l'ordre relatif à l'incorporation des sous-unités dans les fimbriæ en court de synthèse. Les protéines adaptatrices PapF et PapK sont requises pour l'initiation de la pointe de la fibrille et de la tige des fimbriæ. PapK régulerait la longueur de la fibrille. Quant à PapJ, elle serait une protéine co-chaperon (Waksman & Hultgren, 2009).

6- Fimbriæ de type 1

Les fimbriæ de type 1 sont relativement communs et représentent le type de fimbriæ les plus fréquemment retrouvés au sein de la famille des entérobactéries. Ainsi 70 à 80% des souches sauvages d'*E. coli* possèdent les gènes impliqués dans la synthèse de ces fimbriæ. Ce type de fimbriæ est impliqué dans l'attachement des bactéries sur différents types de surface, biotiques ou abiotiques et constituent un élément important dans l'étape initiale du développement du biofilm (Rosen *et al.*, 2007). De plus, les fimbriæ de type 1 jouent également un rôle important dans la virulence des souches UPEC (Müller *et al.*, 2009). En effet, ils favorisent l'adhésion et l'invasion bactérienne dans la vessie en se fixant aux résidus mannose présents à la surface des cellules uroépithéliales (Figure 9). Des études suggèrent également que les fimbriæ de type 1 sont requis dans l'élaboration de communautés bactériennes intracellulaires, suite à l'invasion des cellules uroépithéliales (Justice *et al.*, 2006, Rosen *et al.*, 2007).

La synthèse des fimbriæ de type 1 est sous le contrôle de l'expression de l'opéron *fim*. L'opéron *fim* est composé de deux unités transcriptionnelles indépendantes, l'une codant les recombinases spécifiques FimB et FimE, l'autre

correspondant à un opéron polycistronique à l'origine de la synthèse des composants structuraux (FimA, FimF, FimG et FimH) et du système d'assemblage des pili (FimC et FimD). L'expression de cet opéron dépend de l'orientation d'une séquence nucléotidique de 314 pb contenant le promoteur de *fimA* et entourée de chaque côté par deux séquences inversées répétées de 9 pb. Si l'élément chromosomique contenant le promoteur de *fimA* est dans la bonne orientation, alors la synthèse des fimbriæ peut avoir lieu; dans le cas contraire, l'orientation du promoteur rend la transcription de l'opéron *fim* impossible. L'orientation de cette séquence nucléotidique est sous la dépendance des activités enzymatiques des deux recombinaisons FimB et FimE qui, dans un cas favorise l'orientation dans les deux sens (FimB) et, dans l'autre, catalyse l'orientation uniquement vers la forme défavorisant la transcription de l'opéron *fim* (FimE) (Nielubowicz et al., 2010).

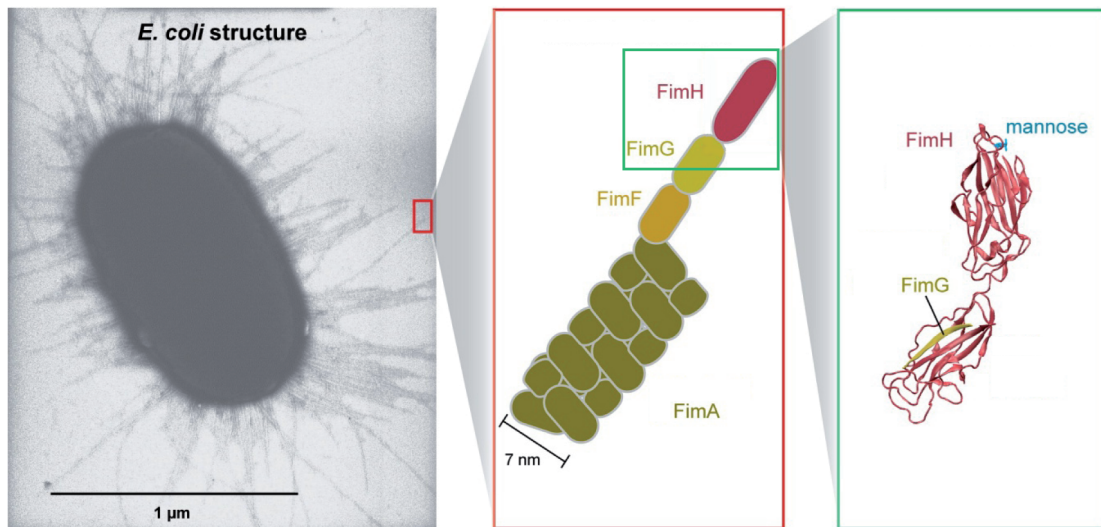


Figure 9. Structure des fimbriæ de type 1 avec notamment la structure de l'adhésine FimH liant le mannose.

7- Les fimbriæ de type P

a) Généralités

La famille des fimbriæ de type P joue un rôle important dans la colonisation des tissus de l'hôte lors d'infections du tractus urinaire supérieur (Lane & Mobley, 2007). Ces fimbriæ reconnaissent un motif Gal α (1→4)Gal présent à la surface des érythrocytes et d'autres types cellulaires dont les cellules uroépithéliales du rein et du tractus urinaire (Dodson et al., 2001, Hung et al., 2002, Lane & Mobley, 2007).

Les fimbriæ Pap sont composés d'une tige longue et rigide de 6 à 8 nm de diamètre constituée d'un homopolymère de la sous-unité majeure PapA pouvant contenir plus de 1.000 copies de cette protéine. La structure fimbriaire se termine par une fibrille courte et mince de 2 à 3 nm constituée des sous-unités PapE, PapF, PapK et de l'adhésine PapG (Waksman & Hultgren, 2009).

b) Structure et organisation génétique de l'opéron *pap*

La formation des fimbriæ P requiert les produits d'environ onze gènes impliqués dans sa biogénèse et sa régulation. Ces gènes codent les six protéines structurales (PapA, PapE, PapF, PapH, PapK, et PapG), les protéines d'assemblage (PapC, PapD et possiblement PapJ) ainsi que les protéines régulatrices PapB et PapI (Waksman & Hultgren, 2009). D'un point de vue génétique, les gènes sont organisés sous forme d'un opéron constitué de l'ensemble des gènes codant les différentes sous-unités (*papBAHCDJKEFG*), à l'exception de *papI* qui est transcrit de façon divergente par rapport au reste de l'opéron (Figure 10) (Casadesús & Low, 2006, Hernday et al., 2002). La transcription de l'opéron *pap* est sous le contrôle du promoteur principal *pBA* qui permet la transcription de l'ensemble des gènes *papBAHCDJKEFG*; la transcription de *papI* étant elle sous le contrôle de son propre promoteur *pI*.

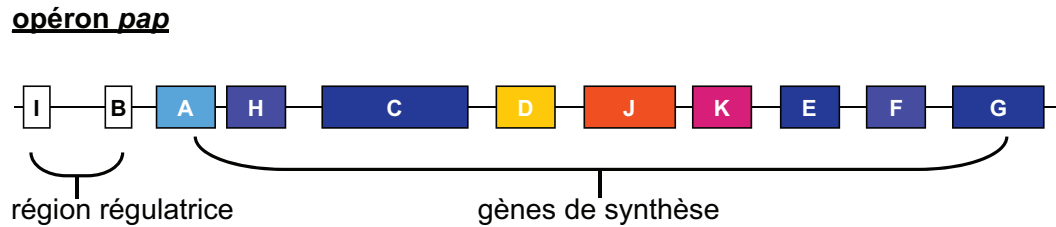


Figure 10. Représentation schématique de la structure des gènes sous forme d'un opéron, impliqué dans la synthèse de l'adhésine fimbriaire Pap. Les gènes impliqués dans la synthèse proprement dite sont représentés par des boîtes de différentes couleurs alors que ceux impliqués dans la régulation transcriptionnelle de l'opéron *pap* correspondent à des cases blanches.

La synthèse des facteurs d'adhésion est soumise à une régulation complexe en réponse à différents stimuli environnementaux. Il existe plusieurs niveaux de régulation. Le premier est transcriptionnel et contrôle l'activité du promoteur et le taux de transcription. Le second niveau de régulation est post-transcriptionnel. Baga *et al.* ont en effet montré que les taux d'ARNm de *papB* et *papA* diffèrent alors que ces 2 gènes sont co-transcrits sous forme d'un long ARNm polycistronique (Baga *et al.*, 1988). Ainsi, la présence d'une séquence terminatrice à l'extrémité 3' de *papA* augmente la stabilité de sa séquence ARNm par rapport à celle de *papB* qui lui est très rapidement dégradé par des exonucléases.

Il existe également un troisième mode de régulation post-traductionnelle qui repose sur la gestion du nombre de sous-unités acheminées dans le périplasme et dans la biosynthèse des pili (Hernday *et al.*, 2004). En effet, la protéase périplasmique DegP dégrade les agrégats de sous-unités non complexées aux protéines chaperons. Or, les souches d'*E. coli* ne possédant pas DegP subissent l'accumulation des sous-unités agrégées, ce qui est létale pour la bactérie. De plus, les altérations tridimensionnelles des sous-unités dans le périplasme jouent un rôle dans le contrôle de la production des pili P. En effet, la

Disulfide oxidoreductase A (DsbA) catalyse la formation de ponts disulfures nécessaires à l'assemblage des pili P. Or la transcription des gènes codant DegP et DsbA est activée par le système à 2 composantes CpxAR, lui-même produit en réponse à l'accumulation périplasmique des sous-unités mal conformées et mal couplées à la protéine chaperon (Hernday et al., 2004). Ainsi, des mécanismes de transduction de signal détectent l'assemblage des fimbriæ de type P. En l'absence de protéine chaperon, l'expression des sous-unités fimbriaires est toxique pour des mutants dépourvus de la protéase périplasmique DegP. Cette toxicité serait attribuable à la formation d'agrégats de sous-unités dans le périplasme, lesquels sont normalement dégradés par DegP (Hernday et al., 2004). Cette agrégation protéique dans le périplasme est détectée par au moins deux systèmes de transduction de signal. Le système à deux composantes CpxA/CpxR, pour lequel CpxA est la protéine senseur associée à la membrane et CpxR est la protéine effectrice se liant à l'ADN (Figure 11); et la voie modulatrice σ^E , pour lequel RseA et RseB régulent négativement le facteur transcriptionnel σ^E . Ces deux voies stimulent la transcription de *degP* ainsi que celle de plusieurs autres gènes codant des protéines de type chaperon. De plus, elles sont toutes deux activées par l'agrégation des sous-unités fimbriaires. Par conséquent, l'assemblage adéquat des fimbriæ de type P nécessite ainsi la présence du régulateur CpxR, lequel influence positivement l'expression du promoteur *pBA*. Par conséquent, le système à deux composantes CpxA/CpxR contrôle la réponse au stress périplasmique et l'assemblage des fimbriæ. Il est suggéré que les voies de signalisation par Cpx et par σ^E contrôlent la biogénèse des fimbriæ en régulant la production de multiples protéases périplasmiques ou autres protéines chaperons, toutes requises pour la formation des fimbriæ.

Enfin, l'ensemble de l'opéron est également régulé par un mécanisme de variation de phase qui régit la quantité de bactéries présentant cette adhésine fimbriaire à leur surface (cf. partie traitant des mécanismes de régulation par variation de phase).

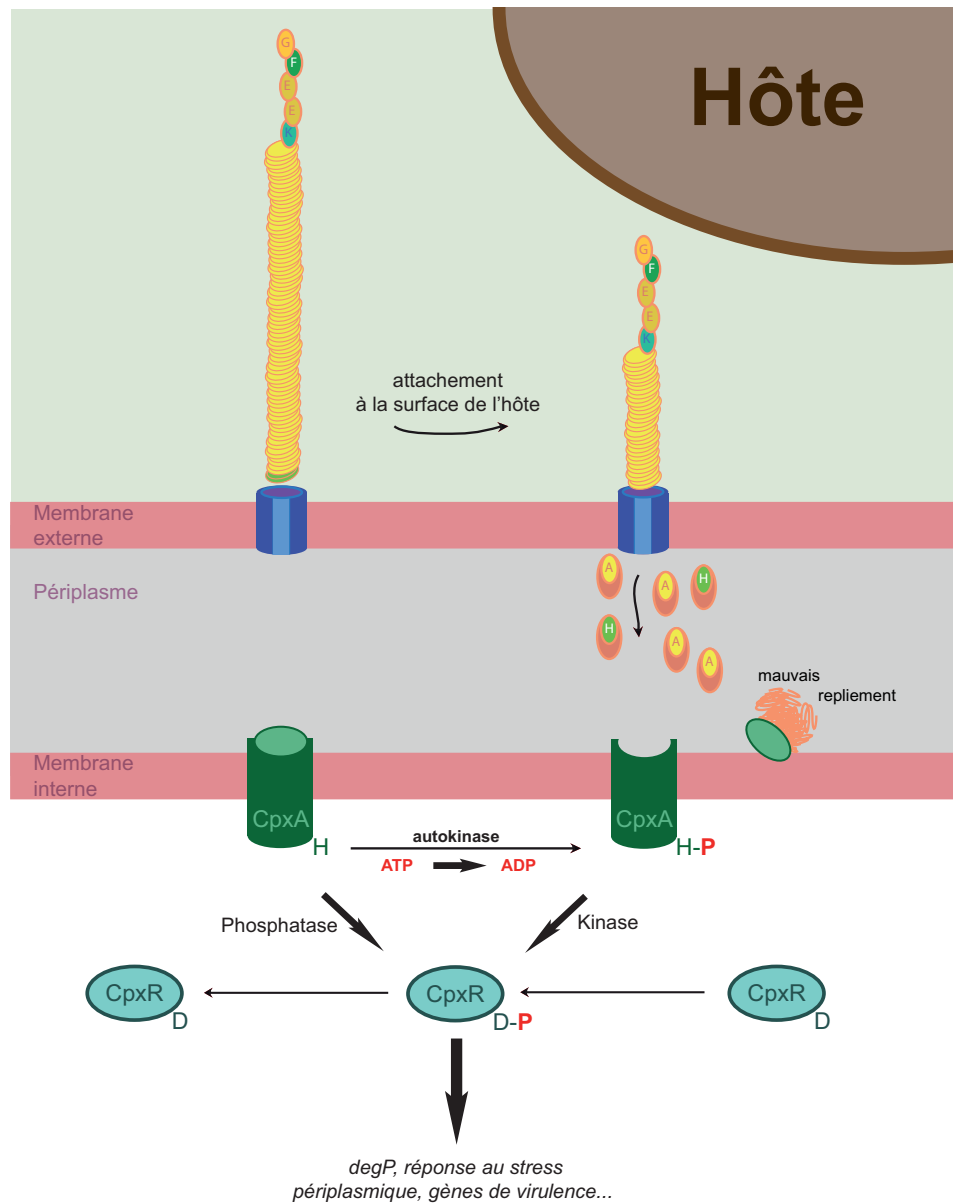


Figure 11. Régulation de la synthèse des fimbriæ Pap via le système CpxAR.

c) Effet de H-NS et de RimJ sur l'expression de l'opéron *pap*

Les fimbriæ de type P sont produits par des cellules se multipliant à 37°C mais pas à des températures plus basses telles que 25°C. En effet, H-NS joue un rôle important dans la thermorégulation de deux gènes codant les régulateurs positifs de l'expression de l'opéron fimbriaire *pap*. Des expériences de retard sur gel ont en effet montré que H-NS se fixe spécifiquement sur la région régulatrice de *pap* sur deux sites GATC et que cette fixation bloque la méthylation de ces deux sites à 23°C alors qu'aucun effet n'est observé à 37°C (White-Ziegler *et al.*, 1998, White-Ziegler *et al.*, 2000). Il en va de même pour la N-acétyltransférase RimJ qui réprime la transcription du promoteur principal *papBA* à basse température selon un mécanisme indépendant de H-NS (White-Ziegler *et al.*, 2002). De plus, H-NS est également impliquée dans la répression transcriptionnelle de l'opéron *pap* en réponse à d'autres conditions environnementales comme une forte pression osmotique, un milieu de croissance riche en nutriment ou en présence de glucose comme seule source de carbone (White-Ziegler *et al.*, 2000).

8- Les fimbriæ F165₁

a) Généralités

F165₁ est un constituant du complexe F165, lui même formé des deux fimbriæ F165₁ et F165₂ (Harel *et al.*, 1991, Harel *et al.*, 1992a, Harel *et al.*, 1995). Tout comme Pap dont il est un homologue, F165₁ se présente sous la forme d'un complexe multi-protéique rigide dont le diamètre fait 4-6 nm pour une longueur supérieure à 200 nm (Fairbrother *et al.*, 1986, Fairbrother *et al.*, 1988). Les fimbriæ F165₁ présentent de fortes homologies avec les fimbriæ P (F11) et possèdent une adhésine G de classe III, permettant la reconnaissance de l'antigène de Forssmann (GB05) (Fairbrother *et al.*, 1986, Harel *et al.*, 1992b).

Initialement, l'antigène de surface F165₁ a été mis en évidence chez des souches non-entérotoxigènes isolées de cas de diarrhées et de septicémies chez le porcelet. Les fimbriæ F165₁ sont associés à 35 % des cas de mortalité due à une bactériémie colibacillaire chez le veau en France, au Canada et en Inde (Contrepolis *et al.*, 1989, Fairbrother *et al.*, 1986, Ngeleka *et al.*, 1993, Ngeleka *et al.*, 1994). Ils sont également retrouvés chez 14% des porcelets infectés au Canada (Contrepolis *et al.*, 1989). Les souches exprimant ces fimbriæ sont généralement associés aux groupes phylogénétiques A et B1 et appartiennent généralement aux sérogroupes O4, O7, O8, O9, O11, O13, O15, O17, O18, O20, O23, O24, O25, O78, O86, O101, O115, O117, O139, O141, O149 et O153 (Contrepolis *et al.*, 1989, Maiti *et al.*, 1993). Cependant, F165₁ a aussi été identifié chez des souches d'*E. coli* septicémiques humaines, majoritairement de type O4 et O6 (Cherifi *et al.*, 1990).

Le rôle exact des fimbriæ F165₁ dans la pathogénie n'est pas connu. Des infections expérimentales de porcelets avec la souche sauvage ou des mutants de délétion pour F165₁ ont suggérés que F165₁ permet aux bactéries de coloniser les organes extra-intestinaux et de résister à la phagocytose (Ngeleka *et al.*, 1993, Ngeleka *et al.*, 1994, Ngeleka & Fairbrother, 1999). Des tests d'adhésion *in vitro* ont également montré que F165₁ permet aux *E. coli* d'adhérer aux cellules MDCK-II (cellules épithéliales rénales canines) et aux cellules LLC-PK1 (cellules tubulo-épithéliales porcines) (Bekal *et al.*, 2003).

b) Structure et organisation génétique de l'opéron *foo*

F165₁ est une adhésine bactérienne de la famille des fimbriæ de type P. Son expression est sous le contrôle de l'opéron *foo* qui partage une organisation génétique homologe à celle des autres membres de cette famille comme *Pap* (Maiti *et al.*, 1994). Ainsi, un alignement des régions régulatrices des opérons *pap* et *foo* montre que ces 2 opérons présentent un pourcentage d'identité supérieur à 92%. En revanche, la régulation de l'adhésine F165₁ se distingue de la plupart des systèmes d'adhésion connus. En effet, grâce à des fusions

transcriptionnelles avec des gènes rapporteurs, des études ont mis en évidence que l'adhésine fimbriaire F165₁ se distingue par des niveaux d'expression significativement plus élevés que tout autre homologue à Pap mais aussi par une sensibilité accrue à la présence d'acides aminés leucine et alanine (Harel *et al.*, 2000, Crost *et al.*, 2004, Berthiaume *et al.*, 2004). Il semble donc que, bien que fortement similaire, ces deux systèmes se distinguent néanmoins par des mécanismes de régulation qui leur sont propres.

9- Adhésines afimbriaires – CS31A

Les adhésines de nature afimbriaire se caractérisent par l'absence de structures visibles autour de la cellule bactérienne. Chez *Escherichia coli*, ce type d'adhésines comprend la famille Afa.

a) CS31A

CS31A est un facteur d'adhésion afimbriaire. En microscopie électronique, la présence de cet antigène de surface se visualise sous forme d'une structure recouvrant l'ensemble de la bactérie, plus proche ainsi d'une structure capsulaire que d'une structure filamenteuse observée dans les cas des adhésines fimbriaires. L'adhésine, ClpG, est la protéine majeure du complexe protéique et permet la reconnaissance de groupements N-acétylneuraminique présents à la surface des cellules de l'hôte (Di Martino *et al.*, 1995, Di Martino *et al.*, 1997, Girardeau *et al.*, 1988).

Initialement, CS31A a été identifié à partir de la souche 31A, dans des fèces de veau diarrhéique (Bertin *et al.*, 1998, Bertin *et al.*, 2000, Contrepois *et al.*, 1993, Girardeau *et al.*, 1988). L'antigène de surface CS31A a été depuis mis en évidence chez d'autres souches d'*E. coli* dans des cas de diarrhées et de septicémies. Les souches exprimant cette adhésine sont associées dans près de 37% des cas de mortalité due à une bactériémie colibacillaire chez le veau en

France, au Canada et en Inde; elles sont également retrouvées chez 1% des porcelets au Canada (Contrepolis et al., 1989). Par ailleurs, CS31A est aussi détecté chez 5% des *E. coli* isolées à partir de selles humaines chez des patients présentant une diarrhée (Jallat *et al.*, 1994, Bertin et al., 2000). Cette adhésine afimbriaire présente des homologies de séquences avec d'autres facteurs d'adhésion comme K88 et F41 (Girardeau *et al.*, 1991, Girardeau et al., 1988, Martin *et al.*, 1991). En revanche, CS31A et K88 présentent tous deux des spécificités dans leur capacité à adhérer à différentes lignées cellulaires (Martin, 1996, Girardeau et al., 1988). Par exemple, K88 est impliqué dans l'adhérence à la bordure en brosse de cellules porcines alors que CS31A ne le permet pas. L'adhésine de CS31A, ClpG, présente aussi des homologies avec l'adhésine CF29K produite par des souches de *Klebsiella pneumoniae* responsables d'infections nosocomiales humaines, avec seulement un seul acide aminé de différence pour la sous-unité majeure (Di Martino et al., 1995, Contrepolis et al., 1993, Korth *et al.*, 1994).

Des tests d'infection expérimentales avec la souche 31A chez des veaux gnotobiotiques et n'ayant pas bu de colostrum ont démontré que cette souche cause une septicémie associée à un œdème rénal qui entraîne la mort des animaux en moins de 48h (Contrepolis *et al.*, 1986). Les souches d'*E. coli* produisant CS31A appartiennent principalement aux sérogroupes O2, O4, O7, O8, O9, O11, O13, O15, O17, O20, O21, O23, O25, O78, O86, O87, O101, O117, O134, O139, O141, O153, O157 et O161, tous présentant une parenté clonale et fréquemment associés à des cas de septicémies chez le veau (Contrepolis et al., 1993, Contrepolis et al., 1989). Cependant, CS31A est également présent chez certaines souches, majoritairement du séro groupe O78, isolées dans des cas de septicémies chez l'homme (Cherifi et al., 1990).

b) Structure et organisation génétique de CS31A

Par comparaison de séquence, il a été de plus montré que cette structure présente des homologies de séquence avec celles d'autres opérons fimbriaires comme l'opéron *pap*. En revanche, À l'exception d'un homologue de *papI*, l'opéron *clp* code l'ensemble des protéines nécessaires à la synthèse de CS31A (Figure 12) (Martin, 1996). De plus, si les régions régulatrices de ces deux opérons présentent une organisation similaire, des alignements de séquence montrent que *pap* n'est pas l'opéron le plus proche de *clp*. Ainsi, à titre d'exemple, ClpB présente 60,6% d'identité avec AfaA, contre 42,4% avec PapB (Martin, 1996).

opéron *clp*



Figure 12. Organisation génétique de l'opéron *clp* codant l'antigène de surface CS31A. Adapté de Martin, 1996.

10- Association des différents fimbriæ entre eux

Généralement, les bactéries pathogènes possèdent plus d'une adhésine, leur permettant ainsi d'interagir avec plusieurs récepteurs distincts. Par conséquent, on peut considérer que les bactéries pathogènes possèdent un éventail d'adhésines leur permettant de répondre à leur environnement de façon très variée. À ce titre, une même souche peut posséder dans son génome les gènes codant les fimbriæ Pap, F165₁ et/ou CS31A (Contrepois et al., 1989, Contrepois et al., 1986, Harel et al., 1991). En effet, Contrepois *et al.* ont montré que, sur plus de 230 souches ExPEC isolées chez des bovins sur trois continents différents, environ 17% présentent simultanément les antigènes F165₁ et CS31A

à leur surface (Contrepois et al., 1989). De même, Harel *et al.* ont montré la présence simultanée de F165₁ et de Pap dans plusieurs souches d'*E. coli* prélevées à partir de différentes espèces animales domestiques (Harel et al., 1991). Ceci laisse suggérer que la présence de plusieurs adhésines, chacune interagissant avec un récepteur spécifique, est largement répandue chez les souches ExPEC. Sur le même principe, il n'est pas surprenant d'imaginer que la présence à la surface bactérienne d'une molécule d'adhésion donnée puisse interférer à la présence d'une seconde non requise à un même site d'infection. De ce fait, il n'est pas surprenant de trouver des mécanismes de régulation croisée entre plusieurs facteurs d'adhésion (Totsika *et al.*, 2008, van der Woude, 2006, van der Woude, 2011, Holden *et al.*, 2006, Simms & Mobley, 2008a, Simms & Mobley, 2008b).

E- RÉGULATION ÉPIGÉNÉTIQUE

1- Introduction

On considère un mécanisme comme étant épigénétique si la capacité d'hériter d'un phénotype est due à un mécanisme autre que la présence d'une mutation ou d'une modification dans la séquence de l'ADN proprement dite. Le terme « épigénétique » dérive du grec *epi-* signifiant « en plus » et de *genno* signifiant « donner naissance » (Casadesús & Low, 2006). De nos jours, ce terme englobe tout mécanisme moléculaire entraînant une modification phénotypique qui se transmet de génération en génération. Ainsi, les modifications biochimiques de l'ADN ou d'un complexe ADN-protéine sont considérées comme des marqueurs d'un mécanisme épigénétique.

Chez les procaryotes, la principale marque épigénétique correspond à une méthylation irréversible de bases nucléotidiques le long de la séquence d'ADN. L'ADN bactérien est sujet à la méthylation de différentes façons, résultant en la présence de N⁴-méthyl-cytosine (m4C), de N⁶-méthyl-Adénine (m6A), et de N⁵-méthyl-Cytosine (m5C) (Cheng, 1995, Wion & Casadesús, 2006). La méthylation du génome bactérien est conservée tout au long du cycle bactérien et se perd par effet de dilution durant les cycles bactériens des générations suivantes (Wion & Casadesús, 2006). Cette modification covalente influence de différentes manières les interactions entre ADN et protéines. Par exemple, la présence de méthylation modifie l'action de certaines enzymes de restriction sur l'ADN (Bickle & Kruger, 1993), elle perturbe la fixation de facteurs transcriptionnels à leurs sites de reconnaissance par des mécanismes d'encombrement stérique (Low et al., 2001) ou encore elle correspond à des sites de reconnaissance importants durant le cycle cellulaire (Reisenauer *et al.*, 1999). Ainsi, chez *Caulobacter crescentus*, CcrM est une méthyltransférase qui joue un rôle clef dans la différenciation entre la forme motile et la forme sessile après chaque cycle de division bactérienne (Zweiger *et al.*, 1994, Marczyński &

Shapiro, 2002). Un autre exemple correspond à la Dam méthyltransférase qui méthyle de façon spécifique l'adénine présente dans les séquences 5'-GATC-3' (Wion & Casadesús, 2006). Ce faisant, la Dam méthyltransférase régule l'expression des gènes en entrant en compétition avec l'interaction de facteurs transcriptionnels sur leur séquence de fixation cible et contenant ce motif GATC.

2- Phénomènes stochastiques et régulation transcriptionnelle

D'une cellule à l'autre, la quantité de facteurs transcriptionnels disponibles varie. En effet, à chaque instant, les facteurs transcriptionnels se trouvent soit fixés à leur séquence cible, soit sous forme libre dans le cytoplasme bactérien. Par conséquent, plus l'affinité d'un régulateur transcriptionnel pour sa séquence cible est élevée, plus la durée de vie de ce dit facteur transcriptionnel sous forme fixée augmente et, par extension, plus la concentration dans le cytoplasme bactérien diminue. Ainsi, de plus en plus d'études considèrent que la quantité de molécules disponibles dans le cytosol varie grandement d'une cellule à l'autre et que le profil transcriptionnel global diffère au sein d'une population clonale, pour une condition environnementale donnée (Acar *et al.*, 2010, Balázs *et al.*, 2011, Mettetal & van Oudenaarden, 2007, Ozbudak *et al.*, 2002, Thattai & van Oudenaarden, 2001, Albano *et al.*, 2005).

Ces fluctuations aléatoires au sein d'une bactérie constituent un élément déterminant dans les variations phénotypiques chez les bactéries et correspondent au bruit de fond. Ce phénomène est appelé un phénomène stochastique. Il correspond à un phénomène aléatoire qui ne se prête qu'à une analyse de probabilité d'apparaître. Cette notion s'oppose donc à d'autres mécanismes que l'on qualifie de phénomènes déterministes. Toutefois, selon Kac *et al.* et Nelson, tout type de développement analysable en terme de probabilité (qu'il soit de nature déterministe ou aléatoire) devrait être considéré comme un processus stochastique (Kac & Logan, 1976, Nelson, 1985).

Dans tout système biologique, les fluctuations menant à un bruit de fond stochastique joue un rôle important dans la force du signal ainsi que dans les boucles de régulation. En effet, de plus en plus d'études considèrent le bruit de fond comme un mécanisme bénéfique plus qu'une nuisance dans l'acquisition de signal. Par exemple, Maamar *et al.* ont montré que la nature aléatoire du choix entre la forme végétative et la compétence chez *Bacillus subtilis* peut être directement considérée en fonction des fluctuations dans la présence d'une protéine particulière d'une cellule à l'autre (Maamar & Dubnau, 2005, Maamar *et al.*, 2007). Sous certaines conditions de croissance, seule une partie de la population bactérienne devient compétente. ComK est à l'origine de ce changement phénotypique. Cette protéine régule son propre niveau de production par une boucle d'autorégulation (Maamar *et al.*, 2007, Veening *et al.*, 2008a). Par conséquent, cette protéine fluctue entre deux états correspondant à une faible ou une forte concentration. Sous forme végétative, les cellules présentent une concentration trop faible en ComK pour activer le gène *comK* (Maamar *et al.*, 2007). À l'inverse, quand elles entrent en compétence, les cellules produisent ComK en grande quantité, maintenant ainsi les cellules en compétence. Il existe un seuil critique dans la quantité de ComK qui permet de différencier ses 2 états. De ce fait, après division bactérienne, une bactérie provenant d'une cellule mère présentant déjà une forte concentration en ComK aura plus de chances d'entrer en compétence qu'une bactérie provenant d'une bactérie ayant de faible concentration en ce régulateur. De plus, seules de rares fluctuations en concentration de ComK permet d'expliquer le fait que certaines bactéries franchissent ce seuil et passent d'un état végétatif à la compétence bactérienne.

En mesurant directement la quantité d'ARNm, Maamar *et al.* ont aussi montré que les fluctuations dans la quantité de ComK présente sont majoritairement dues à une production aléatoire d'ARNm pour *comK* plutôt qu'à la présence d'autres facteurs externes (Maamar *et al.*, 2007). Par conséquent, en favorisant sa diversité phénotypique, *B. subtilis* utiliserait à son avantage la production de

bruit de fond en protéine ComK afin de favoriser sa survie face à des conditions environnementales incertaines.

Comme dans le cas de ComK, l'existence d'une boucle positive d'activation pour l'opéron lactose résulte dans l'établissement d'un système de bistabilité. *In vitro*, les gènes codant l'ensemble des protéines nécessaires à l'incorporation et à l'utilisation du lactose peuvent être artificiellement induits en présence d'isopropyl-D-thio- β -galactopyranoside (IPTG), un analogue non métabolisable du lactose (Veening *et al.*, 2008b). En présence de fortes concentrations en IPTG, l'opéron lactose est sujet à une levée de répression et les cellules expriment alors en grande quantité la perméase LacY. Ceci maintient ainsi les bactéries en mode activé. En présence de faibles concentrations en IPTG, les cellules n'ayant pas été au préalable induites (et ne possédant donc pas de perméase LacY à leur surface) demeurent dans un état réprimé, incapables de répondre à une faible concentration en IPTG. En revanche, une cellule ayant été au préalable induite par de fortes concentrations en IPTG présente toujours des quantités suffisantes de perméase à leur surface, la rendant alors apte à répondre rapidement à l'induction par de l'IPTG (Ozbudak *et al.*, 2004). Elle demeure donc dans un état activé. De ce fait, LacY joue un rôle équivalent à celui joué par ComK dans le cas de *B. subtilis*. En effet, une population bactérienne mise en présence de lactose maintient l'opéron lactose activé même si la concentration à lactose commence à diminuer. La quantité de protéines impliquées dans l'utilisation du lactose a aussi de grandes conséquences sur le profil transcriptionnel d'un individu à l'autre et ce malgré des génotypes et des conditions environnementales identiques. Ce phénomène qualifié d'induction enzymatique sur le principe du « Tout-ou-Rien » indique la présence de 2 sous-populations pouvant coexister simultanément.

3- Variations antigéniques et variations de phase

La première mise en évidence de variations phénotypiques au sein d'une population clonale a été réalisée par F. W. Andrewes en observant les variations morphologiques d'une colonie à l'autre pour une souche de *Salmonella* (Andrewes, 1922). Ces différences phénotypiques sont dues à des variations antigéniques du flagelle chez cette espèce bactérienne et correspond à l'expression d'une protéine qui se trouve sous différentes formes antigéniques d'une cellule à l'autre.

La variation de phase, quant à elle, est un second mécanisme de variation phénotypique qui résulte dans la présence d'un phénotype hétérogène au sein d'une population clonale (van der Woude & Bäumler, 2004). Cette propriété permet à une bactérie de transiter entre deux phénotypes distincts dont l'un correspond à l'expression d'un facteur donné (phase ON) alors que l'autre correspond à sa répression (phase OFF).

Comme pour ComK et l'opéron lactose, la variation de phase présente la particularité d'être héritable et réversible d'une génération à l'autre. Ainsi, une cellule se trouvant en phase ON transmettra ce phénotype à ses descendants avec seulement une faible proportion qui verra son phénotype changer (Jarboe *et al.*, 2004, Zhou *et al.*, 2005). Réciproquement, une cellule en phase OFF engendrera une descendance majoritairement en phase OFF avec seulement une faible population pour laquelle un phénotype en phase ON sera observable.

La fréquence de passage d'un phénotype à l'autre varie grandement entre systèmes, s'étalonnant de 0,001% à 10% de la population totale (Davidson & Surette, 2008). Cette fréquence demeure cependant plus élevée que la fréquence d'apparition d'une mutation spontanée qui provoquerait un changement phénotypique irréversible. Dans le cas d'une fréquence faible, il est nécessaire de

suivre le profil phénotypique sur plusieurs générations pour voir le phénotype changé.

Les mécanismes menant à la variation de phase varient d'un modèle à l'autre. Ils font référence soit à la génétique classique, soit à d'autres mécanismes épigénétiques (Davidson & Surette, 2008, Veening et al., 2008a, Veening et al., 2008b, van der Woude & Bäumler, 2004, Casadesús & Low, 2006, Low & Casadesús, 2008). Dans le cas de la génétique classique, les différences entre une phase ON et une phase OFF impliquent des changements dans la structure même de la séquence de l'ADN. De tels changements peuvent être dus à différents types de modification, comme des changements dans la phase de lecture de la séquence codante ou dans l'orientation d'un promoteur. En revanche, dans tous les cas, ces changements dans la séquence de l'ADN sont réversibles. L'exemple classique de ce type de mécanisme correspond à ceux intervenant durant la variation de phase des fimbriæ de type 1 chez *E. coli* (Casadesús & Low, 2006, van der Woude & Bäumler, 2004).

En revanche, la variation de phase faisant appel à des mécanismes épigénétiques n'implique pas de changements directs dans la séquence proprement dite de l'ADN mais fait référence à des modifications apportées à des nucléotides spécifiques comme le cas de la méthylation. Par conséquent, la Dam méthyltransférase est directement impliquée dans le contrôle de la variation de phase chez plusieurs bactéries. Un des modèles d'étude de l'implication des mécanismes épigénétiques dans la variation de phase correspond ainsi à l'influence des protéines Dam et Lrp dans la régulation par variation de phase de l'opéron *pap*.

a) Variation de phase de l'opéron *pap*

L'expression de l'opéron *pap* est régulée par un mécanisme de variation de phase qui implique l'intervention de Lrp. En effet, les deux phases ON et OFF sont la conséquence directe de la présence de Lrp sur un de ses 2 sites de fixation dans la région régulatrice de l'opéron *pap* (Figure 13). Ce processus fait également intervenir la Dam méthyltransférase ainsi que les 2 régulateurs locaux PapI et PapB (Hernday et al., 2002, Casadesús & Low, 2006). Plus précisément, dans le cas de bactéries en phase OFF, Lrp interagit de manière coopérative avec les sites 1-3, situés à proximité du site d'initiation de la transcription. Ceci provoque alors un encombrement stérique défavorable à l'entrée de l'ARN polymérase. Dans le cas de bactéries en phase ON, Lrp, en combinaison avec le régulateur local PapI, interagit cette fois-ci avec les sites 4-6, situés environ 100 pb plus en amont que les sites 1-3 (Hernday *et al.*, 2003). Cette conformation provoque alors une conformation de l'ADN favorable à l'entrée de l'ARN polymérase qui transcrit l'opéron *pap* dans les deux sens, en un ARNm polycistronique via le promoteur principal *pBA* et en un ARNm simple à partir de *pI*.

Comme toute variation de phase, l'expression de l'opéron *pap* est réversible et se transmet au cours des générations. Ceci est dû à la présence de 2 sites 5'-GATC-3' spécifiques à la Dam méthyltransférase qui sont positionnés sur chacun des sites 2 et 5 et pour lesquels une méthylation complète sur les 2 brins rend la fixation de Lrp impossible (Low et al., 2001). Ainsi, une méthylation du site GATC^{distal} situé dans le site 5 a pour conséquence la fixation de Lrp sur les sites 1-3 et donc une bactérie en phase OFF (Nou *et al.*, 1995, Nou *et al.*, 1993). Réciproquement, une méthylation du site GATC^{proximal} situé dans le site 2 a pour conséquence la fixation de Lrp sur les sites 4-6 et donc une bactérie en phase ON, jusqu'au moins le prochain cycle bactérien. Par conséquent, Lrp et la Dam méthyltransférase présentent des actions exclusives à chacun des sites 1-3 et 4-6.

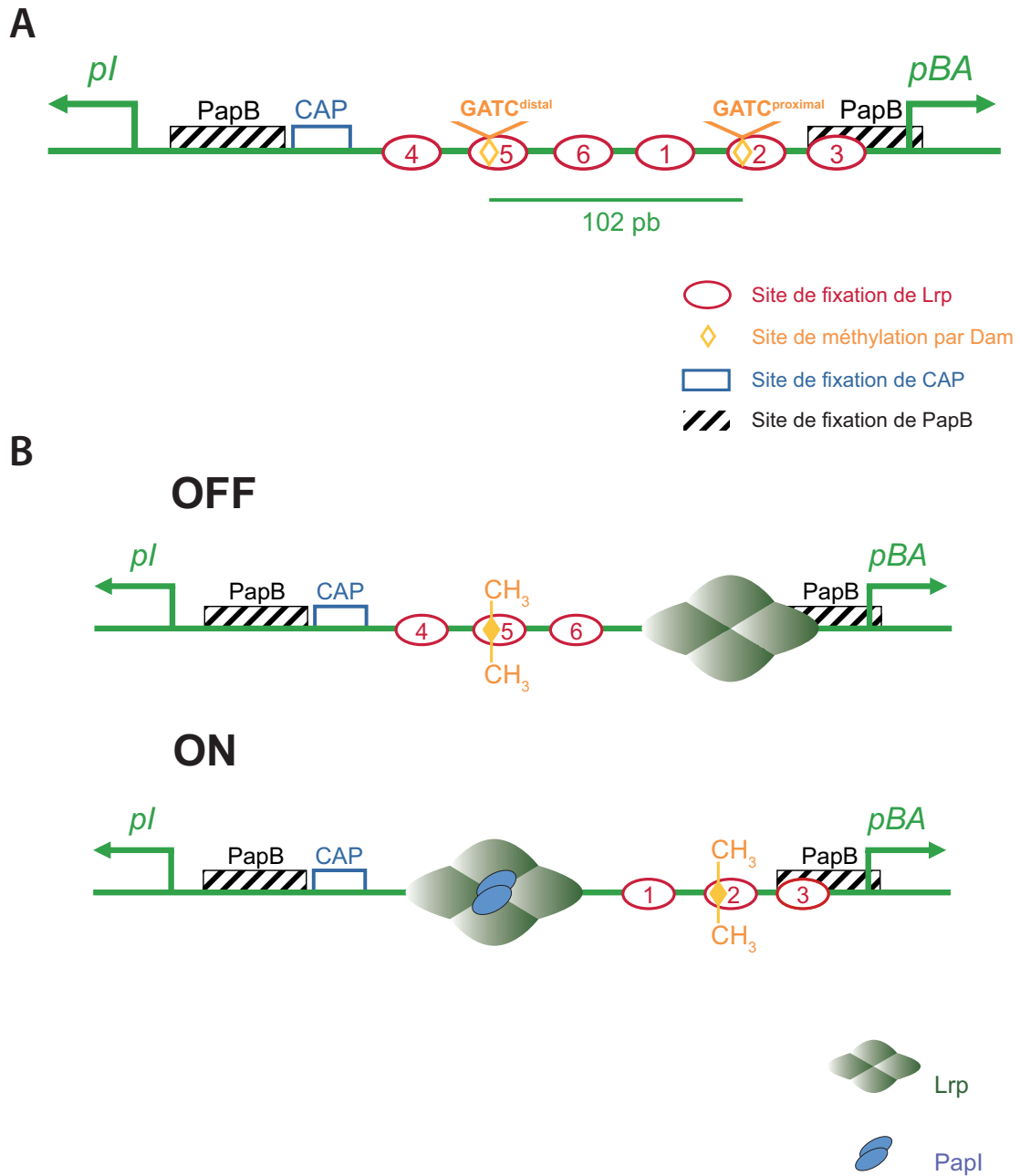


Figure 13. Représentation schématique du mécanisme moléculaire régissant la variation de phase de l'opéron *pap*. *A.* Structure générale de la région régulatrice de l'opéron *pap*. *B.* Représentation des deux scénarios envisageables, correspondant à une phase OFF (majoritaire) et une phase ON.

b) Influence des régulateurs locaux PapI et PapB

-i- PapI

PapI est un cofacteur de Lrp qui favorise la fixation de Lrp sur les sites distaux 4-6 et donc la méthylation du motif GATC^{proximal} (Hernday et al., 2003). Ce faisant, PapI joue un rôle important dans l'apparition de la phase ON en interagissant avec le complexe Lrp-ADN. D'un point de vue structural, cette protéine basique de 8,8 kDa possède un domaine amino-terminal HTH qui jouerait un rôle dans l'interaction avec les molécules d'ADN (Figure 14) (Kawamura *et al.*, 2007). Des études *in vitro* ont de plus montré que l'affinité du complexe PapI/Lrp est supérieure pour les sites 4-6 que pour les sites 1-3 (Hernday et al., 2003, Nou et al., 1995, Kaltenbach *et al.*, 1995). Il semble également que la méthylation du site GATC^{proximal} inhibe fortement les interactions du complexe Lrp-PapI pour ce site particulier, et donc favorise le déplacement du complexe contenant Lrp vers les sites 4-6 (Hernday et al., 2003). À ce jour, 15 allèles du gène *papI* correspondant à 10 variants protéique de ce cofacteur ont été identifiés (Figure 14) et il est à supposer que chacun de ces variants puisse exercer des actions plus ou moins prononcées dans la capacité à orienter l'expression de l'opéron *pap* vers la phase ON (Totsika et al., 2008). La séquence de ce gène est située environ 400pb en amont du promoteur principal *pBA* de l'opéron et est transcrite, à l'aide de son propre promoteur, de façon divergente. L'expression de *papI* est elle-même régulée par la présence du second régulateur spécifique, PapB.

-ii- PapB

PapB est codée par le premier gène de l'opéron *pap*. Cette protéine présente une double fonction. En premier, PapB est un activateur transcriptionnel qui coordonne la transcription des 2 promoteurs *pI* et *pBA* de l'opéron fimbriaire *pap* (Hernday et al., 2002). Ainsi, PapB engendre une boucle positive d'activation de la phase ON en activant la transcription de *papI* (Forsman et al., 1989, Goransson *et al.*, 1989). De ce fait, *papB* est nécessaire pour induire le passage d'une phase OFF à une phase ON (Hernday et al., 2002, Casadesús & Low, 2006, Totsika et al., 2008). En revanche, des concentrations plus élevées en PapB provoquent une inhibition de la transcription de l'opéron *pap* (Totsika et al., 2008, Hernday et al., 2002). Cette double fonction de PapB est due à la présence de 3 sites de fixation pour PapB situés à des positions stratégiques dans la région régulatrice (Totsika et al., 2008). Ainsi, la région régulatrice possède un site de haute affinité pour PapB en aval du promoteur *pI* (impliqué dans la boucle d'activation positive de l'opéron *pap*) ainsi que 2 autres sites à plus faible affinité qui chevauchent la région allant du promoteur principal *pBA* à la séquence codante du gène *papB* (Totsika et al., 2008).

Le site à haute affinité de PapB correspond à l'une des régions les plus variables des divers opérons *pap*. Ce site de reconnaissance consiste en une répétition de 9 pb avec un motif (A/T)₃ à des positions conservées et avec une périodicité de 9 pb. D'un variant à l'autre, ce motif se retrouve répété de quatre à neuf fois (Totsika et al., 2008). Ainsi, ce motif de fixation pour PapB varie de 50 pb jusqu'à un motif d'environ 100 pb et les variations de séquence pourraient influencer sur l'expression transcriptionnelle d'un opéron fimbriaire à l'autre.

c) Variation de phase de F165₁ et de CS31A

Les opérons *foo* et *clp* codent l'ensemble des protéines impliquées dans la biosynthèse des fimbriæ F165₁ et de CS31A, respectivement. Ils possèdent tous les deux une organisation génétique fortement similaire à celle de l'opéron

pap. En particulier, la région régulatrice de *foo* comporte également deux cadres de lectures ouverts orientés de façon divergente qui correspondent aux gènes *fooI* et *fooB* (Maiti et al., 1994). Ces deux gènes codent les protéines FooI et FooB, homologues à PapI et à PapB. Tout comme *pap*, la région intercistronique de *foo* contient également 2 sites GATC séparés l'un de l'autre par 102 pb. Ces caractéristiques sont communes aux membres de la famille régulatrice de l'opéron *pap*. Ainsi, les séquences aux sites de liaison pour PapB dans l'opéron *pap* sont retrouvées également dans l'opéron *foo*. Il en va de même pour les sites de liaison de Lrp. Dans le cas de CS31A, l'opéron *clp* contient également 2 sites GATC séparés l'un de l'autre par 103 pb; en revanche, il se distingue de *pap* et de *foo* par l'absence d'homologue de *papI* (Martin, 1996).

Tout comme *pap*, l'expression de l'opéron *foo* et de *clp* est sujette à la régulation par CRP, Lrp et Dam. En effet, l'expression est négativement régulée par l'ajout de glucose et activée en présence d'AMPc (Crost et al., 2004). De plus, il semblerait que l'expression des opérons *foo* et *clp* implique également l'action d'H-NS dans l'inhibition de l'expression de ces opérons (Crost et al., 2004).

Les opérons *foo* et *clp* sont également soumis à la variation de phase et ce processus requière également la présence de Lrp (Martin, 1996, Crost *et al.*, 2003, Berthiaume et al., 2004, Tessier *et al.*, 2007, Daigle *et al.*, 2000). En revanche, alors que, dans le cas des opérons *pap* et *clp*, la phase OFF est prédominante, les mécanismes conduisant à la variation de phase de l'opéron *foo* conduisent majoritairement à la phase ON (Harel et al., 2000, Crost et al., 2004, Berthiaume et al., 2004).

Ainsi, bien que présentant une organisation similaire et des pourcentages d'identité élevés, chacun des trois opérons *clp*, *foo* et *pap* se distingue dans la variation de phase avec dans un cas un niveau de cellules en phase ON élevé (cas des fimbriæ F165₁), dans l'autre cas un niveau de cellules en phase ON au contraire faible (cas des fimbriæ Pap), et, enfin, un troisième cas pour lequel la

variation de phase n'est observable que si le régulateur local est ajouté en *trans*, comme dans le cas de régulation croisée entre deux opérons fimbriaires (cas de CS31A).

L'objectif principal du présent projet de PhD consiste ainsi à comparer la variation de phase chez ces trois modèles d'études différents, ceci dans le but de mieux comprendre les mécanismes moléculaires impliqués dans la variation de phase de la famille des fimbriæ de type P.

SECTION III – Méthodologie et résultats

Implication de l'auteur dans cet article.

En tant que premier auteur de cet article, j'ai réalisé l'ensemble des constructions bactériennes, ainsi que les expériences de retard de migration sur gel, empreinte de protection à la DNase I, et purification protéique. J'ai également participé au dénombrement des différents types de colonies ainsi qu'aux expériences de résonance plasmonique de surface. Enfin, j'ai analysé l'ensemble des résultats présentés et ai rédigé l'intégralité de cet article.

A. Article I : Lrp-DNA complex stability determines the level of ON cells in type P fimbriae phase variation

Cet article a été publié dans *Molecular Microbiology*, (2011) Vol. **81**(5), p. 1286–1299. Copyright © 2011, WILEY-BLACKWELL. All rights reserved.

Accepted 24 June, 2011.

Authors:

Richard Graveline¹, Michaël Mourez¹, Mark A. Hancock², Christine Martin³,
Stéphanie Boisclair¹ and Josée Harel^{1*}

Addresses:

¹Groupe de Recherche sur les Maladies Infectieuses du Porc (GREMIP) and Centre de Recherche en Infectiologie Porcine (CRIP), Université de Montréal, Saint-Hyacinthe, Quebec, CANADA, ²Sheldon Biotechnology Centre, McGill University, Montreal, Quebec, CANADA, and ³Institut National de la Recherche Agronomique, Saint-Genès-Champanelle, FRANCE

*Corresponding author. Mailing address: Groupe de Recherche en Infectiologie Porcine (GREMIP), Université de Montréal, Faculté de Médecine Vétérinaire, C.P. 5000, Saint-Hyacinthe, Québec, Canada J2S 7C6. Phone: (450) 773-8521, ext. 8233. Fax: (450) 778-8108. E-mail:

Running title:

Lrp control of P fimbriae phase variation switch rate

1- ABSTRACT

F165₁ and the pyelonephritis-associated pili (Pap) are two members of the type P family of adhesive factors that play a key role in the establishment of disease caused by Extra-intestinal *Escherichia coli* (ExPEC) strains. They are both under the control of an epigenetic and reversible switch that defines the number of fimbriated (ON) and afimbriated (OFF) cells within a clonal population. Our present study demonstrates that the high level of ON cells found during F165₁ phase variation is due to altered stability of the DNA complex formed by the leucine-responsive regulatory protein (Lrp) at its repressor binding sites 1-3; after each cell cycle, complex formation is also modulated by the local regulator FooI (homologue to PapI) which promotes the transit of Lrp toward its activator binding sites 4-6. Furthermore, we identified two nucleotides (T490, G508) surrounding the Lrp-binding site 1 that are critical to maintaining a high OFF to ON switch rate during F165₁ phase variation, as well switching Pap fimbriae toward the OFF state.

2- INTRODUCTION

At each stage of bacterial infection, there is a balance between the requirement of virulence factors for survival and the risk of expressing an antigen recognized by the host immune system (Smith *et al.*, 2007, Dobrindt & Hacker, 2008). Pathogenic strains have therefore developed complex regulatory systems that allow them to change phenotype in response to environmental signals. Their ability to generate phenotypic diversity in genetically identical cells is a key mechanism of survival in a hostile environment (Thattai & van Oudenaarden, 2004, Acar *et al.*, 2008). Phase variation is a stochastic and reversible switch between an “all-or-none” (ON/OFF) phenotype, each capable of responding to a specific condition. After cell division, most daughter cells retain the expression phase of the parent whereas a minority switches to the other phenotype. Despite its energetic cost and the risk of being deleterious to

individual cells, phase variation confers a net advantage to the entire population by allowing it to anticipate and survive host environmental changes without having to sense them. Thus, it is not surprising that many virulence factors like fimbrial adhesins are subject to phase variation (van der Woude & Bäumlér, 2004, van der Woude, 2006).

As one of the most studied phase-variation systems (Hernday *et al.*, 2004), Pyelonephritis-associated pili (Pap) fimbriae play an important role in the pathogenesis of urinary tract infections (Plos *et al.*, 1990, Wullt *et al.*, 2000, Wullt *et al.*, 2001, Bergsten *et al.*, 2004). Expression of these adhesins is inherited in an epigenetic manner since it is controlled by a methylation-dependent mechanism (van der Woude & Bäumlér, 2004). Pap phase variation occurs because of methylation-controlled modifications within the regulatory region of the *pap* operon, which consists of an approximately-400 bp intergenic region that is surrounded by two divergently transcribed genes, *papI* and *papB* (Fig. 1). The intergenic region also includes six leucine-responsive regulatory protein (Lrp) binding sites, all containing the consensus sequence GN(2-3)TTT recognized by Lrp (where N is any nucleotide), as well as two GATC sites (referred as GATC^{proximal} and GATC^{distal} respectively) spaced 102 bp apart. The differential methylation status of these sequences determines the binding of Lrp. Indeed, the switch from one phenotype to the other arises from the competition between the binding of Lrp at either its distal or proximal binding sites and methylation by Dam methyltransferase at the opposite GATC site (for reviews, see Blomfield, 2001, Hernday *et al.*, 2002, Casadesús & Low, 2006).

When GATC^{distal} is fully methylated, Lrp cooperatively binds to *pap* sites 1, 2 and 3 to maintain cells in the OFF state (Nou *et al.*, 1995, Hernday *et al.*, 2002). When GATC^{proximal} is fully methylated, however, Lrp binds to sites 4-6 to maintain cells in the ON state (Fig. 1). Control of *pap* expression also requires the action of PapI, a positive regulator that increases the affinity of Lrp for both binding sites 2 and 5 (van der Woude *et al.*, 1996, Hernday *et al.*, 2002). However, methylation at the GATC^{proximal} site provides directionality to the switch by reducing the affinity of PapI-Lrp for sites 1-3 (Hernday *et al.*, 2002). PapB, the

second specific regulator of the *pap* operon, plays an important role at the transcriptional level by coordinating the expression of *pBA* and *pl* promoters primarily (Hernday *et al.*, 2002). Taken together, the combined actions of Lrp, Dam methyltransferase, PapI, and PapB result in a finite probability of each cell to express (ON) or not (OFF) Pap fimbriae following DNA replication.

Similar to Pap mechanistically, F165₁ fimbriae are encoded by the *foo* gene cluster and found in extraintestinal *Escherichia coli* (ExPEC) strains that cause septicemia in young pigs (Harel *et al.*, 1992, Harel *et al.*, 1995). Their role in infection is not fully understood, but previous work has shown that F165₁ fimbriae are essential for virulence and survival during systemic stages of infection (Ngeleka *et al.*, 1994, Ngeleka & Fairbrother, 1999). While both F165₁ and Pap are subject to phase variation control, F165₁ pili are characterized by a higher level of ON cells compared to other members of the type P family of adhesive factors (Harel *et al.*, 2000, Berthiaume *et al.*, 2004, Crost *et al.*, 2004). Similar to *pap*, we have previously shown that *foo* expression relies on the concerted action of Lrp, Dam methylase, FooB, and FooI (Daigle *et al.*, 2000, Harel *et al.*, 2000, Berthiaume *et al.*, 2004). Notably, FooI is a PapI homologue which differs by only a single amino acid modification (D17N) in its sequence.

Unresolved to date is the reason as to why *foo* expression (F165₁ fimbriae) gives rise to a higher level of ON cells compared to *pap* (Pap). In this work, we analyzed and compared the interactions between Lrp, FooI, and/or PapI with the intergenic regions of *foo* and *pap* operons. By combining *in vitro* experiments with genetic studies, we show that subtle sequence differences within the promoter proximal region (when combined with the specific actions of FooI or PapI) can cause a significant change in the level of ON cells. Ultimately, this creates two different phase variation systems: Pap as a slow switcher and F165₁ as a fast switcher.

3- RESULTS

a) Regulatory region of *foo* operon confers preferential ON state of F165₁-positive cells

The *foo* and *pap* operons share more than 96% identity within their regulatory region (Fig. S1), yet previous work has shown that *foo* phase variation is characterized by a higher level of ON cells. To assess the functional role of the *foo* regulatory region in this process, we compared phase variation between strains JH200 (*foo-lacZYA* fusion) and DL4388 (chromosomal *pap17* (referred as *pap* in rest of the text)-*lacZYA* fusion) in an *E. coli* MC4100 background (Table 1). JH200 and DL4388 colonies exhibited distinguishable phenotypes on M9-glycerol plates supplemented with X-Gal (Fig. 2). Since *pap* OFF colonies appeared white and *foo* OFF colonies appeared light blue (Fig. 2A), this indicates that (i) either the basal transcriptional level of OFF individual cells within an OFF colony is not completely shut off, or (ii) there is a high proportion of ON individual cells inside each OFF colony. When the colony phenotypes were followed over three subcultures, phase variation of the JH200 and DL4388 strains was heritably transmitted (Fig. 2B). When dark-blue colonies were plated on fresh M9-glycerol agar, more than half of the progeny remained in the ON state and a significantly higher proportion of dark-blue colonies was observed for JH200 compared to DL4388 (Fig. 2B). In contrast, plating light-blue (JH200) or white (DL4388) colonies on fresh M9-glycerol agar yielded a minority of dark-blue colonies (<10% JH200 and <1% DL4388; Fig. 2B). These findings confirm that, in contrast to phase variation of Pap fimbriae (*pap* operon) for which the majority of cells is in the OFF state, phase variation of F165₁ fimbriae (*foo* operon) is characterized by a higher level of cells in the ON state.

b) Differential phase variation is associated with altered Lrp stability at *foo* or *pap* proximal regions

Because *pap* phase variation requires Lrp (Hernday *et al.*, 2002, Casadesús & Low, 2006), we hypothesized that the high level of F165₁-ON cells could be due to differences in Lrp affinity for either the *foo* distal or *foo* proximal region. In the presence of His₆-Lrp, we performed Electrophoretic Mobility Shift Assays (EMSA) with DNA fragments corresponding to the distal, proximal, and distal-proximal regions of *foo* and *pap* (Fig. S2). In all cases, the specific band shifts that were detected in the presence of increasing Lrp concentration (from 5 nM to 1.6 μM) demonstrated that Lrp can interact with the six DNA regions tested. At low Lrp concentration (10 nM), binding to the *pap* proximal region was notably more intense compared to the *foo* proximal region. This finding provided the first indication that differences between F165₁ and Pap phase variation could be due to differences in Lrp binding between the *foo* and *pap* proximal regions.

To further characterize these interactions, we utilized label-free, real-time Surface Plasmon Resonance (SPR) to quantitate the kinetics of Lrp binding to the *foo* and *pap* regulatory regions. For each set of DNA fragments tested (*i.e.* distal, proximal, distal-proximal), saturable Lrp binding responses exhibited similar association and dissociation kinetics for *foo* and *pap* (Fig. S3). The only major exception was a significantly slower dissociation rate for the *pap* proximal region (Table 2 and compare Fig. 3A and 3C); in turn, the dissociation rate for *pap* distal-proximal was also slower compared to *foo* distal-proximal. While the overall binding affinities for Lrp were low nanomolar in all cases, our data suggests that Lrp complex stability is not different between the *foo* distal and proximal regions whereas it is variable between the *pap* regions. In addition, Lrp binding responses with the *foo* or *pap* distal fragments were ~50% of the total signal changes with the distal-proximal fragments; this observation suggests that each distal and proximal region contains a similar number of binding sites and their cumulative total yields the overall distal-proximal responses observed.

Notably, comparison between the predicted and observed SPR responses indicated that Lrp binds to the distal or proximal regions of *foo* and *pap* as an octamer (data not shown); thus, maximal binding to these immobilized DNA surfaces did not involve the formation of higher-ordered Lrp structure (*e.g.* hexadecamer). The data sets were then analyzed according to previous reports in the literature that Lrp binding is cooperative (*i.e.* sites 1-3 and 4-6). Similar to SPR analyses in which other co-operative protein-DNA kinetics have been analyzed (Kaufman *et al.*, 2007), all Lrp titration series fit well to a “one site model with Hill slope” which indicated positive cooperativity (*i.e.* Hill coefficients >1; Fig. S4). While apparent binding affinities were similar between *foo* and *pap* (Table 2), our data suggests that there is a critical difference in the dissociation rate constant of Lrp for the *foo* or the *pap* proximal regions. This implies that the formation of a stable Lrp complex with the *pap* proximal region could maintain a high level of OFF cells through generations, whereas a less stable complex could allow Lrp to switch from the proximal to distal region of *foo* more often, resulting in a higher level of ON cells.

c) Interaction at Lrp binding site 1 differs between *foo* and *pap*

To identify Lrp binding sites, DNase I footprinting assays were performed with purified His₆-Lrp and DNA fragments encompassing the *foo* and *pap* intergenic regions. Six protected regions were detected on both the *foo* and *pap* fragments (Fig. 4) and, notably, all Lrp binding sites along the *pap* DNA were found at identical positions on the *foo* sequence. We noticed, however, that the protected region corresponding to the *foo* Lrp binding site 1 (proximal region) was smaller than its homologue on *pap* (compare the Top fragments on Fig. 4A and 4B). Specifically, five nucleotides differed between *foo* and *pap* at this site (positions 485, 487, 490, 498 and 508; Fig. S1), including two *foo* DNA nucleotides (positions 490 and 508) that flank the Lrp binding site 1 and create two additional GN(2-3)TTT motifs (the minimal consensus sequence required for Lrp interaction with DNA; (Peterson *et al.*, 2007). It was also observed that equimolar concentrations of Lrp protected the *foo* distal and proximal regions

from DNase I digestion, whereas the *pap* proximal region was protected at a Lrp concentration 10-times lower than that for the *pap* distal region (as previously reported with the complete *pap* intergenic region; (Hernday *et al.*, 2002, Nou *et al.*, 1995). Thus, it is possible that the high level of ON cells that characterizes *foo* phase variation is due to the presence of two additional GN(2-3)TTT motifs along the *foo* proximal region; this could render the binding of Lrp weaker for *foo* proximal compared to *pap* proximal and, in turn, switch Lrp from the *foo* proximal region towards *foo* distal.

d) Stability of Lrp binding to proximal region depends on the number of GN(2-3)TTT motifs

To examine the influence of additional GN(2-3)TTT motifs on Lrp complexation, we utilized our SPR assay to compare the binding of Lrp to proximal fragments in which the wild-type *foo* and *pap* nucleotides at positions 490 and 508 were swapped (Fig 3). As seen earlier (Fig. S3), Lrp binding to the wild-type *foo* proximal fragment exhibited faster dissociation kinetics compared to the wild-type *pap* proximal fragment. When Lrp bound to the “*pap*-swapped”, mutant *foo* fragment (T490A-G508A), the significantly slower dissociation rate resembled the wild-type *pap* proximal fragment (Table 2 and compare Fig. 3B and 3C). Thus, decreasing the number of Lrp consensus sequences in *foo* (equal to that of *pap*) greatly increases the stability of the complex formed with Lrp. In reciprocal experiments, Lrp bound to the “*foo*-swapped”, mutant *pap* fragment (A490T-A508G) with a significantly faster dissociation rate that resembled the wild-type *foo* proximal fragment (Table 2 and compare Fig. 3D and 3A). Thus, increasing the number of Lrp consensus sequences in *pap* (equal to that of *foo*) greatly reduces the stability of the complex formed with Lrp. Taken together, these novel results validate that the altered Lrp-DNA complex stability on the proximal region of the *foo* operon is dependent upon additional GN(2-3)TTT motifs (adjacent to previously characterized Lrp binding sites 1-3 found in *pap*). In contrast, the number of GN(2-3)TTT motifs had no effect on the association kinetics of Lrp for all of the wild-type and mutant proximal fragments tested.

e) FooI and PapI stabilize Lrp binding to both distal and proximal regions

Similar to PapI, FooI expression is essential for phase variation since the deletion of *fooI* results in a uniform phenotype (Croft *et al.*, 2004). Notably, FooI and PapI differ by only a single amino-acid substitution (D17N; Totsika *et al.*, 2008). We monitored for differences in the binding of FooI or PapI at the distal and proximal regions of *foo* and *pap* (Fig. S5 and S6, respectively), but SPR revealed that all of these interactions were characterized by rapid, steady-state kinetics. The apparent binding affinities were low micromolar which correlates well with previous literature demonstrating that PapI binds to *pap* sites 2 (distal) and 5 (proximal) with similar micromolar affinities (Kawamura *et al.*, 2007, Kawamura *et al.*, 2011). While low micromolar injections of FooI or PapI (*i.e.* < 8 μ M) yielded small binding responses to the *foo* or *pap* fragments on their own, Lrp binding was significantly altered in the presence of 6 μ M FooI or 8 μ M PapI (Fig. 5). In a dose-dependent manner, increasing concentrations of FooI or PapI (2-fold dilution series; data not shown) not only enhanced the overall amounts of Lrp bound in all cases (*i.e.* increased association), stability of the resultant *foo* distal, *foo* proximal, and *pap* distal complexes was also increased (*i.e.* dissociation rates were approximately 50% slower by 6 μ M FooI and 8 μ M PapI; Table S1). In contrast, the inherently slow dissociation rate for the *pap* proximal region was not as significantly altered in the presence of increasing FooI or PapI concentration.

f) Influence of specific sequences on phase variation rate

To investigate whether FooI or the two additional GN(2-3)TTT sequences in the *foo* intergenic region are responsible for the high level of ON cells, we swapped the *fooI* and *papI* genes (*i.e.* nucleotides 490 and 508) in JH200 and DL4388, respectively, to generate strains JH201 to JH206 (Table 1). As shown in Fig. 2 and Fig. 6, strain JH200 (*foo* wt) was characterized both by a significantly lower ON to OFF switch frequency and a higher OFF to ON frequency than strain DL4388 (*pap* wt). Exchanging nucleotides at positions 490 and 508 significantly

affected the switch frequency (Fig. 6B-E), as evidenced by OFF colonies becoming light-blue for *pap* and OFF colonies becoming lighter for *foo* (Fig. 6A). Swapping these nucleotides, however, was not sufficient to reach the levels observed with strains JH200 and DL4388. Indeed, only the combination of nucleotide exchange at positions 490 and 508 with the replacement of *fooI* by *papI* rendered the mutant JH200 strain almost indistinguishable from strain DL4388 and vice-versa (Fig. 6). While replacing PapI by FooI did not modify neither the OFF to ON nor the ON to OFF switch rate (Fig. 6C and 6E), replacing FooI by PapI unexpectedly lowered the OFF to ON switch frequency (Fig. 6D). Thus, it appears that FooI contributes to the OFF to ON switch rate and its specific action strictly depends on the stability of the Lrp complex at the proximal region. In addition, the combination of replacing PapI by FooI and swapping nucleotides 490 and 508 increased the *pap* OFF to ON switch frequency to the levels observed for *foo* (compare Fig. 6D and 6E). Taken together, these results indicate that differences in phase variation between *foo* and *pap* is mainly due to the presence of two specific nucleotides at positions 490 and 508, and that FooI promotes a high OFF to ON switch frequency in the presence of T490 and G508.

4- Discussion

The objective of the present study was to determine why F165₁ phase variation is characterized by a higher level of ON cells compared to Pap fimbriae. Using a variety of techniques including gene expression, site-directed mutagenesis, and *in vitro* protein-DNA interaction analysis, we have now demonstrated that altered complex stability between Lrp and the proximal region of the *foo* operon results in the higher level of ON cells observed. We have established that the interaction between Lrp and the *foo* proximal region is more readily dissociated due to the presence of two additional GN(2-3)TTT sequences that likely perturb Lrp binding site 1 in this region. In contrast, absence of the two additional Lrp binding sites and a significantly slower dissociation rate

between Lrp and *pap* proximal likely accounts for the higher level of OFF cells with Pap phase variation.

In Pap and F165₁, phase variation is controlled by methylation of two GATC sites (GATC^{prox} and GATC^{dist}) located within Lrp DNA binding sites 2 and 5, respectively. Methylation controls the binding of Lrp and the coregulatory protein PapI. In phase OFF cells, Lrp binds to sites 1-3 and protects GATC^{prox} from Dam methylation. Since Dam must methylate GATC^{prox} to transit from the OFF to ON phase, the differences in phase variation could involve differential interaction of Lrp as well as the coregulatory protein within the regulatory region. Indeed, binding of Lrp at sites 1-3 reduces the affinity of Lrp for *pap* sites 4-6 by 10-fold, a phenomenon mainly apparent when Lrp binds *pap* DNA carried by a plasmid that has been denoted “mutual exclusion” (Hernday *et al.*, 2002). In addition, we have now shown in our DNase footprinting assay that Lrp does not exhibit any preference to *foo* binding sites 1-3 over sites 4-6, suggesting that the OFF state of *foo* could not be stabilized by a mutual exclusion mechanism. However, since we did not use plasmid DNA during our analysis, we cannot exclude that mutual exclusion occurs *in vivo*. As a result, the OFF to ON switch frequency is almost identical to the ON to OFF one. This result is in agreement with a high switch rate from one phenotype to the other.

In *pap*, the competitive Lrp and Dam assembly on the regulatory region dictates the stability of Lrp to the *pap* proximal region and is important for the heritability of the OFF state (Peterson & Reich, 2008): the kinetics of the reaction suggest that Lrp dissociation from the DNA is the essential control for Dam to gain access and methylate the GATC sites (Fig. 7). Thus, we concluded that nucleotide differences between strains contribute to the differences in phase variation rates by altering the dissociation rate. This suggests that the interaction of Lrp with its binding sites dictates their accessibility to methylation and thus influences the epigenetic process of phase variation. However, as phase variation of *P. fimbriae* requires the action of multiple regulators such as the local regulator PapB, global regulators CpxAR, cAMP-CAP, these may also

contribute to the phase variation difference between strains (van der Woude *et al.*, 1992, Weyand *et al.*, 2001, Hernday *et al.*, 2004, Totsika *et al.*, 2008).

In *E. coli*, Lrp plays a dual role by acting as a key regulator and as a nucleoprotein involved in the organization of the genome (Calvo & Matthews, 1994, Yokoyama *et al.*, 2006). Although a new sequence alignment has been recently proposed (Kawamura *et al.*, 2011), this dual role is due to the property of Lrp to occupy degenerated sequences for which only the minimal consensus sequence GN(2-3)TTT is found in *foo* and *pap* at each Lrp binding site 1-3 and 4-6. It is worth noting that *pap*-related operons in ExPECs show more than three Lrp binding sites at the distal region and a high diversity at the proximal region (Fig. S7). The presence of additional Lrp binding sites, if not directly involved in Lrp binding, probably influences the stability of Lrp. In view of this, we showed that modification of the GN(2-3)TTT motif number within the proximal region of *foo* and *pap* changed not only the binding responses of Lrp but most importantly the switch phenotype of the strains. Similarly, it is possible that the presence of additional GN(2-3)TTT sequences in the vicinity of sites 4-6 might contribute to increasing the dissociation rate of Lrp, favoring the ON to OFF switch. In addition, we could speculate that the absence of one of the three Lrp binding sites 1-3 may increase the level of ON cells.

Fool and PapI are also involved in phase variation. We first showed by Surface Plasmon Resonance that Fool and PapI bind at low micromolar affinities to the distal and proximal regions of the *foo* and *pap* operons. Moreover, we showed that a low concentration of Fool significantly increases the binding and stability of Lrp at both sites 1-3 and 4-6, as it was previously observed for PapI (Kaltenbach *et al.*, 1995, Hernday *et al.*, 2003). Since Lrp alone can form a stable complex with the *pap* proximal region, the action of Fool or PapI is more sensitive at the *pap* distal region for maintaining the cells in the ON state. In contrast, the action of Fool or PapI upon Lrp is similar between *foo* binding sites 1-3 and *foo* binding sites 4-6; this result concurs with a high *foo* switch rate from one phenotype to the other.

Furthermore, we showed by site-directed mutagenesis that the specific

regulator also influences the switch rate by unknown mechanisms. Indeed, swapping FooI for PapI reduced the OFF to ON switch rate of the *foo* operon. However, swapping PapI for FooI did not influence the OFF to ON switch rate of the *pap* operon. In the latter case, it is possible that the presence of a stable Lrp complex at the *pap* proximal region would mask the FooI effect on *pap*. Similarly, Totsika *et al.* (2008) have studied the cross-activation of *pap* variants and suggested that sequence variation among PapI homologues could affect their ability to activate *pap* transcription and they may have evolved in order to prevent cross-activation of related proteins. In addition, our results suggest that the stable PapI-Lrp complex interaction at the proximal region participates in restricting the cross-talk between two *pap* variants by preventing the transactivation of the OFF to ON switch rate. As suggested by Totsika and coworkers, allelic variation in the regulatory regions and regulators could act synergistically to limit coexpression of homologous fimbrial operons (Totsika *et al.*, 2008).

ExPEC strains are recognized as a phylogenetically and epidemiologically distinct group of *E. coli* pathogens that rarely cause enteric diseases. While virulent ExPEC strains mainly belong to the phylogenetic groups B2 and D, F165₁-positive strains fall into the less virulent groups A and B1 and are thus considered to be opportunistic, although isolated from clinical cases (Table S2; Dezfulian *et al.*, 2003). In large populations, switching between two phenotypes creates a bimodal population, even though one phenotype might be more fit than the other in the current environment. However, the unfit subpopulation enables the entire population to hedge its bets in case of future environmental change (Bennett & Hasty, 2007). It is therefore possible that some *E. coli* strains have increased fimbrial OFF to ON switch rates to circumvent host environmental changes.

Acar and coworkers have studied the galactose utilization network of *Saccharomyces cerevisiae* and showed that fast switching populations outgrow slow switchers when the environment fluctuates rapidly, whereas slow-switching phenotypes outgrow fast switchers when the environment changes

rarely (Acar *et al.*, 2008). Similarly, a phylogenetic tree based on the intergenic DNA sequence of each *pap* operon homologues from different ExPEC strains shows that intergenic regions cluster according to the phylogenetic group of the strains (Fig. S8). In phylogenetic group A, all sequenced strains are lacking Lrp binding site 1. In group B1, all strains including strain 4787 bearing *foo* presents the two additional Lrp-binding sites, while strains belonging to groups B2 and D have none or only one additional binding site. ExPEC strains can produce high rates of fimbrial switch depending on the environment within which they have to cope. In this regard, virulent uropathogenic strains, which are found at specific infection sites, may require a more stringent control of Pap fimbriae production. However, the more rapid fimbrial switching of F165₁-ExPEC strains may enhance their adaptation to fluctuating environments and contribute to their fitness. As a result, ExPEC strains with different pathogenic traits do not necessarily require similar levels of expression among *pap* variants; phylogenetically associated strains may have developed similar mechanisms of phase variation and less pathogenic strains (or strains associated with septicaemia) may be more disposed to switching to the ON state.

Overall, the present study has provided novel insights as to why F165₁ is a fast-switcher and Pap is a slow-switcher. Our data emphasize that the dissociation of proteins from their DNA binding sites could represent a generalized mechanism of cell regulation. To this end, several groups (Mulcair *et al.*, 2006, Labib & Hodgson, 2007, Kaplan & Bastia, 2009) have already described that the capacity of a ligand to form a tight complex along DNA is critical for replication. Building upon this theme, we now propose that the stability of protein-DNA complexes is also critical for phase variation.

5- EXPERIMENTAL PROCEDURES

a) Bacterial strains, plasmids and growth conditions

Bacterial strains and plasmids used in this study are listed in Table 1. Luria-Bertani (LB) broth, M63, and M9 minimal broths were prepared as

described previously (Miller, 1972, Thomason *et al.*, 2007). When necessary, media were supplemented with antibiotics at the following concentrations (unless otherwise noted): ampicillin (amp), 100 $\mu\text{g.ml}^{-1}$; chloramphenicol (Cm), 10 $\mu\text{g.ml}^{-1}$; kanamycin, 40 $\mu\text{g.ml}^{-1}$; tetracycline (tet) 12.5 $\mu\text{g.ml}^{-1}$; as well as with 5-bromo-4-chloro-3-indolyl-b-D-galactopyranoside (X-gal) at a final concentration of 40 $\mu\text{g.ml}^{-1}$.

Strain JH200 to JH206 were constructed by first amplifying the *cat-sacB* from strain NC397 (Svenningsen *et al.*, 2005) using primers i-cat2 (forward) and fooUB-sacB (reverse). The PCR product was then recombined into strain DL4388 using the λ red-recombination system as previously described (Murphy & Campellone, 2003). The recombinant strain was selected on LB agar supplemented with chloramphenicol and tested for sensitivity on sucrose minimal plates (M63 salts, 0.2% glycerol, 6% sucrose). The resulting strain was *papI'::cat-sacB-papB-lacZYA*. A single confirmed Cm^R/sucrose-sensitive clone was then used to generate strains JH200 to JH206, using the λ red-recombination system in a second round of recombineering. All recombinant strains were generated using the same procedure, except for the forward primer and the DNA template used. For strain JH200 and 201, the PCR product was amplified on strain 4787 using pI/fl-fwd and fool->papI-F as forward primer, respectively; for strain JH202 and JH203, the PCR product was amplified on plasmid pCR2.1-foo_{T490A-G510A} (Eurofins MWG Operon) using pI/fl-fwd and fool->papI-F as forward primer, respectively; for strain JH204, the PCR product was amplified on strain DL4388 using papI->fool-F as forward primer; for strain JH205 and JH206, the PCR product was amplified on plasmid pCR2.1-pap_{T490A-G510A} (Eurofins MWG Operon) using pI/fl-fwd and papI->fool-F as forward primer, respectively (fooUB40 was used as reverse primer for all the PCR amplification). Recombinant strains were selected for loss of the *cat-sacB* cassette on sucrose M63 minimal plates and at least ten sucrose-resistant colonies were subsequently tested for sensitivity on LB-Cm agar. Candidates were then purified on LB plates and confirmation of the resulting fusions was

then obtained by sequencing with *fool*-fwd and *foopB*-rvs as primers. Oligonucleotides used as probes and for polymerase chain reaction are listed in Table S3. For cloning procedures, PCR amplification was carried out using the Fidelity PCR MasterMix (GE Healthcare).

b) Phase variation and measurement of switching frequencies

The switch frequency was calculated as described previously (Blyn *et al.*, 1989, Crost *et al.*, 2003). Strains were inoculated onto LB-kan agar plates. Colonies were then excised from the agar and resuspended in M9 salts containing no carbon source. Appropriate dilutions were subsequently spread onto M9 X-Gal agar plates containing 0.2% glycerol and digitally photographed after 24 hours of growth at 37°C. Three colonies showing a uniform Lac⁺ or Lac⁻ phenotype were excised from the agar, resuspended in M9 salts, and spread onto M9 X-gal agar as before. The passage from M9 X-gal agar to fresh M9 X-gal agar containing the same source of carbon was repeated two more times to follow the switch frequency through three generations. At least 2,000 colonies were scored for a Lac phenotype as described (Blyn *et al.*, 1989) and their switch frequencies calculated according to the formula $(M/N)/g$, where M is the number of cells that underwent phase variation, N the total number of cells evaluated and g the total number of generations (estimated to be 20 for all the strains tested in our experiments).

c) Surface Plasmon Resonance

Interactions between 5' biotinylated, double-stranded oligonucleotides (*foo* distal, 120 kDa; *foo* proximal, 77 kDa; *foo* distal-proximal, 185 kDa; *pap* distal, 123 kDa; *pap* proximal, 82 kDa; *pap* distal-proximal, 189 kDa) and three DNA-binding proteins (Lrp, 20 kDa monomer; *Fool*, 9 kDa; *PapI*, 9 kDa) were examined using label-free, real-time BIACORE 3000 instrumentation (GE Healthcare Bio-Sciences AB, Uppsala, Sweden). Experiments were performed on research-grade streptavidin-coated (SA) sensor chips (XanTec Bioanalytics GmbH, Muenster, Germany) at 25°C using filtered (0.2 µm) and degassed

CALVO-P running buffer: 50 mM Tris pH 8.0, 150 mM NaCl, 0.4 mM EDTA, 1 mM MgCl₂, 0.1 mM DTT, 0.005% (v/v) Tween-20. As recommended by the manufacturer, SA sensor chips were pre-conditioned with three 1-minute pulses of 1M NaCl in 50mM NaOH. Protein-grade detergents (Tween-20, Empigen) were from Anatrace (Maumee, OH, USA) and fatty acid-free bovine serum albumin (BSA) was from Millipore (Billerica, MA, USA); all other chemicals were reagent grade quality. Purified protein concentrations were determined using the BCA assay and purity (>95%) was assessed by SDS-PAGE (12.5% polyacrylamide) under reducing (5% (v/v) 2-mercaptoethanol) and non-reducing conditions (data not shown).

To examine binding with Lrp, the appropriate oligonucleotides were annealed (see www.piercenet.com protocol) and then captured (20 μ l/min x 10 nM in running buffer containing 0.5M NaCl) to generate low-density surfaces (<100 RU immobilized). Lrp (0 – 100 nM) or BSA (negative control) was injected over reference (*i.e.* SA only) and active (*i.e.* *foo* or *pap* DNA fragments) surfaces in-tandem using 'KINJECT' mode (25 μ l/min x 5 min association + 5 min dissociation). Between sample injections, surfaces were regenerated at 50 μ l/min using two 30 second pulses of solution I (1 M NaCl, 0.02% (v/v) Empigen in CALVO-P) followed by 'EXTRACLEAN' and 'RINSE' procedures.

To examine binding with F_ooI and PapI, medium-density oligonucleotide surfaces (300 – 400 RU immobilized) were prepared as noted above. F_ooI or PapI (0 – 15 μ M) was injected over reference and active surfaces in-tandem using the 'KINJECT' mode (5 μ l/min x 5 min association + 10 min dissociation). Between sample injections, surfaces were regenerated as noted above. In additional experiments, Lrp (50 nM) was also injected in the presence of F_ooI and PapI (0 – 6 or 8 μ M, respectively; 2-fold dilution series) using the 'KINJECT' mode (10 μ l/min x 5 min association + 10 min dissociation).

Mass transport-independent data was doubled-referenced (Myszka, 1999) and is representative of duplicate injections acquired from three independent trials. For each replicate series, a buffer blank was injected first, the

highest titrant concentration second, and then serial dilutions (from the lowest to the highest concentration). Comparing responses between the two highest titrant injections verified consistent DNA surface activity throughout each assay. Since the acquired titration series deviated from a simple 1:1 kinetic model, binding responses (fractional amounts bound at the end of the association phase, RU) were plotted as a function of protein concentration (C); subsequent non-linear regression analysis (“One site – Specific binding with Hill slope”, Prism5 version 5.0c for Mac OS X, GraphPad Software, San Diego, California, USA, www.graphpad.com) yielded apparent equilibrium affinity constants (K_D). To estimate individual dissociation rate constants (k_d), sensorgram profiles were analyzed using the “Fit Kinetics/Separate k_a/k_d ” tool in the BIAevaluation v4.1 software (GE Healthcare). Theoretical binding maxima were predicted using the following equation:

$$R_{\max} = (MWA/MWL) (RL) (n)$$

where R_{\max} is the maximal binding response (RU) at saturating protein concentration; MWA is the molecular weight (kDa) of the protein injected in solution; MWL is the molecular weight (kDa) of the DNA fragment immobilized; RL is the amount (RU) of DNA immobilized; n is the predicted binding stoichiometry (*e.g.* 1:1).

d) Footprint Analysis

DNase I footprinting of free DNA and DNA–protein complexes was performed as described (Kaltenbach *et al.*, 1995). End-labeled ^{32}P -*pap* and *-foo* DNA (125,000 cpm, 0.3 nM) was incubated in a total volume of 20 μL with 3 μg of salmon sperm DNA and 2 μg of acetylated bovine serum albumin (New England Biolabs) in a binding buffer (60 mM Tris-HCl [pH 7.5], 40 mM KCl, 100 mM NaCl, 0.1 mM EDTA, 1 mM DTT). After incubation for 10 min at room temperature, 6 μL of DNase I (2,000 U.ml⁻¹, New England Biolabs) containing 22 mM CaCl₂ and 22 mM MgCl₂ was added for 1 min. The reaction was stopped by

the addition of 200 μ L of stop buffer (10% glycerol, 2.5 mM EDTA, 0.6 M ammonium acetate, 2 mg yeast tRNA per ml) to each sample. DNA fragments were precipitated in ethanol and an equivalent of c.p.m. (5×10^4) from each reaction was loaded onto 6% polyacrylamide/7M urea gels. A + G Maxam Gilbert reactions (Maxam & Gilbert, 1980) were carried out on the appropriate 32 P-labeled DNA fragments and loaded alongside the DNase I footprinting reactions. Gels were dried and analyzed by autoradiography.

6- ACKNOWLEDGMENTS

We are grateful to David Low (University of California) for the generous gift of purified PapI and enlightening discussions. We thank S. Gottesman and P. Mandin (National Cancer Institute) for their kind gift of strain NC397. We are grateful to Cécile Crost (CRIP) and George Szatmari for their invaluable advice during this study. This work was supported by funding from the Centre de Recherche en Infectiologie Porcine (CRIP; Fonds québécois de la recherche sur la nature et les technologies Regroupements stratégiques 111946), the Natural Sciences and Engineering Research Council of Canada (NSERC Discovery Grant 25120 to JH and 262746 to MM), and the Canada Research Chair program to MM. Sheldon Biotechnology Centre is supported by a Research Resource Grant from the Canadian Institutes of Health Research.

7- REFERENCES

- Acar, M., J. T. Mettetal & A. van Oudenaarden, (2008) Stochastic switching as a survival strategy in fluctuating environments. *Nat Genet* **40**: 471-475.
- Bennett, M. R. & J. Hasty, (2007) A DNA methylation-based switch generates bistable gene expression. *Nat Genet* **39**: 146-147.
- Bergsten, G., M. Samuelsson, B. Wullt, I. Leijonhufvud, H. Fischer & C. Svanborg, (2004) PapG-dependent adherence breaks mucosal inertia and triggers the innate host response. *J Infect Dis* **189**: 1734-1742.

- Berthiaume, F., C. Crost, V. Labrie, C. Martin, E. B. Newman & J. Harel, (2004) Influence of L-leucine and L-alanine on Lrp regulation of *foo*, coding for F165₁, a Pap homologue. *J Bacteriol* **186**: 8537-8541.
- Blomfield, I. C., (2001) The regulation of *pap* and type 1 fimbriation in *Escherichia coli*. *Adv Microb Physiol* **45**: 1-49.
- Blyn, L. B., B. A. Braaten, C. A. White-Ziegler, D. H. Rolfson & D. A. Low, (1989) Phase-variation of pyelonephritis-associated pili in *Escherichia coli*: evidence for transcriptional regulation. *EMBO J* **8**: 613-620.
- Calvo, J. M. & R. G. Matthews, (1994) The leucine-responsive regulatory protein, a global regulator of metabolism in *Escherichia coli*. *Microbiol Rev* **58**: 466-490.
- Casadesús, J. & D. Low, (2006) Epigenetic gene regulation in the bacterial world. *Microbiol Mol Biol Rev* **70**: 830-856.
- Crost, C., A. Garrivier, J. Harel & C. Martin, (2003) Leucine-responsive regulatory protein-mediated repression of *clp* (encoding CS31A) expression by L-leucine and L-alanine in *Escherichia coli*. *J Bacteriol* **185**: 1886-1894.
- Crost, C., J. Harel, F. Berthiaume, A. Garrivier, M. C. Tessier, H. Rakotoarivonina & C. Martin, (2004) Influence of environmental cues on transcriptional regulation of *foo* and *clp* coding for F165(1) and CS31A adhesins in *Escherichia coli*. *Res Microbiol* **155**: 475-482.
- Daigle, F., C. Forget, C. Martin, M. Drolet, M. C. Tessier, H. Dezfulian & J. Harel, (2000) Effects of global regulatory proteins and environmental conditions on fimbrial gene expression of F165(1) and F165(2) produced by *Escherichia coli* causing septicaemia in pigs. *Res Microbiol* **151**: 563-574.
- Dezfulian, H., I. Batisson, J. M. Fairbrother, P. C. K. Lau, A. Nassar, G. Szatmari & J. Harel, (2003) Presence and characterization of extraintestinal pathogenic *Escherichia coli* virulence genes in F165-positive *E. coli* strains isolated from diseased calves and pigs. *J Clin Microbiol* **41**: 1375-1385.

- Dobrindt, U. & J. Hacker, (2008) Targeting virulence traits: potential strategies to combat extraintestinal pathogenic *E. coli* infections. *Curr Opin Microbiol* **11**: 409-413.
- Harel, J., F. Daigle, C. Forget, M. C. Tessier, C. Crost & C. Martin, (2000) Phase variation of F165(1) (Prs-like) fimbriae from *Escherichia coli* causing septicaemia in animals. *Can J Microbiol* **46**: 1101-1107.
- Harel, J., C. Forget, J. Saint-Amand, F. Daigle, D. Dubreuil, M. Jacques & J. Fairbrother, (1992) Molecular cloning of a determinant coding for fimbrial antigen F165(1), a Prs-like fimbrial antigen from porcine septicaemic *Escherichia coli*. *J Gen Microbiol* **138**: 1495-1502.
- Harel, J., M. Jacques, J. M. Fairbrother, M. Bossé & C. Forget, (1995) Cloning of determinants encoding F165(2) fimbriae from porcine septicaemic *Escherichia coli* confirms their identity as F1C fimbriae. *Microbiology (Reading, Engl)* **141 (Pt 1)**: 221-228.
- Hernday, A., M. Krabbe, B. Braaten & D. Low, (2002) Self-perpetuating epigenetic pili switches in bacteria. *Proc Natl Acad Sci USA* **99 Suppl 4**: 16470-16476.
- Hernday, A. D., B. A. Braaten, G. Broitman-Maduro, P. Engelberts & D. A. Low, (2004) Regulation of the *pap* epigenetic switch by CpxAR: phosphorylated CpxR inhibits transition to the phase ON state by competition with Lrp. *Mol Cell* **16**: 537-547.
- Hernday, A. D., B. A. Braaten & D. A. Low, (2003) The mechanism by which DNA adenine methylase and PapI activate the *pap* epigenetic switch. *Mol Cell* **12**: 947-957.
- Kaltenbach, L. S., B. A. Braaten & D. A. Low, (1995) Specific binding of PapI to Lrp-*pap* DNA complexes. *J Bacteriol* **177**: 6449-6455.
- Kaplan, D. L. & D. Bastia, (2009) Mechanisms of polar arrest of a replication fork. *Mol Microbiol* **72**: 279-285.
- Kaufman, B. A., N. Durisic, J. M. Mativetsky, S. Costantino, M. A. Hancock, P. Grutter & E. A. Shoubridge, (2007) The mitochondrial transcription factor

- TFAM coordinates the assembly of multiple DNA molecules into nucleoid-like structures. *Mol Biol Cell* **18**: 3225-3236.
- Kawamura, T., L. U. K. Le, H. Zhou & F. W. Dahlquist, (2007) Solution structure of *Escherichia coli* PapI, a key regulator of the *pap* pili phase variation. *J Mol Biol* **365**: 1130-1142.
- Kawamura, T., A. S. Vartanian, H. Zhou & F. W. Dahlquist, (2011) The Design Involved in PapI and Lrp Regulation of the *pap* operon. *J Mol Biol* **409**: 311-332
- Labib, K. & B. Hodgson, (2007) Replication fork barriers: pausing for a break or stalling for time? *EMBO Rep* **8**: 346-353.
- Maxam, A. M. & W. Gilbert, (1980) Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol* **65**: 499-560.
- Miller, J. H., (1972) *Experiments in Molecular Genetics*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Mulcair, M. D., P. M. Schaeffer, A. J. Oakley, H. F. Cross, C. Neylon, T. M. Hill & N. E. Dixon, (2006) A molecular mousetrap determines polarity of termination of DNA replication in *E. coli*. *Cell* **125**: 1309-1319.
- Murphy, K. C. & K. G. Campellone, (2003) Lambda Red-mediated recombinogenic engineering of enterohemorrhagic and enteropathogenic *E. coli*. *BMC Mol Biol* **4**: 11.
- Myszka, D. G., (1999) Improving biosensor analysis. *J Mol Recognit* **12**: 279-284.
- Ngeleka, M. & J. M. Fairbrother, (1999) F165(1) fimbriae of the P fimbrial family inhibit the oxidative response of porcine neutrophils. *FEMS Immunol Med Microbiol* **25**: 265-274.
- Ngeleka, M., B. Martineau-Doizé & J. M. Fairbrother, (1994) Septicemia-inducing *Escherichia coli* O115:K"V165" F165(1) resists killing by porcine polymorphonuclear leukocytes *in vitro*: role of F165(1) fimbriae and K"V165" O-antigen capsule. *Infect Immun* **62**: 398-404.
- Nou, X., B. Braaten, L. Kaltenbach & D. A. Low, (1995) Differential binding of Lrp to two sets of *pap* DNA binding sites mediated by PapI regulates Pap phase variation in *Escherichia coli*. *EMBO J* **14**: 5785-5797.

- Peterson, S. N., F. W. Dahlquist & N. O. Reich, (2007) The role of high affinity non-specific DNA binding by Lrp in transcriptional regulation and DNA organization. *J Mol Biol* **369**: 1307-1317.
- Peterson, S. N. & N. O. Reich, (2008) Competitive Lrp and Dam assembly at the *pap* regulatory region: implications for mechanisms of epigenetic regulation. *J Mol Biol* **383**: 92-105.
- Plos, K., T. Carter, S. Hull, R. Hull & C. Svanborg Eden, (1990) Frequency and organization of *pap* homologous DNA in relation to clinical origin of uropathogenic *Escherichia coli*. *J Infect Dis* **161**: 518-524.
- Smith, J. L., P. M. Fratamico & N. W. Gunther, (2007) Extraintestinal pathogenic *Escherichia coli*. *Foodborne Pathog Dis* **4**: 134-163.
- Svenningsen, S. L., N. Costantino, D. L. Court & S. Adhya, (2005) On the role of Cro in lambda prophage induction. *Proc Natl Acad Sci USA* **102**: 4465-4469.
- Thattai, M. & A. van Oudenaarden, (2004) Stochastic gene expression in fluctuating environments. *Genetics* **167**: 523-530.
- Thomason, L., D. L. Court, M. Bubunencko, N. Costantino, H. Wilson, S. Datta & A. Oppenheim, (2007) Recombineering: genetic engineering in bacteria using homologous recombination. *Curr Protoc Mol Biol* **Chapter 1**: Unit 1.16.
- Totsika, M., S. A. Beatson, N. Holden & D. L. Gally, (2008) Regulatory interplay between *pap* operons in uropathogenic *Escherichia coli*. *Mol Microbiol* **67**: 996-1011.
- van der Woude, M., B. Braaten & D. Low, (1996) Epigenetic phase variation of the *pap* operon in *Escherichia coli*. *Trends Microbiol* **4**: 5-9.
- van der Woude, M. W., (2006) Re-examining the role and random nature of phase variation. *FEMS Microbiol Lett* **254**: 190-197.
- van der Woude, M. W. & A. J. Bäumlner, (2004) Phase and antigenic variation in bacteria. *Clin Microbiol Rev* **17**: 581-611, table of contents.
- van der Woude, M. W., B. A. Braaten & D. A. Low, (1992) Evidence for global regulatory control of pilus expression in *Escherichia coli* by Lrp and DNA

methylation: model building based on analysis of *pap*. *Mol Microbiol* **6**: 2429-2435.

- Weyand, N. J., B. A. Braaten, M. van der Woude, J. Tucker & D. A. Low, (2001) The essential role of the promoter-proximal subunit of CAP in *pap* phase variation: Lrp-and helical phase-dependent activation of *papBA* transcription by CAP from -215. *Mol Microbiol* **39**: 1504-1522.
- Wullt, B., G. Bergsten, H. Connell, P. Röllano, N. Gebratsedik, L. Hang & C. Svanborg, (2001) P-fimbriae trigger mucosal responses to *Escherichia coli* in the human urinary tract. *Cell Microbiol* **3**: 255-264.
- Wullt, B., G. Bergsten, H. Connell, P. Röllano, N. Gebretsadik, R. Hull & C. Svanborg, (2000) P fimbriae enhance the early establishment of *Escherichia coli* in the human urinary tract. *Mol Microbiol* **38**: 456-464.
- Yokoyama, K., S. A. Ishijima, L. Clowney, H. Koike, H. Aramaki, C. Tanaka, K. Makino & M. Suzuki, (2006) Feast/famine regulatory proteins (FFRPs): *Escherichia coli* Lrp, AsnC and related archaeal transcription factors. *FEMS Microbiol Rev* **30**: 89-108.

8- Legends

Figure 1. Schematic representation of *foo* and *pap* phase variation.

Regulatory region shared between the *foo* and the *pap* operons, with binding sites for the main regulators, is depicted. For clarity, only the organization of the *pap* operon is depicted. The Lrp sites are shown as empty circles. The two GATC sites subject to methylation by Dam, shown as empty diamonds, are GATC^{prox} and GATC^{dist}, which are located within Lrp binding sites 2 and 5, respectively. The CAP and PapB binding sites are shown as open and hatched boxes, respectively. Localization of Lrp, PapI, as well as methylation status is shown for either the OFF state (A) or the ON state (B).

Figure 2. F165₁-positive cells present a high percentage of ON cells. A,

visualization of colonies showing a Lac⁺ or a Lac⁻ phenotype on M9 minimal medium supplemented with 0.2% glycerol depending on the starting phenotype. B percentage of Lac⁺ phenotype of colonies grown on M9 minimal agar supplemented with 0.2% glycerol and grown at 37°C.

Figure 3. Association and dissociation kinetics of Lrp binding to wild-type and mutant proximal regions of *foo* and *pap*.

Representative SPR analyses of Lrp (0 – 100 nM; 2-fold dilution series) injected over wild-type or “swapped” proximal DNA fragments (<100 RU each) at 25 μL/min (5 min association + 5 min dissociation): A, wild-type *foo* proximal; B, mutant *foo*_{T490A-G508A} (to mimic *pap*); C, wild-type *pap* proximal; D, mutant *pap*_{A490T-A508G} (to mimic *foo*). Arrows indicate the differences in term of dissociation kinetics between *foo* and *pap*.

Figure 4. Differences in DNA sequence affect Lrp localization on *foo* regulatory region.

DNase I footprinting assays with increasing amounts of Lrp: Lane1, A + G Maxam and Gilbert reactions; Lanes 2-8: DNase I footprinting assays with increasing amounts of Lrp on *foo* (A) and *pap* (B) regulatory regions

end-labeled with ^{32}P on the top (*left*) and bottom (*right*) strands (lane 2, DNA alone; lane 3, Lrp at 1 nM; lane 4, Lrp at 10 nM; lane 5, Lrp at 100 nM; lane 6, Lrp at 1 μM ; lane 7, Lrp at 10 μM ; lane 8, Lrp at 40 μM). The arrow along the *foo* and *pap* DNA sequences indicates the difference observed for the binding of Lrp at its binding site 1; asterisks represent the nucleotides at position 490 and 508.

Figure 5. Fool and PapI increase both association and stability of Lrp at the *foo* and *pap* proximal and distal regions. Representative SPR analyses of 50 nM Lrp binding to immobilized *foo* (400 RU distal, 300 RU proximal) or *pap* (400 RU distal, 300 RU proximal) fragments in the absence (solid lines) and presence (dashed lines) of fixed Fool (6 μM) or PapI (8 μM).

Figure 6. Presence of additional GN(2-3)TTT motifs, combined with the action of Fool, explains the relative expression of *foo* operon. *A.* Lac⁺ and Lac⁻ phenotypes of *foo* wt, *pap* wt and their respective mutants for which Fool or PapI and/or the site 1 were exchanged. *B* to *E.* Switch frequencies of the same construct strains (*B* and *C*, ON to OFF switch frequencies in *foo* and *pap* backgrounds, respectively; *D* and *E*, OFF to ON switch frequencies in *foo* and *pap* backgrounds, respectively). Switch frequencies were calculated using the formula $(M/N)/g$, where M is the number of cells that underwent phase transition, N the total number of cells evaluated and g the number of generations, estimated to be 20, that gave rise to the colony (*P < 0.05).

Figure 7. Schematic model illustrating differences in phase variation between *foo* and *pap*. F165₁ and Pap share similar regulatory mechanisms, however, F165₁ fimbriae expression has its own specificity. While the *pap* phase switch is preferentially found in the OFF state, phase ON cells are largely predominant for F165₁ fimbriae. After each DNA replication, daughter cells can either inherit the parental phenotype or switch to the opposite phenotype. For *foo*, a weak affinity of Lrp for its *foo* binding site 1, due to the presence of two additional sites for Lrp (indicated by two red arrows), facilitates the OFF to ON

switch whereas the formation of a particularly stable complex between Lrp and the *pap* proximal region maintains the OFF state through generations. Compared to PapI, FooI seems also to contribute to the increased level of ON cells through unknown mechanisms. Switch rates indicated with black arrows correspond to measurements of switch frequencies used in this study.

Table 1. Phylogenic and phenotypic characterization of ExPEC strains used in this study.

Table 2. Apparent kinetic rate constants for Lrp, FooI, and PapI binding to *foo* and *pap* DNA fragments, as determined by SPR. ‘Lrp Affinity’ was determined according to “One site – specific binding with Hill slope” model in Prism5 software. ‘Lrp Dissociation’ was determined according to “Fit separate k_a/k_d ” model in BIAevaluation software. ‘FooI or PapI Affinity’ was determined according to “Steady-state affinity” model in BIAevaluation software. Each titration series was analyzed globally using the specified model and estimates represent the average of three independent trials (+/- standard error of the mean).

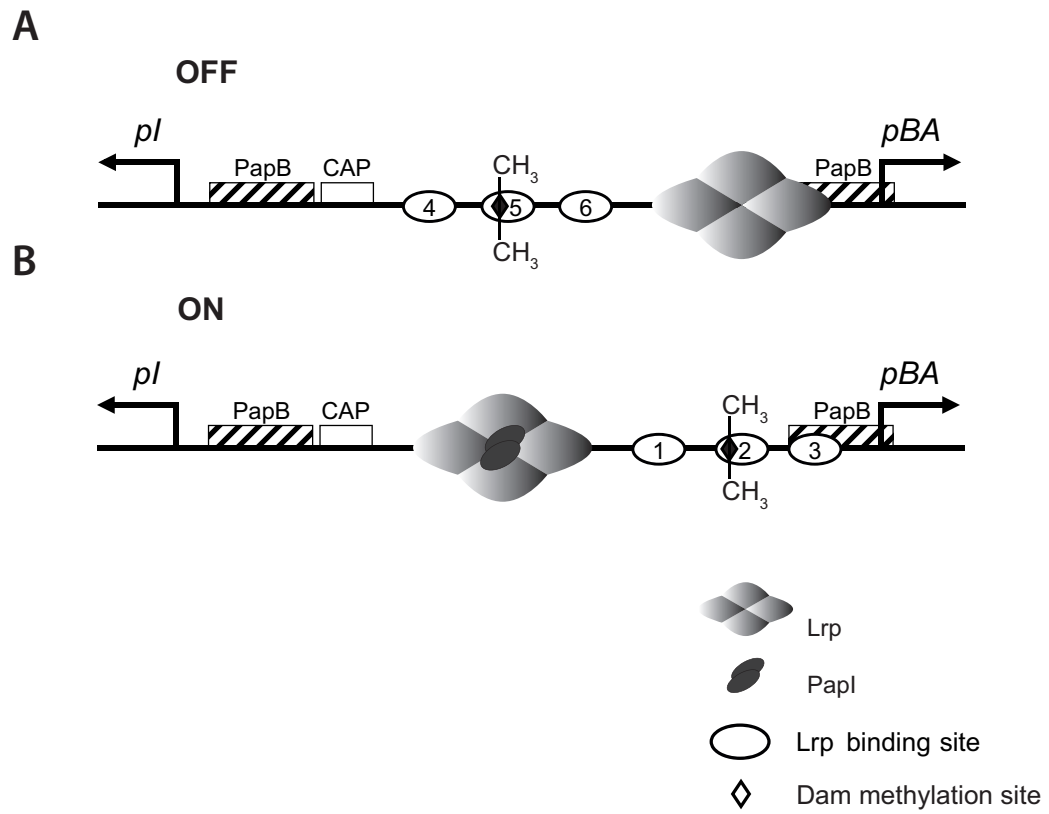


Figure 1

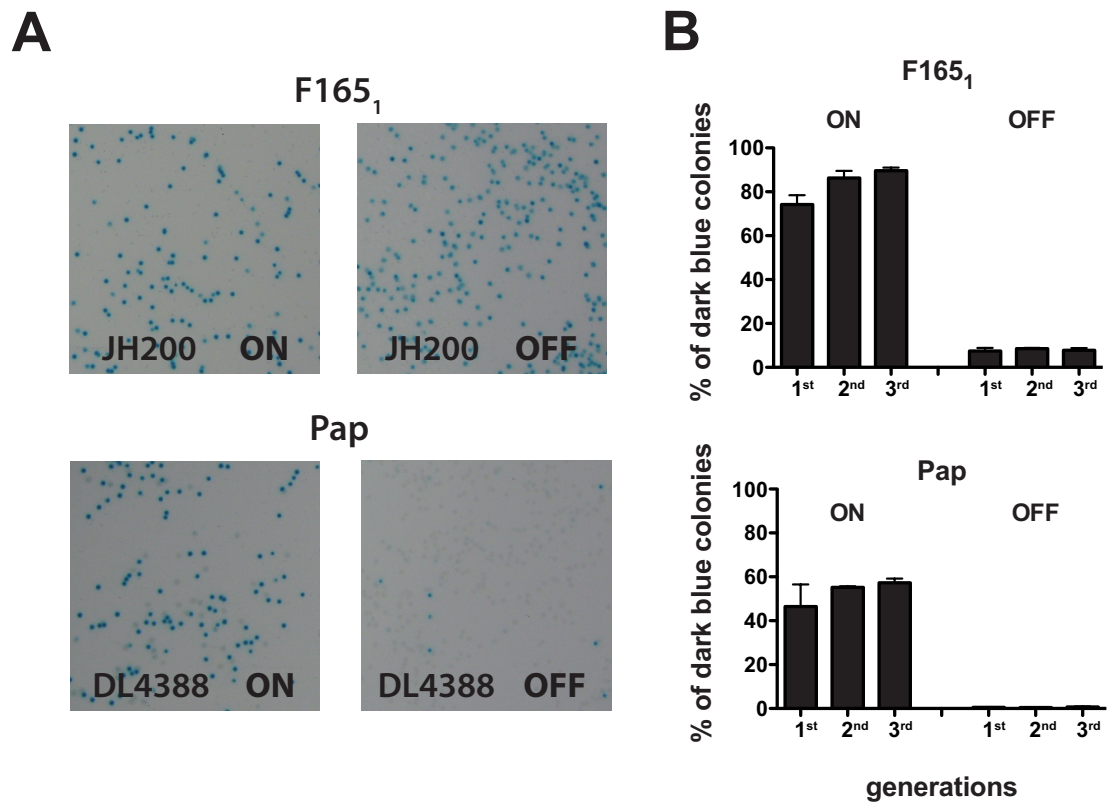


Figure 2

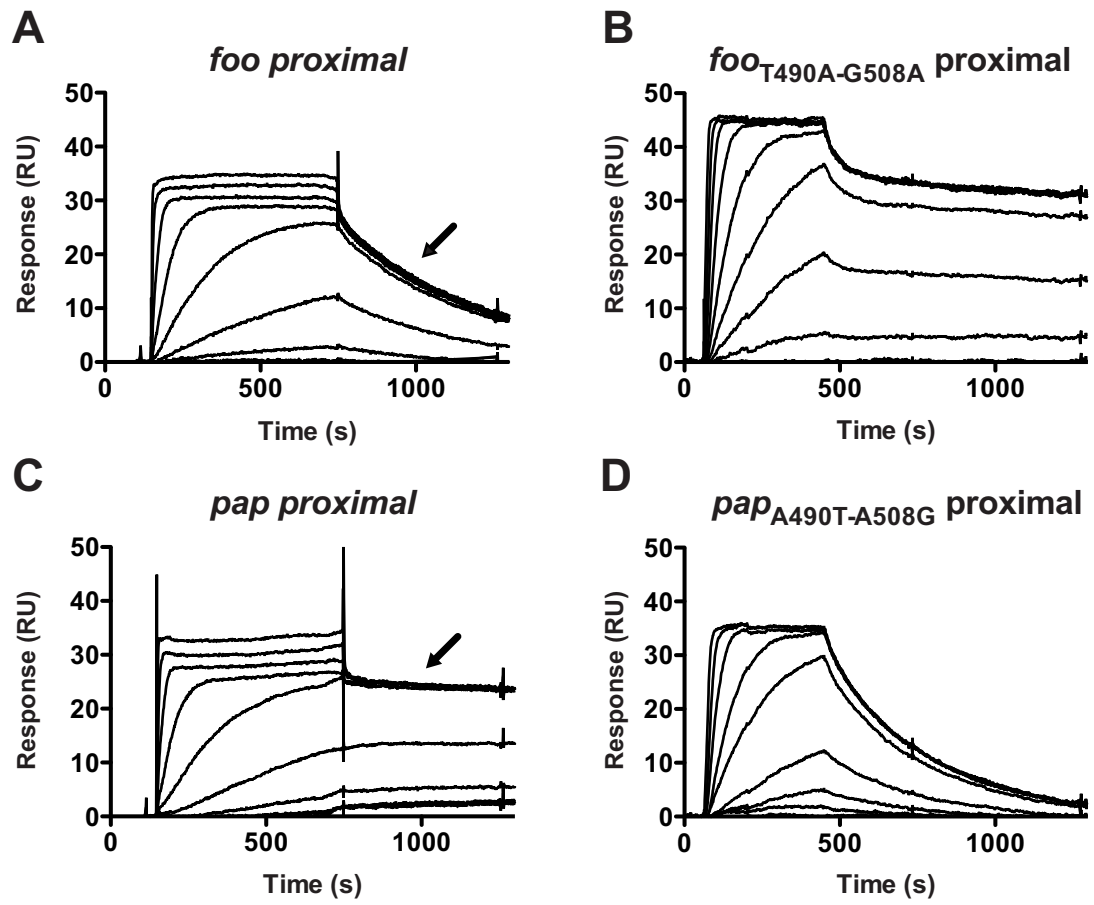


Figure 3

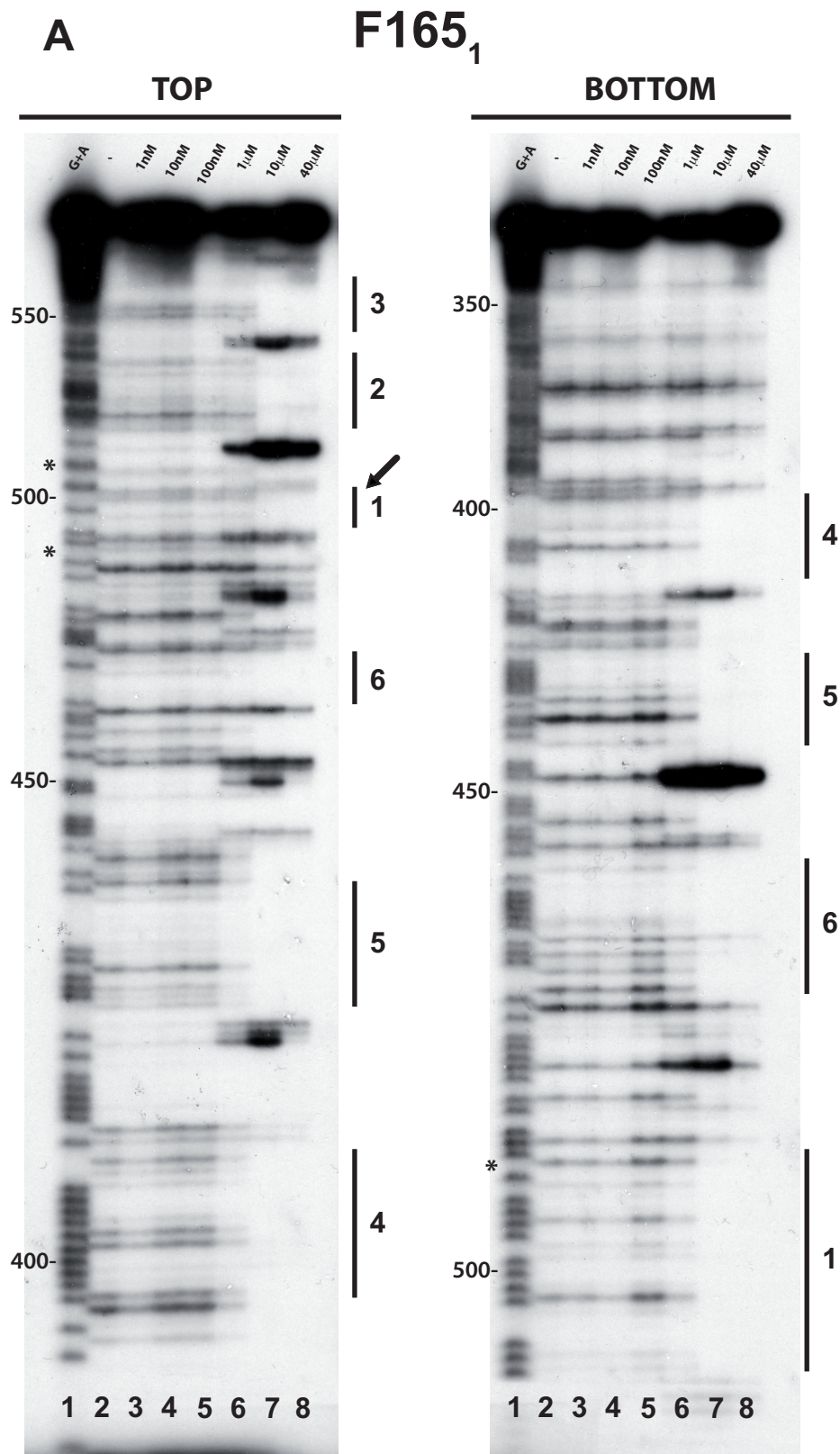


Figure 4A

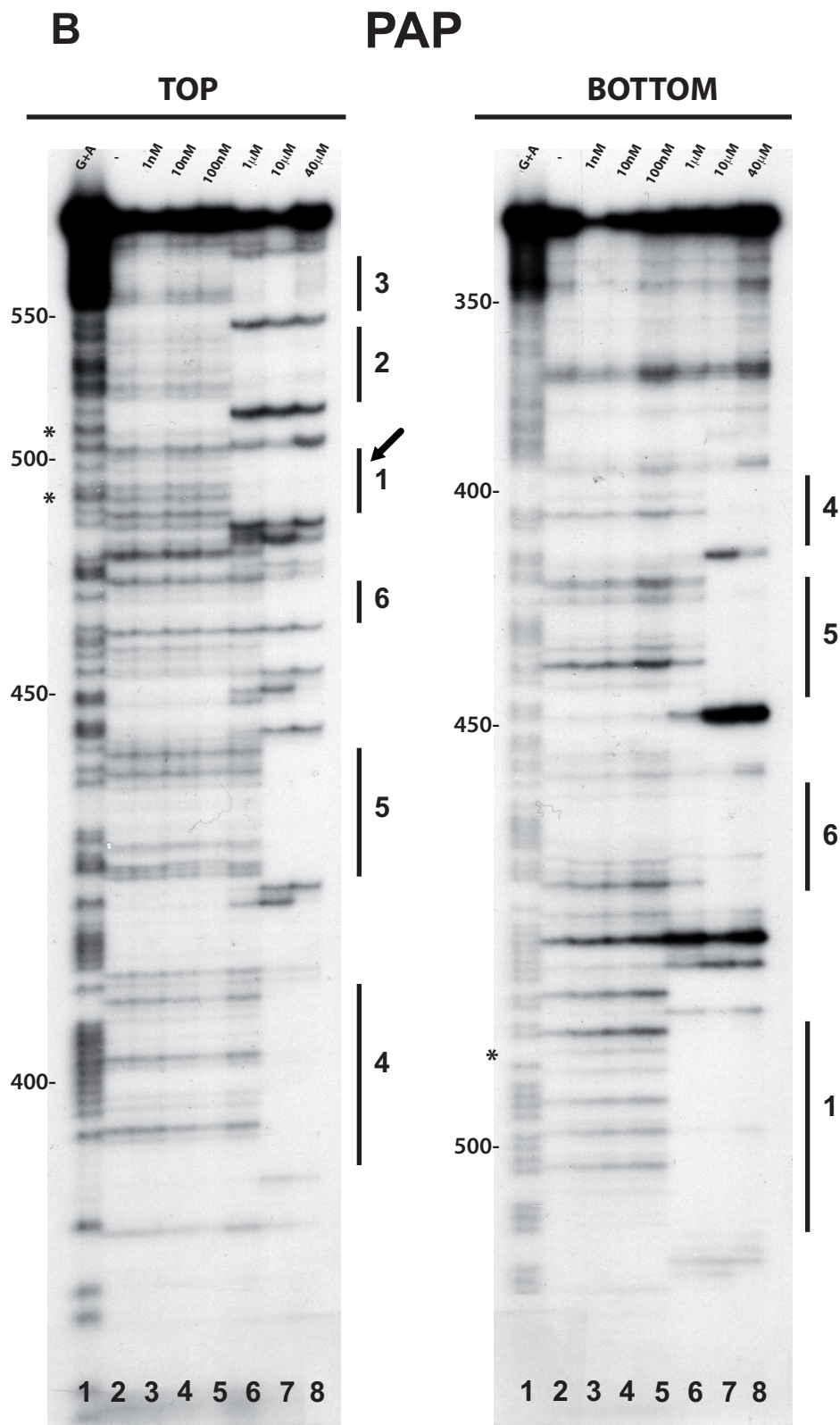


Figure 4B

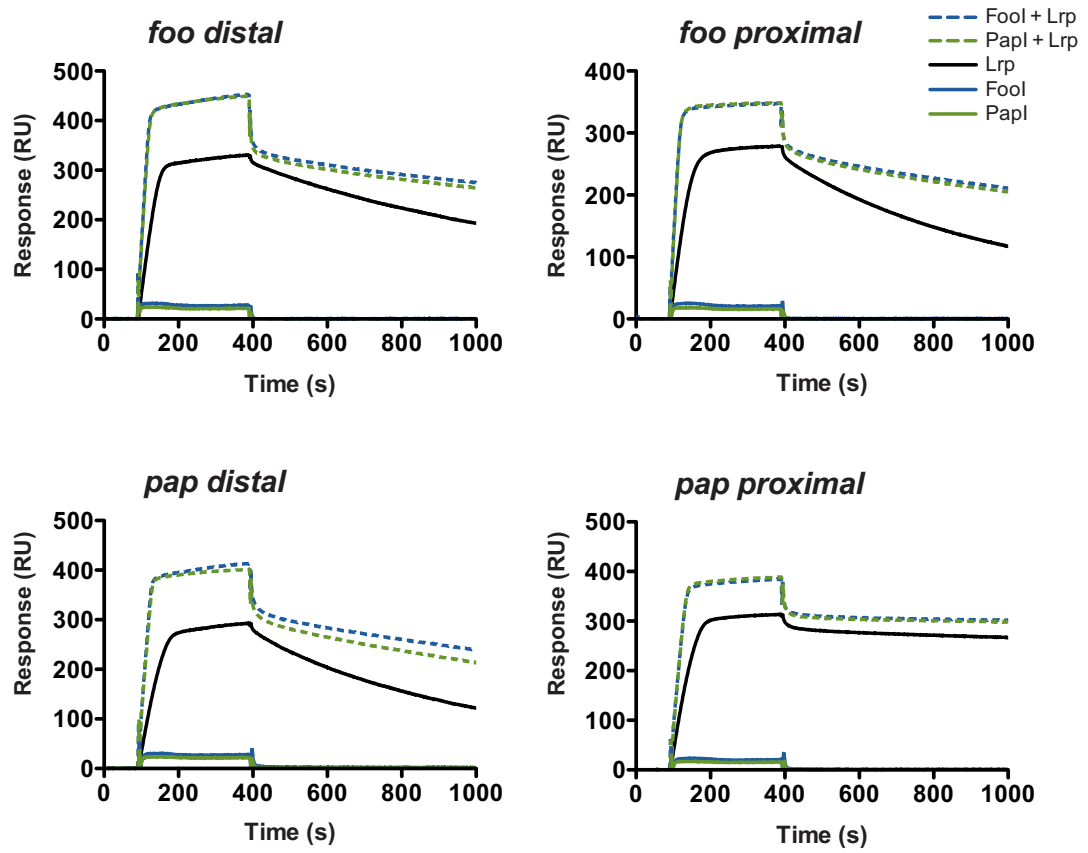


Figure 5

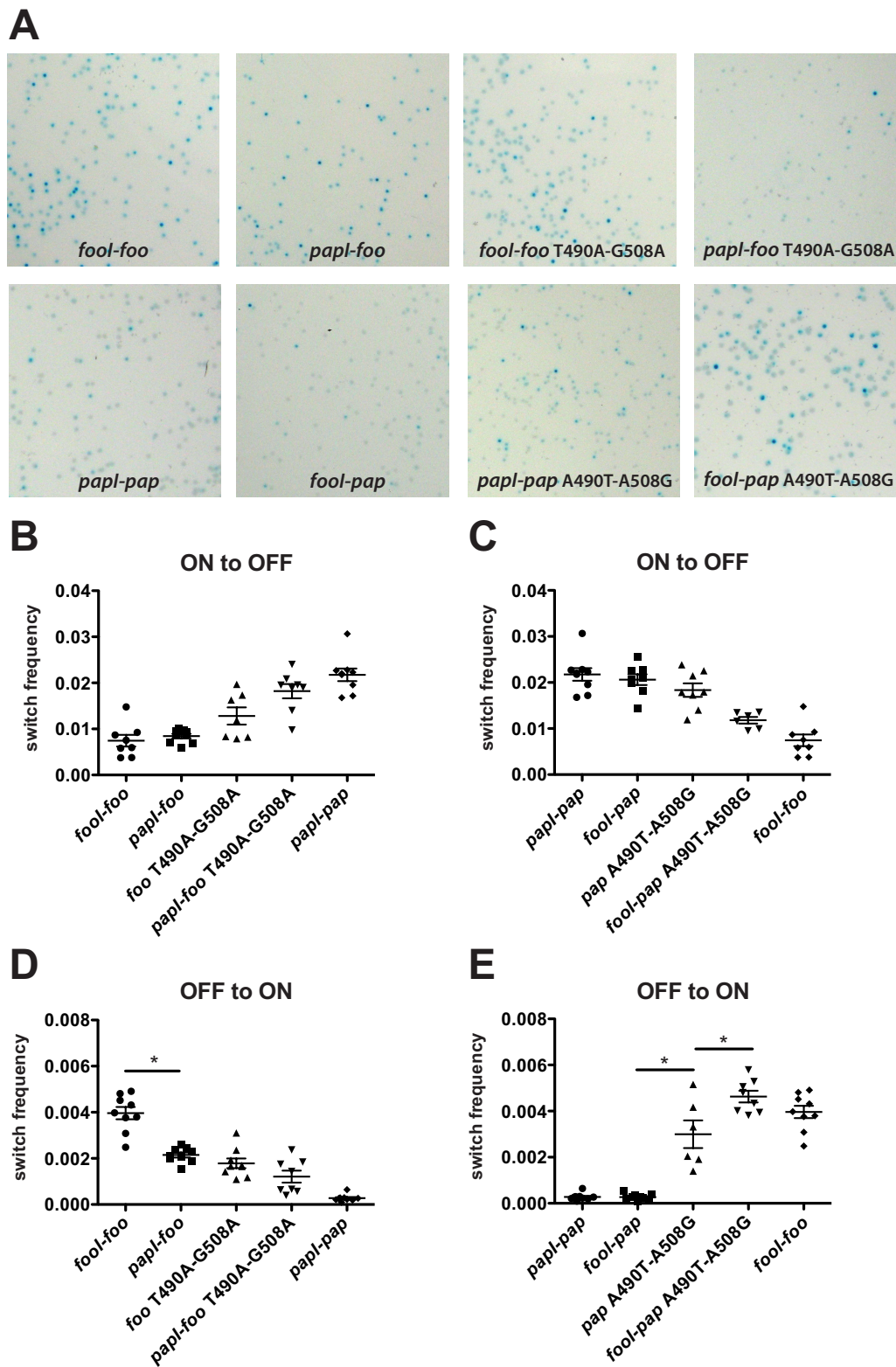


Figure 6

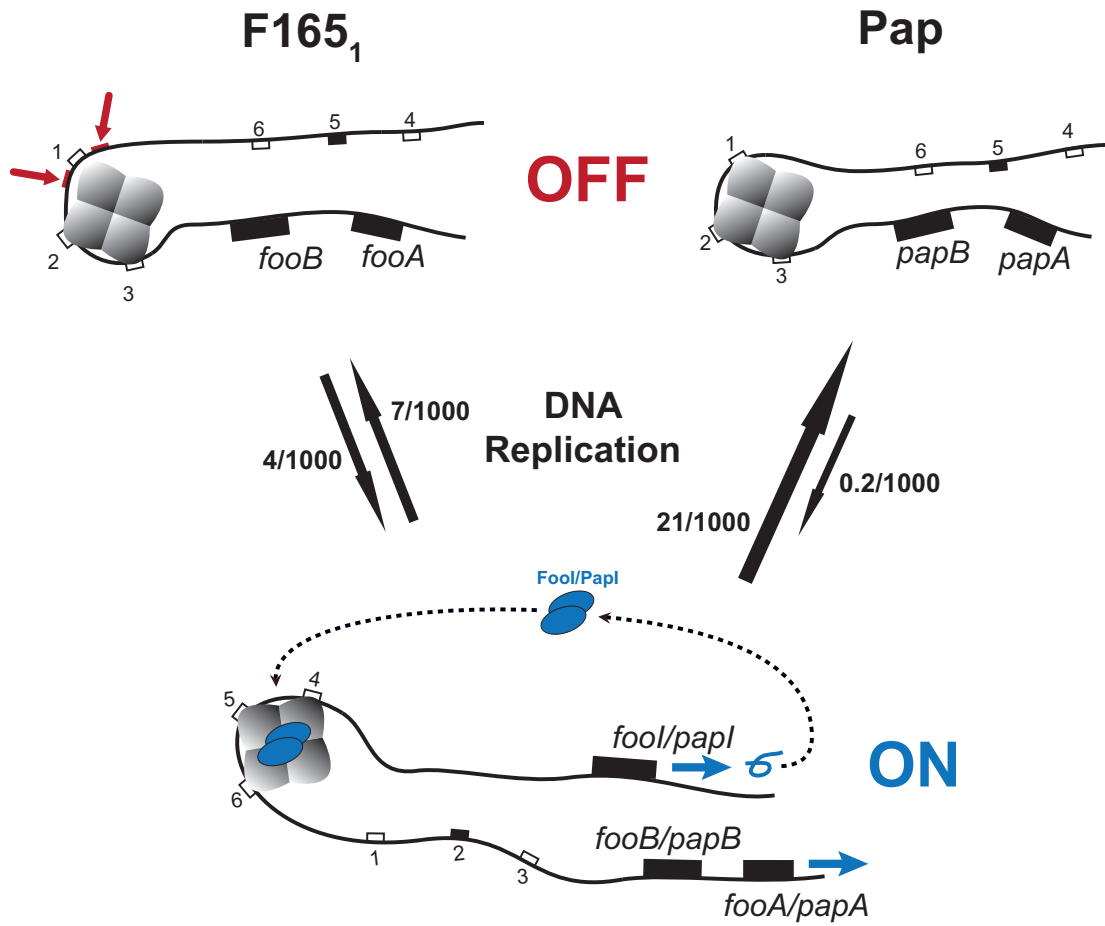


Figure 7

Table 1. Phylogenic and phenotypic characterization of ExPEC strains used in this study.

Strain or plasmid	Description and relevant phenotype	Source or reference
<i>E. coli</i> strains		
CV1494	<i>supE thi Δ(lac-proAB) ilvH::Mud11734 lrp-35::Tn10/F^r traD36 lac^r Δ(lacZ) M15 proA⁺B⁺/pCV294</i>	Matthews <i>et al.</i> (2000)
NC397	<i>lacI::kan-Ter-cat sacB-lacZYA</i>	Svenningsen <i>et al.</i> (2005)
BL21(DE3)pLysS	competent cell for protein expression and purification	Laboratory strain
DL4388	MC4100 λ354 lysogen (wild type <i>papI-pap-papBA-lacZ</i>)	Braaten <i>et al.</i> (1994)
JH100	DL4388 <i>papI::cat-sacB-papBA-lacZ</i>	This study
JH200	JH100 <i>fool-foo-papBA-lacZ</i>	This study
JH201	JH100 <i>papI-foo-papBA-lacZ</i>	This study
JH202	JH100 <i>fool-fooT490A-G508A-papBA-lacZ</i>	This study
JH203	JH100 <i>papI-fooT490A-G508A-papBA-lacZ</i>	This study
JH204	JH100 <i>fool-pap-papBA-lacZ</i>	This study
JH205	JH100 <i>papI-papA490T-A508G-papBA-lacZ</i>	This study
JH206	JH100 <i>fool-papA490T-A508G-papBA-lacZ</i>	This study
ECL2184	O15:KRVC383, clinical isolate (bovine neo-natal septicemia)	Ecl. Laboratory
ECL2186	O15:KRVC383, clinical isolate (bovine instestine)	Ecl. Laboratory
ECL2187	clinical isolate (bovine)	Ecl. Laboratory
ECL2214	O11:KF12, clinical isolate (bovine neo-natal septicemia and diarrhea)	Ecl. Laboratory
ECL2244	O115:KV165, clinical isolate (porcine atrophic rhinitis)	Ecl. Laboratory
ECL3337	O9:K32, clinical isolate (porcine)	Ecl. Laboratory
ECL3641	O9:K28, clinical isolate (porcine diarrhea)	Ecl. Laboratory
Plasmids		
pKM201	Lambda genes <i>red</i> and <i>gam</i> under the control of the IPTG-inducible Ptac promoter (AmpR)	Murphy <i>et al.</i> (2003)
pET32a-Fool	expression vector for production of the His6-Trx-Fool fusion protein	This study
pET32a-PapI	expression vector for production of the His6-Trx-PapI fusion protein	This study
pCR2.1-fool ^{T490A-G510A}	F ⁻ <i>mcrA Δ(mrr-hsdRMS-mcrBC) Φ80 lacZΔM15 Δlac X74 recA1 araD139 Δ(ara-leu)7697 galU galK rpsL (StrR) endA1 nupG</i>	This study
pCR2.1-pap ^{A490T-A510G}	F ⁻ <i>mcrA Δ(mrr-hsdRMS-mcrBC) Φ80 lacZΔM15 Δlac X74 recA1 araD139 Δ(ara-leu)7697 galU galK rpsL (StrR) endA1 nupG</i>	This study

Table 2. apparent kinetic rate constants for Lrp, Foli, PapI binding to foo and pap DNA fragments, as determined by SPR.

Fragment	Lrp		Foli	PapI
	Affinity	Dissociation	Affinity	Affinity
	(K_D , nM)	($k_d \times 10^{-4} \text{ s}^{-1}$)	(K_D , μM)	(K_D , μM)
<i>foo</i> distal	3.4 \pm 0.1	12 \pm 0.1	11.6 \pm 0.7	13.2 \pm 0.6
<i>foo</i> proximal	3.6 \pm 0.1	23 \pm 0.1	11.8 \pm 0.3	16.1 \pm 1.1
<i>foo</i> _{T490A-G508A}	3.2 \pm 0.1	2 \pm 0.1	N.D.	N.D.
<i>foo</i> distal-proximal	4.1 \pm 0.2	21 +/- 0.1	N.D.	N.D.
<i>pap</i> distal	4.7 \pm 0.2	32 \pm 0.2	17.1 \pm 0.2	16.1 \pm 2.0
<i>pap</i> proximal	3.3 \pm 0.2	3 \pm 0.1	13.6 \pm 0.6	18.8 \pm 2.0
<i>pap</i> _{T490A-G508A}	3.6 \pm 0.2	32 \pm 0.1	N.D.	N.D.
<i>pap</i> distal-proximal	5.3 \pm 0.4	23 +/- 0.2	N.D.	N.D.

SUPPORTING INFORMATION**Title :**

Lrp-DNA complex stability determines the level of ON cells in type P fimbriæ phase variation

Authors :

Richard Graveline¹, Michaël Mourez¹, Mark A. Hancock², Christine Martin³, Stéphanie Boisclair¹ and Josée Harel^{1*}

Addresses :

¹Groupe de Recherche sur les Maladies Infectieuses du Porc (GREMIP) and Centre de Recherche en Infectiologie Porcine (CRIP), Université de Montréal, Saint-Hyacinthe, Québec, CANADA, ²Sheldon Biotechnology Centre, McGill University, Montréal, Québec, Canada, and ³Institut National de la Recherche Agronomique, Saint-Genès-Champanelle, France

*Corresponding author. Mailing Address : Groupe de Recherche en Infectiologie Porcine (GREMIP), Université de Montréal, Faculté de Médecine Vétérinaire, C.P. 5000, Saint-Hyacinthe, Québec, Canada, J2S 7C6. Phone : (450) 773-8521, ext. 8233. Fax : (450) 778-8108. E-mail :

9- SUPPORTING INFORMATION

a) Supplemental Experimental Procedures

-i- Plasmid construction

For overexpression and purification of the Fool protein, the pET32a-fool plasmid was constructed by first amplifying the *fool* sequence from the chromosome of strain 4787, using the NcoI-foolrg1F and XhoI-foolrg1R primers (containing the NcoI and XhoI restriction sites, respectively). After enzymatic digestion, the fragment was ligated into the NcoI/XhoI restriction sites of the expression plasmid pET32a. Inserts were sequenced for verification after transformation of *E. coli* BL21(DE3)pLysS by electroporation. Plasmid pET32a-papI was constructed using the QuickChange Site-Directed Mutagenesis System (Stratagene). Briefly, expression plasmid pET32a-fool was first extracted from strain *E. coli* BL21(DE3)pLysS, purified and then used as a Template for amplification using the foolIN17D-F and foolIN17D-R primers. PCR amplified plasmids were purified and digested with the restriction enzyme DpnI to remove any trace of the pET32a-fool Template. PCR amplified plasmids were then dialyzed and transformed into strain *E. coli* BL21(DE3)pLysS by electroporation.

-ii- Purification of recombinant proteins

E. coli strain CV1494 was used to purify His6-tagged Lrp as previously described (Matthews *et al.*, 2000). A one-liter culture of CV1494 was grown in LB broth at 37°C, with shaking, to an OD600 of 0.7-0.8. Isopropyl β -D-1-thiogalactopyranoside (IPTG) was then added to the culture at a final concentration of 0.5mM for 6 hours. Bacteria were then harvested, resuspended in 30 mL of Tris-buffered saline 920 mM Tris-HCl [pH 7.4], 0.5 M NaCl), and lysed with a French press and an ultrasonic processor. The soluble fraction was

clarified by centrifugation at 15,000 g for 20 min at 4°C. The proteins were purified using the ÄKTA purifier system with a 1 mL HisTrap HP column (Amersham Biosciences) according to manufacturer instructions. Purified protein was subsequently dialyzed against TG50ED buffer (10 mM Tris-HCl [pH 8.0], 1 mM EDTA, 0.1 mM DTT, 0.1 M NaCl). Purity of the proteins was confirmed by SDS-PAGE and staining was achieved with Coomassie blue.

For purification of FooI and PapI, a single *E. coli* BL21(DE3)pLysS colony carrying plasmid pET32a-fooI or pET32a-papI was incubated in LB broth with shaking at 37°C, to an OD600 of 0.4-0.6. The culture was incubated for 3 more hours in the presence of IPTG at a final concentration of 1 mM. Bacteria were harvested, resuspended in 30 mL of Tris-buffered saline (20 mM Tris-HCl [pH 7.4], 0.5 M NaCl), and lysed with a French press and an ultrasonic processor. The soluble fraction was clarified by centrifugation at 15,000 g for 20 min at 4°C. The His6-Trx-FooI and His6-Trx-PapI fusion proteins were purified using an ÄKTA purifier system using the same procedure as for His6-tagged Lrp. FooI and PapI proteins were removed from the His6-Trx tag after enzymatic digestion using Enterokinase (New England BioLabs) in digestion buffer (20 mM Tris-HCl [pH 8.0], 50 mM NaCl, 2 mM CaCl₂). Proteins were then dialyzed against 20 mM 2-[N-morpholino]-ethanesulphonic acid (Mes), pH 6.0 and purified using an ÄKTA purifier system with a mono S 5/50 GL column (Amersham Biosciences) according to manufacturer instructions. Purity of FooI and PapI proteins was confirmed by SDS-PAGE and staining was completed with Coomassie blue.

-iii- Electrophoretic mobility shift assays

DNA-binding reactions for gel retardation were prepared as previously described. Six different DNA probes were used in gel retardation experiments. The digoxigenin-labeled DNA probes were generated by PCR using the Fidelity PCR MasterMix (GE Healthcare). Primers used are listed in Table S2. Each binding reaction containing 8 ng of DIG-labeled *foo* or *pap* DNA fragments and 3

µg of poly-dIdC (GE Healthcare), and the total volume were adjusted to 20 µL with binding buffer (60 mM Tris-HCl [pH 7.5], 100 mM NaCl, 40 mM KCl, 1 mM EDTA, 1 mM DTT, 12.5% glycerol). A range of 9 Lrp concentrations was tested. The reaction mixture was incubated at 4°C for 15 min and then loaded onto a 4% acrylamide non denaturing gel (1X TBE, 2.5% glycerol). Electrophoresis was carried out at 180 V for 50 min at 4°C. DNA was then transferred onto a nylon membrane (GE Healthcare). Detection of His6-Lrp in shifted DNA bands was revealed using an anti-poly histidine conjugated to alkaline phosphatase (Sigma Aldrich) after Western blot on nitrocellulose membrane (Biorad) with ECF Western blot detection reagent.

-iv- Multiple alignment and phylogenetic tree

A multiple alignment and a phylogenetic tree of 25 unique *pap*-like intergenic regions from 23 strains were performed using the Geneious alignment (global alignment with free ends gap) and the Geneious Tree Builder (« Jukes-Cantor Genetic Distance Model » with Neighbor-Joining method) programs from Geneious Pro v5.1.6 (Drummond *et al.*, 2010).

b) Supplementary Legends

Figure S1. Comparison of the nucleotide sequences of the regulatory regions of *foo* and *pap* operons. The putative promoter regions (-10 and -35 sites) are represented by dotted lines. The identification of *foo* Lrp binding sites (represented as boxed regions on the DNA sequences), PapB and CAP binding sites (underlined) is based on the strong similarity with the *pap* operon. The open reading frames of the two specific regulators *fooI/papI* and *fooB/papB* are translated into the corresponding amino acid sequences. Differences in either nucleotides or amino acids are depicted in red.

Figure S2. DNA-binding pattern of Lrp to either *foo* or *pap* intergenic region. Representative electrophoretic mobility shift analysis of His6-Lrp (lane 0, DNA alone; lane 1, Lrp at 5 nM; lane 2, Lrp at 10 nM; lane 3, Lrp at 20 nM; lane 4, Lrp at 40 nM; lane 5, Lrp at 80 nM; lane 6, Lrp at 160 nM; lane 7, Lrp at 300 nM; lane 8, Lrp at 600 nM; lane 9, Lrp at 1.6 μ M) in the presence of the distal, proximal or the entire distal-proximal intergenic region of *foo* and *pap* operons.

Figure S3. Association and dissociation kinetics of His6-Lrp binding to the *foo* and *pap* regions. Representative SPR analysis of Lrp (0-100 nM, 2-fold dilution series) injected over immobilized distal, proximal and distal-proximal fragments (<100 RU each) at 25 μ L/min (5 min association + 5 min dissociation).

Figure S4. Representative curve fitting on SPR data. Comparison of Lrp binding to the regulatory regions of *foo* and *pap* as analyzed using the « One site – Specific binding with Hill slope » model (Prism5 software).

Figure S5. Association and dissociation kinetics of FooI binding to the distal or proximal regions of *foo* and *pap* operons. Representative SPR analysis of FooI (0-15 μ M; 2-fold dilution series) binding to immobilized *foo*

(400 RU distal, 300 RU proximal) or *pap* (400 RU distal, 300 RU proximal) at 5 $\mu\text{L}/\text{min}$ (5 min association + 5 min dissociation).

Figure S6. Association and dissociation kinetics of PapI binding to the distal or proximal regions of *foo* and *pap* operons. Representative SPR analysis of PapI (0-15 μM ; 2-fold dilution series) binding to immobilized *foo* (400 RU distal, 300 RU proximal) or *pap* (400 RU distal, 300 RU proximal) at 5 $\mu\text{L}/\text{min}$ (5 min association + 5 min dissociation).

Table S1. Apparent dissociation rate constants for Lrp binding to *foo/pap* DNA fragments in the absence or presence of F_{oo}I and PapI, as determined by SPR. Estimates (\pm standard error of the mean; n=3) represent local analysis of each injection series using the « Fit separate k_a/k_d » tool in the BIAevaluation software.

Figure S7. Presence of Lrp binding sites within the intergenic region of Pap variants. Multiple sequence alignment of 25 intergenic regions found in 22 different ExPEC strains. Arrows indicate the presence of every GN(2-3)TTT motif found between the two divergent promoters *pl* and *pBA*. Orange arrows correspond to Lrp binding sites 4-6 and 1-3, based on the strong sequence similarity with the well-described *pap* operon (corresponding here to *pap17*). All intergenic regions have more than six GN(2-3)TTT motifs. While all Lrp binding sites 4-6 within the distal region are conserved, seven *pap* homologues lack one of the three GN(2-3)TTT binding sites within the proximal region. In addition, at least one of the two additional GN(2-3)TTT motifs surrounding Lrp binding site 1 is found in ten of the strains tested, including strain 5131 harboring *foo*. Thus, the number of GN(2-3)TTT motifs in the vicinity of Lrp binding sites might influence the dissociation kinetics of Lrp binding, and influence the fimbrial switch rate.

Table S2. Phylogenetic and phenotypic characterization of ExPEC strains used in this study.

Figure S8. DNA sequence of *pap*-like intergenic regions among various ExPEC strains correlates with phylogenetic analysis. An unrooted phylogenetic tree using the neighbor-joining method was generated after alignment of the nucleotide sequences of the 25 intergenic regions found. ExPEC strains are recognized as a phylogenetically and epidemiologically distinct group of *E. coli* pathogens that are responsible for a large Spectrum of infections (Dobrindt & Hacker, 2008; Smith *et al.*, 2007), and can be grouped into four main phylogenetic groups. Virulent ExPEC strains mainly clustered into groups B2 and D, whereas commensal or opportunistic strains belonged to phylogenetic groups B1 and A (Clermont *et al.*, 2000). The colors indicate the four phylogenetic groups to which the strains belonged (green=A; blue=B1; red=B2; yellow=D). Group A strains clustered together, as did group B1 strains. Group B2 strains were found to cluster into more than one subgroup and group D strains did not cluster. Since the nucleotide sequence within the intergenic region influences the phase variation rate, clustering analysis indicates that phylogenetic associated strains may have developed similar mechanisms of phase variation, and less pathogenic strains of groups A and B1 may be more prone to increasing their OFF to ON Switch frequency.

Table S3. Primers used in this study.

```

1      10     20     30     40     50     60     70     80     90     100    110    120
foo   TTAAGTTGTGGAAGAACAGCTTTGCCCGCCTGCTTCTCTCCTTTCAGAAACCAGTATGTTGCCATGCCCGTCTTAATGGTGAGCGCTGAACCATACCTGCTTTTCCAGTAATAACAG
* T T S S C S Q G A Q K E G K L F W Y T A M G R R L P S R Q V M G A K E L L L L
pap17 TTAAGTTGTGGAAGAACAGCTTTGCCCGCCTGCTTCTCTCCTTTCAGAAACCAGTATGTTGCCATGCCCGTCTTAATGGTGAGCGCTGAACCATACCTGCTTTTCCAGTAATAACAG
* T T S S C S Q G A Q K E G K L F W Y T A M G R R L P S R Q V M G A K E L L L L

130    140    150    160    170    180    190    200    210    220    230    240
foo   GTAA TAGCGGCGCTGGTAATCCGTTACCGCAGCGCCTCTGCAATTTCTGCAAGTTTCCCTCCATTAATGCCTGTCAGAAATTCAGTATTTCACTCTCATGATTCACTCATCTCACT
Y Y R A Q Y D T V A L A E A I E A T K G G N H R N L F E L I E N K M Y E S M
pap17 GTAA TAGCGGCGCTGGTAATCCGTTACCGCAGCGCCTCTGCAATTTCTGCAAGTTTCCCTCCATTAATGCCTGTCAGAAATTCAGTATTTCACTCTCATGATTCACTCATCTCACT
Y Y R A Q Y D T V A L A E A I E A T K G G D H R N L F E L I E N K M Y E S M

250    260    270    280    290    300    310    320    330    340    350    360
foo   GTAACAAAGTTTCTTCGAAATAATAAAAATCATGCTCTCTGTTATCAACGGAAAGGTATTTTATTTCTCTGATTTGCTTTATTTGCGAAATTTAGTGAATTTGCTTTTCTTTGGCTTTAT
-10 -35 PapB site 1
pap17 GTAACAAAGTTTCTTCGAAATAATAAAAATCATGCTCTCTGTTATCAACGGAAAGGTATTTTATTTCTCTGATTTGCTTTATTTGCTCAATTTAGTGAATTTGCTTTTATTTGGATTTAT

370    380    390    400    410    420    430    440    450    460    470    480
foo   TTGATGTGTGCACATTTTGTGTTATTTTATGTAAGAAAGGAAAGTCCATAAAATCATTTAGACGATCTTTATGCTGTAAATCAATTCACCATGATGTTTTATCTGAGTGATTT
CAP binding site 4 5 6
pap17 TTGATGTGTGCACATTTTGTGTTATTTTCTGCGAAAGGAAAGTCCATAAAATCATTTAGACGATCTTTATGCTGTAAATCAATTTGCCATGATGTTTTATCTGAGTACCCCT

490    500    510    520    530    540    550    560    570    580    590    600
foo   CTTGTTGTTGTTTTATTTGTTAGTTTGAATTTTGTGTAAGTAAAGATCGTTTAAATTAATATTTTCAAATATAAAACATAAAACATAAAATTAACTTATGCTGAAGAGTATTTCCGGGCC
1 2 -35 3 -10 PapB site 2
pap17 CTTGCTATTAATGTTTTGTTCTGTTAATTTTGTGTAAGTAAAGATCGTTTAAATTAATATTTTCAAATATAAAACATAAAACATAAAATTAACTTATGCTGAAGAGTATTTCCGGGCC

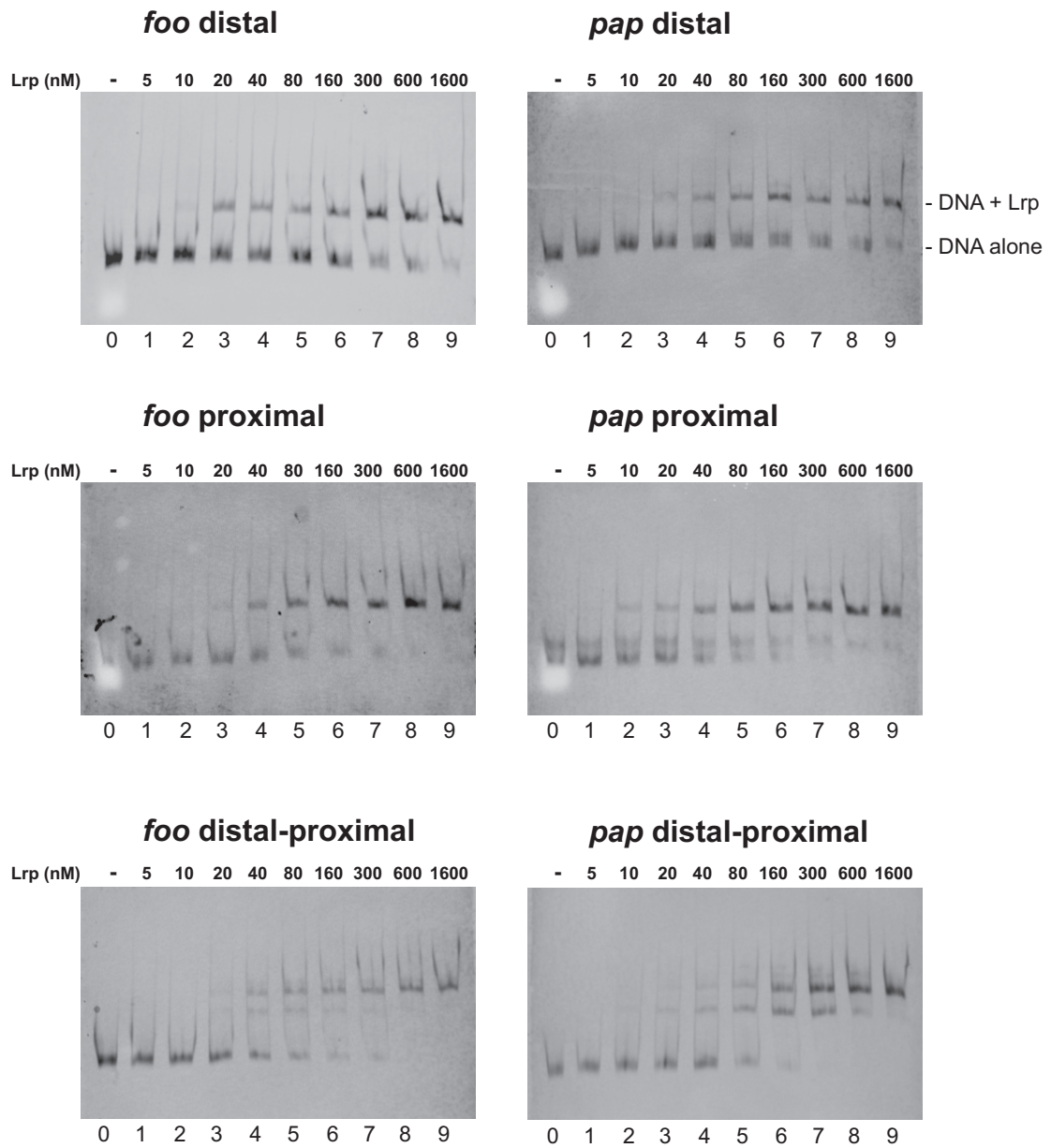
610    620    630    640    650    660    670    680    690    700    710    720
foo   GGAA GCATATATTCAGGGGCCCGACAGAAAGGGGAAACATGGCCCATCATGAAATCATGTCAGTCAATCAGGAAATGTTGTTTGC TGAATAACCGGAGAGCGTACTGTTGCCCGCTCTA
PapB site 3 M A H H E V I S Q S G N V F L L N I R E S V L L P G S
pap17 GGAA GCATATATTCAGGGGCCCGACAGAAAGGGGAAACATGGCCCATCATGAAATCATGTCAGTCAATCAGGAAATGCGTTTGC TGAATAACCGGAGAGCGTACTGTTGCCCGCTCTA
M A H H E V I S R S G N A F L L N I R E S V L L P G S

730    740    750    760    770    780    790    800    810    820    830    840
foo   TGTCTGAAATGCAATTTTTTTTACTGATAGGTATTTCTTCTATTACAGTGCAGGGTCAATTTCTGGCTATGAAAGACTATCTGGTAGTGGGCACTCCCGTAAGGAGGCTGCGAGAAAT
M S E M H F F L L I G I S S I H S D R V I L A M K D Y L V G G H S R K E V C E K
pap17 TGTCTGAAATGCAATTTTTTTTACTGATAGGTATTTCTTCTATTACAGTGCAGGGTCAATTTCTGGCTATGAAAGACTATCTGGTAGTGGGCACTCCCGTAAGGAGGCTGCGAGAAAT
M S E M H F F L L I G I S S I H S D R V I L A M K D Y L V G G H S R K E V C E K

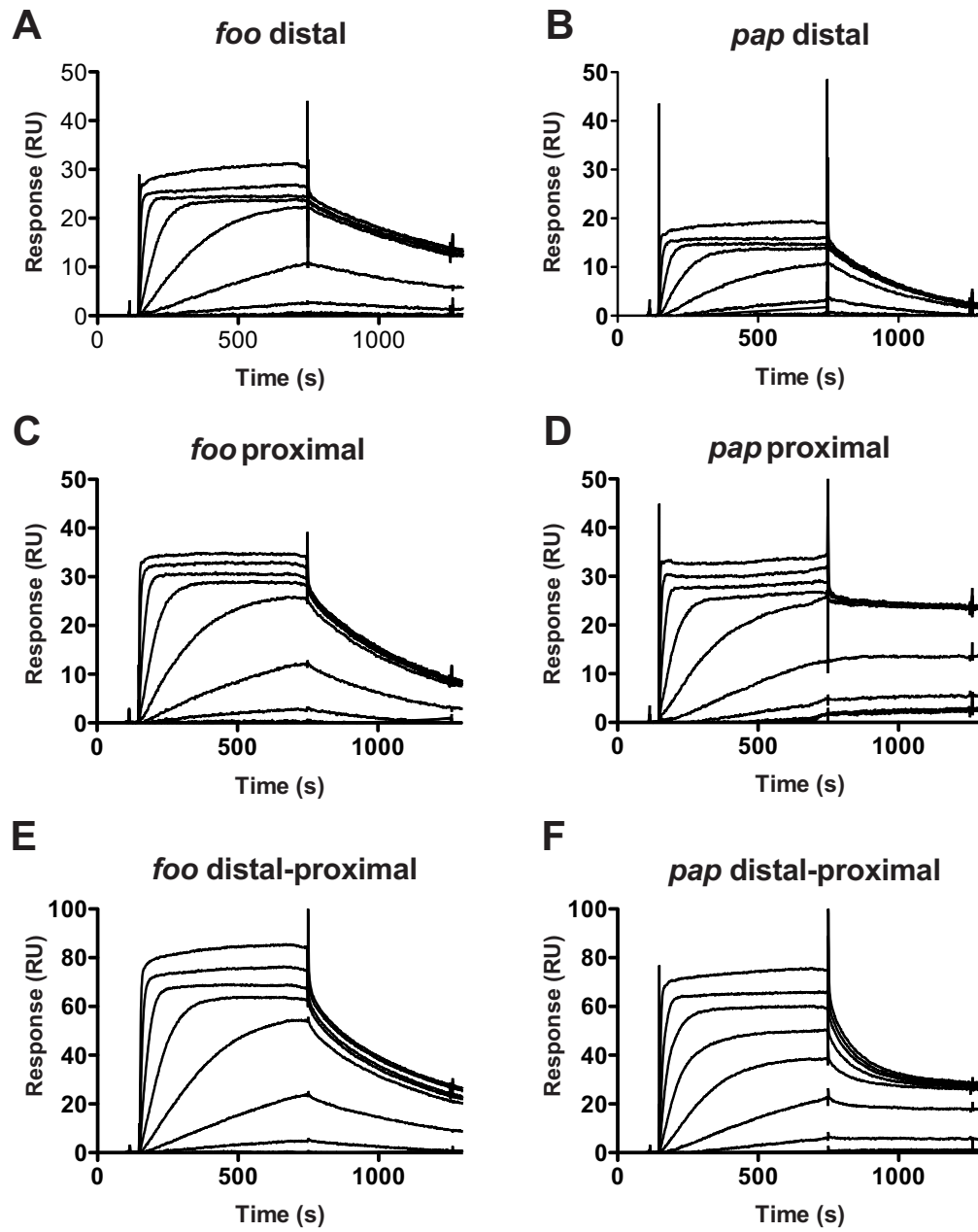
850    860    870    880    890    900    910    920    930    940    950 953
foo   ACCAGATGAAATAATGGGTATTTTCAGTACAACCGCTGGGGAGACTTATACGGCTGAATGCTCTGTCAGCAAGGCTGCACCTTATTATACAGATGAGTCTCGGCAATTTGACTAA
Y Q M N N G Y F S T T L G R L I R L N A L A A R L A P Y Y T D E S S A F D *
pap17 ACCAGATGAAATAATGGGTATTTTCAGTACAACCGCTGGGGAGACTTATACGGCTGAATGCTCTGTCAGCAAGGCTGCACCTTATTATACAGATGAGTCTCGGCAATTTGACTAA
Y Q M N N G Y F S T T L G R L I R L N A L A A R L A P Y Y T D E S S A F D *

```

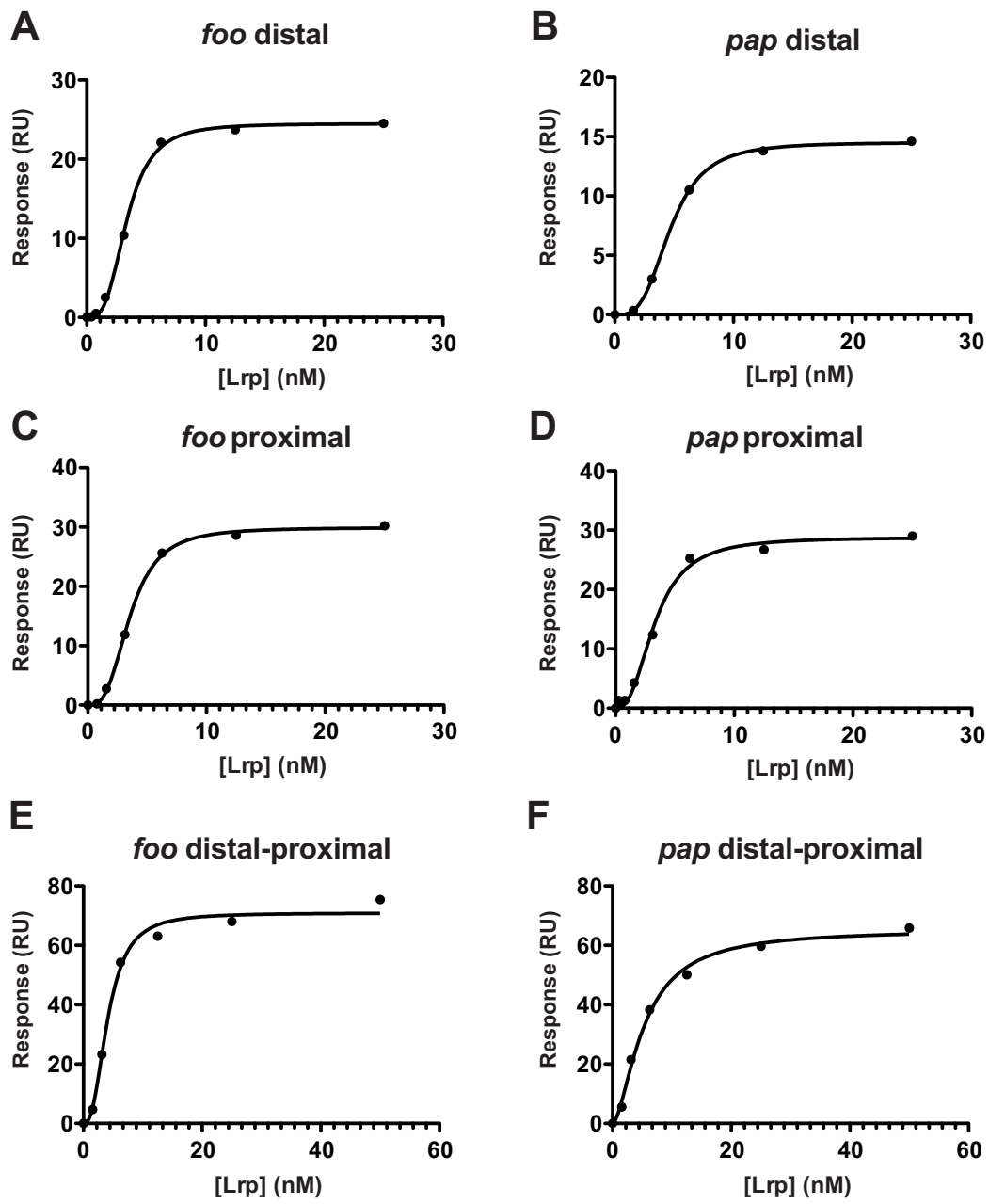
Supplemental figure S1



Supplemental figure S2

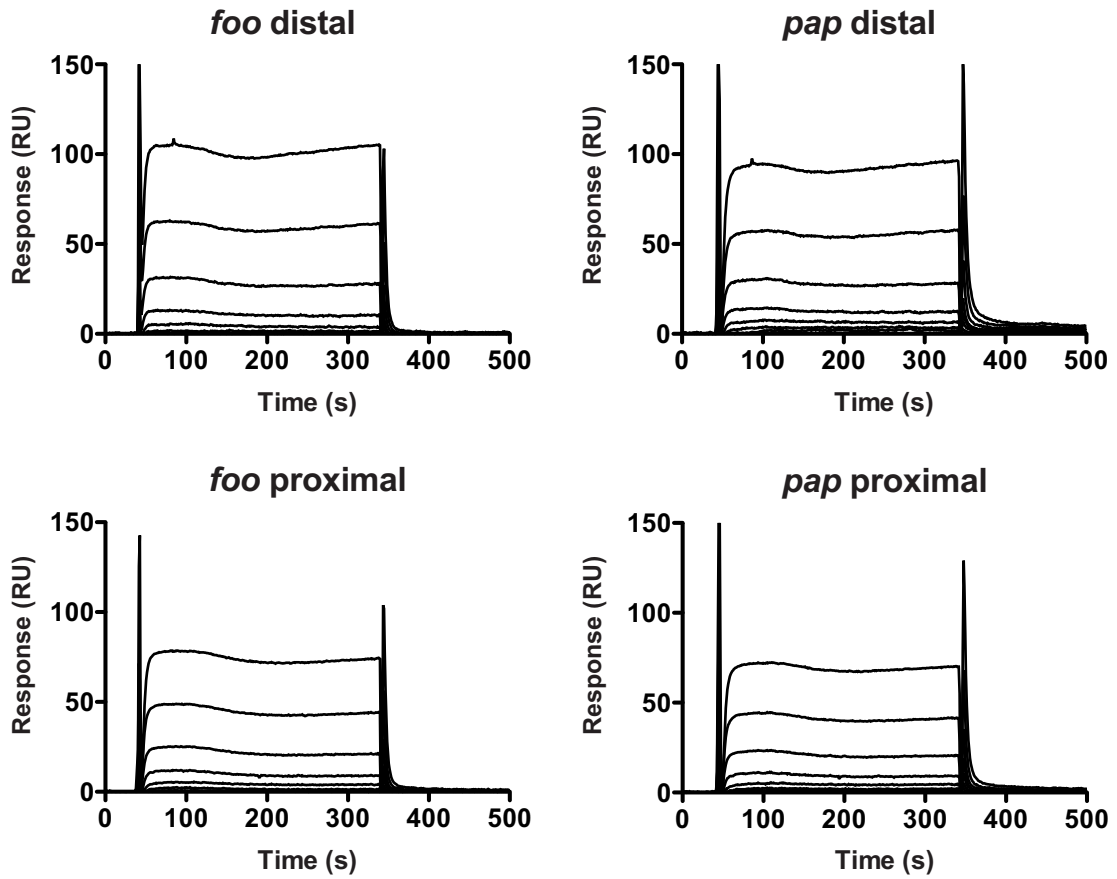


Supplemental figure S3



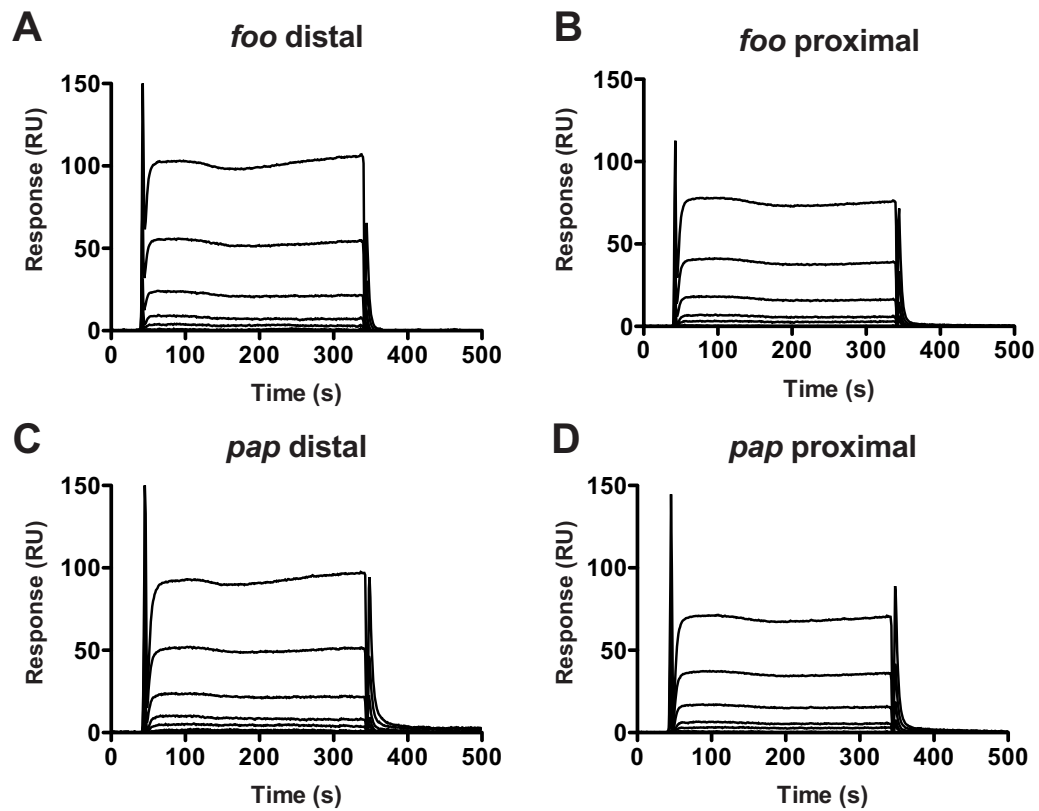
Supplemental figure S4

Fool



Supplemental figure S5

PapI

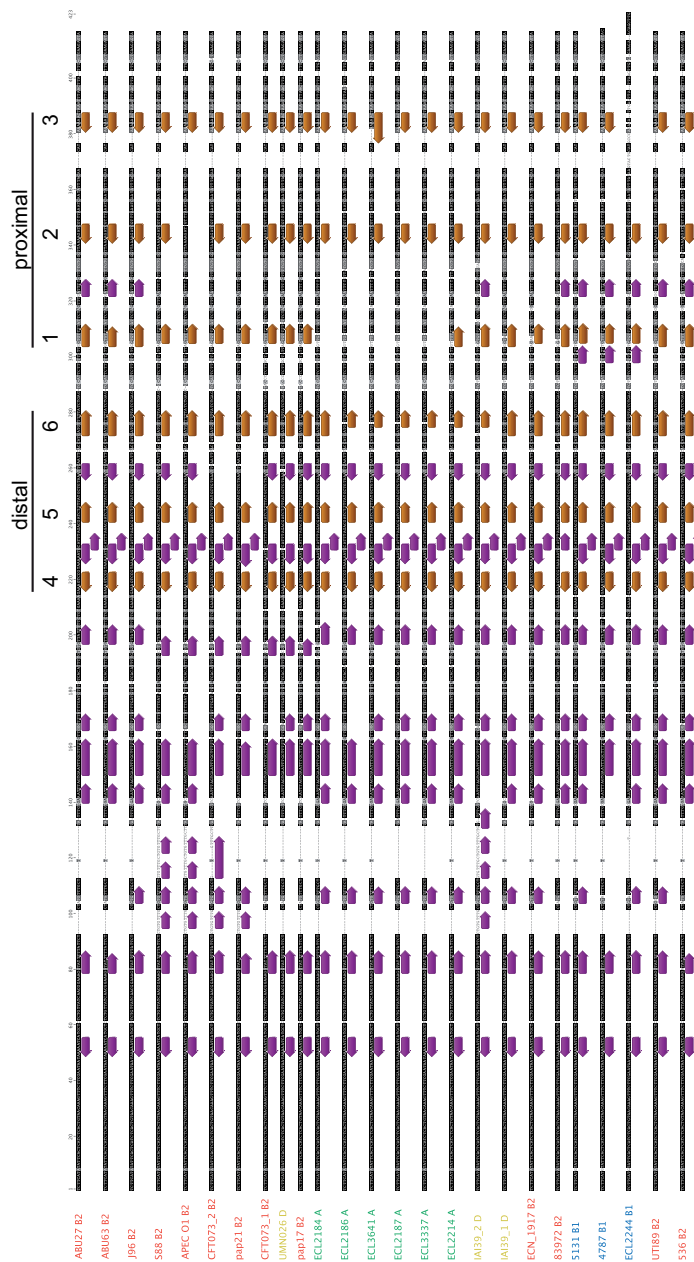


Supplemental figure S6

Table S1. Apparent dissociation rate constants for Lrp binding to *foo/pap* DNA fragments in the absence of presence of Fool and PapI, as determined by SPR.

Fragment	50 nM Lrp*	+ 6 μ M Fool	+ 8 μ M PapI
	($k_d \times 10^{-4} \text{ s}^{-1}$)	($k_d \times 10^{-4} \text{ s}^{-1}$)	($k_d \times 10^{-4} \text{ s}^{-1}$)
<i>foo distal</i>	7.8 \pm 0.01	4.0 \pm 0.01	4.1 \pm 0.01
<i>foo proximal</i>	12.8 \pm 0.02	5.6 \pm 0.01	4.9 \pm 0.01
<i>pap distal</i>	13.1 \pm 0.01	5.8 \pm 0.01	6.3 \pm 0.01
<i>pap proximal</i>	0.94 \pm 0.01	0.55 \pm 0.01	0.46 \pm 0.01

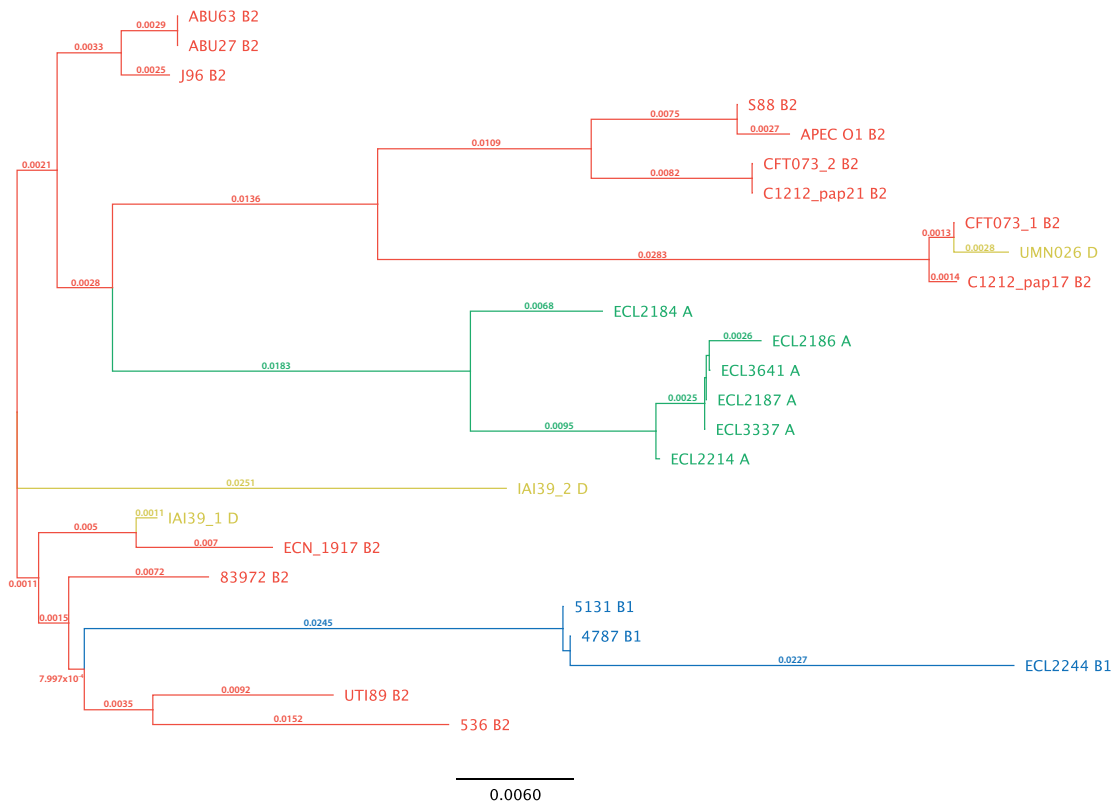
* Compared to the Lrp dissociation rate constants noted in Table 1 (i.e. optimal fitting with low-density oligo surfaces (<100 RU) and global analysis of the multi-concentration series), higher-density oligo surfaces (300 – 400 RU) were required to detect the binding of Fool or PapI alone (Fig. S5 or S6, respectively). For the subsequent single-concentration injection series (Fig. 5), local analysis yielded modestly different Lrp dissociation rate estimates. Regardless of the absolute values noted in Table S2, general trends in the Lrp-only dissociation rates remained unaltered: e.g. *foo proximal* > *foo distal*, *pap distal* > *pap proximal*. Except for *pap proximal*, dissociation rates for the fixed, 50 nM Lrp injections were significantly slower in the presence of 6 μ M Fool or 8 μ M PapI.



Supplemental figure S7

Table S2. Phylogenic and phenotypic characterization of strains used in this study.

Strain	Phylogenetic Group	Clinical conditions (Pathotype)	Accession number	Reference
ECL2184	A	Septicemia	This study	(Dezfulian <i>et al.</i> , 2003)
ECL2186	A	Diarrhea	This study	(Dezfulian <i>et al.</i> , 2003)
ECL2187	A	Nd	This study	(Dezfulian <i>et al.</i> , 2003)
ECL2214	A	Septicemia and diarrhea	This study	(Dezfulian <i>et al.</i> , 2003)
ECL3337	A	Nd	This study	(Dezfulian <i>et al.</i> , 2003)
ECL3641	A	Diarrhea	This study	(Dezfulian <i>et al.</i> , 2003)
4787	B1	Diarrhea	L07420	(Harel <i>et al.</i> , 1992, Dezfulian <i>et al.</i> , 2003)
5131	B1	Diarrhea	AF109675	(Harel <i>et al.</i> , 1992, Dezfulian <i>et al.</i> , 2003)
ECL2244	B1	Diarrhea	This study	(Dezfulian <i>et al.</i> , 2003)
ABU 63	B2	Asymptomatic bacteremia	AM690766	(Zdziarski <i>et al.</i> , 2008)
ABU 27	B2	Asymptomatic bacteremia	AM690765	(Zdziarski <i>et al.</i> , 2008)
J96	B2	Pyelonephritis	X61239	(Minshew <i>et al.</i> , 1978, Bingen-Bidois <i>et al.</i> , 2002)
S88	B2	New born meningitis	CU928161	(Touchon <i>et al.</i> , 2009)
APEC O1	B2	Colisepticemia	NC_008563	(Johnson <i>et al.</i> , 2007, Touchon <i>et al.</i> , 2009)
CFT073	B2	Pyelonephritis	AE014075	(Moblely <i>et al.</i> , 1990, Bingen-Bidois <i>et al.</i> , 2002)
C1212	B2	Pyelonephritis	X14471 ; X14472	(Low <i>et al.</i> , 1988)
ECN_1917	B2	Commensal strain	AJ586888	(Grozdanov <i>et al.</i> , 2004, Toh <i>et al.</i> , 2010)
83972	B2	Asymptomatic bacteremia	DQ010312	(Andersson <i>et al.</i> , 1991, Zdziarski <i>et al.</i> , 2008)
UTI89	B2	Cystitis	CP000243	(Touchon <i>et al.</i> , 2009)
536	B2	Pyelonephritis	CP000247	(Berger <i>et al.</i> , 1982, Bingen-Bidois <i>et al.</i> , 2002)
UMN026	D	Cystitis	CU928163	(Touchon <i>et al.</i> , 2009)
IAI39	D	Pyelonephritis	CU928164	(Touchon <i>et al.</i> , 2009)



Supplemental figure S8

Table S3. Primers used in this study.

Name	Function	Sequence
foopB_rvs	Sequencing of the intergenic regions	ACTTCATGATGCGCCATGTT TCCC
pI/fli-fwd	Amplify the regulatory region for replacing the <i>cat-sacB</i> cassette using the lambda red integration system.	CCTGGTAATCCGTTACCGCC AG
fool_fwd	Confirmation and sequencing of the intergenic region. Used with foopB_rvs as reverse primer.	AGTTGTGGAAGAACAGCTTT GCCC
fooUB40	Amplify the regulatory region for replacing the <i>cat-sacB</i> cassette using the lambda red integration system. Used with fool/papI_fwd as forward primer.	TACTCTTCACGCAATAAGTT AAATTTAGTTTTTTATGTTG
fooUB	Reverse primer used for the synthesis of DNA containing the <i>foo</i> proximal region.	TACTCTTCACGCAATAAG
FooDI	Forward primer used to amplify both the <i>foo</i> distal and the <i>foo</i> distal-proximal region.	GTATTTGCTTTATTTGCG
FooR1	Reverse primer used to amplify the <i>foo</i> distal region.	CAAACAACAAGAATACAC
FooR2	Forward primer used to amplify the <i>foo</i> proximal region.	GTGTATTCTTGTTGTTTG
FooUB-sacB	Amplify the <i>cat-sacB</i> cassette to insert it into strain DL4388 using the lambda red integration system. Used with i-cat2 as forward primer.	TACTCTTCACGCAATAAGTT AAATTTAGTTTTTTATGTTG ATCAAAGGGAAAACGTCCA TATGCACAGATG
i-cat_2	Amplify the <i>cat-sacB</i> cassette to insert it into strain DL4388 using the lambda red integration system.	CCTGGTAATCCGTTACCGCC AGCGCCTCTGCAATTTCTGC AAAATGAGACGTTGATCGGC ACGT
Pap4-6F	Forward primer used to amplify both the <i>pap</i> distal and the <i>pap</i> distal-proximal region.	TTCTCTATGTTTGCTTTATT TGTT
Pap4-6R	Reverse primer used to amplify the <i>pap</i> distal region.	TAATAGCAAGAGGGTACTCA GATA
Pap1-3F	Forward primer used to amplify the <i>pap</i> proximal region.	TATCTGAGTACCCTCTTGCT ATTA

Pap1-3R	Reverse primer used for the synthesis of DNA containing the <i>pap</i> proximal region.	TACTCTTCACGCAATAAGTT AAAT
FooDI_5PCBio	Biotinylated forward primer used to amplify both the <i>foo</i> distal and the <i>foo</i> distal-proximal region for SPR experiments.	/5PCBio/GTATTTGCTTTAT TTGCG
FooR2_5PCBio	Biotinylated forward primer used to amplify the <i>foo</i> proximal region for SPR experiments.	/5PCBio/GTGTATTCTTGTT GTTTG
Pap4-6F_5PCBio	Biotinylated forward primer used to amplify both the <i>pap</i> distal and the <i>pap</i> distal-proximal region for SPR experiments.	/5PCBio/TTCTCTATGTTTG CTTTATTTGTT
Pap1-3F_5PCBio	Biotinylated forward primer used to amplify the <i>pap</i> proximal region for SPR experiments.	/5PCBio/TATCTGAGTACCC TCTTGCTATTA
FooT490A-G508A_5PCBio	Biotinylated forward primer used to amplify the <i>foo</i> proximal region containing the two mutations at position 490 and 508 for SPR experiments.	/5PCBio/TATCTGAGTGTAT TCTTGTTGTTA
PapA490T-A508G_5PCBio	Biotinylated forward primer used to amplify the <i>pap</i> proximal region containing the two mutations at position 490 and 508 for SPR experiments.	/5PCBio/ GTACCCTCTTGCTATTTG
Fool->PapI-F	Forward primer used for replacing the <i>cat-sacB</i> cassette using the lambda red integration system and to exchange <i>fool</i> with <i>papI</i> .	AATCCGTTACCGCCAGCGCC TCTGCAATTTCTGCCGTTTT CCCTCCATCATGCCTGTTCA GAAATTCC
PapI->Fool-F	Forward primer used for replacing the <i>cat-sacB</i> cassette using the lambda red integration system and to exchange <i>papI</i> with <i>fool</i> .	AATCCGTTACCGCCAGCGCC TCTGCAATTTCTGCAGTTTT CCCTCCATTATGCCTGTTCA GAAATTCC
NcoI-foolrg1F	Ligation of <i>fool</i> into the NcoI/XhoI restriction sites of the expression plasmid pET32a	CCCCCATGGCTATGAGTGA ATACATGAAGAA
XhoI-foolrg1R	Ligation of <i>fool</i> into the NcoI/XhoI restriction sites of the expression plasmid pET32a	GGGGCTCGAGTTTAAGTTGT GGAAGAACAGCT
foolIN17D-F	Construction of the expression plasmid pET32a-papI from pET32a-fool	TTCTGCAGTTTTCCCTCCAT CATGCCTGTTCCAGAAATTCC

		A
fooIN17D-R	Construction of the expression plasmid pET32a-papI from pET32a-fool	TGGAATTTCTGAACAGGCAT GATGGAGGGAAAACGCAGA A

c) Supplemental references

- Andersson, P., I. Engberg, G. Lidin-Janson, K. Lincoln, R. Hull, S. Hull & C. Svanborg, (1991) Persistence of *Escherichia coli* bacteriuria is not determined by bacterial adherence. *Infect Immun* **59**: 2915-2921.
- Berger, H., J. Hacker, A. Juarez, C. Hughes & W. Goebel, (1982) Cloning of the chromosomal determinants encoding hemolysin production and mannose-resistant hemagglutination in *Escherichia coli*. *J Bacteriol* **152**: 1241-1247.
- Bingen-Bidois, M., O. Clermont, S. Bonacorsi, M. Terki, N. Brahimi, C. Loukil, D. Barraud & E. Bingen, (2002) Phylogenetic analysis and prevalence of urosepsis strains of *Escherichia coli* bearing pathogenicity island-like domains. *Infect Immun* **70**: 3216-3226.
- Dezfulian, H., I. Batisson, J. M. Fairbrother, P. C. K. Lau, A. Nassar, G. Szatmari & J. Harel, (2003) Presence and characterization of extraintestinal pathogenic *Escherichia coli* virulence genes in F165-positive *E. coli* strains isolated from diseased calves and pigs. *J Clin Microbiol* **41**: 1375-1385.
- Grozdanov, L., C. Raasch, J. Schulze, U. Sonnenborn, G. Gottschalk, J. Hacker & U. Dobrindt, (2004) Analysis of the genome structure of the nonpathogenic probiotic *Escherichia coli* strain Nissle 1917. *J Bacteriol* **186**: 5432-5441.
- Harel, J., C. Forget, J. Saint-Amand, F. Daigle, D. Dubreuil, M. Jacques & J. Fairbrother, (1992) Molecular cloning of a determinant coding for fimbrial antigen F165(1), a Prs-like fimbrial antigen from porcine septicaemic *Escherichia coli*. *J Gen Microbiol* **138**: 1495-1502.
- Johnson, T. J., S. Kariyawasam, Y. Wannemuehler, P. Mangiamele, S. J. Johnson, C. Doetkott, J. A. Skyberg, A. M. Lynne, J. R. Johnson & L. K. Nolan, (2007) The genome sequence of avian pathogenic *Escherichia coli* strain O1:K1:H7 shares strong similarities with human extraintestinal pathogenic *E. coli* genomes. *J Bacteriol* **189**: 3228-3236.

- Low, D. A., B. A. Braaten, G. V. Ling, D. L. Johnson & A. L. Ruby, (1988) Isolation and comparison of *Escherichia coli* strains from canine and human patients with urinary tract infections. *Infect Immun* **56**: 2601-2609.
- Minschew, B. H., J. Jorgensen, G. W. Counts & S. Falkow, (1978) Association of hemolysin production, hemagglutination of human erythrocytes, and virulence for chicken embryos of extraintestinal *Escherichia coli* isolates. *Infect Immun* **20**: 50-54.
- Mobley, H. L., D. M. Green, A. L. Trifillis, D. E. Johnson, G. R. Chippendale, C. V. Lockatell, B. D. Jones & J. W. Warren, (1990) Pyelonephritogenic *Escherichia coli* and killing of cultured human renal proximal tubular epithelial cells: role of hemolysin in some strains. *Infect Immun* **58**: 1281-1289.
- Toh, H., K. Oshima, A. Toyoda, Y. Ogura, T. Ooka, H. Sasamoto, S.-H. Park, S. Iyoda, K. Kurokawa, H. Morita, K. Itoh, T. D. Taylor, T. Hayashi & M. Hattori, (2010) Complete genome sequence of the wild-type commensal *Escherichia coli* strain SE15, belonging to phylogenetic group B2. *J Bacteriol* **192**: 1165-1166.
- Touchon, M., C. Hoede, O. Tenaillon, V. Barbe, S. Baeriswyl, P. Bidet, E. Bingen, S. Bonacorsi, C. Bouchier, O. Bouvet, A. Calteau, H. Chiapello, O. Clermont, S. Cruveiller, A. Danchin, M. Diard, C. Dossat, M. E. Karoui, E. Frapy, L. Garry, J. M. Ghigo, A. M. Gilles, J. Johnson, C. Le Bouguéneq, M. Lescat, S. Mangenot, V. Martinez-Jéhanne, I. Matic, X. Nassif, S. Oztas, M. A. Petit, C. Pichon, Z. Rouy, C. S. Ruf, D. Schneider, J. Tourret, B. Vacherie, D. Vallenet, C. Médigue, E. P. C. Rocha & E. Denamur, (2009) Organised genome dynamics in the *Escherichia coli* species results in highly diverse adaptive paths. *PLoS Genet* **5**: e1000344.
- Zdziarski, J., C. Svanborg, B. Wullt, J. Hacker & U. Dobrindt, (2008) Molecular basis of commensalism in the urinary tract: low virulence or virulence attenuation? *Infect Immun* **76**: 695-703.

Implication de l'auteur dans cet article.

En tant que premier auteur de cet article, j'ai réalisé l'ensemble des constructions bactériennes, ainsi que les expériences de retard de migration sur gel, empreinte de protection à la DNase I, et purification protéique. J'ai également participé au dénombrement des différents types de colonies ainsi qu'aux expériences de résonance plasmonique de surface. J'ai finalement analysé l'ensemble des résultats présentés et ai rédigé l'intégralité de cet article.

B. Article II : Leucine-responsive regulatory protein (Lrp) and PapI homologues control during phase variation of CS31A surface antigen

Cet article a été soumis à *Journal of bacteriology*

Authors:

Richard Graveline,¹ Michaël Mourez,¹ Mark A. Hancock,² Christine Martin,³ Rémi Lavoie¹ and Josée Harel^{1*}

Addresses:

Groupe de Recherche sur les Maladies Infectieuses du Porc (GREMIP) and Centre de Recherche en Infectiologie Porcine (CRIP), Université de Montréal, Saint-Hyacinthe, Québec, CANADA;¹ Sheldon Biotechnology Centre, McGill University, Montréal, Québec, CANADA; ² and Institut National de la Recherche Agronomique, Saint-Genès-Champanelle, FRANCE ³

*Corresponding author. Mailing address: Groupe de Recherche en Infectiologie Porcine (GREMIP), Université de Montréal, Faculté de Médecine Vétérinaire, C.P. 5000, Saint-Hyacinthe, Québec, Canada J2S 7C6. Phone: (450) 773-8521, ext. 8233. Fax: (450) 778-8108. E-mail:

Running title:

Lrp control of CS31A surface antigen phase variation

1- ABSTRACT

CS31A is a K88-related surface antigen that belongs to the type P family of adhesive factors. It is under the control of methylation-dependent transcriptional regulation for which the leucine-responsive regulatory protein (Lrp) plays an essential role. While CS31A, similarly to Pap, is preferentially in the OFF state, distinct regulatory features are observed. Here, we combined DNase I protection assay with Surface Plasmon Resonance (SPR) to analyse the interaction between Lrp and the *clp* regulatory region. SPR assays showed that Lrp presents a low micromolar affinity for the distal fragment and a strong nanomolar affinity for the proximal fragment. DNase I protection assay confirmed these results, as Lrp strongly interacted with the proximal region but weakly with the distal region. Moreover, the addition of a PapI homologue increased the affinity of Lrp for the distal and proximal regions, and was necessary to induce phase variation. Taken together, these results suggest that the preferential OFF state observed in CS31A-positive cells is mainly due to a weak interaction of Lrp with the *clp* distal region and that the presence of a PapI homologue within the cell plays a key role in CS31A production.

2- INTRODUCTION

Adhesins are important virulence factors that mediate specific attachment of bacteria to different host cells and tissue sites. The presence of multiple adhesin gene clusters in one single cell often determines the tissue tropism. Thus, pathogenic bacteria have acquired specific mechanisms to coordinately express more than one adhesin at the same time.

CS31A, encoded by the *clp* gene cluster, is a K88-related antigen found in septicaemic and enterotoxigenic *Escherichia coli* strains of bovine and human origin (7, 10). It is required for full bacterial virulence and its production is controlled by phase variation (8, 9, 18).

Phase variation is a stochastic and reversible switch between an “all-or-none” (ON/OFF) phenotype, each fitting to a specific condition (for review, see (5, 28)).

After cell division, most of daughter cells retain the expression phase of the parent, whereas some switch to the other expression phase. Thus, phase variation allows bacteria to anticipate and survive environmental changes without the need to sense it.

Mechanistically, Pap fimbriae are one of the most studied phase-variable systems (5, 13, 17). Expression of these adhesins is inherited in an epigenetic manner (28) because of methylation-controlled modifications within the regulatory region of the *pap* operon (28). This region consists of an approximately-400 bp DNA sequence surrounded by two divergently transcribed genes, *papI* and *papB* (Fig. 1). It also includes six leucine-responsive regulatory protein (Lrp) binding sites, all containing the consensus sequence GN(2-3)TTT (where N is any nucleotide), although a new alignment has been recently characterized (16), as well as two GATC sites (referred to as GATC^{prox} and GATC^{dist}, respectively) spaced 102 bp apart.

The differential methylation status of these sequences determines the binding of Lrp. Indeed, the switch from one phenotype to the other arises from the competition between the binding of Lrp at sites 1-3 or sites 4-6 and methylation by Dam methyltransferase at the opposite GATC site (Fig. 1) (for reviews, see (3, 5, 12)). Thus, when GATC^{dist} is fully methylated, Lrp cooperatively binds to sites 1-3, maintaining the cells in the OFF state, and when GATC^{prox} is fully methylated, Lrp binds to sites 4-6, maintaining the ON state (Fig. 1). Control of *pap* expression also requires the action of PapI, a positive regulator that increases the affinity of Lrp for sites 4-6 *in vivo* (12, 27). PapB, the second specific regulator of the *pap* operon, plays an important role at the transcriptional level, mainly by coordinating the expression of the *pBA* and *pI* promoters (12). Taken together, the combined actions of Lrp, Dam methyltransferase, PapI, and PapB result in a finite probability of each cell to express (ON) or not (OFF) Pap fimbriae just after DNA replication.

CS31A belongs to the Pap-regulatory family of adhesive factors. Previous works showed that the expression of *clp* is also subject to phase variation and that CS31A presents a much higher level of OFF cells than *pap*. This is also mediated

by the global regulator Lrp, the methylation protection of the two *clp* GATC sites as well as by the presence of two specific regulatory proteins homologous to PapB and PapI (8, 18). However, sequence analysis of the *clp* operon showed that it doesn't carry any *papI* homologue.

Previously, we have showed that in the absence of a PapI homologue, Lrp and ClpB (PapB homologue) repress *clp* transcription. Thus, the absence of a PapI homologue results in a moderate level of *clp* transcription. Yet, the presence of a PapI homologue is required during *clp* phase variation, as evidenced by the addition in a K12 *E. coli* strain of *afaF* in *trans*, the *papI* homologue belonging to the *afa-3* gene cluster (8, 18). Moreover, *afaF*⁺ cells in the OFF state show a lower expression of *clp* than levels obtained for cells lacking *afaF*. In their study, Bertin *et al.* also showed the reference strain 31A encodes P fimbriae (1). Thus, PapI homologues may act as negative regulators of CS31A production although *clp* phase variation depends on the presence of such regulators into the chromosome of CS31A-positive *E. coli* isolates. Whether PapI encoded by this *pap* operon cross-regulates *clp* transcription has not been investigated yet. We also previously showed that both GATC are methylated in the absence of Lrp and that methylation protection of these sites required the presence of Lrp (8). Based on these results it was hypothesized that the lower affinity of Lrp for the *clp* distal region than for the *clp* proximal region is responsible for *clp* transcriptional repression and in the presence of a PapI homologue strong bias towards the OFF phase.

The objective of the present study was to determine the molecular interactions between the Lrp and PapI, or a second homologue FooI and the intergenic region of *clp* operon during CS31A phase variation. By combining *in vitro* experiments with genetic studies, we show that the preferential OFF state observed in CS31A-positive cells is mainly due to a weak interaction of Lrp to the distal region of *clp*. Moreover, the addition of a PapI homologue promotes phase variation by increasing the affinity of Lrp for both GATC^{dist} and GATC^{prox}. Furthermore, the

affinity of the ternary complex is so high for the *clp* proximal region that only a small fraction of cells can switch toward the ON state.

3- MATERIALS AND METHODS

a) Bacterial strains, plasmids and growth conditions

E. coli strain MC4100.λ6 contains a single chromosomal copy of a *clp-lacZYA* fusion (see Fig. 1). ptrf5 and ptrf6 are multicopy recombinant plasmids derived from ptrc99A (Pharmacia Biotech) and express *fool* and *papI*, respectively, under the control of the *trc* promoter. Luria-Bertani (LB) broth and M9 minimal broth were prepared as described previously (21, 25). When necessary, media were supplemented with antibiotics at the following concentrations (unless otherwise noted): ampicillin, 100 $\mu\text{g}\cdot\text{ml}^{-1}$; kanamycin, 40 $\mu\text{g}\cdot\text{ml}^{-1}$; as well as with isopropyl- β -D-thiogalactopyranoside (IPTG), 1 mM and 5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside (X-gal), 40 $\mu\text{g}\cdot\text{ml}^{-1}$. Cultures of strains harboring ptrf5 or ptrf6 were grown in the absence of IPTG to avoid a too high expression of *fool* or *papI*.

b) Plasmid construction

For overexpression and purification of the Fool protein, the pET32a-fool plasmid was constructed by first amplifying the *fool* sequence from the chromosome of strain 4787, using the NcoI-foolrg1F and XhoI-foolrg1R primers (containing the NcoI and XhoI restrictions sites, respectively). After enzymatic digestion, the fragment was ligated into the NcoI/XhoI restriction sites of the expression plasmid pET32a. Inserts were sequenced for verification after transformation of *E. coli* BL21(DE3)pLysS by electroporation. Plasmid pET32a-papI was constructed using the QuickChange Site-Directed Mutagenesis System (Stratagene). Briefly, expression plasmid pET32a-fool was first extracted from strain *E. coli* BL21(DE3)pLysS, purified and then used as template for amplification using the foolN17D-F and foolN17D-R primers. PCR amplified plasmids were purified and digested with the restriction enzyme DpnI to remove

any trace of the pET32a-fool template. PCR amplified plasmids were then dialyzed and transformed into strain *E. coli* BL21(DE3)pLysS by electroporation.

c) Purification of recombinant proteins

E. coli strain CV1494 was used to purify His₆-tagged Lrp as previously described (19). One-liter culture of CV1494 was grown in LB broth at 37°C, with shaking, to an OD₆₀₀ of 0.7-0.8. Isopropyl β-D-1-thiogalactopyranoside (IPTG) was then added to the culture at a final concentration of 0.5 mM for 6 hours. Bacteria were harvested, resuspended in 30 ml of Tris-buffered saline (20 mM Tris-HCl [pH 7.4], 0.5 M NaCl), and lysed with a French press and an ultrasonic processor. The soluble fraction was clarified by centrifugation at 15,000 g for 20 min at 4°C. Proteins were purified using an ÄKTA purifier system with a 1-ml His Trap HP column (Amersham Biosciences) according to the instructions of the manufacturer. Purified proteins were subsequently dialyzed against TG₅₀ED buffer (10 mM Tris-HCl [pH 8.0], 1 mM EDTA, 0.1 mM DTT, 0.1 M NaCl) and purity of the proteins was confirmed by SDS-PAGE and staining with Coomassie blue.

For purification of Fool and PapI, a single *E. coli* BL21(DE3)pLysS colony carrying plasmid pET32a-fool or pET32a-PapI was incubated in LB broth with shaking at 37°C, to an OD₆₀₀ of 0.4-0.6. The culture was incubated for 3 more hours in the presence of IPTG at a final concentration of 1 mM. Bacteria were harvested, resuspended in 30 ml of Tris-buffered saline (20 mM Tris-HCl [pH 7.4], 0.5 M NaCl), and lysed with a French press and an ultrasonic processor. The soluble fraction was clarified by centrifugation at 15,000 g for 20 min at 4°C. The His₆-Trx-Fool and His₆-Trx-PapI fusion proteins were purified using an ÄKTA purifier system using the same procedure than for His₆-tagged Lrp. Fool and PapI protein were removed from their His₆-Trx tag after enzymatic digestion using Enterokinase (New England BioLabs) in digestion buffer (20 mM Tris-HCl [pH 8.0], 50 mM NaCl, 2 mM CaCl₂). Proteins were then dialyzed against 20 mM 2-[N-morpholino]ethanesulphonic acid (Mes), pH 6.0 and purified using an

ÄKTA purifier system with a mono S 5/50 GL column (Amersham Biosciences) according to the instructions of the manufacturer. Purity of F_ooI and P_apI proteins was confirmed by SDS-PAGE and staining with Coomassie blue.

d) Surface Plasmon resonance

Interactions between 5' biotinylated, double-stranded oligonucleotides (*clp* distal, 120 kDa; *clp* proximal, 77 kDa; *clp* distal-proximal, 185 kDa) and three DNA-binding proteins (Lrp, 20 kDa monomer; F_ooI, 9 kDa; P_apI, 9 kDa) were examined using label-free, real-time BIACORE 3000 instrumentation (GE Healthcare Bio-Sciences AB, Uppsala, Sweden). Experiments were performed on research-grade streptavidin-coated (SA) sensor chips (XanTec Bioanalytics GmbH, Muenster, Germany) at 25°C using filtered (0.2 µm) and degassed CALVO-P running buffer: 50 mM Tris pH 8.0, 150 mM NaCl, 0.4 mM EDTA, 1 mM MgCl₂, 0.1 mM DTT, 0.005% (v/v) Tween-20. As recommended by the manufacturer, SA sensor chips were pre-conditioned with three 1-minute pulses of 1M NaCl in 50mM NaOH. Protein-grade detergents (Tween-20, Empigen) were from Anatrace (Maumee, OH, USA) and fatty acid-free bovine serum albumin (BSA) was from Millipore (Billerica, MA, USA); all other chemicals were reagent grade quality. Purified protein concentrations were determined using the BCA assay and purity (>95%) was assessed by SDS-PAGE (12.5% polyacrylamide) under reducing (5% (v/v) 2-mercaptoethanol) and non-reducing conditions (data not shown).

To examine binding with Lrp, the appropriate oligonucleotides were annealed (see www.piercenet.com protocol) and then captured (20 ml/min x 10 nM in running buffer containing 0.5M NaCl) to generate low-density surfaces (<100 RU immobilized). Lrp (0 – 100 nM) or BSA (negative control) was injected over reference (*i.e.* SA only) and active (*i.e.* *foo* or *pap* DNA fragments) surfaces in-tandem using 'KINJECT' mode (25 ml/min x 5 min association + 5 min dissociation). Between sample injections, surfaces were regenerated at 50 ml/min using two 30 sec pulses of solution I (1 M NaCl, 0.02% (v/v) Empigen in CALVO-P) followed by 'EXTRACLEAN' and 'RINSE' procedures.

To examine binding with FcoI and PspI, medium-density oligonucleotide surfaces (300 – 400 RU immobilized) were prepared as noted above. FcoI or PspI (0 – 15 μ M) was injected over reference and active surfaces in-tandem using ‘KINJECT’ mode (5 ml/min x 5 min association + 10 min dissociation). Between sample injections, surfaces were regenerated as noted above. In additional experiments, Lrp (50 nM) was also injected in the presence of FcoI and PspI (0-5.6 and 0-8.1 mM, respectively) using ‘KINJECT’ mode (10 ml/min x 5 min association + 10 min dissociation).

Mass transport-independent data was doubled-referenced (22) and is representative of duplicate injections acquired from three independent trials. For each replicate series, a buffer blank was injected first, the highest titrant concentration second, and serial dilutions followed (from the lowest to the highest concentration); comparing responses between the two highest titrant injections verified consistent DNA surface activity throughout each assay. To estimate apparent equilibrium dissociation constants (K_D), steady-state binding responses (R_{eq} ; average RU near the end of the association phase) were plotted as a function of protein concentration (C) and then subjected to non-linear regression (“One site – Specific binding with Hill slope”; using Prism5 version 5.0c for Mac OS X, GraphPad Software, San Diego California USA, www.graphpad.com). To estimate individual dissociation rate constants (kd), sensorgram profiles were analyzed using the “Fit Kinetics / Separate ka/kd” tool. Theoretical binding maxima were predicted using the following equation:

$$R_{max} = (MWA / MWL) (RL) (n)$$

where R_{max} is the maximal binding response (RU) at saturating protein concentration; MWA is the molecular weight (kDa) of the protein injected in solution; MWL is the molecular weight (kDa) of the DNA fragment immobilized; RL is the amount (RU) of DNA immobilized; n is the predicted binding stoichiometry (*e.g.* 1:1).

e) Footprint Analysis

DNase I footprinting of free DNA and DNA–protein complexes was performed as described (14). End-labeled ^{32}P -*clp* DNA (125,000 cpm, 0.3 nM) was incubated in a total volume of 20 μL with 3 μg of salmon sperm DNA and 2 μg of acetylated bovine serum albumin (New England Biolabs) in binding buffer (60 mM Tris-HCl [pH 7.5], 40 mM KCl, 100 mM NaCl, 0.1 mM EDTA, 1 mM DTT). After incubation for 10 min at room temperature, 6 μL of DNase I (2,000 U.ml⁻¹, New England Biolabs) containing 22 mM CaCl₂ and 22 mM MgCl₂ was added for 1 min. The reaction was stopped by addition of 200 μL of stop buffer (10% glycerol, 2.5 mM EDTA, 0.6 M ammonium acetate, 2 mg yeast tRNA per ml) to each sample. DNA fragments were precipitated in ethanol and an equivalent of c.p.m. (5×10^4) from each reaction was loaded onto 6% polyacrylamide/7M urea gels. A + G Maxam Gilbert reactions (20) were carried out on the appropriate ^{32}P -labeled DNA fragments and loaded alongside the DNase I footprinting reactions. Gels were dried and analyzed by autoradiography.

f) Phase variation and measurement of switching frequencies

The switch frequency was calculated as described previously (4, 8). Strains were inoculated onto LB-kan agar. Colonies were then excised from the agar and resuspended in M9 salts containing no carbon source. Appropriate dilutions were subsequently spread onto M9 X-Gal agar containing 0.2% glycerol and, after 36 h of growth at 37 °C, plates were digitally photographed for record. Three colonies showing a uniform Lac⁺ or Lac⁻ phenotype were excised from the agar, resuspended in M9 salts, and spread onto M9 X-gal agar as before. The passage from M9 X-gal agar to fresh M9 X-gal agar containing the same source of carbon was repeated two more times to follow the switch frequency through three generations. At least 2,000 colonies were scored for a Lac phenotype as described (4) and to calculate the switch frequencies according to the formula $(M/N)/g$, where M is the number of cells that underwent phase variation, N the

total number of cells evaluated and g the total number of generations (estimated to be 20 for all the strains tested in our experiments).

4- RESULTS

a) Differences in stability of Lrp interacting with either *clp* proximal or distal region

Similar to the intensively studied *pap* operon, *clp* phase variation requires Lrp. The Lrp target regulatory region contains two GATC sites that are spaced 102 or 103 bp apart and methylation protection of these sites requires Lrp (8, 18). We hypothesized that the high level of OFF cells observed with CS31A could be also due to a specific profile of Lrp affinity for the *clp* distal or proximal region, by analogy to *pap*.

In order to characterize these interactions, a BIAcore biosensor was used to monitor in real-time the dynamic response of Lrp for two DNA fragments corresponding to the *clp* proximal and distal regulatory regions. Biotinylated oligonucleotides duplexes containing the *clp* distal, *clp* proximal, or the *clp* distal-proximal region were immobilized on a streptavidin sensor chip, and His6-Lrp, at concentrations ranging from 0.4 to 100 nM, was injected over the DNA surface. As shown in Fig. 2, injection of Lrp over the DNA surface containing the *clp* regions resulted in a net mass accumulation, indicating that Lrp binds to the distal and proximal regions of *clp* as it has already been described for *pap* (12, 23). However, what mainly characterized *clp* was an important decrease in affinity for the *clp* distal region when compared to the *clp* proximal region. While fitting the titration series data, the intergenic regions comprising the proximal region exhibited low nanomolar and highly cooperative binding affinities for Lrp, with a Hill coefficient greater than 3 (data not shown). However, the *clp* distal region exhibited a significant lower affinity for Lrp, with estimated K_D values in a micromolar range (Fig. 2 and table 1). Taken together, our data suggest that Lrp has a better affinity for the *clp* proximal region than for the *clp*

distal region, and this could block the entry of the RNA polymerase at its *clpBA* promoter. Thus, this result is in agreement with a predominant phase OFF.

b) Lrp binding site occupancies within the *clp* regulatory region correlates with difference in affinity between the *clp* distal and proximal regions

Because our SPR analysis showed that Lrp has a better affinity for *clp* proximal region than for the *clp* distal region, we hypothesized that the differences in affinity between the *clp* distal and the *clp* proximal regions could be due to different Lrp-binding sites, and/or to divergent nucleotide sequences surrounding the GATC sites. Thus, we used DNase footprint analysis to identify the Lrp binding sites within the full-length *clp* regulatory region. As shown in Fig. 3, the addition of His6-Lrp was associated with the presence of multiple protected regions. In addition, protections were observed at a lower concentration of His6-Lrp for the upper DNA region than for the bottom region, in agreement with a better affinity of Lrp for the *clp* proximal region. Contrary to *pap* for which six Lrp binding sites has been identified, the presence of Lrp protected nine regions in *clp*. This included five Lrp binding sites corresponding to the *clp* proximal region and four to the *clp* distal region (Fig. 1). Moreover, each Lrp binding site contained at least one GN₂₋₃TTT motif as well as the presence of a GATC at two Lrp binding sites. Taken together, these results indicate that although *clp* shares common regulatory features with *pap* such as a greater affinity for the proximal region than for the distal region and the presence of GATC sites at two distinct Lrp binding sites, it presents a specific pattern of Lrp localization that results in a very weak affinity, especially for the *clp* distal region.

c) PapI and Foli increase Lrp affinity to both distal and proximal regions

In *E. coli* K12 strain, phase variation of *clp* operon requires the presence in *trans* of a PapI homologue (8, 9, 18). Since the reference strain 31A carries a *pap*-fimbrial operon in its genome, *clp* phase variation may result from the cross-regulation between *clp* and *pap* fimbriae through the supply of PapI in *trans*.

Thus, we wondered if a PapI homologue can influence *clp* phase variation as it has already been demonstrated with AfaF. First, we monitored the binding of FooI or PapI at the *clp* distal and proximal regions. The binding of FooI and PapI to the different sequences was characterized by a rapid-on, rapid-off kinetics and low micromolar binding affinities for all tested fragments (Table 1), confirming the previous work of Kawamura *et al.* showing the direct binding of PapI to *pap* sites 2 and 5 with similar low micromolar affinities (15). However, no noticeable difference was observed between FooI and PapI kinetics. We then incubated the *clp* distal and proximal regions with either His6-Lrp alone at 50 nM or with His6-Lrp at the same concentration in presence of increasing amounts of PapI or FooI (ranging from 0 to 8.1 nM or 0 to 5.6 nM, respectively). As shown in Fig. 4, the presence of FooI, a PapI homologue greatly influenced the overall affinity of Lrp for both the *clp* distal and proximal regions. Moreover, the addition of PapI, or its homologue FooI, increased the affinity of Lrp for the DNA fragments tested to a level higher than the sum of the actions of Lrp and of PapI (data not shown). This was particularly the case for the *clp* distal region. When PapI was compared to FooI, no differences were observed in their ability to promote Lrp interaction. PapI and FooI also prevented the dissociation of Lrp from the DNA, especially from the *clp* proximal region, resulting in a stable complex. Because the transcription of *papI* homologues on ptrc99A vector is independent of that of *clp* operon is, a PapI homologue may contribute to enhance the binding of Lrp at the *clp* proximal region. Thus, these results suggest that a PapI homologue has two opposite functions. First, it contributes to switch the *clp* operon toward the ON state by increasing the kinetics association of Lrp for the *clp* distal region; second, it maintains cells in the OFF states by forming a particularly stable complex between Lrp and the *clp* proximal region.

The influence of PapI homologues on *clp* phase variation was also measured *in vivo* by providing FooI and PapI in *trans*. Strain MC4100.λ6 carrying a chromosomal fusion between the *clp* regulatory region and the *lacZYA* reporter system was transformed with plasmids ptrf5 or ptrf6, expressing *fooI* or *papI*,

respectively. The two transformed strains were then plated onto X-gal M9 agar without IPTG (to avoid a too high level of PapI or FooI protein in the cell) and the [Lac] phenotype was followed for up to two generations. While strain MC4100.λ6 led to a uniform phenotype characterized by light-blue colonies on minimal M9 X-gal plates (data not shown), transformation with either ptrf5 or ptrf6 restored *clp* phase variation. Thus, PapI or FooI produced in *trans* induces phase variation of the *clp* operon as it has already been described with AfaF. Moreover, phase variation of *clp* was always toward the OFF state, even when the starting phenotype was the ON state. This is in agreement with the observation that Lrp shows a better affinity for the *clp* proximal region than for the *clp* distal region whether a PapI homologue is present or not (8, 18). As a result, the low affinity of Lrp for the *clp* distal region greatly contributes to the low level of ON cells. Interestingly, differences in the ability of PapI and FooI to switch the entire bacterial population toward the ON state were observed. Notably, the level of ON cells was higher in the presence of FooI than in the presence of PapI. Thus it seems that, even if PapI homologues are all able to promote phase variation, they still can differently influence the level of the ON population.

5- DISCUSSION

During phase variation of P fimbriae, the competitive Lrp and Dam assembly on the regulatory region is important for the heritability of the OFF state. In phase OFF cells, Lrp binds the proximal region, which blocks the entry of the RNA polymerase and protects GATC^{prox} from Dam methylation. In phase ON cells, Lrp binds to the distal region and protects GATC^{dist} from methylation. Thus, for P fimbriae to switch from the OFF to the ON state, Dam must methylate GATC^{prox}, a process for which PapI contributes by moving Lrp from the proximal region to the distal region, thus making GATC^{prox} available for methylation.

Regulation of the *clp* operon, encoding CS31A, shares similar mechanisms with the regulation of the expression for other P related fimbrial operons (18).

Notably, in strain 31A CS31A production is under phase variation control. However, the *clp* gene cluster does not carry any *papI* homologue, although a PapI homologue is required for phase variation to occur. Thus, it has been suggested that the presence of an additional *pap* operon on the chromosome of CS31A-positive strains may supply the PapI required in *trans*, resulting in *clp* phase variation (8). Using several techniques including gene expression, and *in vitro* protein–DNA interaction analysis, we have now demonstrated that similar to the effect of AfaF, the addition of a PapI homologue induces phase variation of the *clp* operon, mainly by its capacity to promote Lrp interaction with the distal region of the *clp* operon. However we show here that the affinity of Lrp for the distal region is 10^3 -fold lower than for the proximal region, and thus this altered interaction between Lrp and the distal region greatly contributes to the overall repression of *clp*.

In the absence of a PapI homologue supplied in *trans*, *i.e.* in the absence of phase variation, Martin showed that Lrp protects the GATC sites from methylation and cannot bind methylated GATC sites (18). Here, we show that Lrp presents a lower affinity for GATC^{dist} than for GATC^{prox} of the *clp* regulatory region (8, 18).

In Pap, Lrp affinity is only 2-fold higher for the proximal sites 1-3 than for the distal sites 4-6 (in a similar low nanomolar range) (12). Moreover, binding of Lrp at the proximal sites 1-3 reduces the affinity of Lrp for *pap* at the distal sites 4-6 by 10-fold, a phenomenon apparent when Lrp binding on plasmid DNA is analyzed and that has been denoted “mutual exclusion” (12). However, no “mutual exclusion” *sensu stricto* could be observed in CS31A since, Lrp affinity in CS31A is 10^3 -fold higher for the proximal region than for the distal region. Because of the differential affinities of Lrp for the *clp* proximal and distal regions, this is similar to the comparison between two high Lrp interactions, one specific *clp* proximal with low nanomolar affinity (in the range similar to what is shown when Lrp is considered as a specific regulator), and one *clp* distal non-specific with low micromolar affinity (in the same range than what is shown

when Lrp is considered as a nucleoprotein) as previously reported (24). As a result, Lrp protects GATC^{prox} from methylation and negatively controls the basal *clp* expression.

By comparing phase variation control of F165₁ another member of the P regulatory family to that of Pap, we have recently shown that higher level of ON cells of F165₁ is due to altered stability of the DNA complex formed by the Lrp at its repressor binding sites 1-3 after each cell cycle (11). Furthermore, we identified nucleotides surrounding the Lrp-binding site 1 critical to maintaining a high OFF to ON switch rate during F165₁ phase variation, as well switching Pap fimbriae toward the OFF state by influencing Lrp dissociation rate to the proximal site. In the case of Clp, the predominance of the OFF state is mainly due to a weak affinity of Lrp with the *clp* distal region (11). This suggests that the interaction of Lrp with its binding sites dictates their accessibility to methylation and thus influences the epigenetic process of phase variation. However, as phase variation of P fimbriae requires the action of multiple regulators such as the local regulators PapB and PapI, global regulators CpxAR, cAMP-CAP, these may also contribute to the phase variation difference between strains (13, 26, 29, 30).

In *E. coli* reference strain 31A, CS31A is encoded by a high-molecular-weight plasmid called p31A that does not harbor a *clp* specific homologue to *papI*. It was suggested that the phase variation control observed in wild-type strains could result from a cross talk between *clp* and a chromosomal *pap*-related operon as strain 31A encodes P fimbriae (8, 9). In this study, alike AfaF, PapI and the homologue Fool when supplied in *trans* restore the bistable state of phase variation. Similar to *pap*, the affinity of Lrp for both the distal and proximal regions increases in the presence of PapI or Fool. Moreover, the affinity of the PapI/Fool-Lrp complex being significantly higher for the proximal than for the distal region, this could explain why, during phase variation, only a minority of *clp* cells is in the ON phase. Furthermore, the strong affinity of the Lrp/PapI (or Lrp/Fool complex) for the proximal region and the decreased dissociation of the complex to the DNA target could also explain why, in the OFF phase, the level of

expression of the *clp* operon is even lower than the moderate level observed in the absence of a PapI or an homologue, since the PapI-Lrp stable complex blocks *clp* transcription more efficiently than Lrp alone (8). Taken together, our results show that, in the case of *clp*, PapI homologues can play a dual role by first allowing phase variation and, second, by promoting the OFF phase. How PapI homologues promote the switch from OFF to ON remains however unclear.

Previously, we showed that AfaF, another PapI homologue, induces phase variation of *clp* (8). Here, we show that this ability to induce *clp* phase variation is not specific to AfaF but PapI and FooI, another PapI homologue, can act as well on *clp* phase variation. It is noteworthy that levels of ON colonies can vary greatly depending on which PapI homologue is present in the cell. Indeed, depending on what PapI homologue is present in a given strain, the level of CS31A can significantly vary. Previous studies using AfaF showed higher level on ON cells than what we observed in this study (18) and our comparison between FooI and PapI functions also demonstrated slight but still significant differences in the level of ON cells. Considering the fact that at least sixteen *papI* alleles have been identified so far (26) and that CS31A is often associated with the presence of other type P-related fimbriae such as Pap or F165₁ (1, 2, 6, 7), this suggests that PapI homologues can be seen as specific regulators that co-ordinately regulate the expression of more than one fimbrial operon (26, 31). Similarly, Totsika *et al.*, studying the cross-activation of *pap* variants by FooI and PapI, suggested that sequence variation among PapI homologues could affect their ability to activate *pap* transcription and that they may have evolved in order to prevent cross-activation of related proteins (26).

Overall, the present study has provided novel insights as to why Clp is a slow-switcher like Pap due to the very low affinity of Lrp to *clp* distal site. Our data emphasize that the fine molecular interaction of regulatory proteins from their DNA binding sites influence phase variation and expression of fimbriae. In the large repertoire of fimbrial variants of the P family, there is also an exquisite fine

tuning distinction in their regulatory expression that arms the bacterium with strategies for adapting to more than one particular environment.

6- ACKNOWLEDGEMENTS

We are grateful to David Low (University of California) for the generous gift of purified PapI and enlightening discussions. We are grateful to Cécile Crost (CRIP) and George Szatmari for their invaluable advice during this study. This work was supported by funding from the Centre de Recherche en Infectiologie Porcine (CRIP; Fonds québécois de la recherche sur la nature et les technologies Regroupements stratégiques 111946), the Natural Sciences and Engineering Research Council of Canada (NSERC Discovery Grant 25120 to JH and 262746 to MM), and the Canada Research Chair program to MM. Sheldon Biotechnology Centre is supported by a Research Resource Grant from the Canadian Institutes of Health Research.

7- REFERENCES

1. **Bertin, Y., J. P. Girardeau, A. Darfeuille-Michaud, and C. Martin.** 2000. Epidemiological study of *pap* genes among diarrheagenic or septicemic *Escherichia coli* strains producing CS31A and F17 adhesins and characterization of Pap(31A) fimbriae. *J Clin Microbiol* **38**:1502-1509.
2. **Bertin, Y., C. Martin, J. P. Girardeau, P. Pohl, and M. Contrepois.** 1998. Association of genes encoding P fimbriae, CS31A antigen and EAST 1 toxin among CNF1-producing *Escherichia coli* strains from cattle with septicemia and diarrhea. *FEMS Microbiol Lett* **162**:235-239.
3. **Blomfield, I. C.** 2001. The regulation of *pap* and type 1 fimbriation in *Escherichia coli*. *Adv Microb Physiol* **45**:1-49.
4. **Blyn, L. B., B. A. Braaten, C. A. White-Ziegler, D. H. Rolfson, and D. A. Low.** 1989. Phase-variation of pyelonephritis-associated pili in *Escherichia coli*: evidence for transcriptional regulation. *EMBO J* **8**:613-620.
5. **Casadesús, J., and D. Low.** 2006. Epigenetic gene regulation in the bacterial world. *Microbiol Mol Biol Rev* **70**:830-856.
6. **Cherifi, A., M. Contrepois, B. Picard, P. Goulet, J. de Rycke, J. Fairbrother, and J. Barnouin.** 1990. Factors and markers of virulence in *Escherichia coli* from human septicemia. *FEMS Microbiol Lett* **58**:279-283.
7. **Contrepois, M., J. M. Fairbrother, Y. K. Kaura, and J. P. Girardeau.** 1989. Prevalence of CS31A and F165 surface antigens in *Escherichia coli* isolates from animals in France, Canada and India. *FEMS Microbiol Lett* **50**:319-323.
8. **Crost, C., A. Garrivier, J. Harel, and C. Martin.** 2003. Leucine-responsive regulatory protein-mediated repression of *clp* (encoding CS31A) expression by L-leucine and L-alanine in *Escherichia coli*. *J Bacteriol* **185**:1886-1894.

9. **Crost, C., J. Harel, F. Berthiaume, A. Garrivier, M. C. Tessier, H. Rakotoarivonina, and C. Martin.** 2004. Influence of environmental cues on transcriptional regulation of *foo* and *clp* coding for F165(1) and CS31A adhesins in *Escherichia coli*. *Res Microbiol* **155**:475-482.
10. **Girardeau, J. P., M. Der Vartanian, J. L. Ollier, and M. Contrepois.** 1988. CS31A, a new K88-related fimbrial antigen on bovine enterotoxigenic and septicemic *Escherichia coli* strains. *Infect Immun* **56**:2180-2188.
11. **Graveline, R., M. Mourez, M. A. Hancock, C. Martin, S. Boisclair, and J. Harel.** 2011. Lrp-DNA complex stability determines the level of ON cells in type P fimbriae phase variation. *Mol Microbiol* **81**:1286-1299.
12. **Hernday, A., M. Krabbe, B. Braaten, and D. Low.** 2002. Self-perpetuating epigenetic pili switches in bacteria. *Proc Natl Acad Sci USA* **99 Suppl 4**:16470-16476.
13. **Hernday, A. D., B. A. Braaten, G. Broitman-Maduro, P. Engelberts, and D. A. Low.** 2004. Regulation of the *pap* epigenetic switch by CpxAR: phosphorylated CpxR inhibits transition to the phase ON state by competition with Lrp. *Mol Cell* **16**:537-547.
14. **Kaltenbach, L. S., B. A. Braaten, and D. A. Low.** 1995. Specific binding of PapI to Lrp-pap DNA complexes. *J Bacteriol* **177**:6449-6455.
15. **Kawamura, T., L. U. K. Le, H. Zhou, and F. W. Dahlquist.** 2007. Solution structure of *Escherichia coli* PapI, a key regulator of the *pap* pili phase variation. *J Mol Biol* **365**:1130-1142.
16. **Kawamura, T., A. S. Vartanian, H. Zhou, and F. W. Dahlquist.** 2011. The Design involved in PapI and Lrp regulation of the *pap* operon. *Journal of molecular biology* **409**:311-332.
17. **Low, D. A., and J. Casadesús.** 2008. Clocks and switches: bacterial gene regulation by DNA adenine methylation. *Curr Opin Microbiol* **11**:106-112.
18. **Martin, C.** 1996. The *clp* (CS31A) operon is negatively controlled by Lrp, ClpB, and L-alanine at the transcriptional level. *Mol Microbiol* **21**:281-292.

19. **Matthews, R. G., Y. Cui, D. Friedberg, and J. M. Calvo.** 2000. Wild-type and hexahistidine-tagged derivatives of leucine-responsive regulatory protein from *Escherichia coli*. *Methods Enzymol* **324**:322-329.
20. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol* **65**:499-560.
21. **Miller, J. H.** 1972. *Experiments in Molecular Genetics*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
22. **Myszka, D. G.** 1999. Improving biosensor analysis. *J Mol Recognit* **12**:279-284.
23. **Nou, X., B. Braaten, L. Kaltenbach, and D. A. Low.** 1995. Differential binding of Lrp to two sets of pap DNA binding sites mediated by Pap I regulates Pap phase variation in *Escherichia coli*. *EMBO J* **14**:5785-5797.
24. **Peterson, S. N., F. W. Dahlquist, and N. O. Reich.** 2007. The role of high affinity non-specific DNA binding by Lrp in transcriptional regulation and DNA organization. *J Mol Biol* **369**:1307-1317.
25. **Thomason, L., D. L. Court, M. Bubunenko, N. Costantino, H. Wilson, S. Datta, and A. Oppenheim.** 2007. Recombineering: genetic engineering in bacteria using homologous recombination. *Curr Protoc Mol Biol* **Chapter 1**:Unit 1.16.
26. **Totsika, M., S. A. Beatson, N. Holden, and D. L. Gally.** 2008. Regulatory interplay between *pap* operons in uropathogenic *Escherichia coli*. *Mol Microbiol* **67**:996-1011.
27. **van der Woude, M., B. Braaten, and D. Low.** 1996. Epigenetic phase variation of the *pap* operon in *Escherichia coli*. *Trends Microbiol* **4**:5-9.
28. **van der Woude, M. W., and A. J. Bäumlner.** 2004. Phase and antigenic variation in bacteria. *Clin Microbiol Rev* **17**:581-611, table of contents.
29. **van der Woude, M. W., B. A. Braaten, and D. A. Low.** 1992. Evidence for global regulatory control of pilus expression in *Escherichia coli* by Lrp and DNA methylation: model building based on analysis of *pap*. *Mol Microbiol* **6**:2429-2435.

30. **Weyand, N. J., B. A. Braaten, M. van der Woude, J. Tucker, and D. A. Low.** 2001. The essential role of the promoter-proximal subunit of CAP in *pap* phase variation: Lrp- and helical phase-dependent activation of *papBA* transcription by CAP from -215. *Mol Microbiol* **39**:1504-1522.
31. **Xia, Y., D. Gally, K. Forsman-Semb, and B. E. Uhlin.** 2000. Regulatory cross-talk between adhesin operons in *Escherichia coli*: inhibition of type 1 fimbriae expression by the PapB protein. *EMBO J* **19**:1450-1457.

8- LEGENDS

Figure 1. Comparison of the nucleotide sequences of the regulatory regions of the *clp* and *pap* operons. The putative promoter regions (-10 and -35 sites) are represented by dotted lines. The identification of *pap* and *clp* Lrp binding sites are represented as boxed regions on the DNA sequences (numbers 1-6 indicate the positions of Lrp binding sites in *pap* sequence and greek letters indicate the positions of Lrp binding sites in *clp* sequence, as identified by DNase I footprint), and the PapB and CAP binding sites are underlined.

Figure 2. Association and dissociation kinetics of His₆-Lrp binding to the *clp* proximal regions. Lrp (0-100 nM) was injected over biotinylated distal and proximal fragments of *clp* (95 and 40RU, respectively) DNA captured to streptavidin-coated sensor chips. Notably, *clp* distal exhibited a significantly low affinity for Lrp. All binding data are representative of duplicate injections acquired from three independent trials. Lrp binding responses (Req; average 710-730 sec) were plotted using Prism5 version 5.0c for Mac OS X (GraphPad Software, San Diego California USA, www.graphpad.com) to estimate the equilibrium dissociation constants (K_D).

Table 1. Apparent Surface Plasmon Resonance (SPR) rate constants for Lrp, FooI, and PapI binding to *clp* binding sites.

Figure 3. Lrp localization on *clp* regulatory region. DNase I footprinting assays with increasing amounts of Lrp: Lane1, A + G Maxam and Gilbert reactions; Lanes 2-8: DNase I footprinting assays with increasing amounts of Lrp on *clp* regulatory regions end-labeled with ³²P on the top strand (lane 2, DNA alone; lane 3, Lrp at 1nM; lane 4, Lrp at 10nM; lane 5, Lrp at 100nM; lane 6, Lrp at 1μM; lane 7, Lrp at 10μM; lane 8, Lrp at 40μM). The arrow along the foo DNA sequence indicates the presence of GN(2-3)TTT minimal motif recognized by Lrp. Orientation of the arrow depends on the orientation of the GN(2-3)TTT

motifs. Greek letters indicate the positions of Lrp binding sites in the *clp* sequence.

Figure 4. PapI and Fool increase both affinity and stability of Lrp at the *clp* proximal and distal regions. Representative label-free, real-time binding of 50 nM Lrp to immobilized *clp* (95 RU distal, 40 RU proximal) oligonucleotides in the absence (solid lines) and presence (dashed lines) of Fool (0-5.6 nM) or PapI (0-8.1 nM).

Figure 5. The presence of a PapI homologue leads to the bistable expression of the *clp* operon. Lac⁺ and Lac⁻ phenotypes of *clp* strains transformed with either *papI* or *fool*. Switch frequencies are calculated using the formula $(M/N)/g$, where M is the number of cells that underwent phase transition, N the total number of cells evaluated, and g the number of generations, estimated to be 20, that gave rise to the colony (*, $p < 0.001$).

Table 2. Effects of PapI homologues on phase variation of *E. coli clp-lacZYA* fusion lysogen.

```

clp -----GTTCTACTAATTGGATGCGCTACClGTTTTTGA
pap CTCACGTAAACAAGTTTCTTCGAATAATAAAAAATCATGCTCT-CT-GTTATCAACGGAAAGGTATTTTTA

clp CTCTCCCTGTTTGATTTCTθGTGCTTTAATAATTTTAACTT--TAAATTCATAATTTATATGAAAAGTGCT
pap TTCTCTATGTT-----TGCTT---TATTTGTTCAATTTAGTGAATT--TGCTTTTTAT-----TG-G

                                     PapB site 1
clp ηACATTTTTTTGATGAAAATCTAAATTAACGCAAAAACTTTATCTG-GA-ATGCAATTTGATTTTTTTGT
pap --ATTTATTTGATGTGTATCACATTTTGCCTTTA---TTTTTCTGC4GAAAAGAAAGTCCGTAAAAATTC

                                     CAP binding site
clp TTTAGACζGATCTTTTATCTGTGATTTTGTGTTTTTTTGGTGCTTTTGCT--εGTGTTGTTGTTTTTG
pap TTTAGACGATCTTTTATGCTGTAAATTCAT-TTGCCATγGATGTTTTTATCTGAGTACCTCTTGCTATTA

                                     5           δ           6           β
clp TGGTTTTATTTGTGTTGATGTTTGTGTTTGTAGTTGTAAGATCGTT-AAATTAGTGTGTTTCGC-TATA
pap GTGTTTTG-TTCCTAGTTTAAATTTGTT-TTGTGGTTAAAAGATCGTTTAAATCAATATTTACAA35CATAAA

                                     1           α           2           3
clp AAAC---AT-----CGTG-ATA--ATCTTCG---CTGCAGAAA---CAGGACGGCAACAGA--
pap AAAC10TAAATTTAACTTATTCGTTGAAGAGTATTTCCGGGCCGGAAGCATATATCCAGGGGCCGACAGAA

                                     PapB site 2
clp GTGAATACGTTTTTGTATCGGGGTATGTCGTCTTTTTGGTCTGGTTATATTTCTGGAAATGGATATAA
pap GGGGAAAC-----

```

Figure 1

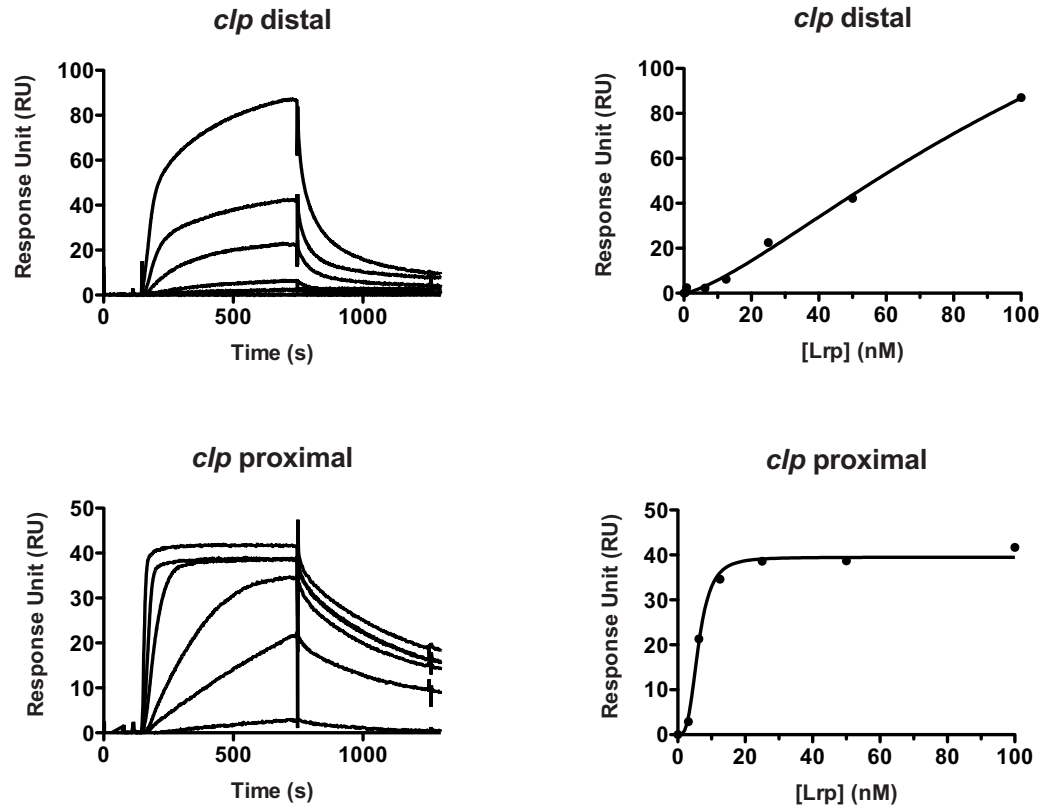


Figure 2

Fragment	Affinity (K_D)		
	Lrp	PapI	Foal
<i>clp</i> distal	11 $\mu\text{M} \pm 1.0$	12.90 $\mu\text{M} \pm 0.98$	14.17 $\mu\text{M} \pm 1.4$
<i>clp</i> proximal	10 nM ± 1.0	16.67 $\mu\text{M} \pm 1.14$	14.13 $\mu\text{M} \pm 0.73$

Table 1

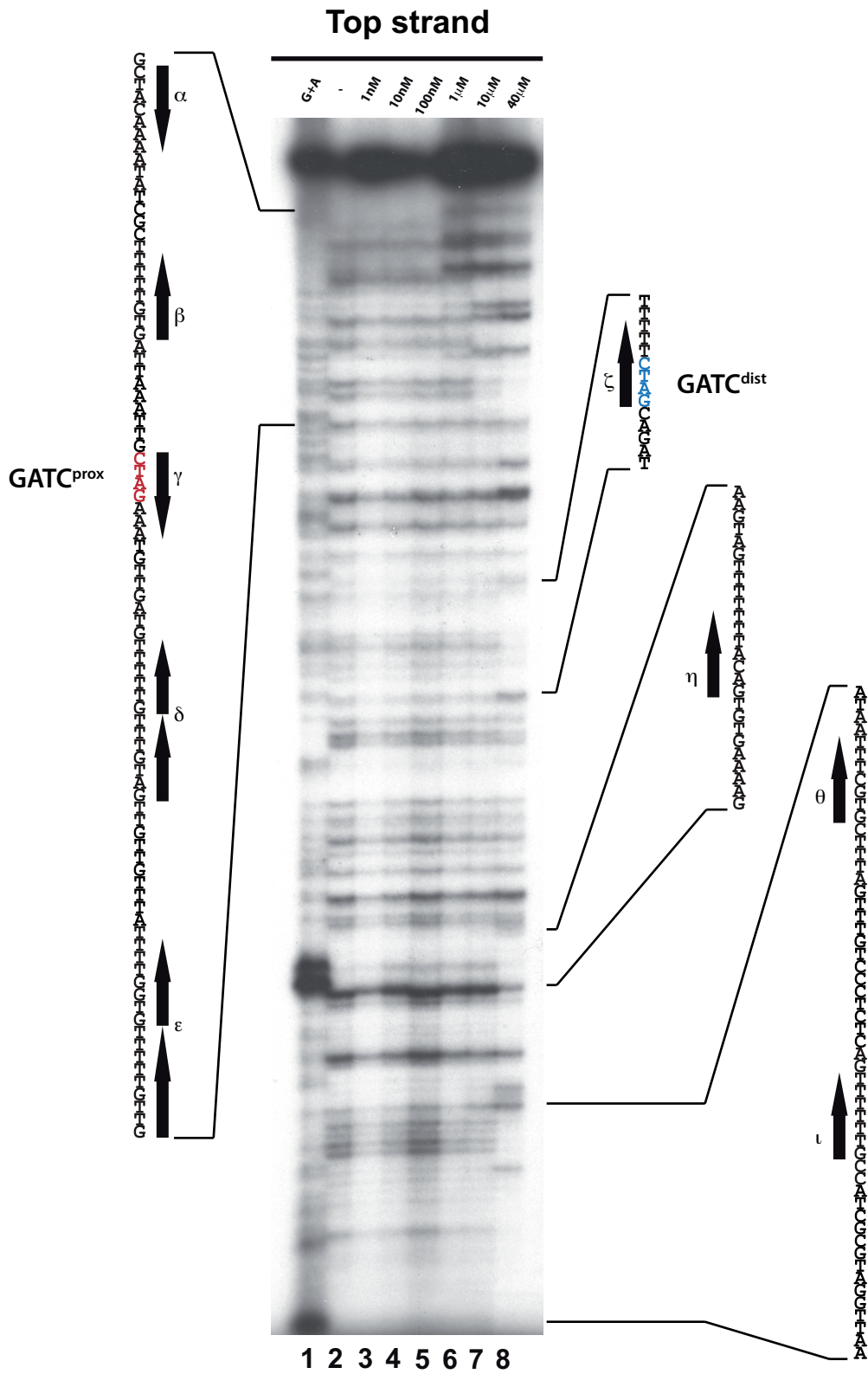


Figure 3

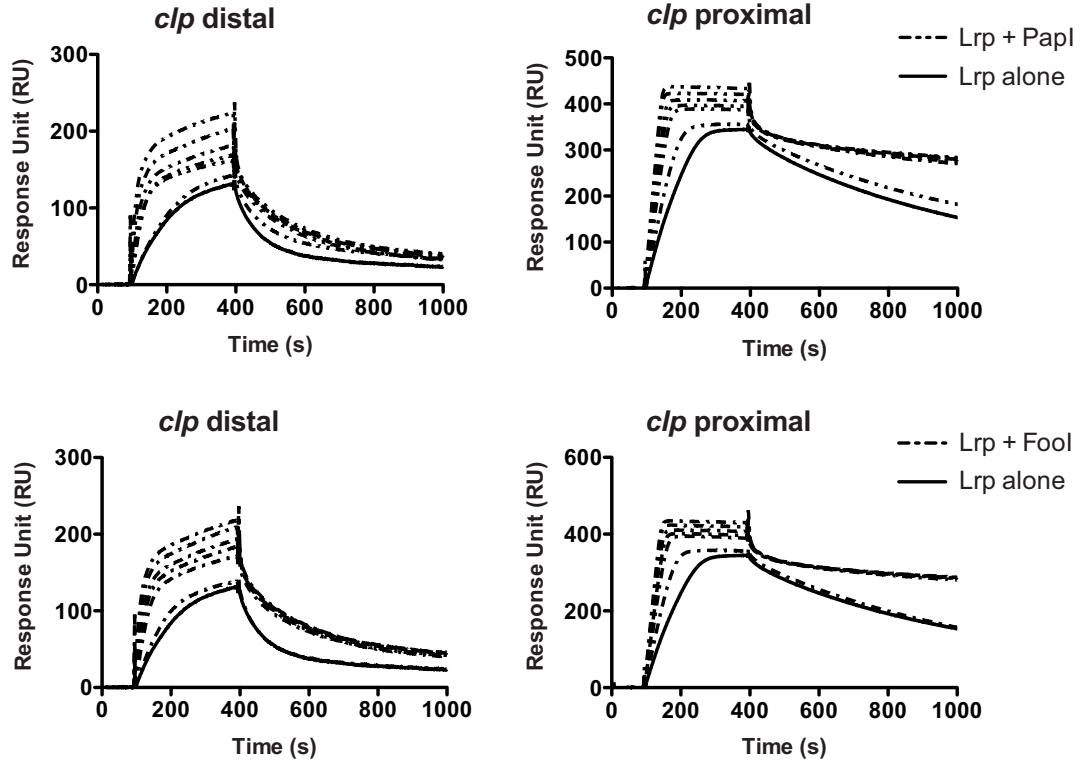


Figure 4

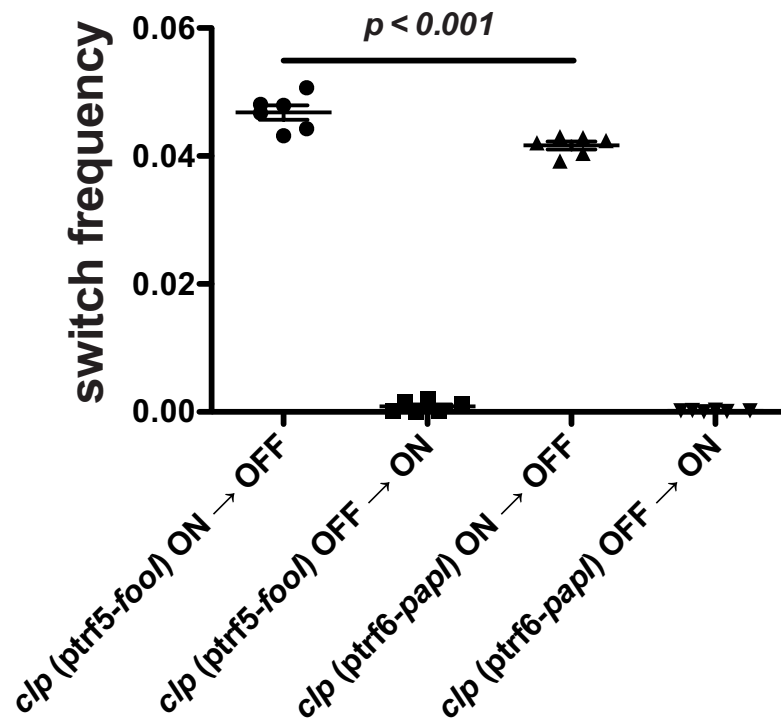


Figure 5

Strain/phenotype	Phenotypic switch			
	Total no. of colonies counted	No. of [Lac ⁺]/no. Of [Lac ⁻] colonies	% of [Lac ⁺] colonies	Switch frequency (α or β)
MC4100.λ6 (ptrf5-fool)				
[Lac ⁻]	5889	239/5650	4.05%	$\alpha \Rightarrow 1.0 \times 10^{-3}$
Second colony as above	4838	57/4781	1.18%	$\alpha \Rightarrow 0.6 \times 10^{-3}$
[Lac ⁺]	4864	155/4709	3.18%	$\beta \Rightarrow 4.9 \times 10^{-2}$
Second colony as above	6664	575/6089	8.62%	$\beta \Rightarrow 4.4 \times 10^{-2}$
MC4100.λ6 (ptrf6-papl)				
[Lac ⁻]	7774	19/7755	0.2%	$\alpha \Rightarrow 8.2 \times 10^{-5}$
Second colony as above	6141	29/6112	0.4%	$\alpha \Rightarrow 0.2 \times 10^{-3}$
[Lac ⁺]	8701	72/8629	0.8%	$\beta \Rightarrow 4.0 \times 10^{-2}$
Second colony as above	7099	149/6950	2.1%	$\beta \Rightarrow 4.2 \times 10^{-2}$

Table 2

Implication de l'auteur dans cet article.

En tant que premier auteur de cet article, j'ai réalisé l'ensemble des constructions bactériennes, ainsi qu'au dénombrement des différents types de colonies. J'ai également participé aux expériences de cytométrie en flux. J'ai analysé l'ensemble des résultats présentés et ai rédigé l'intégralité de cet article.

C. Article III: Single cell measurements of type P fimbriae expression in *Escherichia coli* show a higher level of complexity during phase variation

Authors:

Richard Graveline,¹ Serge Sénéchal,² Rémi Lavoie,¹ France Daigle,^{1,2} Christine Martin,³ and Josée Harel^{1*}

Addresses:

Groupe de Recherche sur les Maladies Infectieuses du Porc (GREMIP) and Centre de Recherche en Infectiologie Porcine (CRIP), Université de Montréal, Saint-Hyacinthe, Québec, CANADA;¹ Département de Microbiologie et Immunologie, Université de Montréal, Montréal, Québec, CANADA;² and Institut National de la Recherche Agronomique, Saint-Genès-Champanelle, FRANCE³

*Corresponding author. Mailing address: Groupe de Recherche en Infectiologie Porcine (GREMIP), Université de Montréal, Faculté de Médecine Vétérinaire, C.P. 5000, Saint-Hyacinthe, Québec, Canada J2S 7C6. Phone: (450) 773-8521, ext. 8233. Fax: (450) 778-8108. E-mail:

1- Abstract

Fimbriae F165₁ synthesized by septicemic *Escherichia coli* strains belong to the type P family of adhesive factors and are under the control of a methylation-dependent transcriptional regulation, called phase variation. Phase variation is defined as the ability of bacteria to switch between two phenotypes, resulting in a heterogeneous phenotype among individual cells within a clonal population. If F165₁ and Pap share similar regulatory mechanisms, F165₁ fimbriae expression however has its own specificity, with notably a more heterogeneous phenotype and a predominant ON state. Recently, we showed that the highly heterogeneous phenotype in F165₁ phase variation is due to an impairment in the complex formed by the leucine-responsive regulatory protein (Lrp) at its repressor binding sites 1-3, in combination with a specific action of the local regulator Fool (homologue to PapI), which promotes the transit of Lrp toward its activator binding sites 4-6. In order to further characterize phase variation of type P fimbriae, we monitored the phase variation of Pap and F165₁ in individual *E. coli* cells in real time, using an amplification system that enables single cells in the ON and OFF states of the *foo* and *pap* strains to be distinguished by flow cytometry. Our results confirmed that *pap* cells are mostly in the OFF state, whereas *foo* cells are preferentially ON, showing a heterogeneous phenotype. However, we observed that ON cells displayed a higher level of diversity and that conversion of Fool to PapI generated an intermediate Pap phase variation phenotype of the *foo* cells. These results constitute an important step in the understanding of F165₁ phase variation and its physiological role.

2- Introduction

F165₁ fimbriae, encoded by the *foo* gene cluster, are related to the Pyelonephritis-associated pili (Pap) that are involved in the pathogenesis of urinary tract infections (Plos, *et al.*, 1990, Wullt, *et al.*, 2000, Wullt, *et al.*, 2001, Bergsten, *et al.*, 2004). Although their role is not fully understood, previous work

showed that F165₁ fimbriae are essential for virulence and survival during systemic stages of infection (Ngeleka, *et al.*, 1994, Ngeleka & Fairbrother, 1999). As members of the type P family, both F165₁ and Pap are subject to phase variation control. Phase variation is a stochastic and reversible switch between an “all-or-none” (ON/OFF) phenotype, each capable of responding to a specific condition. After cell division, most daughter cells retain the expression phase of the parent, but a minority switches to the other expression phase.

Mechanistically, Pap phase variation occurs because of methylation-controlled modifications within the regulatory region of the *pap* operon. This region consists of an approximately-400 bp intergenic region that is surrounded by two divergently transcribed genes, *papI* and *papB*. The intergenic region also includes six leucine-responsive regulatory protein (Lrp) binding sites, as well as two GATC sites (referred to as GATC^{prox} and GATC^{dist} respectively) spaced 102 bp apart. The differential methylation status of these sequences determines the binding of Lrp. Indeed, the switch from one phenotype to the other arises from the competition between the binding of Lrp at its distal or proximal binding sites and methylation by Dam methyltransferase at the opposite GATC site (for reviews, see (Blomfield, 2001, Hernday, *et al.*, 2002, Casadesús & Low, 2006)). Thus, when GATC^{dist} is fully methylated, Lrp cooperatively binds to sites 1-3, maintaining the cells in the OFF state and when GATC^{prox} is fully methylated, Lrp binds to sites 4-6, maintaining the ON state (Fig. 1). As a result, the probability of Lrp translocation from one site to the other is affected by the DNA methylation pattern. Hemimethylation of GATC does not prevent Lrp movement, but reduces the translocation probability compared to unmethylated GATC.

Control of *pap* expression also requires the action of PapI, a positive regulator that increases the affinity of Lrp for sites 4-6 *in vivo* (van der Woude, *et al.*, 1996, Hernday, *et al.*, 2002). PapB, the second specific regulator of the *pap* operon, plays an important role at the transcriptional level, mainly by coordinating the expression of the *pBA* and *pI* promoters (Hernday, *et al.*, 2002). Taken together, the combined actions of Lrp, Dam methyltransferase, PapI, and PapB result in a

finite probability of each cell to express (ON) or not express (OFF) Pap fimbriae just after DNA replication.

Previously, we showed that, as with *pap*, *foo* expression relies on the concerted action of Lrp, Dam methylase, FooB, and of FooI (a PapI homologue, which differs by a single modification in its amino acid sequence) (Daigle, *et al.*, 2000, Harel, *et al.*, 2000, Berthiaume, *et al.*, 2004). However, F165₁ pili are characterized by a higher level of ON cells than other members of the type P family of adhesive factors. In this work, we adapted a single-cell measurement approach to distinguish changes during phase variation of *foo* and *pap* operons. Since PapI and FooI play a key role during phase variation, we also analyzed their influence by exchanging them in both regulatory systems. Our results show a higher level of complexity than previously thought using the *lacZYA* reporter system.

3- Materials and methods

a) Media and growth conditions

Luria-Bertani (LB) broth, M63, and M9 minimal broths were prepared as described previously (Miller, 1972, Thomason, *et al.*, 2007). When necessary, media were supplemented with antibiotics at the following concentrations (unless otherwise noted): ampicillin (amp), 100 $\mu\text{g.ml}^{-1}$; chloramphenicol (Cm), 10 $\mu\text{g.ml}^{-1}$; gentamycin, 10 $\mu\text{g.ml}^{-1}$; kanamycin, 40 $\mu\text{g.ml}^{-1}$; tetracycline (tet) 12.5 $\mu\text{g.ml}^{-1}$; as well as with 5-bromo-4-chloro-3-indolyl-b-D-galactopyranoside (X-gal) at a final concentration of 40 $\mu\text{g.ml}^{-1}$.

b) Strains and plasmids

Bacterial strains and plasmids used in this study are listed in Table 1. Strain JH200, JH201 and JH204 were constructed by first amplifying the *cat-sacB* from strain NC397 (Svenningsen, *et al.*, 2005) using primers i-cat2

(5'CCTGGTAATCCGTTACCGCCAGCGCCTCTGCAATTTCTGCAAAAATGAGACGTTGATCGGCACGT-3') and fooUB-sacB (5'TACTCTTCACGCAATAAGTTAAATTTAGTTTTTTATGTTGATCAAAGGGAAAAGTGTCCATATGCACAGATG-3'). The PCR product was then recombined into strain DL4388 using the λ red-recombination system as previously described (Murphy & Campellone, 2003). The recombinant strain was selected on LB agar supplemented with chloramphenicol and tested for sensitivity on sucrose minimal plates (M63 salts, 0.2% glycerol, 6% sucrose). The resulting strain was *papI'::cat-sacB-papB-lacZYA*. A single confirmed Cm^R/sucrose-sensitive clone was then used to generate strains JH200, JH201 and JH206, using the λ red-recombination system in a second round of recombineering. All recombinant strains were generated using the same procedure, except for the forward primer and the DNA template used. For strain JH200 and 201, the PCR product was amplified on strain 4787 using pI/fli-fwd (5'-CCTGGTAATCCGTTACCGCCAG-3') and fooI->papI-F (5'-AATCCGTTACCGCCAGCGCCTCTGCAATTTCTGCCGTTTTCCCTCCATCATGCCTGTT CAGAAATTCC-3') as forward primer, respectively; for strain JH204, the PCR product was amplified on strain DL4388 using papI->fooI-F (5'AATCCGTTACCGCCAGCGCCTCTGCAATTTCTGCAGTTTTCCCTCCATTATGCCTG TTCAGAAATTCC-3') as forward primer. fooUB40 (5'-TACTCTTCACGCAATAAGTTAAATTTAGTTTTTTATGTTG-3') was used as reverse primer for all the PCR amplification. Recombinant strains were selected for loss of the *cat-sacB* cassette on sucrose M63 minimal plates and at least ten sucrose-resistant colonies were subsequently tested for sensitivity on LB-Cm agar. Candidates were then purified on LB plates and confirmation of the resulting fusions was then obtained by sequencing with fooI-fwd (5'-AGTTGTGGAAGAAGACAGCTTTGCC-3') and fooP-rvs (5'-ACTTCATGATGCGCCATGTTTCCC-3') as primers.

For strains JH300, JH301, JH303 and JH304, a PCR fragment corresponding to the T7 RNA polymerase gene (*T7pol*) and the chloramphenicol resistant gene (*Cam^R*) was first amplified from strain cHNL135 using lacZ-T7pol_fwd (5'-

GCATTTTAACTTTCTTTATCACACAGGAAACAGCTACACGATTAACATCGCTAAGAA
 C-3') and lacZ-CAM_R (5'-
 CGGTCATCGCCGGTAGCCAGCGCGGATCATCGGTCAGACGATGGAGTTCTGAGGTCA
 TTAAGT-3'). The PCR product was then recombined into strains JH200, JH201,
 DL4388 and strain JH204 to generate strains JH300, JH301, JH303, and strain
 JH304, using the same λ red-recombination system as before. Recombinant
 strains were selected for resistance on LB-Cm agar.

Since strains JH300, JH301, JH303 and strain JH304 already contained in their
 genome a kanamycin-resistance cassette, we constructed the reporter plasmid
 pHL40 by replacing the kanamycin-resistance cassette from plasmid pHL32 by a
 gentamycin-resistance cassette. Briefly, a gentamycin-resistance cassette was
 first amplified from plasmid pUCGM using primers SacI-genta-F (5'-
 TACCGAGCTCGAATTGGCCGCGGCG-3') and AatII-genta-R (5'-
 ATTAGACGTCGGAATTGCCAGCTGGGGCGCCCTCTGGTAAGGTTGGGAAGCCCATTG
 ACATAAGCCTGTTTCGGTTCGTAAACTGTAATGC-3'). Plasmid pHL32 was then
 digested using the endonucleases AatII and SacI (Neb) and, after purification, the
 DNA fragment containing the T7 RNA polymerase-specific promoter and *gfp* was
 ligated with the PCR-amplified gentamycin-resistance cassette to generate
 plasmid pHL40. For cloning procedures, PCR amplification was carried out using
 the FidelityTaq PCR MasterMix (GE Healthcare).

c) Measurement of phase variation using the *lacZYA* reporter system

The switch frequency was calculated as described previously (Blyn, *et al.*, 1989,
 Crost, *et al.*, 2003). Strains were inoculated onto LB-kan agar plates. Colonies
 were then excised from the agar and resuspended in M9 salts containing no
 carbon source. Appropriate dilutions were subsequently spread onto M9 X-Gal
 agar plates containing 0.2% glycerol and digitally photographed after 24 hours
 of growth at 37°C. Three colonies showing a uniform Lac⁺ or Lac⁻ phenotype were
 excised from the agar, resuspended in M9 salts, and spread onto M9 X-gal agar
 as before. The passage from M9 X-gal agar to fresh M9 X-gal agar containing the
 same source of carbon was repeated two more times to follow the switch

frequency through three generations. At least 2,000 colonies were scored for a Lac phenotype as described (Blyn, *et al.*, 1989) and their switch frequencies calculated according to the formula $(M/N)/g$, where M is the number of cells that underwent phase variation, N the total number of cells evaluated and g the total number of generations (estimated to be 20 for all the strains tested in our experiments).

d) Flow cytometry acquisition and analysis

Cells were streaked on LB plates with appropriate antibiotics. After 24 hours of growth at 37°C, a single colony was subsequently spread onto M9 X-Gal agar plates containing 0.2% glycerol and 10 µg/ml of gentamycin. After another 24 h of growth, cells were inoculated into 1 mL sterile PBS, and ON and OFF single cells were separated using a BectonDickinson FACSort flow cytometer (488-nm excitation laser; 525-nm emission filter), according to the level of fluorescence intensity. ON and OFF (corresponding to a low level of fluorescence intensity due to the basal expression of *gfp*) populations were then resuspended in 1 mL of fresh M9 media containing 0.2% glycerol and let for growth at 37°C for 20 hours. Data from at least 20,000 individual cells were collected for each sample and analysed with a BectonDickinson FACScalibur flow cytometer (488-nm excitation laser; 525-nm emission filter). FACScalibur data were converted to ASCII format using MFI (E. Martz, University of Massachusetts, Amherst) and analyzed with FlowJo software (Tree Star, Inc.).

4- Results and discussion

a) The *foo* and *pap* regulatory regions confer different level of ON colonies.

foo and *pap* operons share more than 96% identity within their regulatory region; yet, previous work showed that a high level of ON cells characterizes *foo* phase variation. In order to determine the levels of expression during *foo* and *pap* phase variation, we plated strains DL4388 and JH200 on M9 X-Gal agar plates containing 0.2% glycerol. DL4388 and JH200 contain respectively a

chromosomal *pap-lacZYA* fusion and a *foo-lacZYA* fusion in the *E. coli* MC4100 background. To characterize the influence of these specific regulators FooI and PapI, we also included in this study two strains for which *fooI* was replaced by *papI* (strain JH201) and *papI* was replaced by *fooI* (strain JH204). When the phenotype of the colonies was followed through generations, phase variation of strains DL4388 and JH200 was heritably transmitted (Table 2). Indeed, when ON colonies were plated on M9-glycerol agar, more than half of the progeny remained in the ON state, with approximately 55% and 86% of ON colonies for strain DL4388 and strain JH200, respectively (Table 2). Similarly, plating OFF colonies on fresh M9-glycerol agar gave rise to a minority of ON colonies, with less than 1% and 10% for DL4388 and JH200, respectively (Table 2). Notably, colonies in the OFF state showed two distinguishable phenotypes on M9-glycerol plates supplemented with X-Gal between strain DL4388 and strain JH200. Indeed, while *pap* OFF colonies appeared white on M9-glycerol plates, *foo* OFF colonies were characterized by a light blue phenotype (data not shown). This indicates that either the basal transcriptional level of OFF individual cells within an OFF colony is not completely shut off or that there is a high proportion of ON individual cells inside each OFF colony. Moreover, replacing *fooI* with *papI* influenced the level of ON cells, especially when the starting phenotype corresponded to OFF colonies (Table 2, compare strains JH200 and JH201). This suggests that differences in sequence between PapI homologues may be involved in the relative level of ON cells. However, replacing *papI* with *fooI* did not influence the switch rate from OFF to ON or from ON to OFF (Table 2, compare strains DL4388 and JH204). Recently, we showed that FooI and PapI bind at low micromolar affinities to the distal and proximal regions of the *foo* and *pap* operons (Graveline, *et al.*, 2011). Moreover, we showed that if replacing PapI by FooI did not modify either the OFF to ON or the ON to OFF switch rate, replacing FooI by PapI unexpectedly lowered the OFF to ON switch frequency (Graveline, *et al.*, 2011). Thus, it appears that FooI contributes to the OFF to ON switch rate. Taken together, these results confirm the observation that phase

variation of type P fimbriae is inherited in an epigenetic manner and that F165₁ fimbriae is characterized by a high level of cells in the ON state.

b) Real time measurement of *foo* and *pap* phase variation reveals a third sub-population at the single-cell level.

Although the use of a *lacZ* reporter system allowed us to characterize *foo* and *pap* phase variation, the different phenotypes between *foo* and *pap* OFF colonies prompted us to develop a new reporter system that would facilitate single-cell measurements. We therefore adapted a genetic amplifier reporter system to better characterize *foo* and *pap* phase variation (Lim & van Oudenaarden, 2007). This amplification system replaced the coding sequence of *lacZ* in the bacterial chromosome of strains JH200 and DL4388 with the coding sequence of the T7 RNA polymerase (*T7pol*), resulting in strains JH300 and JH303. Strains were also transformed with the reporter plasmid pHL40 that express *gfp* under the control of the T7pol-specific promoter. Thus, the expression of multiple *gfp* transcripts by each T7pol molecule, which is under the control of the *foo* or *pap* regulatory region, and the placement of *gfp* on a multicopy plasmid increases the fluorescence signal and improves the signal-to-noise ratio. The amplification system enabled colonies and single cells in the ON and OFF states of each strains to be distinguished by fluorescence microscopy and flow cytometry. As controls, we also transformed strain JH200 with plasmid pHL40 (negative control for basal expression of GFP) and the laboratory strain BL21DE3 that express T7pol for protein expression and purification purpose (positive control). Similarly to strains JH300 and JH303, we also transformed strains JH201 and JH204 to replace *lacZ* with *T7pol*, resulting in strains JH301 and JH304 (Table 1).

As shown in Fig. 3, the histograms of single-cell fluorescence measured by flow cytometry showed different levels in *gfp* expression between strains. For the two control strains, single-cell measurement of the negative control strain JH200 transformed with the reporter plasmid pHL40 always gave rise to a low level of fluorescence, whether the initial bacterial population corresponded to a low level, or a high level of GFP. Similarly, single-cell measurement of the positive

control strain BL21DE3 transformed with the reporter plasmid pHL40 always gave rise to a high level of fluorescence, whether the initial bacterial population was selected according to their level of GFP production. This confirmed to us that the expression of *gfp* was strictly depending on the presence of T7pol and that all the differences in fluorescence for the four remaining strains would be due to the expression of *T7pol*, under the control of phase variation. For most of the four remaining strains, the levels in *gfp* expression strongly differed depending on the starting phenotype. Moreover, selecting an initial bacterial population with a high level of fluorescence resulted in a higher level of GFP production than when initial bacterial population showed a low level of GFP, and vice versa. Thus, the expression of *gfp* was inherited in an epigenetic manner, as it was the case for the expression of *lacZ*.

When strain DL4388 was compared with strain JH200, we observed a higher repression in the presence of the *pap* regulatory region than in the presence of the *foo* regulatory region, as already shown with the use of the *lacZ* reporter system. Moreover, replacing *fool* with *papI* had a great impact on the intensity of fluorescence and replacing *papI* with *fool* resulted in the loss of the inheritability of phase variation. Since Fool and PapI only differ from each other by one amino acid, the reason why swapping these two homologous regulators causes such change in the intensity of fluorescence remains unclear.

When we compared the histogram profiles between two strains, more than two different populations were observed as evidenced by the intensity of fluorescence. Indeed, each profile showed three bacterial populations, corresponding to a low, an intermediate and a high level of fluorescence (Fig. 3). Thus, our results suggest that phase variation of type P fimbriae is not, as thought before, strictly a random transition in promoter configurations that result either in a highly transcriptionally active state or an inactive state but intermediate levels also exist.

We then estimated the relative proportion of each of the three bacterial populations observed according to the fluorescence intensity. As shown in Table

3, half of the bacterial population showed an intermediate level of fluorescence intensity for most of the strains, ranging from 58% (*foo wt*) to more than 68% (*fooI-pap*) when initial bacterial population showed a low level of fluorescence intensity, and from 48% (*foo wt*) to 67% (*fooI-pap*), when initial bacterial population showed the higher level of fluorescence intensity. In comparison, the two other low and high levels of fluorescence intensity greatly varied, representing from 13% to 27% for the low level of fluorescence intensity, and from 14% to 38% for the high level of fluorescence intensity, depending on the strain and the starting phenotype. Strain JH403, carrying the full-length *pap* regulatory region, however, showed a low level of fluorescence intensity more than 60% of bacterial cells. In contrast, the proportion of cells showing an intermediate level of *gfp* expression represented from 13 to 36% and cells showing a high level of *gfp* expression never reached a percentage greater than 1.87%. Since the ON state is inherited in an epigenetic manner as for the OFF state, this finding appears different from our previous experiments (Graveline, *et al.*, 2011). Thus, we concluded that the bacterial population showing an intermediate level of fluorescence intensity could represent at least a part of the ON population. Lim *et al.*, while studying switching mechanism of *agn43*, also observed that the transition bias away from the hemimethylated state and the intermediate unmethylated states towards the fully methylated and off states in *agn43*, respectively, acts as a buffer that 'captures' cells as they leave the on and off states and returns them to their original state. As a consequence, cells that leave the on state owing to DNA replication, or that leave the off state owing to fluctuations in repression of the partial state, rarely switch to the opposite state (Lim & van Oudenaarden, 2007). Similarly, we propose that the intermediate population could play the role of a buffer, avoiding switching directly from an OFF phase to the opposite ON phase, and vice versa (Fig. 4). However, other systems may be applied. In the case of type 1 fimbriæ, the mechanism of *fim* (fimbriæ) switching is characterized by a reversible DNA inversion event, in which a *cis* regulatory element that contains the promoter of the *fim* operon changes its orientation. The DNA inversion process involves the formation of a

DNA recombinase (FimB and/or FimE) complex forming a loop of the 314-bp invertible region, the subsequent breaking and rejoining of single DNA strands, and the dissipation of the synapse complex to allow transcription of the *fim* operon. Unlike other phase variation control mechanisms such as slipped-strand mispairing or specific Dam methylation sites, the inversion of *fimS* is not expected to require DNA replication as an essential step in the switching process (Adicptaningrum, *et al.*, 2009). However, the OFF-to-ON *fim* switching occurs preferentially at the beginning of the cell cycle, before replication of the *fim* genes.

5- Concluding remarks

Phase variation of type P fimbriae arises from competition between two global regulatory proteins, Lrp and Dam methyltransferase, within the regulatory region.

Since Dam methylation and Lrp binding are mutually exclusive, they must compete during these events. Thus, to switch from one configuration to the other, Lrp must bind to hemi-methylation region. As a result, 4 different situations can be seen: -i- Lrp bound to unmethylated GATC^{dist}, -ii- Lrp bound to hemi-methylated GATC^{dist}, -iii- Lrp bound to unmethylated GATC^{prox}, and -iv- Lrp bound to hemi-methylated GATC^{prox}. Previous works always considered that cells presenting Lrp bound to unmethylated or hemi-methylated GATC would show the same phase. However, the presence in our experiments of an intermediate population corresponding to half of the bacterial population suggest that we cannot consider these two situations as similar (Fig. 3). Thus, the switch from the OFF phase to the ON phase (or from the ON phase to the OFF phase) would require transiting to an intermediate level of expression. As a result, multistep epigenetic switches with biased transition rates may exist in various configurations.

6- References

- [1] Adiciptaningrum AM, Blomfield IC & Tans SJ (2009) Direct observation of type 1 fimbrial switching. *EMBO Rep* **10**: 527-532.
- [2] Bergsten G, Samuelsson M, Wullt B, Leijonhufvud I, Fischer H & Svanborg C (2004) PapG-dependent adherence breaks mucosal inertia and triggers the innate host response. *J Infect Dis* **189**: 1734-1742.
- [3] Berthiaume F, Crost C, Labrie V, Martin C, Newman EB & Harel J (2004) Influence of L-leucine and L-alanine on Lrp regulation of *foo*, coding for F1651, a Pap homologue. *J Bacteriol* **186**: 8537-8541.
- [4] Blomfield IC (2001) The regulation of *pap* and type 1 fimbriation in *Escherichia coli*. *Adv Microb Physiol* **45**: 1-49.
- [5] Blyn LB, Braaten BA, White-Ziegler CA, Rolfson DH & Low DA (1989) Phase-variation of pyelonephritis-associated pili in *Escherichia coli*: evidence for transcriptional regulation. *EMBO J* **8**: 613-620.
- [6] Casadesús J & Low D (2006) Epigenetic gene regulation in the bacterial world. *Microbiol Mol Biol Rev* **70**: 830-856.
- [7] Crost C, Garrivier A, Harel J & Martin C (2003) Leucine-responsive regulatory protein-mediated repression of *clp* (encoding CS31A) expression by L-leucine and L-alanine in *Escherichia coli*. *J Bacteriol* **185**: 1886-1894.
- [8] Daigle F, Forget C, Martin C, Drolet M, Tessier MC, Dezfulian H & Harel J (2000) Effects of global regulatory proteins and environmental conditions on fimbrial gene expression of F165(1) and F165(2) produced by *Escherichia coli* causing septicaemia in pigs. *Res Microbiol* **151**: 563-574.
- [9] Graveline R, Mourez M, Hancock MA, Martin C, Boisclair S & Harel J (2011) Lrp-DNA complex stability determines the level of ON cells in type P fimbriae phase variation. *Mol Microbiol* **81**: 1286-1299.
- [10] Harel J, Daigle F, Forget C, Tessier MC, Crost C & Martin C (2000) Phase variation of F165(1) (Prs-like) fimbriae from *Escherichia coli* causing septicaemia in animals. *Can J Microbiol* **46**: 1101-1107.

- [11] Hernday A, Krabbe M, Braaten B & Low D (2002) Self-perpetuating epigenetic pili switches in bacteria. *Proc Natl Acad Sci USA* **99 Suppl 4**: 16470-16476.
- [12] Lim HN & van Oudenaarden A (2007) A multistep epigenetic switch enables the stable inheritance of DNA methylation states. *Nat Genet* **39**: 269-275.
- [13] Miller JH (1972) *Experiments in Molecular Genetics*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- [14] Murphy KC & Campellone KG (2003) Lambda Red-mediated recombinogenic engineering of enterohemorrhagic and enteropathogenic *E. coli*. *BMC Mol Biol* **4**: 11.
- [15] Ngeleka M & Fairbrother JM (1999) F165(1) fimbriae of the P fimbrial family inhibit the oxidative response of porcine neutrophils. *FEMS Immunol Med Microbiol* **25**: 265-274.
- [16] Ngeleka M, Martineau-Doizé B & Fairbrother JM (1994) Septicemia-inducing *Escherichia coli* O115:K"V165" F165(1) resists killing by porcine polymorphonuclear leukocytes in vitro: role of F165(1) fimbriae and K"V165" O-antigen capsule. *Infect Immun* **62**: 398-404.
- [17] Plos K, Carter T, Hull S, Hull R & Svanborg Eden C (1990) Frequency and organization of *pap* homologous DNA in relation to clinical origin of uropathogenic *Escherichia coli*. *J Infect Dis* **161**: 518-524.
- [18] Svenningsen SL, Costantino N, Court DL & Adhya S (2005) On the role of Cro in lambda prophage induction. *Proc Natl Acad Sci USA* **102**: 4465-4469.
- [19] Thomason L, Court DL, Bubunenko M, Costantino N, Wilson H, Datta S & Oppenheim A (2007) Recombineering: genetic engineering in bacteria using homologous recombination. *Curr Protoc Mol Biol* **Chapter 1**: Unit 1.16.
- [20] van der Woude M, Braaten B & Low D (1996) Epigenetic phase variation of the *pap* operon in *Escherichia coli*. *Trends Microbiol* **4**: 5-9.
- [21] Wullt B, Bergsten G, Connell H, Röllano P, Gebretsadik N, Hull R & Svanborg C (2000) P fimbriae enhance the early establishment of *Escherichia coli* in the human urinary tract. *Mol Microbiol* **38**: 456-464.

[22] Wullt B, Bergsten G, Connell H, Röllano P, Gebratsedik N, Hang L & Svanborg C (2001) P-fimbriae trigger mucosal responses to *Escherichia coli* in the human urinary tract. *Cell Microbiol* **3**: 255-264.

7- Legends

Figure 1. Schematic representation of *foo* and *pap* phase variation.

Organization of the regulatory region shared between the *foo* and *pap* operons, with binding sites for the main regulators depicted. To avoid making the schematic diagram too complex, only the organization of the *pap* operon is depicted. The Lrp sites are shown as empty circles. The two GATC sites subject to methylation by Dam, shown as empty diamonds, are GATC^{prox} and GATC^{dist}, which are located within Lrp binding sites 2 and 5, respectively. The CAP and PapB binding sites are shown as open and hatched boxes, respectively. Localization of Lrp, PapI, as well as methylation status is also shown for either the OFF state or the ON state.

Table 1. Strains and plasmids used in this study.

Table 2. Frequency of phenotypic switch. Switch frequency was measured using the formula $(M/N)/g$, where M is the number of colonies having changed of phenotype, N the total number of colonies evaluated, and g the number of generations having produced the colonies (evaluated at 20 during our experimentations).

Figure 2. Schematic representation of the reporter system used in this study.

In order to characterize *foo* and *pap* phase variation, a reporter system was constructed by replacing the chromosomal coding sequence of the fimbrial operon by the sequence of the T7 RNA polymerase (T7pol). The expression of T7pol is under the control of phase variation and, in turn, dictates the expression of plasmidic *gfp* (under the control of a T7-specific promoter; Lim & van

Oudenaarden, 2007). This system allows an amplified signal and thus improves the signal vs background ratio.

Figure 3. Monitoring of phase variation for individualized bacterial cells.

Histograms of single-cell fluorescence for each bacterial population carrying the *gfp*-based genetic amplification system. Strains BL21DE3 and JH200 are used in this study as positive and negative controls, respectively. The remaining strains correspond to four different constructs harbouring the coding sequence of T7pol under the control of the *pap* or *foo* regulatory region, in the presence of *papI* or *fooI*. All strains also contain the reporter plasmid pHL40 that express *gfp* under the control of the T7pol-specific promoter. Bacterial colonies are first isolated from a rich media plate, then streaked on M9 minimal media. After 24h of incubation at 37°C, colonies are resuspended in PBS and bacteria corresponding to the OFF (red) and ON (green) populations are sorted, using a flow cytometer. Cells from strains BL21DE3 and JH200 are similarly sorted to test the efficiency of the procedure. Cells are then reinoculated into liquid M9 minimal media and analyzed by flow cytometry after a second round of incubation. Histograms of single-cell fluorescence for each strains show gaussian fits with either a low level (a), a partial level (b) or a high level (c) of fluorescence intensity, which is thought to correspond to OFF, intermediate and ON bacterial populations, respectively. This results is representative of four independent experiments.

Table 3. Percentages of cells depending on their level of fluorescent intensity analyzed by flow cytometry.

Figure 4. Hypothetical model for phase variation of fimbrial operons of P fimbriae (adapted from Hernday et al, 2002). The phase variation is an epigenetic mechanism. Thus the passage from a phenotype to another requires at least one round of replication.

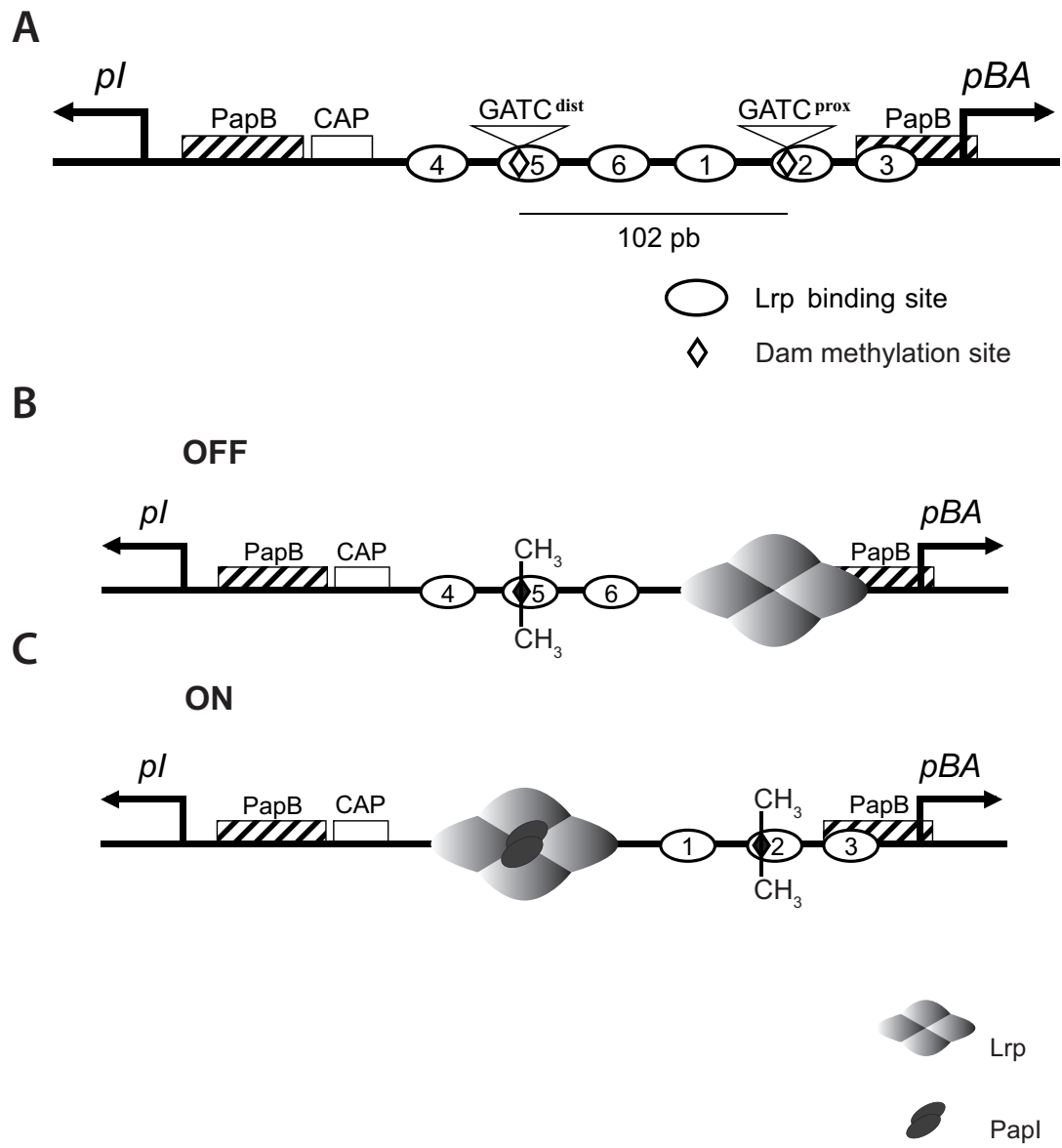


Figure 1

Strain or plasmid	Description and relevant phenotype	Source or reference
<i>E. coli</i> strains		
NC397	<i>lacI::kan-Ter-cat sacB-lacZYA</i>	Svenningsen <i>et al.</i> (2005)
cHNL135	MC4100 <i>agn43::T7pol-Cam^R</i>	Lim <i>et al.</i> (2007)
BL21(DE3)	competent cell for protein expression and purification	Laboratory strain
DL4388	MC4100 λ 354 lysogen (wild type <i>papI-pap-papBA-lacZYA</i>)	Braaten <i>et al.</i> (1994)
JH100	DL4388 <i>papI'::cat-sacB-papBA-lacZYA</i>	This study
JH200	MC4100 λ 354 lysogen (<i>fool-foo-papBA-lacZYA</i>)	This study
JH201	MC4100 λ 354 lysogen (<i>papI-foo-papBA-lacZYA</i>)	This study
JH204	MC4100 λ 354 lysogen (<i>fool-pap-papBA-lacZYA</i>)	This study
JH300	MC4100 λ 354 lysogen (<i>fool-foo-papBA-T7pol-Cam^R</i>)	This study
JH301	MC4100 λ 354 lysogen (<i>papI-foo-papBA-T7pol-Cam^R</i>)	This study
JH303	MC4100 λ 354 lysogen (<i>papI-pap-papBA-T7pol-Cam^R</i>)	This study
JH304	MC4100 λ 354 lysogen (<i>fool-pap-papBA-T7pol-Cam^R</i>)	This study
pKM201	Lambda genes <i>red</i> and <i>gam</i> under the control of the IPTG-inducible Ptac promoter (Amp ^R)	Murphy <i>et al.</i> (2003)
pHL32	Reporter plasmid that express <i>gfp</i> under the control of a T7 RNA polymerase-specific promoter and that possess a kanamycin resistance gene	This study
pHL40	Reporter plasmid that express <i>gfp</i> under the control of a T7 RNA polymerase-specific promoter and that possess a gentamycin resistance gene	This study

Table 1

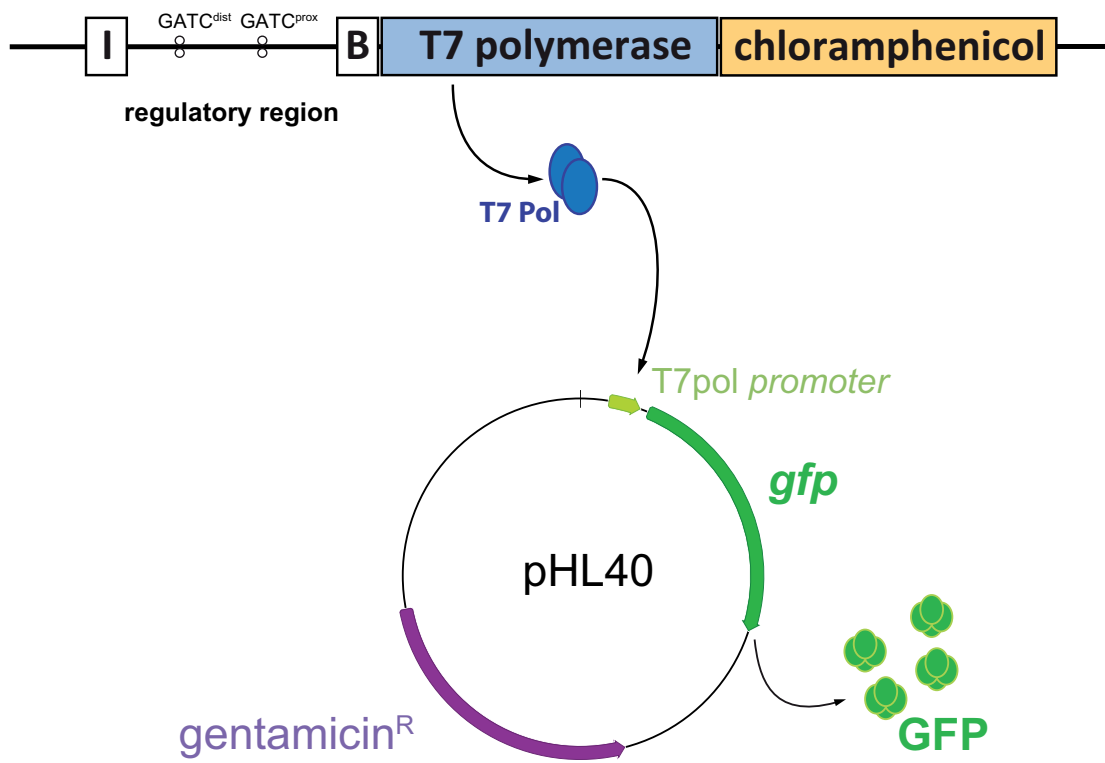


Figure 2

Strain/phenotype	Phenotypic switch			
	Total no. of colonies counted	No. of [Lac ⁺]/no. Of [Lac ⁻] colonies	% of [Lac ⁺] colonies	Switch frequency (α or β)
JH200 (<i>fool-foo-lacZ</i>)				
[Lac ⁻]	8011	677/7334	8.45%	$\alpha \Rightarrow 4.2 \times 10^{-3}$
Second colony as above	7151	561/6590	7.84%	$\alpha \Rightarrow 3.9 \times 10^{-3}$
[Lac ⁺]	7717	6663/1054	86.34%	$\beta \Rightarrow 6.8 \times 10^{-2}$
Second colony as above	6132	5476/656	89.30%	$\beta \Rightarrow 5.2 \times 10^{-2}$
JH201 (<i>papl-foo-lacZ</i>)				
[Lac ⁻]	6727	308/6419	4.58%	$\alpha \Rightarrow 2.3 \times 10^{-3}$
Second colony as above	6206	259/5947	4.17%	$\alpha \Rightarrow 2.0 \times 10^{-3}$
[Lac ⁺]	6757	5636/1121	83.40%	$\beta \Rightarrow 8.3 \times 10^{-3}$
Second colony as above	6761	5590/1171	82.68%	$\beta \Rightarrow 8.7 \times 10^{-3}$
DL4388 (<i>papl-pap-lacZ</i>)				
[Lac ⁻]	6823	29/6794	0.42%	$\alpha \Rightarrow 2.1 \times 10^{-4}$
Second colony as above	7181	49/7132	0.68%	$\alpha \Rightarrow 3.5 \times 10^{-4}$
[Lac ⁺]	7600	4193/3407	55.17%	$\beta \Rightarrow 2.2 \times 10^{-2}$
Second colony as above	6692	3827/2865	57.19%	$\beta \Rightarrow 2.1 \times 10^{-2}$
JH204 (<i>fool-pap-lacZ</i>)				
[Lac ⁻]	6926	42/6884	0.60%	$\alpha \Rightarrow 3.0 \times 10^{-4}$
Second colony as above	6988	33/6965	0.47%	$\alpha \Rightarrow 1.7 \times 10^{-4}$
[Lac ⁺]	6735	3953/2782	58.69%	$\beta \Rightarrow 2.0 \times 10^{-2}$
Second colony as above	6759	4181/2578	61.86%	$\beta \Rightarrow 1.9 \times 10^{-2}$

Table 2

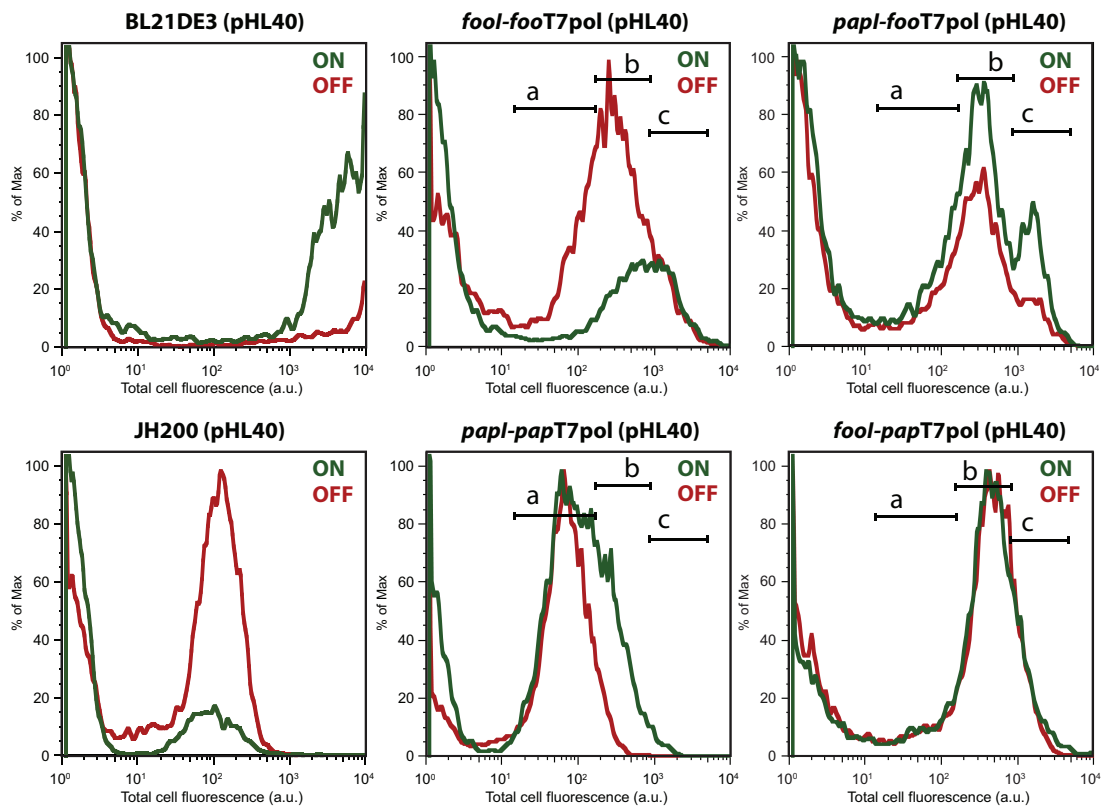


Figure 3

Starting phenotype		strains			
		<i>fool-fooT7pol</i>	<i>papl-fooT7pol</i>	<i>papl-papT7pol</i>	<i>fool-papT7pol</i>
OFF	% OFF cells	27,11	26,08	86,22	13,16
	% intermediate	58,30	58,93	13,64	68,66
	% ON cells	13,95	14,27	0,12	16,74
ON	% OFF cells	13,00	23,19	61,86	13,03
	% intermediate	47,59	54,97	36,09	67,00
	% ON cells	38,26	21,44	1,87	18,61

Table 3

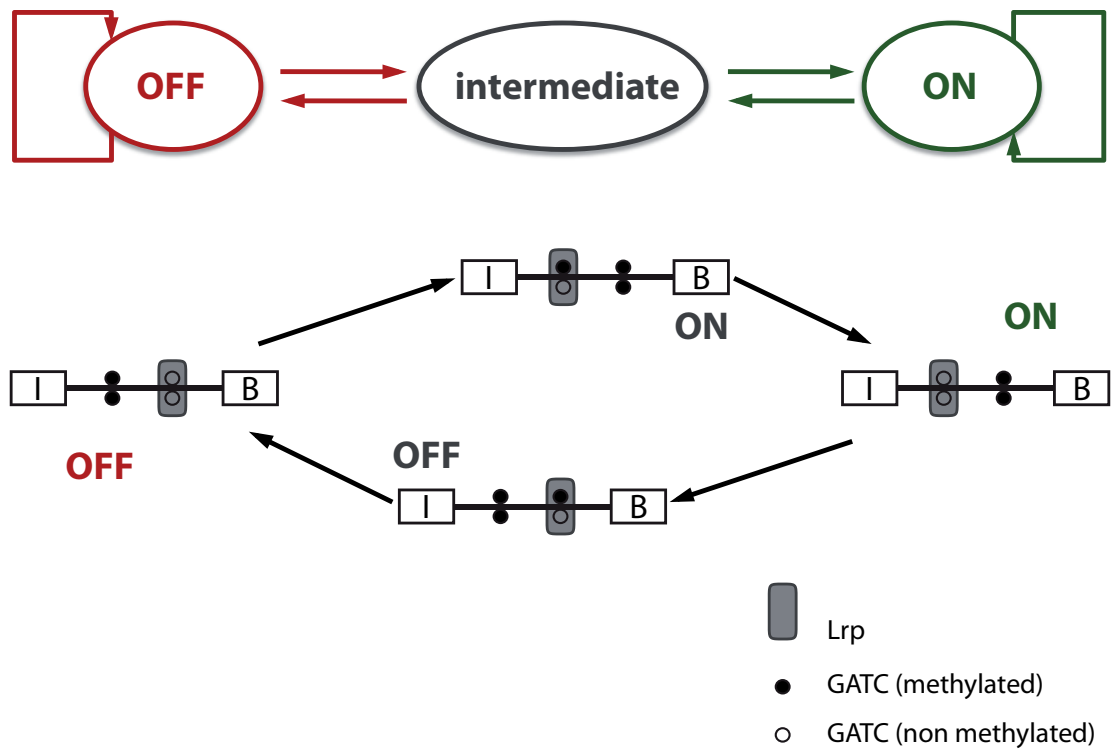


Figure 4

SECTION IV : Discussion – Conclusions.

A- DISCUSSION

L'objectif principal du présent projet de PhD a consisté à comparer la variation de phase chez trois modèles d'études différents : les adhésines fimbriaires F165₁, Pap, et l'antigène fibrillaire de surface CS31A, ceci dans le but de mieux comprendre les mécanismes moléculaires impliqués dans la variation de phase de la famille des fimbriæ de type P. L'utilisation de ces trois systèmes présente en effet l'avantage de pouvoir comparer trois modèles différents dont l'un se caractérise par un niveau de cellules en phase ON élevé (cas des fimbriæ F165₁), l'autre par un niveau de cellules en phase ON au contraire faible (cas des fimbriæ Pap), et, enfin, un troisième pour lequel la variation de phase n'est observable que si le régulateur local est ajouté en *trans*, comme dans le cas de régulation croisée entre deux opérons fimbriaires (cas de l'antigène de surface CS31A).

En utilisant différentes techniques, telles que l'étude de l'expression génétique, la mutagenèse dirigée et plusieurs types d'analyse *in vitro* des interactions entre ADN et protéines, nous avons ainsi démontré qu'une altération dans la stabilité du complexe formé par Lrp sur la région proximale de l'opéron *foo* est à l'origine même des différences de phase entre les deux modèles représentés par F165₁ et Pap. Plus précisément, nous avons démontré que l'interaction de Lrp avec la région proximale de l'opéron *foo* se caractérise par une cinétique de dissociation plus rapide. Ces différences de cinétique sont dues à la présence chez *foo* de deux sites GN(2-3)TTT qui vraisemblablement perturbent le maintien de Lrp sur son site de reconnaissance 1. Dans le cas de Pap, en revanche, l'absence de sites de fixation additionnels pour Lrp, résulte en une cinétique de dissociation plus lente, une fois Lrp fixée sur la région proximale de *pap*, et contribueraient ainsi à maintenir la population de cellules bactériennes en phase OFF. Enfin, dans le cas de CS31A, nous avons montré que la phase majoritaire OFF est due à une affinité de Lrp particulièrement faible

pour la région distale. De plus, l'utilisation de ce dernier système a mis en valeur le rôle joué par le régulateur local PapI, ainsi que par les homologues de PapI dans la régulation croisée entre différents systèmes d'adhésion au sein d'une même bactérie.

1- Les bases moléculaires de la variation de phase

Chez les fimbriæ Pap, tout comme chez les fimbriæ F165₁, la variation de phase est régie selon la région sur laquelle Lrp et le corégulateur PapI se trouvent (Casadesús & Low, 2006, Hernday et al., 2002). La fixation de ce complexe protéique dépend lui-même de l'état de méthylation de deux sites GATC (GATC^{prox} et GATC^{dist}) dans les sites de fixation 2 et 5 de Lrp. Ainsi, chez des cellules considérées en phase OFF, Lrp se retrouve sur les sites 1-3 et protège le site GATC^{prox} de la méthylation par la protéine Dam (Casadesús & Low, 2006, Hernday et al., 2002). De façon réciproque, l'expression des fimbriæ Pap chez des cellules en phase ON intervient lorsque Lrp se retrouve sur les sites 4-6, en combinaison avec PapI, empêchant l'action de la protéine Dam sur le site GATC^{dist}. Dans la mesure où, et Pap, et F165₁, partagent ce système de contrôle par méthylation, nous avons postulé que les différences de variation de phase entre ces deux modèles sont vraisemblablement dues à des différences de séquence dans la région régulatrice, modifiant ainsi la capacité d'interagir avec les régulateurs Lrp et/ou le corégulateur local PapI (ou FooI dans le cas des fimbriæ F165₁).

Dans le cas de l'opéron *pap*, la fixation de Lrp sur les sites 1-3 diminue d'un facteur 10 l'affinité de Lrp pour les sites 4-6, par un mécanisme qui a été qualifié d'exclusion mutuelle (Nou et al., 1995, Hernday et al., 2002). Par une expérience d'empreinte de protection à la DNase I, nous avons montré que, dans le cas de *foo*, Lrp ne présente aucune préférence vis à vis des sites 1-3 ou 4-6. Par conséquent, la phase OFF de l'opéron *foo* ne semble pas être stabilisée par un mécanisme d'exclusion mutuelle sur le même principe que pour l'opéron *pap*. Il

sera toutefois nécessaire de poursuivre d'avantage l'étude de l'absence d'exclusion mutuelle chez *foo* à l'aide de molécules d'ADN circulaires, sur le même mode expérimental que précédemment utilisé pour *pap*, afin de confirmer cette hypothèse (Nou et al., 1995, Hernday et al., 2002). Il n'en demeure pas moins que, l'absence d'exclusion mutuelle observée par empreinte de protection à la DNase I pour *foo* est en accord avec des fréquences de passage d'un état OFF à un état ON et, réciproquement, d'un état ON à un état OFF du même ordre de grandeur, résultant en une population fortement hétérogène.

Dans le cas de l'opéron *clp*, codant l'antigène de surface CS31A, les mécanismes de régulation sont similaires à ceux impliqués dans l'expression des fimbriæ appartenant à la famille de type P. En particulier, la production de l'antigène CS31A est également soumise à la variation de phase (Crost et al., 2003, Martin, 1996). Ainsi, en l'absence d'un homologue de PapI fourni en *trans*, *i.e.* en l'absence de variation de phase, Martin a montré que Lrp protège les sites GATC de la méthylation et que, réciproquement, la méthylation d'un de ces sites empêche la fixation de Lrp (Martin, 1996). Durant cette étude, nous avons montré que Lrp présente une plus faible affinité pour la région distale de *clp*. En effet, dans le cas de *clp*, l'affinité de Lrp pour la région distale est tellement faible que le système est systématiquement orienté vers la population OFF. Ainsi, quand une différence d'un facteur 2 est observable dans la constante d'affinité de Lrp pour les sites 1-3 et 4-6 de *pap*, la même comparaison montre déjà une différence d'un facteur 10^3 entre les régions proximale et distale de *clp*. Alors que la majorité des protéines régulatrices présentent des différences d'affinité entre sites spécifiques et sites non-spécifiques de l'ordre de 10^3 à 10^7 , Peterson *et al.* ont montré que la double fonction de Lrp en tant que régulateur et nucléoprotéine résiderait dans la particularité de cette protéine à présenter des différences d'affinité entre sites spécifiques et sites non-spécifiques de l'ordre d'un facteur 20 à 400 (Peterson *et al.*, 2007). Par conséquent, la comparaison de l'affinité de Lrp pour les différentes régions de *clp* s'apparente plus à la comparaison de l'action de Lrp comme régulateur transcriptionnel (au site

proximal) et comme nucléoprotéine (au site distal) qu'à une compétition entre la fixation de Lrp à un de ses deux sites opérateurs. Ainsi, le rôle majeur de Lrp consisterait à protéger le site GATC^{prox} de la méthylation et ainsi contrôle négativement l'expression basale de l'expression de *clp*.

En conclusion, l'ensemble de ces résultats suggèrent que l'héritage d'une phase majoritaire ON ou OFF, au cours de générations, dépend directement des affinités de Lrp pour les régions proximale et distale de *clp*, *foo* et *pap*. De ce fait, des différences de séquence dans la région régulatrice des opérons fimbriaires ont des conséquences importantes sur la variation de phase des adhésines correspondantes.

2- L'importance d'être stable

Chez *Pap*, il est communément admis que la compétition entre la fixation de Lrp et la méthylation des sites GATC par la protéine Dam joue un rôle prépondérant dans la transmission de la phase OFF, principalement en favorisant la présence de Lrp sur la région proximale de l'opéron *pap* (Peterson & Reich, 2008, Peterson & Reich, 2006). Au regard des constantes biochimiques obtenues durant nos expériences de résonance plasmonique de surface, il apparaît en revanche que la dissociation de Lrp sur le brin d'ADN est le principal élément contrôlant l'action enzymatique de la protéine Dam sur les sites GATC. Ainsi, les différences nucléotidiques d'une souche à l'autre contribuent grandement aux différences dans les taux de phases ON et OFF observés, en altérant la cinétique de dissociation de Lrp aux régions proximale et/ou distale. Il semble donc que la fixation de Lrp sur l'ADN, mécanisme « classique » de régulation génétique, dicte le profil de méthylation de sites GATC, et non l'inverse. Toutefois, d'autres régulateurs tels que PapB, CpxAR, H-NS et la répression catabolique jouent également un rôle important dans la variation de phase des fimbriæ de type P, et pourraient dès lors contribuer aux différences de phase majoritaire d'une souche à l'autre.

3- La différence entre *pap* et *foo* ne tient donc qu'à peu de choses

Chez *E. coli*, Lrp joue le rôle à la fois de régulateur global et de nucléoprotéine impliquée dans l'organisation du génome bactérien (D'Ari et al., 1993, Newman & Lin, 1995). Dans les deux cas, l'action de cette protéine est liée à sa capacité d'interagir avec l'ADN pour des séquences dites dégénérées, *i.e.* non conservées, qui se caractérisent néanmoins par la présence de séquences riches en A/T ainsi que par celle d'une séquence consensus minimale GN(2-3)TTT (Peterson et al., 2007). Chez *pap* comme chez *foo*, chacun des sites 1-6 possède un motif GN(2-3)TTT; de plus, la région intergénique de tous les homologues de *pap* étudiés durant ce projet présente plus de 6 motifs, majoritairement dans la région distale, avec une diversité également importante pour la région proximale (Figure 15). Il serait dès lors possible d'envisager que la présence de sites cryptiques de fixation pour Lrp le long de la région intergénique influence, directement ou non, la stabilité de Lrp à ses sites distal et/ou proximal. Les résultats obtenus lors de cette étude vont dans ce sens. En effet, nous avons montré que changer le nombre de motifs GN(2-3)TTT dans la région proximale modifie non seulement la constante de dissociation de Lrp pour cette région, mais aussi la fréquence de passage d'un phénotype à l'autre et ce, sur plusieurs générations, aussi bien dans le cas de l'opéron *foo* que de l'opéron *pap*. Réciproquement, la présence d'un grand nombre de motifs GN(2-3)TTT à proximité des sites 4-6 dans la région distale pourrait expliquer les différences de cinétique de dissociation observée entre les fragments *pap* distal et *pap* proximal. Ainsi, même s'ils ne sont pas directement impliqués dans la fixation de Lrp sur la région distale, ces sites faciliteraient la dissociation de Lrp et, par la même occasion, contribuerait à orienter les souches possédant l'opéron *pap* vers une phase OFF (Figure 16). De même, il est tout aussi possible d'envisager que l'absence d'un motif GN(2-3)TTT à l'un des trois sites 1 à 3 chez certaines souches ExPECs contribuerait à favoriser une population ON.

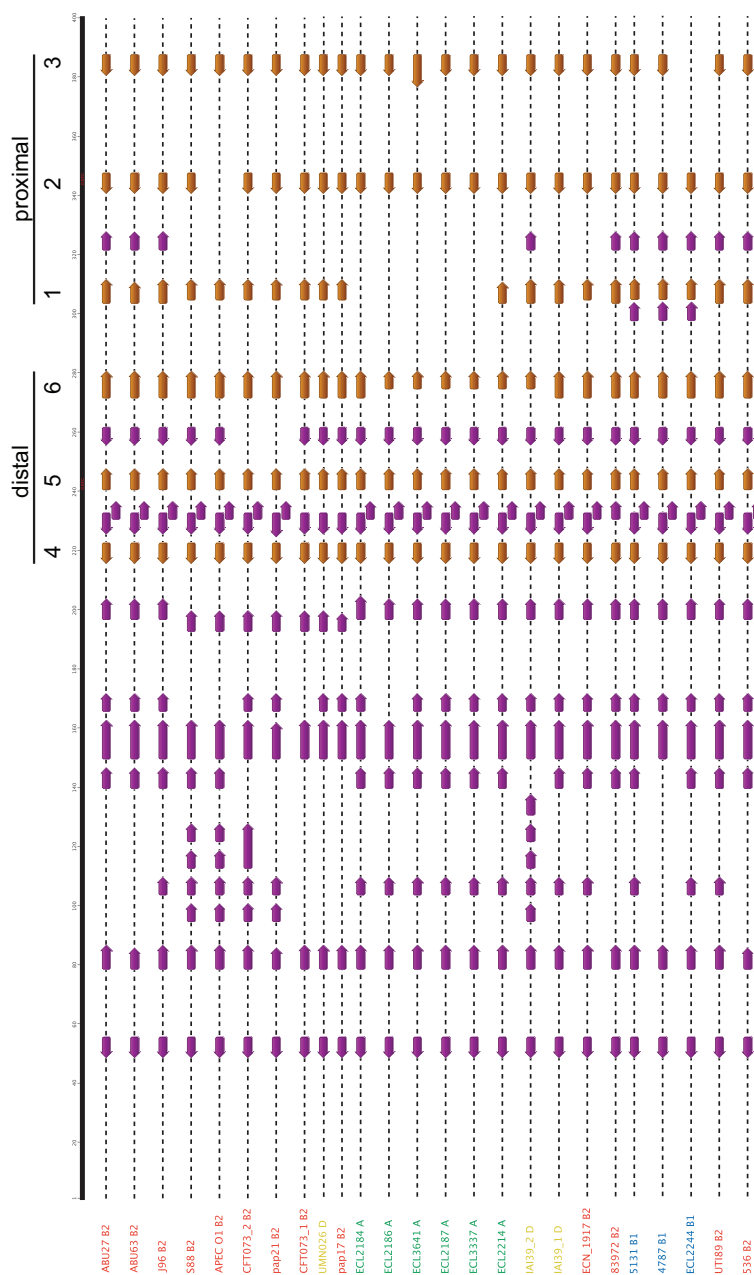


Figure 15. Représentation schématique de l'ensemble des motifs GN(2-3)TTT présents au niveau de la région régulatrice de différents homologues de *pap*. Chaque flèche indique la présence d'un motif GN(2-3)TTT compris entre les 2 promoteurs divergents *pI* et *pBA*. Les motifs représentés en orange correspondent aux sites 1-6 par comparaison avec le modèle *pap* (correspondant dans le cas présent à *pap17*). Les flèches en mauve correspondent à tous les autres motifs GN(2-3)TTT non directement impliqués dans la fixation de Lrp. Le code de couleur associé au nom des 22 souches correspond à leur appartenance à l'un des 4 groupes phylogénétiques majeurs: vert pour A, bleu pour B1, rouge pour B2 et jaune pour D.

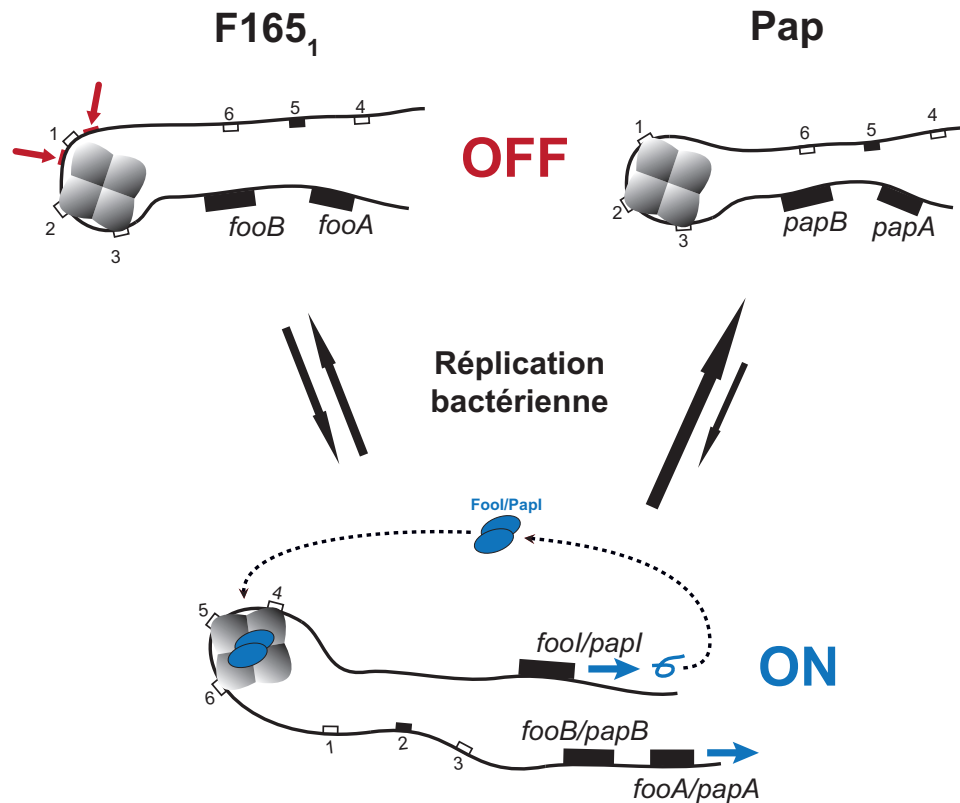


Figure 16. Modèle représentant la configuration des deux phases ON et OFF des opérons *foo* et *pap*. Le principe de la variation de phase des fimbriæ de type P repose sur la compétition entre la fixation de Lrp aux sites 1-3 ou 4-6 et la méthylation par la dam méthylase d'une séquence nucléotidique GATC, située à l'un des sites 2 ou 5. Dans le cas de *pap*, majoritairement en phase OFF, la stabilité de Lrp sur les sites 1-3 est telle qu'elle défavorise le passage, après répllication bactérienne, de Lrp sur les sites 4-6. À l'inverse, dans le cas de *foo*, la présence de deux sites additionnels situés de part et d'autre du site 1 (indiqués par deux flèches rouges) a pour conséquence de ne pas maintenir le complexe sur les sites 1-3, et donc de favoriser davantage le passage vers une phase ON et la transcription de l'opéron (schématisée par une flèche bleue). Dans le cas du passage d'une phase ON à OFF, les interactions de Lrp avec les sites 4-6 des opérons *pap* et *foo* sont équivalentes; cependant la présence du régulateur local PapI pourrait favoriser, par un mécanisme mal défini, davantage le passage vers une phase OFF.

4- F165₁ et Pap partagent plus de points en commun qu'il ne semble

Les acides aminés aliphatiques leucine et alanine jouent un rôle important dans l'état d'oligomérisation de Lrp (Chen & Calvo, 2002, Chen et al., 2001a, Chen *et al.*, 2005, Chen et al., 2001b). En effet, Lrp possède deux sites de fixation pour la leucine, un à haute affinité qui favorise Lrp sous forme de dimère et un second site à moindre affinité qui favorise le passage de l'hexadécamère à l'octamère (Chen & Calvo, 2002).

Outre la forte prévalence de la phase ON chez les fimbriæ F165₁, il est communément admis que les fimbriae F165₁ se distinguent également des autres membres de cette famille d'adhésines, y compris l'opéron *pap*, par une répression en présence de leucine et d'alanine dans le milieu environnemental (Croft et al., 2004). Il est de ce fait à supposer que la régulation de l'adhésine fimbriaire F165₁ implique des mécanismes de régulation liés plus intimement au métabolisme bactérien que les autres membres de la famille des fimbriae de type P. Durant ce projet, nous avons ainsi également étudié le rôle joué par ces deux acides aminés dans la répression de l'opéron *foo* (Annexe I). Or, les résultats que nous avons obtenus, aussi bien par retard de migration sur gel qu'en observant la variation de phase sur plusieurs générations, montrent que l'opéron *pap* est tout aussi sensible à la présence d'acides aminés que l'opéron *foo*. Par conséquent, les mécanismes de régulation des opérons fimbriaires apparentés à la famille des adhésines de type P partagent plus de mécanismes de régulation que ce qu'il était soupçonné jusqu'alors. En particulier, alors que la leucine et l'alanine inhibent de façon similaire la fixation de Lrp aux différentes régions régulatrices *in vitro*, l'alanine inhibe de façon beaucoup plus importante que la leucine l'expression de *pap* et de *foo in vivo*. Dans la mesure où l'on trouve l'alanine en quantité non négligeable dans l'urine, que cet acide aminé joue plusieurs rôles (Anfora *et al.*, 2007), notamment dans la synthèse du peptidoglycane et dans le métabolisme bactérien, et du fait de l'importance de la famille des fimbriae de type P dans le processus infectieux, il apparaît dès lors

nécessaire de mieux définir les relations existant entre le métabolisme bactérien et cette famille d'adhésine.

5- Un régulateur pas si local

Fool, tout comme PapI, joue également un rôle important dans la variation de phase. Nous avons en premier montré par résonance plasmonique de surface que ces deux régulateurs locaux se fixent sur les régions distale et proximale de *pap* et de *foo*, avec des constantes d'affinité de l'ordre du micromolaire. En revanche, nous ne pouvons exclure que ces interactions ne soient pas aspécifiques et il sera nécessaire de réaliser le même type d'expérience en utilisant d'autres fragments d'ADN ne correspondant pas à des régions régulatrices d'opérons fimbriaires pour confirmer la spécificité de ces interactions. Nous avons en revanche montré qu'à faible concentration, Fool favorise déjà de façon significative la fixation tout comme la stabilité de Lrp aux sites 1-3 et 4-6, en accord avec ce qui a déjà été montré pour PapI. Dans la mesure où Lrp forme déjà un complexe particulièrement stable avec la région proximale de *pap*, l'influence de Fool et de PapI sur la fixation de Lrp se montre par conséquent plus importante pour les régions distales de *pap* et de *foo*, en cohérence avec une action positive de ces 2 régulateurs locaux dans la propagation de la phase ON. Dans le cas de *foo*, toutefois, l'influence de Fool et PapI sur la fixation de Lrp aux sites 1-3 ou aux sites 4-6 est similaire, *in vitro*. Ce résultat est à nouveau en accord avec des fréquences de passage d'un phénotype à l'autre très rapide chez les bactéries possédant les fimbriæ F165₁. De plus, en interchangeant par mutagenèse dirigée Fool et PapI, aussi bien dans l'opéron *pap* que dans l'opéron *foo*, nous avons montré que ces deux régulateurs jouent des rôles spécifiques dans la variation de phase, par des mécanismes encore mal définis. Par exemple, remplacer Fool par PapI dans l'opéron *foo* ralentit la fréquence de passage d'une phase OFF à une phase ON. Cependant, remplacer PapI par Fool dans l'opéron *pap* n'a aucune influence sur le passage d'un phénotype à l'autre. Il semble donc que l'action spécifique de PapI ou d'un de ses

homologues dépend de la séquence nucléotidique de l'opéron auquel il est associé. Ainsi, dans le cas de l'opéron *pap*, la stabilité de Lrp sur la région proximale masque l'effet de Fool sur la transcription de cet opéron. De façon similaire, Totsika *et al.* ont montré l'existence d'activation croisée entre différents homologues de *pap* (Totsika et al., 2008). Il est possible que les variations de séquences entre homologues de PapI affectent leur capacité à activer la transcription des opérons fimbriaires; ces différents variants pourraient avoir été sélectionnés au cours de l'évolution de façon à prévenir des activations croisées lorsque deux opérons fimbriaires sont présents au sein d'une même souche. Réciproquement, nos résultats suggèrent que la stabilité du complexe PapI-Lrp sur la région proximale de *pap* minimise en partie l'action d'autres homologues de PapI, susceptibles d'être présents dans une cellule bactérienne.

Dans le cas de l'opéron *clp*, une des différences majeures réside dans l'absence d'homologue de *papI* associé à l'opéron *clp*. Cependant, sa présence est nécessaire pour que CS31A soit régulé par variation de phase. Du fait du pourcentage élevé d'identité entre *clp* et *afa* codant l'adhésine Afa/Dr, les résultats obtenus jusqu'alors a été réalisés en présence d'*afaF*, un autre homologue de *papI*. Or, il est communément admis que la présence d'un second opéron *pap* dans le génome des souches positives pour CS31A fourni en *trans* le régulateur local PapI, résultant alors dans la variation de phase de *clp* via une régulation croisée (Crost et al., 2003, Martin, 1996, Crost et al., 2004). En effet, la souche de référence 31A, contient un opéron *pap* situé dans le chromosome bactérien (Bertin et al., 1998). Durant ce projet, nous avons montré que PapI et Fool induisent la variation de phase de *clp* au même titre que ce qu'il a déjà été montré avec AfaF. En effet, tout comme dans le cas de Pap et de F165₁, Fool et PapI induisent l'apparition d'une population ON en favorisant l'interaction de Lrp avec la région distale de l'opéron *clp*. Par conséquent, PapI et Fool présentent tous les deux la capacité de restaurer la variation de phase de *clp*. De plus, l'affinité de Lrp pour les régions distale et proximale de *clp* augmente en

présence de PapI ou FooI et l'affinité du complexe constitué de Lrp et de PapI (ou celui formé de Lrp et FooI) est plus importante pour la région proximale que pour la région distale. De ce fait, seule une minorité de bactéries est en phase ON. L'absence d'expression concertée entre l'opéron *clp* et *papI* pourrait favoriser la formation d'un complexe stable Lrp-PapI (ou Lrp-FooI) sur la région proximale qui expliquerait pourquoi le niveau d'expression des bactéries en phase OFF est encore plus faible que le niveau basal d'expression de l'opéron *clp*, observé en absence de PapI, FooI ou AfaF. En effet, ce complexe au niveau de la région proximale de *clp* bloquerait la transcription de *clp* de façon plus efficace que celle de Lrp seule. Ainsi, dans le cas de *clp*, un homologue de PapI jouerait un double rôle en 1) permettant la variation de phase de *clp* et, paradoxalement, en 2) favorisant la phase OFF (Figure 17). Dans le cas de la phase ON de *pap*, il est communément admis que la présence des 2 promoteurs divergents *pI* et *pBA* contribue grandement à un repliement de l'ADN favorable à l'entrée et à l'activité de l'ARNpolymérase. Chez l'opéron *foo*, nous avons également montré que la topologie de l'ADN dans la région régulatrice de l'opéron joue un rôle important dans l'expression transcriptionnelle. Notamment, le remplacement de *fooI* par un promoteur transcriptionnel fort comme *pTrc* est suffisant pour restaurer l'expression de *fooB* (cf annexe II, Tessier *et al.*, 2007). Pour *clp*, il se pourrait que l'absence d'homologue de *papI* puisse défavoriser un tel repliement de l'ADN et rendre ainsi la transcription de *clp* plus difficile. Par conséquent, le fait de fournir en *trans papI* ne semble pas équivalent à la présence de ce gène à l'intérieur de l'opéron.

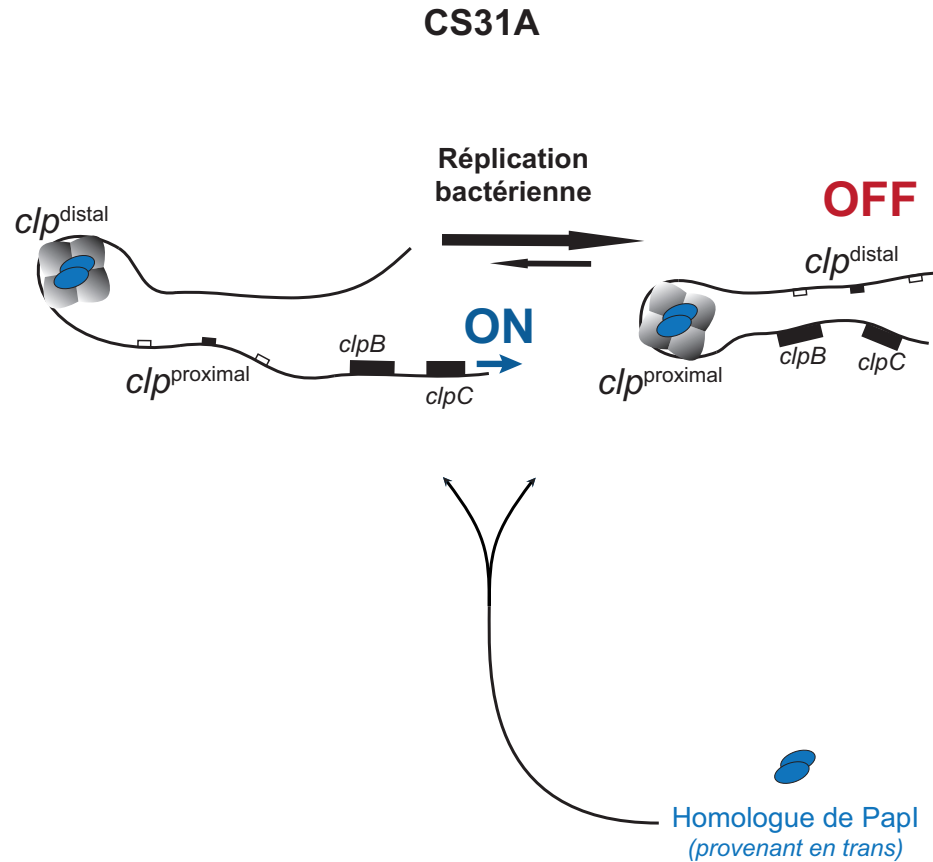


Figure 17. Modèle de la variation de phase de *clp*. Tout comme *pap*, le principe de la variation de phase de CS31A repose sur la compétition entre la fixation de Lrp sur une des deux régions distale et proximale de l'opéron et la méthylation par la dam méthylase d'une séquence nucléotidique GATC, située dans l'une des deux régions proximale et distale. En revanche, à l'inverse de *pap*, Lrp présente une affinité plus importante pour la région proximale que pour la même région distale de *clp*, maintenant ainsi l'opéron *clp* majoritairement en phase OFF. De plus, l'opéron *clp* est dépourvu de tout homologue à PapI, bien que ce dernier est nécessaire à la variation de phase de *clp*. Par conséquent, l'expression de *clp* est tributaire de la présence d'un autre opéron fimbriaire dans le chromosome bactérien. Ne présentant donc pas de transcription coordonnée avec l'opéron *clp*, il est de ce fait possible d'envisager qu'un homologue de PapI puisse favoriser et la fixation de Lrp sur la région distale, et la fixation de Lrp sur la région proximale.

Durant ce projet d'étude, nous avons ainsi montré que la capacité d'induire la variation de phase de *clp* n'est pas spécifique à AfaF et que d'autres homologues de PapI ont également la capacité d'influencer sur l'expression de *clp*. En revanche, il paraît intéressant de noter que la proportion de bactéries en phase ON varie grandement selon l'homologue de PapI étudié. En effet, les études précédentes utilisant AfaF ont montré que la présence de cette protéine induit un plus grand nombre de cellules en phase ON que ce que nous avons pu observer en utilisant F00I ou PapI. Dans la mesure où pas moins de dix variants de PapI ont été recensés à ce jour et que la présence de CS31A dans une souche d'*E. coli* est fréquemment associée à celle de fimbriæ appartenant à la famille de type P, un homologue de PapI serait à même de réguler de façon coordonnée l'expression de plus d'un opéron à la fois.

En conclusion, des variations alléliques dans la région intergénique ou du régulateur local PapI pourrait agir de façon simultanée pour coordonner positivement ou négativement l'expression de plus d'un opéron fimbriaire à la fois.

6- Analyse de la variation de phase en temps réel

Dans la mesure où l'utilisation d'une fusion transcriptionnelle entre la région régulatrice des opérons fimbriaires et le gène rapporteur *lacZ* ne permet pas de suivre la variation de phase pour des cellules différenciées, nous avons également adapté un système rapporteur afin d'appréhender d'avantage la variation de phase des fimbriæ F165₁ et Pap (Lim & van Oudenaarden, 2007). Par cytométrie en flux, nous avons été en mesure de mettre en valeur un niveau de complexité supplémentaire dans les mécanismes de régulation de la variation de phase de ces opérons fimbriaires. Ainsi, chaque cellule s'est caractérisée par un profil d'expression spécifique qui laisse suggérer au moins trois populations distinctes. En effet, nos résultats suggèrent que la variation de phase des fimbriæ de type P n'est pas la simple transition entre 2 configurations distinctes, ON et

OFF, pour la séquence promotrice, mais qu'un niveau intermédiaire de transcription existe également. Cette population caractérisée par des intensités de fluorescence intermédiaire pourrait ainsi comprendre des bactéries considérées comme ON tout comme OFF et jouer ainsi le rôle de tampon. En effet, dans la mesure où les brins d'ADN sont bicaténaires, la transition d'une phase à l'autre nécessite plusieurs cycles de réplication avec, au moins un cycle durant lequel une des deux régions proximale ou distale se trouve dans un état hémiméthylé (Figure 18). Par conséquent, 4 scénarios différents sont à envisager: -i- Lrp fixé au site $GATC^{dist}$ non-méthylé, -ii- Lrp fixé sur le site $GATC^{dist}$ sous forme héli-méthylé, -iii- Lrp fixé au site $GATC^{prox}$ non-méthylé, -iv- Lrp fixé sur le site $GATC^{prox}$ sous forme héli-méthylé. Alors qu'il a été toujours considéré que les cellules présentant Lrp fixé aux sites non-méthylés ou héli-méthylés présentent des phases identiques, nos résultats suggèrent que ces 2 situations ne sont pas nécessairement équivalentes et que le passage d'une phase ON à OFF, ou d'une phase OFF à ON, requière de transiter par un niveau intermédiaire d'expression.

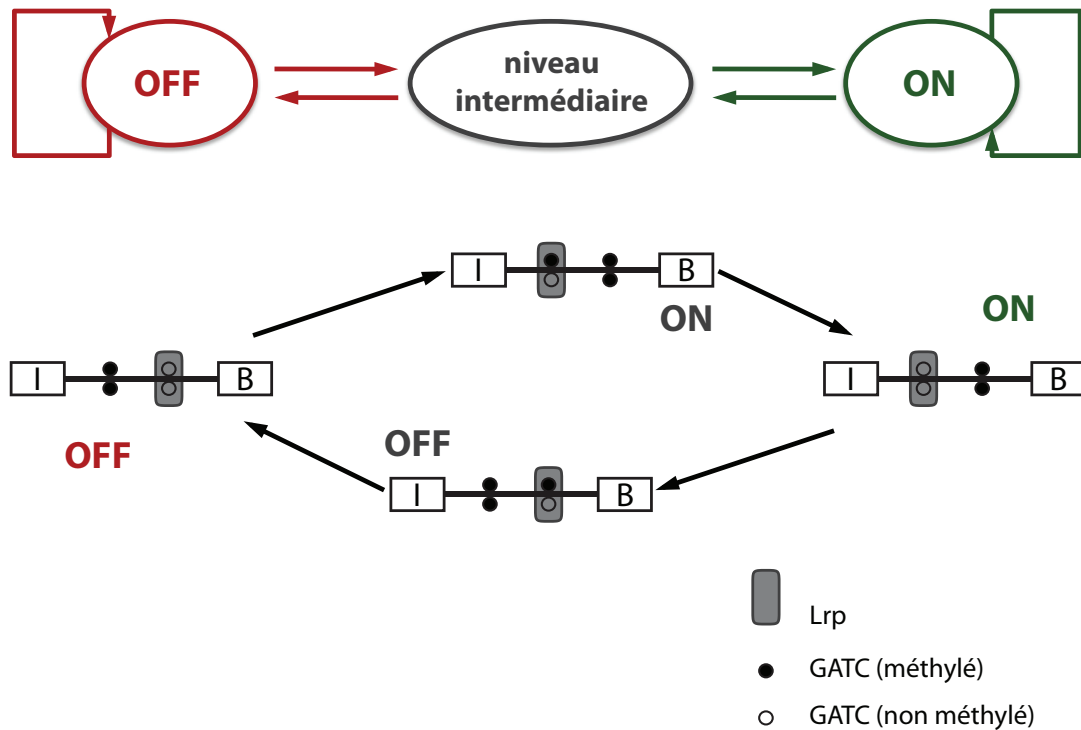


Figure 18. Modèle hypothétique de la variation de phase des opérons fimbriaires codant les adhésines de type P. La variation de phase étant un mécanisme épigénétique, le passage d'un phénotype à l'autre nécessite au moins un cycle de réplication.

Ces résultats nécessitent toutefois d'être confirmés par d'autres techniques. En particulier, par microscopie à fluorescence, nous pourrions observer directement la variation de phase de fimbriae de cellules individuelles qui se diviseront en temps réel. Ainsi, nous pourrions vérifier le patron d'expression des opérons fimbriaires en suivant les lignées généalogiques. Sur le même principe, Adiciptaningrum *et al.* ont mis en évidence l'existence d'une corrélation entre la variation de phase et l'âge de la cellule qui interpelle la notion de stochasticité, en suivant l'expression de l'opéron *fim* codant les fimbriae de type 1 par microscopie à fluorescence (Adiciptaningrum *et al.*, 2009).

En conclusion, l'utilisation de nouvelles approches pour suivre la variation de phase des opérons fimbriaires montre que ces systèmes de régulation présentent des degrés de complexité importants et qu'ils s'apparentent peut-être plus à la régulation dite « classique » qu'il n'y paraît.

7- Une infection est une chose trop sérieuse pour être laissée à des facteurs de virulence

Inspiré d'une citation de Georges Clémenceau :

« La guerre est une chose trop sérieuse pour être laissée aux soldats. »

Les souches ExPEC sont des souches pathogènes, occasionnant rarement des maladies entériques et qui appartiennent à des groupes phylogénétiques et épidémiologiques distincts (Johnson & Russo, 2002). À l'inverse des souches ExPEC virulentes, associées dans la grande majorité des cas aux groupes phylogénétiques B2 et D, les souches exprimant à leur surface les fimbriæ F165₁ se trouvent classées parmi les souches considérées comme moins virulentes, voire commensales, des groupes phylogénétiques A et B1 (Picard et al., 1999). De ce fait, bien qu'étant isolées à partir de cas cliniques, ces souches sont considérées comme des bactéries opportunistes.

Chez une population clonale, transiter entre deux états distincts, et donc créer simultanément une population bimodale, présente l'avantage d'avoir au moins une de ces deux populations susceptible d'être adaptée à la condition environnementale de croissance dans laquelle les bactéries se trouvent. De plus, la présence d'une sous-population bactérienne inadaptée à la condition de croissance actuelle peut tout autant présenter des avantages pour la population totale lors de changements environnementaux plus ou moins soudains. Il est de ce fait envisageable que certaines souches d'*E. coli* privilégient une malléabilité dans la fréquence de passage d'un phénotype à l'autre par rapport à un contrôle strict d'un facteur de virulence, quitte à sacrifier une plus grande proportion de la population bactérienne (Figure 19).

Sur le même principe, Acar *et al.* ont étudié le système d'utilisation du galactose chez deux souches de *Saccharomyces cerevisiæ*, une présentant la capacité de passer d'une phase à l'autre rapide et l'autre lente (Acar *et al.*, 2008, Acar *et al.*, 2010). Ce faisant, ils ont montré que la souche se caractérisant par une fréquence de variation rapide se caractérise par un avantage de croissance sur celle plus lente lorsque les conditions environnementales fluctuent rapidement. À l'inverse, pour des fluctuations environnementales lentes, la souche passant d'un phénotype à l'autre de façon occasionnelle présente un avantage certain sur celle caractérisée par des fréquences de passage d'un phénotype à l'autre rapide.

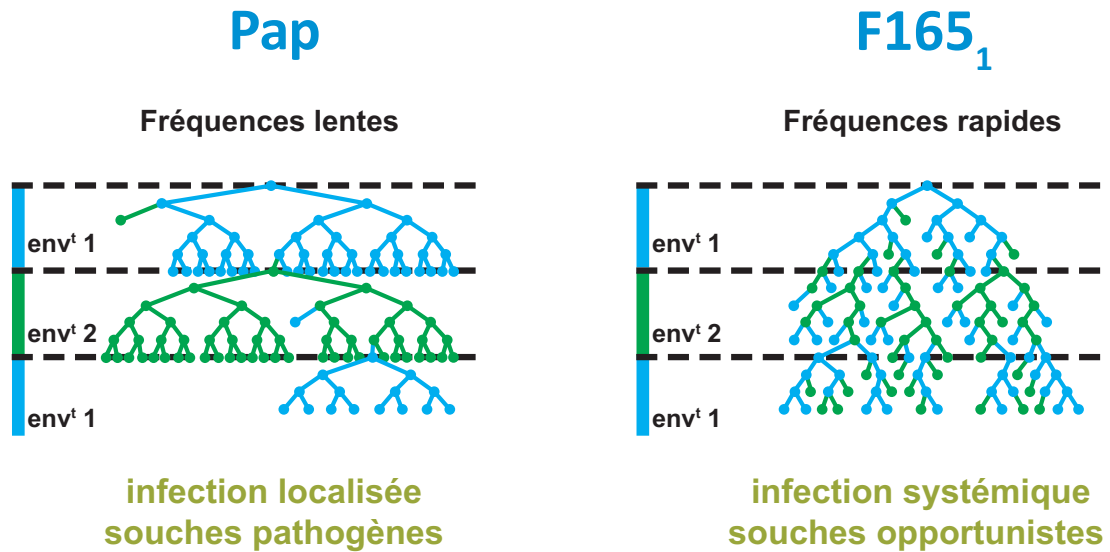


Figure 19. Suivi du devenir de 2 lignées de cellules (une se caractérisant par une fréquence de passage d'un phénotype à l'autre lente, l'autre rapide) mises de façon alternée dans deux conditions de culture distinctes. Le changement de couleur au sein d'une même lignée correspond au changement d'expression phénotypique de la cellule considérée. Si une cellule se trouve dans un phénotype non favorable à la condition environnementale de croissance, alors cette cellule est éliminée. Dans le cas contraire, elle pourra alors continuer à croître et à se diviser. Une cellule avec fréquence de passage d'un phénotype à l'autre lente se caractérisera par une tendance à ne pas présenter de grandes variabilités phénotypique pour chacune des conditions environnementales rencontrées. Adapté d'Acar *et al.*, 2008.

En réalisant un arbre phylogénétique basé uniquement sur la région intergénique d'homologues de *pap* à partir d'un large panel de souches ExPEC, nous avons observé que les régions régulatrices de ces opérons fimbriaires se regroupent ensemble, selon l'appartenance à l'un des 4 groupes phylogénétiques A, B1, B2 et D (Figure 20). Par ailleurs, en analysant plus en détail ces différentes régions régulatrices, nous avons constaté que toutes les souches séquencées appartenant au groupe phylogénétique A sont dépourvues du motif GN(2-3)TTT dans le site 1 et que toutes les souches appartenant au groupe B1 présente les 2 motifs GN(2-3)TTT situés de part et d'autre du site 1 de fixation pour Lrp (Figure 15). À l'inverse, toutes les souches appartenant aux groupes phylogénétiques B2 et D sont dépourvus de tel motif surnuméraire dans la région proximale, voire n'en possèdent qu'un seul.

Il est dès lors possible d'envisager que les souches ExPEC adaptent leur fréquence de passage d'un phénotype à l'autre selon le type d'environnement, et donc par extension selon le type d'infection qu'elles provoquent. Ainsi, une souche virulente occasionnant des infections localisées du tractus urinaire, se caractériserait par un contrôle strict dans la production de fimbriæ de type P alors que les souches opportunistes, comme celles exprimant les fimbriæ F165₁, favoriserait une souplesse dans la régulation par variation de phase de ces fimbriæ, afin de pouvoir se disséminer plus facilement dans l'hôte.

En conclusion, toutes les souches ExPEC ne nécessitent pas obligatoirement le même degré d'expression de l'opéron fimbriaire et les souches appartenant à un même groupe phylogénétique pourraient partager des mécanismes similaires de régulation de la variation de phase avec les souches moins pathogènes plus disposées à augmenter leur proportion de cellules ON.

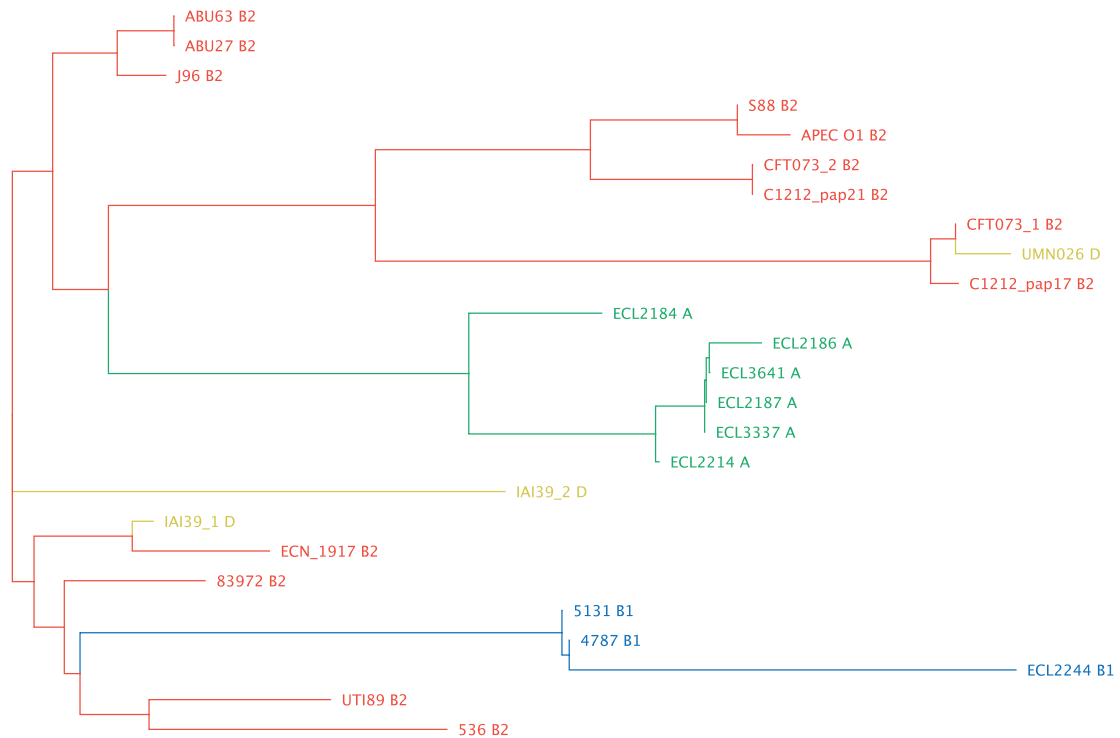


Figure 20. Analyse phylogénétique basée sur la comparaison de séquence de la région intergénique de 22 homologues de *pap*. Chaque couleur représente un des quatre groupe phylogénétique (vert=A; bleu=B1; rouge=B2; jaune=D).

B- CONCLUSIONS

Durant ce projet de PhD, nous avons montré que, fondamentalement, de faibles modifications de séquence nucléotidique ont des répercussions importantes sur les interactions entre ADN et protéines et que la capacité à former un complexe stable le long de la séquence d'ADN est un facteur qui doit entrer en considération. Ainsi, la comparaison entre *pap* et *foo* a montré que la stabilité des interactions de Lrp sur l'ADN avait de grandes répercussions sur la quantité de cellules en phase ON. Par conséquent, cette observation amène à se demander si les mécanismes dits stochastiques ne seraient finalement peut-être pas différents des mécanismes de régulation « dits classiques », dans la mesure où la fixation d'un régulateur tel que Lrp influence directement le profil de méthylation de l'ADN.

Nous avons également montré que la notion de régulateur local appliqué aux homologues de PapI était peut-être à remettre en question dans la mesure où cette protéine coordonne l'expression de plusieurs opérons simultanément. Dans la mesure où Lrp présente un nombre important de sites de fixation, il serait de ce fait intéressant de voir si l'action de PapI et de ses homologues a un impact sur des unités transcriptionnelle autres que les opérons fimbriaires.

Enfin, nous avons aussi montré qu'un même facteur de virulence ne nécessite pas nécessairement le même degré d'expression selon le potentiel pathogène de la souche considérée. Ainsi, comme l'a souligné Charles Darwin, ce n'est pas nécessairement toujours l'espèce la plus intelligente ou la plus forte qui est la mieux adaptée mais celle qui sera plus à même de résister à des changements environnementaux.

SECTION V : Bibliographie

- Acar, M., J. T. Mettetal & A. van Oudenaarden, (2008) Stochastic switching as a survival strategy in fluctuating environments. *Nat Genet* **40**: 471-475.
- Acar, M., B. F. Pando, F. H. Arnold, M. B. Elowitz & A. van Oudenaarden, (2010) A general mechanism for network-dosage compensation in gene circuits. *Science* **329**: 1656-1660.
- Adicptaningrum, A. M., I. C. Blomfield & S. J. Tans, (2009) Direct observation of type 1 fimbrial switching. *EMBO Rep* **10**: 527-532.
- Aki, T., H. E. Choy & S. Adhya, (1996) Histone-like protein HU as a specific transcriptional regulator: co-factor role in repression of gal transcription by GAL repressor. *Genes Cells* **1**: 179-188.
- Albano, M., W. K. Smits, L. T. Ho, B. Kraigher, I. Mandic-Mulec, O. P. Kuipers & D. Dubnau, (2005) The Rok protein of *Bacillus subtilis* represses genes for cell surface and extracellular functions. *J Bacteriol* **187**: 2010-2019.
- Andrewes, F. W., (1922) Studies in group-agglutination. I. The *Salmonella* group and its antigenic structure. *J. Pathol. Bacteriol.* **25**: 505-521.
- Anfora, A. T., B. J. Haugen, P. Roesch, P. Redford & R. A. Welch, (2007) Roles of serine accumulation and catabolism in the colonization of the murine urinary tract by *Escherichia coli* CFT073. *Infect Immun* **75**: 5298-5304.
- Annane, D., E. Bellissant & J.-M. Cavillon, (2005) Septic shock. *Lancet* **365**: 63-78.
- Baek, C. H., S. Wang, K. L. Roland & R. Curtiss, 3rd, (2009) Leucine-responsive regulatory protein (Lrp) acts as a virulence repressor in *Salmonella enterica* serovar Typhimurium. *J Bacteriol* **191**: 1278-1292.
- Baga, M., M. Goransson, S. Normark & B. E. Uhlin, (1988) Processed mRNA with differential stability in the regulation of *E. coli* pilin gene expression. *Cell* **52**: 197-206.
- Balázsi, G., A. van Oudenaarden & J. J. Collins, (2011) Cellular decision making and biological noise: from microbes to mammals. *Cell* **144**: 910-925.
- Balsalobre, C., J. Johansson & B. E. Uhlin, (2006) Cyclic AMP-dependent osmoregulation of *crp* gene expression in *Escherichia coli*. *J Bacteriol* **188**: 5935-5944.

- Barnard, A., A. Wolfe & S. Busby, (2004) Regulation at complex bacterial promoters: how bacteria use different promoter organizations to produce different regulatory outcomes. *Curr Opin Microbiol* **7**: 102-108.
- Bekal, S., R. Brousseau, L. Masson, G. Prefontaine, J. Fairbrother & J. Harel, (2003) Rapid identification of *Escherichia coli* pathotypes by virulence gene detection with DNA microarrays. *J Clin Microbiol* **41**: 2113-2125.
- Berthiaume, F., C. Crost, V. Labrie, C. Martin, E. B. Newman & J. Harel, (2004) Influence of L-leucine and L-alanine on Lrp regulation of *foo*, coding for F165₁, a Pap homologue. *J Bacteriol* **186**: 8537-8541.
- Bertin, Y., J. P. Girardeau, A. Darfeuille-Michaud & C. Martin, (2000) Epidemiological study of *pap* genes among diarrheagenic or septicemic *Escherichia coli* strains producing CS31A and F17 adhesins and characterization of Pap(31A) fimbriae. *J Clin Microbiol* **38**: 1502-1509.
- Bertin, Y., C. Martin, J. P. Girardeau, P. Pohl & M. Contrepois, (1998) Association of genes encoding P fimbriae, CS31A antigen and EAST 1 toxin among CNF1-producing *Escherichia coli* strains from cattle with septicemia and diarrhea. *FEMS Microbiol Lett* **162**: 235-239.
- Bickle, T. A. & D. H. Kruger, (1993) Biology of DNA restriction. *Microbiol Rev* **57**: 434-450.
- Bower, J. M., D. S. Eto & M. A. Mulvey, (2005) Covert operations of uropathogenic *Escherichia coli* within the urinary tract. *Traffic* **6**: 18-31.
- Boye, E., A. Lobner-Olesen & K. Skarstad, (2000) Limiting DNA replication to once and only once. *EMBO Rep* **1**: 479-483.
- Browning, D. F. & S. J. Busby, (2004) The regulation of bacterial transcription initiation. *Nat Rev Microbiol* **2**: 57-65.
- Brusés, J. L. & U. Rutishauser, (2001) Roles, regulation, and mechanism of polysialic acid function during neural development. *Biochimie* **83**: 635-643.
- Buchler, N. E., U. Gerland & T. Hwa, (2003) On schemes of combinatorial transcription logic. *Proc Natl Acad Sci U S A* **100**: 5136-5141.

- Calvo, J. M. & R. G. Matthews, (1994) The leucine-responsive regulatory protein, a global regulator of metabolism in *Escherichia coli*. *Microbiol Rev* **58**: 466-490.
- Campbell, J. L. & N. Kleckner, (1990) *E. coli oriC* and the *dnaA* gene promoter are sequestered from dam methyltransferase following the passage of the chromosomal replication fork. *Cell* **62**: 967-979.
- Casadesús, J. & D. Low, (2006) Epigenetic gene regulation in the bacterial world. *Microbiol Mol Biol Rev* **70**: 830-856.
- Chen, S. & J. M. Calvo, (2002) Leucine-induced dissociation of *Escherichia coli* Lrp hexadecamers to octamers. *J Mol Biol* **318**: 1031-1042.
- Chen, S., Z. Hao, E. Bieniek & J. M. Calvo, (2001a) Modulation of Lrp action in *Escherichia coli* by leucine: effects on non-specific binding of Lrp to DNA. *J Mol Biol* **314**: 1067-1075.
- Chen, S., M. Iannolo & J. M. Calvo, (2005) Cooperative binding of the leucine-responsive regulatory protein (Lrp) to DNA. *J Mol Biol* **345**: 251-264.
- Chen, S., M. H. Rosner & J. M. Calvo, (2001b) Leucine-regulated self-association of leucine-responsive regulatory protein (Lrp) from *Escherichia coli*. *J Mol Biol* **312**: 625-635.
- Cheng, X., (1995) Structure and function of DNA methyltransferases. *Annu Rev Biophys Biomol Struct* **24**: 293-318.
- Cherifi, A., M. Contrepolis, B. Picard, P. Gouillet, J. de Rycke, J. Fairbrother & J. Barnouin, (1990) Factors and markers of virulence in *Escherichia coli* from human septicemia. *FEMS Microbiol Lett* **58**: 279-283.
- Cho, B.-K., C. L. Barrett, E. M. Knight, Y. S. Park & B. Ø. Palsson, (2008) Genome-scale reconstruction of the Lrp regulatory network in *Escherichia coli*. *Proc Natl Acad Sci USA* **105**: 19462-19467.
- Ciampi, M. S., (2006) Rho-dependent terminators and transcription termination. *Microbiology (Reading, England)* **152**: 2515-2528.
- Cieslewicz, M. & E. Vimr, (1997) Reduced polysialic acid capsule expression in *Escherichia coli* K1 mutants with chromosomal defects in *kpsF*. *Mol Microbiol* **26**: 237-249.

- Collado-Vides, J., B. Magasanik & J. D. Gralla, (1991) Control site location and transcriptional regulation in *Escherichia coli*. *Microbiol Rev* **55**: 371-394.
- Contrepois, M., Y. Bertin, J. P. Girardeau, B. Picard & P. Goulet, (1993) Clonal relationships among bovine pathogenic *Escherichia coli* producing surface antigen CS31A. *FEMS Microbiol Lett* **106**: 217-222.
- Contrepois, M., H. C. Dubourguier, A. L. Parodi, J. P. Girardeau & J. L. Ollier, (1986) Septicaemic *Escherichia coli* and experimental infection of calves. *Vet Microbiol* **12**: 109-118.
- Contrepois, M., J. M. Fairbrother, Y. K. Kaura & J. P. Girardeau, (1989) Prevalence of CS31A and F165 surface antigens in *Escherichia coli* isolates from animals in France, Canada and India. *FEMS Microbiol Lett* **50**: 319-323.
- Crost, C., A. Garrivier, J. Harel & C. Martin, (2003) Leucine-responsive regulatory protein-mediated repression of *clp* (encoding CS31A) expression by L-leucine and L-alanine in *Escherichia coli*. *J Bacteriol* **185**: 1886-1894.
- Crost, C., J. Harel, F. Berthiaume, A. Garrivier, M. C. Tessier, H. Rakotoarivonina & C. Martin, (2004) Influence of environmental cues on transcriptional regulation of *foo* and *clp* coding for F165(1) and CS31A adhesins in *Escherichia coli*. *Res Microbiol* **155**: 475-482.
- Cui, Y., Q. Wang, G. D. Stormo & J. M. Calvo, (1995) A consensus sequence for binding of Lrp to DNA. *J Bacteriol* **177**: 4872-4880.
- D'Ari, R., R. T. Lin & E. B. Newman, (1993) The leucine-responsive regulatory protein: more than a regulator? *Trends Biochem Sci* **18**: 260-263.
- Daigle, F., C. Forget, C. Martin, M. Drolet, M. C. Tessier, H. Dezfulian & J. Harel, (2000) Effects of global regulatory proteins and environmental conditions on fimbrial gene expression of F165(1) and F165(2) produced by *Escherichia coli* causing septicaemia in pigs. *Res Microbiol* **151**: 563-574.
- Davidson, C. J. & M. G. Surette, (2008) Individuality in bacteria. *Annu Rev Genet* **42**: 253-268.

- de los Rios, S. & J. J. Perona, (2007) Structure of the *Escherichia coli* leucine-responsive regulatory protein Lrp reveals a novel octameric assembly. *J Mol Biol* **366**: 1589-1602.
- Dezfulian, H., D. Tremblay & J. Harel, (2004) Molecular characterization of extraintestinal pathogenic *Escherichia coli* (ExPEC) pathogenicity islands in F165-positive *E. coli* strain from a diseased animal. *FEMS Microbiol Lett* **238**: 321-332.
- Di Martino, P., Y. Bertin, J. P. Girardeau, V. Livrelli, B. Joly & A. Darfeuille-Michaud, (1995) Molecular characterization and adhesive properties of CF29K, an adhesin of *Klebsiella pneumoniae* strains involved in nosocomial infections. *Infect Immun* **63**: 4336-4344.
- Di Martino, P., J. P. Girardeau, M. Der Vartanian, B. Joly & A. Darfeuille-Michaud, (1997) The central variable V2 region of the CS31A major subunit is involved in the receptor-binding domain. *Infect Immun* **65**: 609-616.
- Dodson, K. W., F. Jacob-Dubuisson, R. T. Striker & S. J. Hultgren, (1993) Outer-membrane PapC molecular usher discriminately recognizes periplasmic chaperone-pilus subunit complexes. *Proc Natl Acad Sci U S A* **90**: 3670-3674.
- Dodson, K. W., J. S. Pinkner, T. Rose, G. Magnusson, S. J. Hultgren & G. Waksman, (2001) Structural basis of the interaction of the pyelonephritic *E. coli* adhesin to its human kidney receptor. *Cell* **105**: 733-743.
- Dorman, C. J., (2009a) Global regulators and environmental adaptation in Gram-negative pathogens. *Clin Microbiol Infect* **15 Suppl 1**: 47-50.
- Dorman, C. J., (2009b) Nucleoid-associated proteins and bacterial physiology. *Adv Appl Microbiol* **67**: 47-64.
- Dorman, C. J., (2009c) Regulatory integration of horizontally-transferred genes in bacteria. *Front Biosci* **14**: 4103-4112.
- Dorman, C. J. & C. P. Corcoran, (2009) Bacterial DNA topology and infectious disease. *Nucleic Acids Res* **37**: 672-678.

- Dorman, C. J. & K. A. Kane, (2009) DNA bridging and antibridging: a role for bacterial nucleoid-associated proteins in regulating the expression of laterally acquired genes. *FEMS Microbiol Rev* **33**: 587-592.
- Drekonja, D. M. & J. R. Johnson, (2008) Urinary tract infections. *Prim Care* **35**: 345-367, vii.
- Escobar-Páramo, P., O. Clermont, A.-B. Blanc-Potard, H. Bui, C. Le Bouguéneq & E. Denamur, (2004) A specific genetic background is required for acquisition and expression of virulence factors in *Escherichia coli*. *Mol Biol Evol* **21**: 1085-1094.
- Espinosa, V., A. D. Gonzalez, A. T. Vasconcelos, A. M. Huerta & J. Collado-Vides, (2005) Comparative studies of transcriptional regulation mechanisms in a group of eight gamma-proteobacterial genomes. *J Mol Biol* **354**: 184-199.
- Ewers, C., E.-M. Antão, I. Diehl, H.-C. Philipp & L. H. Wieler, (2009) Intestine and environment of the chicken as reservoirs for extraintestinal pathogenic *Escherichia coli* strains with zoonotic potential. *Appl Environ Microbiol* **75**: 184-192.
- Ewers, C., G. Li, H. Wilking, S. Kiessling, K. Alt, E.-M. Antão, C. Laternus, I. Diehl, S. Glodde, T. Homeier, U. Böhnke, H. Steinrück, H.-C. Philipp & L. H. Wieler, (2007) Avian pathogenic, uropathogenic, and newborn meningitis-causing *Escherichia coli*: how closely related are they? *Int J Med Microbiol* **297**: 163-176.
- Fabret, C., V. A. Feher & J. A. Hoch, (1999) Two-component signal transduction in *Bacillus subtilis*: how one organism sees its world. *Journal of bacteriology* **181**: 1975-1983.
- Fairbrother, J. M., S. Lariviere & W. M. Johnson, (1988) Prevalence of fimbrial antigens and enterotoxins in nonclassical serogroups of *Escherichia coli* isolated from newborn pigs with diarrhea. *Am J Vet Res* **49**: 1325-1328.
- Fairbrother, J. M., S. Lariviere & R. Lallier, (1986) New fimbrial antigen F165 from *Escherichia coli* serogroup O115 strains isolated from piglets with diarrhea. *Infect Immun* **51**: 10-15.

- Fairbrother, J. M., E. Nadeau & C. L. Gyles, (2005) *Escherichia coli* in postweaning diarrhea in pigs: an update on bacterial types, pathogenesis, and prevention strategies. *Anim Health Res Rev* **6**: 17-39.
- Ferry, S. A., S. E. Holm, H. Stenlund, R. Lundholm & T. J. Mosen, (2007) Clinical and bacteriological outcome of different doses and duration of pivmecillinam compared with placebo therapy of uncomplicated lower urinary tract infection in women: the LUTIW project. *Scand J Prim Health Care* **25**: 49-57.
- Forsman, K., M. Göransson & B. E. Uhlin, (1989) Autoregulation and multiple DNA interactions by a transcriptional regulatory protein in *E. coli* pili biogenesis. *EMBO J* **8**: 1271-1277.
- Forsman, K., B. Sondén, M. Göransson & B. E. Uhlin, (1992) Antirepression function in *Escherichia coli* for the cAMP-cAMP receptor protein transcriptional activator. *Proc Natl Acad Sci USA* **89**: 9880-9884.
- Foxman, B., (2002) Epidemiology of urinary tract infections: incidence, morbidity, and economic costs. *Am J Med* **113 Suppl 1A**: 5S-13S.
- Foxman, B., (2010) The epidemiology of urinary tract infection. *Nature reviews Urology* **7**: 653-660.
- Foxman, B. & L. Riley, (2001) Molecular epidemiology: focus on infection. *Am J Epidemiol* **153**: 1135-1141.
- Garcia-Del Portillo, F., M. G. Pucciarelli & J. Casadesus, (1999) DNA adenine methylase mutants of *Salmonella typhimurium* show defects in protein secretion, cell invasion, and M cell cytotoxicity. *Proc Natl Acad Sci U S A* **96**: 11578-11583.
- Gertz, J. & B. A. Cohen, (2009) Environment-specific combinatorial cis-regulation in synthetic promoters. *Mol Syst Biol* **5**: 244.
- Giacomodonato, M. N., S. H. Sarnacki, R. L. Caccuri, D. O. Sordelli & M. C. Cerquetti, (2004) Host response to a dam mutant of *Salmonella enterica* serovar enteritidis with a temperature-sensitive phenotype. *Infect Immun* **72**: 5498-5501.

- Girardeau, J. P., Y. Bertin, C. Martin, M. Der Vartanian & C. Boeuf, (1991) Sequence analysis of the *clpG* gene, which codes for surface antigen CS31A subunit: evidence of an evolutionary relationship between CS31A, K88, and F41 subunit genes. *J Bacteriol* **173**: 7673-7683.
- Girardeau, J. P., M. Der Vartanian, J. L. Ollier & M. Contrepois, (1988) CS31A, a new K88-related fimbrial antigen on bovine enterotoxigenic and septicemic *Escherichia coli* strains. *Infect Immun* **56**: 2180-2188.
- Girardeau, J. P., L. Lalioui, A. M. O. Said, C. De Champs & C. Le Bouguéneq, (2003) Extended virulence genotype of pathogenic *Escherichia coli* isolates carrying the *afa-8* operon: evidence of similarities between isolates from humans and animals with extraintestinal infections. *J Clin Microbiol* **41**: 218-226.
- Goransson, M., P. Forsman, P. Nilsson & B. E. Uhlin, (1989) Upstream activating sequences that are shared by two divergently transcribed operons mediate cAMP-CRP regulation of pilus-adhesin in *Escherichia coli*. *Mol Microbiol* **3**: 1557-1565.
- Göransson, M., B. Sondén, P. Nilsson, B. Dagberg, K. Forsman, K. Emanuelsson & B. E. Uhlin, (1990) Transcriptional silencing and thermoregulation of gene expression in *Escherichia coli*. *Nature* **344**: 682-685.
- Griebing, T. L., (2005) Urologic diseases in America project: trends in resource use for urinary tract infections in women. *J Urol* **173**: 1281-1287.
- Gunther, N. W., J. A. Snyder, V. Lockett, I. Blomfield, D. E. Johnson & H. L. T. Mobley, (2002) Assessment of virulence of uropathogenic *Escherichia coli* type 1 fimbrial mutants in which the invertible element is phase-locked on or off. *Infect Immun* **70**: 3344-3354.
- Gyles, C. L. & J. M. Fairbrother, (2010) *Escherichia coli*. In: Pathogenesis of Bacterial Infections in Animals. G. C. L, P. J. F, S. J. G & T. C. O (eds). Ames, Iowa, USA: Wiley-Blackwell, pp. 267-308.
- Hacker, J., G. Blum-Oehler, I. Mühldorfer & H. Tschäpe, (1997) Pathogenicity islands of virulent bacteria: structure, function and impact on microbial evolution. *Mol Microbiol* **23**: 1089-1097.

- Hantke, K., (2001) Iron and metal regulation in bacteria. *Current opinion in microbiology* **4**: 172-177.
- Harel, J., F. Daigle, C. Forget, M. C. Tessier, C. Crost & C. Martin, (2000) Phase variation of F165(1) (Prs-like) fimbriae from *Escherichia coli* causing septicaemia in animals. *Can J Microbiol* **46**: 1101-1107.
- Harel, J., F. Daigle, S. Maiti, C. Désautels, A. Labigne & J. M. Fairbrother, (1991) Occurrence of *pap*-, *sfa*-, and *afa*-related sequences among F165-positive *Escherichia coli* from diseased animals. *FEMS Microbiol Lett* **66**: 177-182.
- Harel, J., C. Forget, M. Ngeleka, M. Jacques & J. M. Fairbrother, (1992a) Isolation and characterization of adhesin-defective TnphoA mutants of septicaemic porcine *Escherichia coli* of serotype O115:K:F165. *J Gen Microbiol* **138**: 2337-2345.
- Harel, J., C. Forget, J. Saint-Amand, F. Daigle, D. Dubreuil, M. Jacques & J. Fairbrother, (1992b) Molecular cloning of a determinant coding for fimbrial antigen F165(1), a Prs-like fimbrial antigen from porcine septicaemic *Escherichia coli*. *J Gen Microbiol* **138**: 1495-1502.
- Harel, J., M. Jacques, J. M. Fairbrother, M. Bossé & C. Forget, (1995) Cloning of determinants encoding F165(2) fimbriae from porcine septicaemic *Escherichia coli* confirms their identity as F1C fimbriae. *Microbiology (Reading, Engl)* **141 (Pt 1)**: 221-228.
- Harel, J. & C. Martin, (1999) Virulence gene regulation in pathogenic *Escherichia coli*. *Vet Res* **30**: 131-155.
- Heithoff, D. M., R. L. Sinsheimer, D. A. Low & M. J. Mahan, (1999) An essential role for DNA adenine methylation in bacterial virulence. *Science* **284**: 967-970.
- Henkin, T. M., (2000) Transcription termination control in bacteria. *Current opinion in microbiology* **3**: 149-153.
- Hermsen, R., S. Tans & P. R. ten Wolde, (2006) Transcriptional regulation by competing transcription factor modules. *PLoS Comput Biol* **2**: e164.

- Hernday, A., M. Krabbe, B. Braaten & D. Low, (2002) Self-perpetuating epigenetic pili switches in bacteria. *Proc Natl Acad Sci USA* **99 Suppl 4**: 16470-16476.
- Hernday, A. D., B. A. Braaten, G. Broitman-Maduro, P. Engelberts & D. A. Low, (2004) Regulation of the *pap* epigenetic switch by CpxAR: phosphorylated CpxR inhibits transition to the phase ON state by competition with Lrp. *Mol Cell* **16**: 537-547.
- Hernday, A. D., B. A. Braaten & D. A. Low, (2003) The mechanism by which DNA adenine methylase and PapI activate the *pap* epigenetic switch. *Mol Cell* **12**: 947-957.
- Herzer, P. J., S. Inouye, M. Inouye & T. S. Whittam, (1990) Phylogenetic distribution of branched RNA-linked multicopy single-stranded DNA among natural isolates of *Escherichia coli*. *J Bacteriol* **172**: 6175-6181.
- Holden, N. J., M. Totsika, E. Mahler, A. J. Roe, K. Catherwood, K. Lindner, U. Dobrindt & D. L. Gally, (2006) Demonstration of regulatory cross-talk between P fimbriae and type 1 fimbriae in uropathogenic *Escherichia coli*. *Microbiology (Reading, Engl)* **152**: 1143-1153.
- Hommais, F., E. Krin, C. Laurent-Winter, O. Soutourina, A. Malpertuy, J. P. Le Caer, A. Danchin & P. Bertin, (2001) Large-scale monitoring of pleiotropic regulation of gene expression by the prokaryotic nucleoid-associated protein, H-NS. *Mol Microbiol* **40**: 20-36.
- Hooton, T. M., D. Scholes, J. P. Hughes, C. Winter, P. L. Roberts, A. E. Stapleton, A. Stergachis & W. E. Stamm, (1996) A prospective study of risk factors for symptomatic urinary tract infection in young women. *N Engl J Med* **335**: 468-474.
- Huang, S. H. & A. Y. Jong, (2001) Cellular mechanisms of microbial proteins contributing to invasion of the blood-brain barrier. *Cell Microbiol* **3**: 277-287.
- Huang, S. H., M. F. Stins & K. S. Kim, (2000) Bacterial penetration across the blood-brain barrier during the development of neonatal meningitis. *Microbes Infect* **2**: 1237-1244.

- Hultgren, S. J., F. Jacob-Dubuisson, C. H. Jones & C. I. Branden, (1993) PapD and superfamily of periplasmic immunoglobulin-like pilus chaperones. *Adv Protein Chem* **44**: 99-123.
- Hultgren, S. J., S. Normark & S. N. Abraham, (1991) Chaperone-assisted assembly and molecular architecture of adhesive pili. *Annu Rev Microbiol* **45**: 383-415.
- Hung, C. S., J. Bouckaert, D. Hung, J. Pinkner, C. Widberg, A. DeFusco, C. G. Augustine, R. Strouse, S. Langermann, G. Waksman & S. J. Hultgren, (2002) Structural basis of tropism of *Escherichia coli* to the bladder during urinary tract infection. *Mol Microbiol* **44**: 903-915.
- Hyde, T. B., T. M. Hilger, A. Reingold, M. M. Farley, K. L. O'Brien, A. Schuchat & A. B. C. s. A. o. t. E. I. P. Network, (2002) Trends in incidence and antimicrobial resistance of early-onset sepsis: population-based surveillance in San Francisco and Atlanta. *Pediatrics* **110**: 690-695.
- Jackson, L. A., P. Benson, K. M. Neuzil, M. Grandjean & J. L. Marino, (2005) Burden of community-onset *Escherichia coli* bacteremia in seniors. *J Infect Dis* **191**: 1523-1529.
- Jallat, C., A. Darfeuille-Michaud, J. P. Girardeau, C. Rich & B. Joly, (1994) Self-transmissible R plasmids encoding CS31A among human *Escherichia coli* strains isolated from diarrheal stools. *Infect Immun* **62**: 2865-2873.
- Jarboe, L. R., D. Beckwith & J. C. Liao, (2004) Stochastic modeling of the phase-variable *pap* operon regulation in uropathogenic *Escherichia coli*. *Biotechnol Bioeng* **88**: 189-203.
- Johnson, J. R., (1991) Virulence factors in *Escherichia coli* urinary tract infection. *Clin Microbiol Rev* **4**: 80-128.
- Johnson, J. R., J. J. Brown, U. B. Carlino & T. A. Russo, (1998) Colonization with and acquisition of uropathogenic *Escherichia coli* as revealed by polymerase chain reaction-based detection. *J Infect Dis* **177**: 1120-1124.
- Johnson, J. R. & T. A. Russo, (2002) Extraintestinal pathogenic *Escherichia coli*: "the other bad *E coli*". *J Lab Clin Med* **139**: 155-162.

- Johnson, J. R. & T. A. Russo, (2005) Molecular epidemiology of extraintestinal pathogenic (uropathogenic) *Escherichia coli*. *Int J Med Microbiol* **295**: 383-404.
- Jones, C. H., J. S. Pinkner, R. Roth, J. Heuser, A. V. Nicholes, S. N. Abraham & S. J. Hultgren, (1995) FimH adhesin of type 1 pili is assembled into a fibrillar tip structure in the Enterobacteriaceae. *Proc Natl Acad Sci U S A* **92**: 2081-2085.
- Justice, S. S., S. R. Lauer, S. J. Hultgren & D. A. Hunstad, (2006) Maturation of intracellular *Escherichia coli* communities requires SurA. *Infect Immun* **74**: 4793-4800.
- Kac, M. & J. Logan, (1976) *Fluctuation Phenomena*. Amsterdam.
- Kaltenbach, L. S., B. A. Braaten & D. A. Low, (1995) Specific binding of PapI to Lrp-pap DNA complexes. *J Bacteriol* **177**: 6449-6455.
- Kanamaru, S., H. Kurazono, M. Nakano, A. Terai, O. Ogawa & S. Yamamoto, (2006) Subtyping of uropathogenic *Escherichia coli* according to the pathogenicity island encoding uropathogenic-specific protein: comparison with phylogenetic groups. *Int J Urol* **13**: 754-760.
- Kaper, J. B., J. P. Nataro & H. L. Mobley, (2004) Pathogenic *Escherichia coli*. *Nat Rev Microbiol* **2**: 123-140.
- Kawamura, T., L. U. K. Le, H. Zhou & F. W. Dahlquist, (2007) Solution structure of *Escherichia coli* PapI, a key regulator of the *pap* pili phase variation. *J Mol Biol* **365**: 1130-1142.
- Kawamura, T., A. S. Vartanian, H. Zhou & F. W. Dahlquist, (2011) The Design involved in PapI and Lrp regulation of the *pap* operon. *Journal of molecular biology* **409**: 311-332.
- Kennedy, R. P., J. J. Plorde & R. G. Petersdorf, (1965) Studies on the epidemiology of *Escherichia coli* infections. IV. Evidence for a nosocomial flora. *J Clin Invest* **44**: 193-201.
- Kim, K. S., (2001) *Escherichia coli* translocation at the blood-brain barrier. *Infect Immun* **69**: 5217-5222.

- Korth, M. J., J. C. Lara & S. L. Moseley, (1994) Epithelial cell invasion by bovine septicemic *Escherichia coli*. *Infect Immun* **62**: 41-47.
- Kucheria, R., P. Dasgupta, S. H. Sacks, M. S. Khan & N. S. Sheerin, (2005) Urinary tract infections: new insights into a common problem. *Postgrad Med J* **81**: 83-86.
- Kuehn, M. J., J. Heuser, S. Normark & S. J. Hultgren, (1992) P pili in uropathogenic *E. coli* are composite fibres with distinct fibrillar adhesive tips. *Nature* **356**: 252-255.
- Landgraf, J. R., J. Wu & J. M. Calvo, (1996) Effects of nutrition and growth rate on Lrp levels in *Escherichia coli*. *J Bacteriol* **178**: 6930-6936.
- Landini, P. & A. J. Zehnder, (2002) The global regulatory *hns* gene negatively affects adhesion to solid surfaces by anaerobically grown *Escherichia coli* by modulating expression of flagellar genes and lipopolysaccharide production. *J Bacteriol* **184**: 1522-1529.
- Lane, M. C. & H. L. T. Mobley, (2007) Role of P-fimbrial-mediated adherence in pyelonephritis and persistence of uropathogenic *Escherichia coli* (UPEC) in the mammalian kidney. *Kidney Int* **72**: 19-25.
- Lässig, M., (2007) From biophysics to evolutionary genetics: statistical aspects of gene regulation. *BMC bioinformatics* **8 Suppl 6**: S7.
- Li, H., L. Qian, Z. Chen, D. Thibault, G. Liu, T. Liu & D. G. Thanassi, (2004) The outer membrane usher forms a twin-pore secretion complex. *J Mol Biol* **344**: 1397-1407.
- Lim, H. N. & A. van Oudenaarden, (2007) A multistep epigenetic switch enables the stable inheritance of DNA methylation states. *Nat Genet* **39**: 269-275.
- Lindberg, F., J. M. Tennent, S. J. Hultgren, B. Lund & S. Normark, (1989) PapD, a periplasmic transport protein in P-pilus biogenesis. *J Bacteriol* **171**: 6052-6058.
- Lloyd, G., P. Landini & S. Busby, (2001) Activation and repression of transcription initiation in bacteria. *Essays Biochem* **37**: 17-31.
- Low, D. A. & J. Casadesús, (2008) Clocks and switches: bacterial gene regulation by DNA adenine methylation. *Curr Opin Microbiol* **11**: 106-112.

- Low, D. A., N. J. Weyand & M. J. Mahan, (2001) Roles of DNA adenine methylation in regulating bacterial gene expression and virulence. *Infect Immun* **69**: 7197-7204.
- Maamar, H. & D. Dubnau, (2005) Bistability in the *Bacillus subtilis* K-state (competence) system requires a positive feedback loop. *Molecular microbiology* **56**: 615-624.
- Maamar, H., A. Raj & D. Dubnau, (2007) Noise in gene expression determines cell fate in *Bacillus subtilis*. *Science (New York, NY)* **317**: 526-529.
- Maiti, S. N., L. DesGroseillers, J. M. Fairbrother & J. Harel, (1994) Analysis of genes coding for the major and minor fimbrial subunits of the Prs-like fimbriae F165(1) of porcine septicemic *Escherichia coli* strain 4787. *Microb Pathog* **16**: 15-25.
- Maiti, S. N., J. Harel & J. M. Fairbrother, (1993) Structure and copy number analyses of *pap*-, *sfa*-, and *afa*-related gene clusters in F165-positive bovine and porcine *Escherichia coli* isolates. *Infect Immun* **61**: 2453-2461.
- Marczynski, G. T. & L. Shapiro, (2002) Control of chromosome replication in *Caulobacter crescentus*. *Annu Rev Microbiol* **56**: 625-656.
- Marrs, C. F., L. Zhang & B. Foxman, (2005) *Escherichia coli* mediated urinary tract infections: are there distinct uropathogenic *E. coli* (UPEC) pathotypes? *FEMS Microbiol Lett* **252**: 183-190.
- Martin, C., (1996) The *clp* (CS31A) operon is negatively controlled by Lrp, ClpB, and L-alanine at the transcriptional level. *Mol Microbiol* **21**: 281-292.
- Martin, C., C. Boeuf & F. Bousquet, (1991) *Escherichia coli* CS31A fimbriae: molecular cloning, expression and homology with the K88 determinant. *Microb Pathog* **10**: 429-442.
- Maynard, C., S. Bekal, F. Sanschagrín, R. C. Levesque, R. Brousseau, L. Masson, S. Larivière & J. Harel, (2004) Heterogeneity among virulence and antimicrobial resistance gene profiles of extraintestinal *Escherichia coli* isolates of animal and human origin. *J Clin Microbiol* **42**: 5444-5452.

- McBean, M. & S. Rajamani, (2001) Increasing rates of hospitalization due to septicemia in the US elderly population, 1986-1997. *J Infect Dis* **183**: 596-603.
- Merckel, M. C., J. Tanskanen, S. Edelman, B. Westerlund-Wikstrom, T. K. Korhonen & A. Goldman, (2003) The structural basis of receptor-binding by *Escherichia coli* associated with diarrhea and septicemia. *J Mol Biol* **331**: 897-905.
- Mettetal, J. T. & A. van Oudenaarden, (2007) Microbiology. Necessary noise. *Science* **317**: 463-464.
- Moreno, E., I. Planells, G. Prats, A. M. Planes, G. Moreno & A. Andreu, (2005) Comparative study of *Escherichia coli* virulence determinants in strains causing urinary tract bacteremia versus strains causing pyelonephritis and other sources of bacteremia. *Diagn Microbiol Infect Dis* **53**: 93-99.
- Moreno-Campuzano, S., S. C. Janga & E. Perez-Rueda, (2006) Identification and analysis of DNA-binding transcription factors in *Bacillus subtilis* and other Firmicutes-a genomic approach. *BMC Genomics* **7**: 147.
- Morschhauser, J., B. E. Uhlin & J. Hacker, (1993) Transcriptional analysis and regulation of the *sfa* determinant coding for S fimbriae of pathogenic *Escherichia coli* strains. *Mol Gen Genet* **238**: 97-105.
- Müller, C. M., A. Aberg, J. Strasevičienė, L. Emody, B. E. Uhlin & C. Balsalobre, (2009) Type 1 fimbriae, a colonization factor of uropathogenic *Escherichia coli*, are controlled by the metabolic sensor CRP-cAMP. *PLoS Pathog* **5**: e1000303.
- Nagy, G., A. Altenhoefer, O. Knapp, E. Maier, U. Dobrindt, G. Blum-Oehler, R. Benz, L. Emody & J. Hacker, (2006) Both alpha-haemolysin determinants contribute to full virulence of uropathogenic *Escherichia coli* strain 536. *Microbes Infect* **8**: 2006-2012.
- Nakayama, J., M. N. Fukuda, B. Fredette, B. Ranscht & M. Fukuda, (1995) Expression cloning of a human polysialyltransferase that forms the polysialylated neural cell adhesion molecule present in embryonic brain. *Proc Natl Acad Sci USA* **92**: 7031-7035.

- Nataro, J. P. & J. B. Kaper, (1998) Diarrheagenic *Escherichia coli*. *Clin Microbiol Rev* **11**: 142-201.
- Nelson, E., (1985) *Quantum Fluctuations*. Princeton University Press, Princeton.
- Newman, E. B. & R. Lin, (1995) Leucine-responsive regulatory protein: a global regulator of gene expression in *E. coli*. *Annu Rev Microbiol* **49**: 747-775.
- Ngeleka, M. & J. M. Fairbrother, (1999) F165(1) fimbriae of the P fimbrial family inhibit the oxidative response of porcine neutrophils. *FEMS Immunol Med Microbiol* **25**: 265-274.
- Ngeleka, M., M. Jacques, B. Martineau-Doizé, F. Daigle, J. Harel & J. M. Fairbrother, (1993) Pathogenicity of an *Escherichia coli* O115:K"V165" mutant negative for F165(1) fimbriae in septicemia of gnotobiotic pigs. *Infect Immun* **61**: 836-843.
- Ngeleka, M., B. Martineau-Doizé & J. M. Fairbrother, (1994) Septicemia-inducing *Escherichia coli* O115:K"V165" F165(1) resists killing by porcine polymorphonuclear leukocytes *in vitro*: role of F165(1) fimbriae and K"V165" O-antigen capsule. *Infect Immun* **62**: 398-404.
- Nicolle, L. E., (2003) Asymptomatic bacteriuria: when to screen and when to treat. *Infect Dis Clin North Am* **17**: 367-394.
- Nielubowicz, G. R. & H. L. T. Mobley, (2010) Host-pathogen interactions in urinary tract infection. *Nature reviews Urology*.
- Nou, X., B. Braaten, L. Kaltenbach & D. A. Low, (1995) Differential binding of Lrp to two sets of *pap* DNA binding sites mediated by Pap I regulates Pap phase variation in *Escherichia coli*. *EMBO J* **14**: 5785-5797.
- Nou, X., B. Skinner, B. Braaten, L. Blyn, D. Hirsch & D. Low, (1993) Regulation of pyelonephritis-associated pili phase-variation in *Escherichia coli*: binding of the PapI and the Lrp regulatory proteins is controlled by DNA methylation. *Mol Microbiol* **7**: 545-553.
- Nowrouzian, F. L., I. Adlerberth & A. E. Wold, (2006) Enhanced persistence in the colonic microbiota of *Escherichia coli* strains belonging to phylogenetic group B2: role of virulence factors and adherence to colonic cells. *Microbes Infect* **8**: 834-840.

- Nowrouzian, F. L., A. E. Wold & I. Adlerberth, (2005) *Escherichia coli* strains belonging to phylogenetic group B2 have superior capacity to persist in the intestinal microflora of infants. *J Infect Dis* **191**: 1078-1083.
- Nudler, E. & M. E. Gottesman, (2002) Transcription termination and anti-termination in *E. coli*. *Genes to cells : devoted to molecular & cellular mechanisms* **7**: 755-768.
- Olsen, A., M. J. Wick, M. Morgelin & L. Bjorck, (1998) Curli, fibrous surface proteins of *Escherichia coli*, interact with major histocompatibility complex class I molecules. *Infect Immun* **66**: 944-949.
- Ozbudak, E. M., M. Thattai, I. Kurtser, A. D. Grossman & A. van Oudenaarden, (2002) Regulation of noise in the expression of a single gene. *Nat Genet* **31**: 69-73.
- Ozbudak, E. M., M. Thattai, H. N. Lim, B. I. Shraiman & A. van Oudenaarden, (2004) Multistability in the lactose utilization network of *Escherichia coli*. *Nature* **427**: 737-740.
- Peterson, S. N., F. W. Dahlquist & N. O. Reich, (2007) The role of high affinity non-specific DNA binding by Lrp in transcriptional regulation and DNA organization. *J Mol Biol* **369**: 1307-1317.
- Peterson, S. N. & N. O. Reich, (2006) GATC flanking sequences regulate Dam activity: evidence for how Dam specificity may influence *pap* expression. *J Mol Biol* **355**: 459-472.
- Peterson, S. N. & N. O. Reich, (2008) Competitive Lrp and Dam assembly at the *pap* regulatory region: implications for mechanisms of epigenetic regulation. *J Mol Biol* **383**: 92-105.
- Picard, B., J. S. Garcia, S. Gouriou, P. Duriez, N. Brahimi, E. Bingen, J. Elion & E. Denamur, (1999) The link between phylogeny and virulence in *Escherichia coli* extraintestinal infection. *Infect Immun* **67**: 546-553.
- Pong, A. & J. S. Bradley, (1999) Bacterial meningitis and the newborn infant. *Infect Dis Clin North Am* **13**: 711-733, viii.

- Pourbakhsh, S. A., M. Boulianne, B. Martineau-Doizé, C. M. Dozois, C. Desautels & J. M. Fairbrother, (1997) Dynamics of *Escherichia coli* infection in experimentally inoculated chickens. *Avian Dis* **41**: 221-233.
- Pucciarelli, M. G., A. I. Prieto, J. Casadesus & F. Garcia-del Portillo, (2002) Envelope instability in DNA adenine methylase mutants of *Salmonella enterica*. *Microbiology* **148**: 1171-1182.
- Reisenauer, A., L. S. Kahng, S. McCollum & L. Shapiro, (1999) Bacterial DNA methylation: a cell cycle regulator? *J Bacteriol* **181**: 5135-5139.
- Riber, L. & A. Lobner-Olesen, (2005) Coordinated replication and sequestration of *oriC* and *dnaA* are required for maintaining controlled once-per-cell-cycle initiation in *Escherichia coli*. *J Bacteriol* **187**: 5605-5613.
- Rodionov, D. A., (2007) Comparative genomic reconstruction of transcriptional regulatory networks in bacteria. *Chem Rev* **107**: 3467-3497.
- Rodionov, D. A., P. S. Novichkov, E. D. Stavrovskaya, I. A. Rodionova, X. Li, M. D. Kazanov, D. A. Ravcheev, A. V. Gerasimova, A. E. Kazakov, G. Y. Kovaleva, E. A. Permina, O. N. Laikova, R. Overbeek, M. F. Romine, J. K. Fredrickson, A. P. Arkin, I. Dubchak, A. L. Osterman & M. S. Gelfand, (2011) Comparative genomic reconstruction of transcriptional networks controlling central metabolism in the *Shewanella* genus. *BMC Genomics* **12 Suppl 1**: S3.
- Rodriguez-Siek, K. E., C. W. Giddings, C. Doetkott, T. J. Johnson, M. K. Fakhr & L. K. Nolan, (2005a) Comparison of *Escherichia coli* isolates implicated in human urinary tract infection and avian colibacillosis. *Microbiology (Reading, Engl)* **151**: 2097-2110.
- Rodriguez-Siek, K. E., C. W. Giddings, C. Doetkott, T. J. Johnson & L. K. Nolan, (2005b) Characterizing the APEC pathotype. *Vet Res* **36**: 241-256.
- Ron, E. Z., (2006) Host specificity of septicemic *Escherichia coli*: human and avian pathogens. *Curr Opin Microbiol* **9**: 28-32.
- Rosen, D. A., T. M. Hooton, W. E. Stamm, P. A. Humphrey & S. J. Hultgren, (2007) Detection of intracellular bacterial communities in human urinary tract infection. *PLoS Med* **4**: e329.

- Russo, T. A., B. A. Davidson, S. A. Genagon, N. M. Warholc, U. Macdonald, P. D. Pawlicki, J. M. Beanan, R. Olson, B. A. Holm & P. R. Knight, (2005) *E. coli* virulence factor hemolysin induces neutrophil apoptosis and necrosis/lysis in vitro and necrosis/lysis and lung injury in a rat pneumonia model. *Am J Physiol Lung Cell Mol Physiol* **289**: L207-216.
- Russo, T. A. & J. R. Johnson, (2000) Proposal for a new inclusive designation for extraintestinal pathogenic isolates of *Escherichia coli*: ExPEC. *J Infect Dis* **181**: 1753-1754.
- Russo, T. A. & J. R. Johnson, (2003) Medical and economic impact of extraintestinal infections due to *Escherichia coli*: focus on an increasingly important endemic problem. *Microbes Infect* **5**: 449-456.
- Salyers, A. A., A. Gupta & Y. Wang, (2004) Human intestinal bacteria as reservoirs for antibiotic resistance genes. *Trends Microbiol* **12**: 412-416.
- Silva-Rocha, R. & V. de Lorenzo, (2008) Mining logic gates in prokaryotic transcriptional regulation networks. *FEBS Lett* **582**: 1237-1244.
- Simms, A. N. & H. L. T. Mobley, (2008a) Multiple genes repress motility in uropathogenic *Escherichia coli* constitutively expressing type 1 fimbriae. *J Bacteriol* **190**: 3747-3756.
- Simms, A. N. & H. L. T. Mobley, (2008b) PapX, a P fimbrial operon-encoded inhibitor of motility in uropathogenic *Escherichia coli*. *Infect Immun* **76**: 4833-4841.
- Smith, J. L., P. M. Fratamico & N. W. Gunther, (2007) Extraintestinal pathogenic *Escherichia coli*. *Foodborne Pathog Dis* **4**: 134-163.
- Smits, W. K., T. T. Hoa, L. W. Hamoen, O. P. Kuipers & D. Dubnau, (2007) Antirepression as a second mechanism of transcriptional activation by a minor groove binding protein. *Mol Microbiol* **64**: 368-381.
- So, S. S. & D. G. Thanassi, (2006) Analysis of the requirements for pilus biogenesis at the outer membrane usher and the function of the usher C-terminus. *Mol Microbiol* **60**: 364-375.
- Soto, G. E., K. W. Dodson, D. Ogg, C. Liu, J. Heuser, S. Knight, J. Kihlberg, C. H. Jones & S. J. Hultgren, (1998) Periplasmic chaperone recognition motif of

- subunits mediates quaternary interactions in the pilus. *EMBO J* **17**: 6155-6167.
- Stamm, W. E. & T. M. Hooton, (1993) Management of urinary tract infections in adults. *N Engl J Med* **329**: 1328-1334.
- Stülke, J. & W. Hillen, (1999) Carbon catabolite repression in bacteria. *Current opinion in microbiology* **2**: 195-201.
- Tal, S., V. Guller, S. Levi, R. Bardenstein, D. Berger, I. Gurevich & A. Gurevich, (2005) Profile and prognosis of febrile elderly patients with bacteremic urinary tract infection. *J Infect* **50**: 296-305.
- Taylor, V. L., R. W. Titball & P. C. Oyston, (2005) Oral immunization with a dam mutant of *Yersinia pseudotuberculosis* protects against plague. *Microbiology* **151**: 1919-1926.
- Tessier, M.-C., R. Graveline, C. Crost, J. A. Desabrais, C. Martin, M. Drolet & J. Harel, (2007) Effects of DNA supercoiling and topoisomerases on the expression of genes coding for F165(1), a P-like fimbriae. *FEMS Microbiol Lett* **277**: 28-36.
- Thattai, M. & A. van Oudenaarden, (2001) Intrinsic noise in gene regulatory networks. *Proc Natl Acad Sci USA* **98**: 8614-8619.
- Totsika, M., S. A. Beatson, N. Holden & D. L. Gally, (2008) Regulatory interplay between *pap* operons in uropathogenic *Escherichia coli*. *Mol Microbiol* **67**: 996-1011.
- van der Woude, M. W., (2006) Re-examining the role and random nature of phase variation. *FEMS Microbiol Lett* **254**: 190-197.
- van der Woude, M. W., (2011) Phase variation: how to create and coordinate population diversity. *Curr Opin Microbiol*.
- van der Woude, M. W. & A. J. Bäumlner, (2004) Phase and antigenic variation in bacteria. *Clin Microbiol Rev* **17**: 581-611, table of contents.
- van Hijum, S. A. F. T., M. H. Medema & O. P. Kuipers, (2009) Mechanisms and evolution of control logic in prokaryotic transcriptional regulation. *Microbiol Mol Biol Rev* **73**: 481-509, Table of Contents.

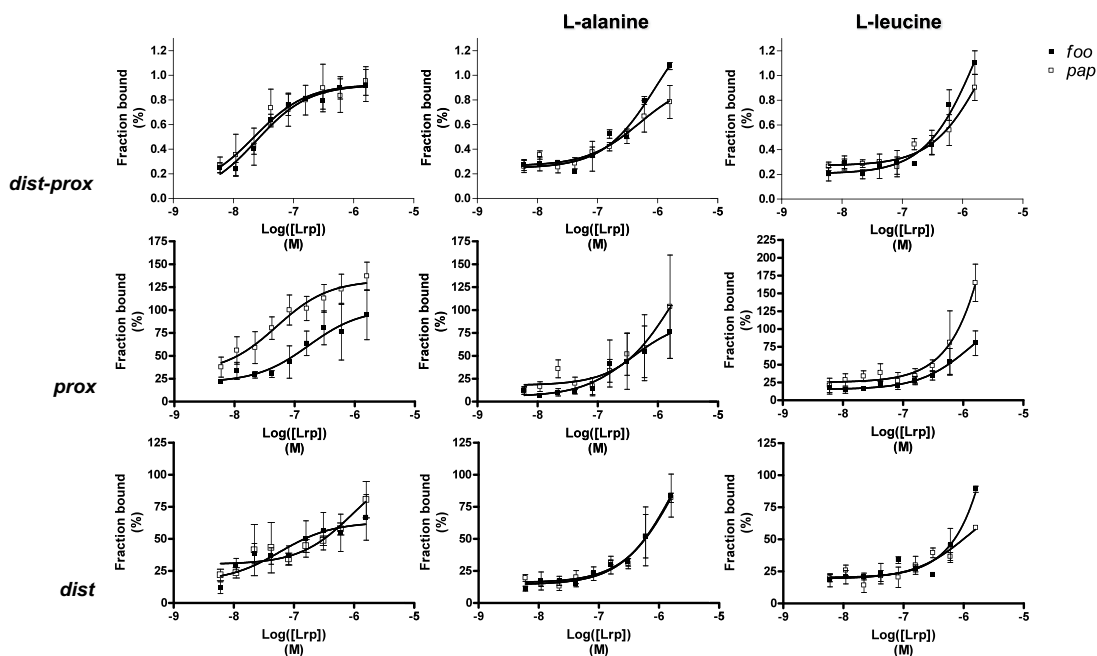
- Veening, J.-W., W. K. Smits & O. P. Kuipers, (2008a) Bistability, epigenetics, and bet-hedging in bacteria. *Annu Rev Microbiol* **62**: 193-210.
- Veening, J.-W., E. J. Stewart, T. W. Berngruber, F. Taddei, O. P. Kuipers & L. W. Hamoen, (2008b) Bet-hedging and epigenetic inheritance in bacterial cell development. *Proc Natl Acad Sci USA* **105**: 4393-4398.
- Vergnano, S., M. Sharland, P. Kazembe, C. Mwansambo & P. T. Heath, (2005) Neonatal sepsis: an international perspective. *Arch Dis Child Fetal Neonatal Ed* **90**: F220-224.
- Waksman, G. & S. J. Hultgren, (2009) Structural biology of the chaperone-usher pathway of pilus biogenesis. *Nature reviews Microbiology* **7**: 765-774.
- Watson, M. E., Jr., J. Jarisch & A. L. Smith, (2004) Inactivation of deoxyadenosine methyltransferase (dam) attenuates *Haemophilus influenzae* virulence. *Mol Microbiol* **53**: 651-664.
- Westerlund-Wikstrom, B. & T. K. Korhonen, (2005) Molecular structure of adhesin domains in *Escherichia coli* fimbriae. *Int J Med Microbiol* **295**: 479-486.
- White-Ziegler, C. A., M. L. Angus Hill, B. A. Braaten, M. W. van der Woude & D. A. Low, (1998) Thermoregulation of *Escherichia coli* *pap* transcription: H-NS is a temperature-dependent DNA methylation blocking factor. *Mol Microbiol* **28**: 1121-1137.
- White-Ziegler, C. A., A. M. Black, S. H. Eliades, S. Young & K. Porter, (2002) The N-acetyltransferase RimJ responds to environmental stimuli to repress *pap* fimbrial transcription in *Escherichia coli*. *J Bacteriol* **184**: 4334-4342.
- White-Ziegler, C. A., A. Villapakkam, K. Ronaszeki & S. Young, (2000) H-NS controls *pap* and *daa* fimbrial transcription in *Escherichia coli* in response to multiple environmental cues. *J Bacteriol* **182**: 6391-6400.
- Wiles, T. J., R. R. Kulesus & M. A. Mulvey, (2008) Origins and virulence mechanisms of uropathogenic *Escherichia coli*. *Exp Mol Pathol* **85**: 11-19.
- Wion, D. & J. Casadesús, (2006) N6-methyl-adenine: an epigenetic signal for DNA-protein interactions. *Nat Rev Microbiol* **4**: 183-192.

- Xie, J., B. Foxman, L. Zhang & C. F. Marrs, (2006) Molecular epidemiologic identification of *Escherichia coli* genes that are potentially involved in movement of the organism from the intestinal tract to the vagina and bladder. *J Clin Microbiol* **44**: 2434-2441.
- Xie, Y., K. J. Kim & K. S. Kim, (2004) Current concepts on *Escherichia coli* K1 translocation of the blood-brain barrier. *FEMS Immunol Med Microbiol* **42**: 271-279.
- Yamamoto, S., T. Tsukamoto, A. Terai, H. Kurazono, Y. Takeda & O. Yoshida, (1997) Genetic evidence supporting the fecal-perineal-urethral hypothesis in cystitis caused by *Escherichia coli*. *J Urol* **157**: 1127-1129.
- Yan, F. & D. B. Polk, (2004) Commensal bacteria in the gut: learning who our friends are. *Curr Opin Gastroenterol* **20**: 565-571.
- Yokoyama, K., S. A. Ishijima, L. Clowney, H. Koike, H. Aramaki, C. Tanaka, K. Makino & M. Suzuki, (2006) Feast/famine regulatory proteins (FFRPs): *Escherichia coli* Lrp, AsnC and related archaeal transcription factors. *FEMS Microbiol Rev* **30**: 89-108.
- Yokoyama, K., H. Nogami, M. Kabasawa, S. Ebihara, A. Shimowasa, K. Hashimoto, T. Kawashima, S. A. Ishijima & M. Suzuki, (2009) The DNA-recognition mode shared by archaeal feast/famine-regulatory proteins revealed by the DNA-binding specificities of TvFL3, FL10, FL11 and Ss-LrpB. *Nucleic Acids Res* **37**: 4407-4419.
- Zhang, L. & B. Foxman, (2003) Molecular epidemiology of *Escherichia coli* mediated urinary tract infections. *Front Biosci* **8**: e235-244.
- Zhou, B., D. Beckwith, L. R. Jarboe & J. C. Liao, (2005) Markov Chain modeling of pyelonephritis-associated pili expression in uropathogenic *Escherichia coli*. *Biophys J* **88**: 2541-2553.
- Zupan, J., (2005) Perinatal mortality in developing countries. *N Engl J Med* **352**: 2047-2048.
- Zweiger, G., G. Marczynski & L. Shapiro, (1994) A Caulobacter DNA methyltransferase that functions only in the predivisive cell. *J Mol Biol* **235**: 472-485.

ANNEXES

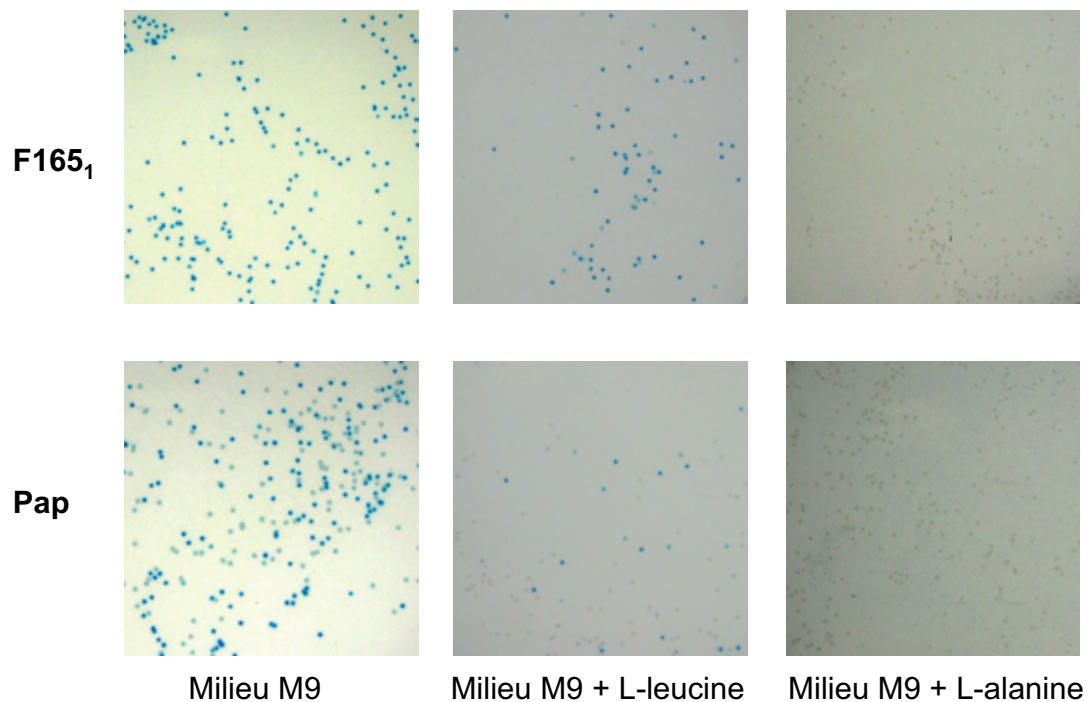
ANNEXE I: Influence de la leucine et de l'alanine sur la variation de phase de Pap et de F165₁, ainsi que sur la fixation de Lrp aux sites 1-3 et 4-6 des opérons *pap* et *foo*.

La leucine et l'alanine inhibent la fixation de Lrp sur la région régulatrice des opérons *foo* et *pap* *in vitro*.



Influence de la L-leucine et de la L-alanine sur la fixation de Lrp sur les régions régulatrices des opérons *foo* et *pap*. Les interactions de Lrp avec les régions distale et proximale des opérons *foo* et *pap* ont été mesurées par retard de migration sur gel, en absence (colonne de gauche) ou en présence de L-alanine ou de L-leucine (colonne du milieu et colonne de droite respectivement) à 10 mM final.

La L-alanine et la L-leucine inhibent également la variation de phase des souches exprimant une fusion transcriptionnelle entre la région régulatrice des opérons *foo* et *pap* et le système rapporteur *lacZYA*.



Suivi de la variation de phase à l'aide du système rapporteur *lacZYA*. Les souches contenant une fusion transcriptionnelle entre la région régulatrice de l'opéron *foo* ou de l'opéron *pap* et le système rapporteur *lacZYA* ont été ensemencées sur milieu M9 normal (colonne de gauche) ou supplémenté avec de la L-leucine (10 mM, colonne du milieu) ou de la L-alanine (10mM, colonne de droite). Après approximativement 20 heures de croissance à 37°C, les gélules sont alors sorties de l'incubateur et une photographie numérique a été prise. Chaque expérience a été réalisée au moins à trois reprises différentes.

ANNEXE II: Effects of DNA supercoiling and topoisomerases on the expression of genes coding for F165(1), a P-like fimbriae.

Effects of DNA supercoiling and topoisomerases on the expression of genes coding for F165₁, a P-like fimbriae

Marie-Catherine Tessier¹, Richard Graveline¹, Cécile Crost¹, Julie Annick Desabrais², Christine Martin³, Marc Drolet² & Josée Harel¹

¹Département de pathologie et microbiologie, Université de Montréal, St-Hyacinthe, QC, Canada; ²Département de microbiologie et immunologie, Université de Montréal, Montréal, QC, Canada; and ³Laboratoire de microbiologie, INRA Clermont-Ferrand-Theix, 63122 St-Genès-Champanelle, France

Correspondence: Josée Harel, Département de pathologie et microbiologie, Université de Montréal, St-Hyacinthe, QC Canada J2S 7C6.

Received 3 July 2007; accepted 10 August 2007.

First published online November 2007.

DOI:10.1111/j.1574-6968.2007.00919.x

Editor: Mark Schembri

Keywords

fimbriae; regulation; topoisomerase.

Abstract

Pathogenic *Escherichia coli* 4787 (O115:KV165) causes septicemia in pigs and expresses the fimbriae F165₁ encoded by the *foo* operon that belongs to the P fimbrial family. *fooI* and *fooB*, encoding specific *foo* regulators, are divergently transcribed; their intergenic region is responsible for the regulation of *foo* expression. The role of global and local supercoiling (transcription-induced supercoiling within the intergenic region) on the regulation of *foo* expression was investigated. Expression of *fooB* was significantly altered when global negative supercoiling was reduced by a mutation that decreases DNA gyrase activity. Deletion of the *topA* gene, encoding for topoisomerase I that relaxes local negative supercoiling, further reduced *fooB* expression. This suggests that both global and local supercoiling can significantly affect *fooB* expression. Moreover, *FooI*, a positive regulator of *fooB* expression, has no effect on *fooB* expression in the *topA* null mutant. This study showed that divergent transcription from a strong promoter can significantly enhance *fooB* expression and compensate for the absence of *FooI* in a wild-type strain.

Introduction

F165₁ fimbriae are produced by *Escherichia coli* 4787, a strain associated with septicemia in pigs, and involved in pathogenesis (Ngeleka *et al.*, 1992, 1993). The F165₁ adhesin from *E. coli*, encoded by the *foo* operon, belongs to the family of Pap-related fimbriae. The expression of *foo* is controlled at the transcriptional level by two superimposed mechanisms that modulate both the percentage of cells producing fimbriae in a single colony (phase variation) and the amount of antigen at the surface of a single cell [steady-state transcription (Daigle *et al.*, 2000; Harel *et al.*, 2000)]. As in Pap (Hernday *et al.*, 2002), F165₁ expression is mediated by regulatory proteins such as Lrp, Dam methylase, and by *FooI* and *FooB*, the *PapI* and *PapB* homologs, respectively (Daigle *et al.*, 2000; Berthiaume *et al.*, 2004). *FooI* is encoded by the *fooI* gene that is divergently transcribed with respect to *fooB*, the first gene of the *foo* operon (Daigle *et al.*, 2000; Hernday *et al.*, 2002). Similar to the *pap* operon, there are two 102 bp-spaced GATC sites within the

intergenic region between *fooI* and *foo* operon. Both the presence of regulatory sites for global regulators and the fact that *fooI* and *fooB* are divergently transcribed may suggest an important role for DNA supercoiling in *foo* operon expression. Indeed, both global supercoiling, which reflects the average superhelical density of all supercoiling domains, and local supercoiling, which is generated in the wake of moving RNA polymerases during transcription (Liu & Wang, 1987), have been recognized as important determinants for gene expression (Opel *et al.*, 2001). In *E. coli*, the level of global supercoiling is mainly regulated by DNA gyrase, made of two subunits encoded by *gyrA* and *gyrB*, that has a supercoiling activity, and by the cellular ATP/ADP ratio (Drlica, 1992). DNA topoisomerase I, which relaxes negative supercoiling, plays at best only a minor role in the regulation of global supercoiling, but is responsible for regulating transcription-induced negative supercoiling (Masse & Drolet, 1999). To study the influence of DNA supercoiling on *foo* operon expression, chromosomal *lacZ* fusions with the *fooB* gene were constructed and a set of isogenic strains was used,

allowing the modulation both of global and local supercoiling. This study shows that both global and transcription-induced supercoiling (local) can significantly affect *fooB* expression. These results suggest that DNA topology within the *fooIB* intergenic region can significantly affect the multiple protein–DNA interactions that must take place to properly regulate *foo* expression.

Materials and methods

Bacterial strains and plasmids

The bacterial strains used in this study are listed in Table 1. P1vir transduction (Miller, 1992) was used to introduce the $\Delta(\text{argF-lac})\text{U169 zah735}::\text{Tn10}$ region into strains DM800 $\Delta\text{topA gyrB225}$, EM205 $\text{topA}^+\text{gyrB225}$ and EM209 $\text{topA}^+\text{gyrB}^+$ to obtain, respectively, strains PH617, PH615 and PH613.

The chromosomal *foo-lac* fusions (Fig. 1) were obtained using λRS45 as described by Simons *et al.* (1987). Pfu polymerase (Stratagene, CA) was used for PCR to obtain various DNA fragments from *E. coli* 4787 strain carrying the *foo* operon (Daigle *et al.*, 2000; Crost *et al.*, 2004). The PCR-generated fragments, which were all sequenced, were inserted into the pGEM-T vector (Promega, WI) before cloning into the BamHI site of pRS551 to generate the *lac* fusions. Each recombinant plasmid was introduced into strains PH617, PH615 and PH613 and transformants were transfected with the lambda phage λRS45 , which carries homology to pRS551 as described by Simons *et al.* Single λ phage integrations were verified by PCR as described by Powell *et al.* (1994). To obtain fusion 1 (*fooIB-lac*), a 1039-bp DNA fragment with BamHI and BglII linkers was generated

by PCR with 5'-GGATCCTTTTATCTGTTTCAGT-3' and 5'-AGATCTAAGGGCTTCTTTTTTTT-3' primers. For fusion 2 (*fooIp-fooB-lac*), a 776-bp DNA fragment with BamHI and BglII linkers was generated by PCR with 5'-GGATCCGTATTCACTCATCTCACTG-3' and 5'-AGATCTAAGGGCTTCTTTTTTTT-3' primers. The PCR fragment with BamHI and BglII linkers carrying the *Ptrc* promoter in fusion 3 (*tetA-trcP-fooB-lac*) and 4 (*trcP-fooB-lac*) was obtained using pTrc99a (Amersham Biosciences), and 5'-GGATCCGATCTGGTTTGA-3' and 5'-AGATCTCCGA GCTCGAAT-3' primers. A DNA fragment with BamHI and BglII linkers carrying the *tetA* gene from pACYC184 without its promoter was generated by PCR using 5'-GGATCCGCGGTAGTTTAT-3' and 5'-AGATCTATTCCG TTAGCGAGG-3' primers. This fragment was cloned under the control of *Ptrc* to generate fusion 3.

Media and growth conditions

Luria–Bertani (LB) and M9 minimal media were prepared as described previously (Miller, 1992). LB agar was supplemented with 50 $\mu\text{g mL}^{-1}$ cysteine and bromo-4-chloro-3-indolyl- β -D-galactopyranoside (X-gal) at a final concentration of 40 $\mu\text{g mL}^{-1}$. When needed, antibiotics were added to the media at the following concentration: kanamycin, 16.7 $\mu\text{g mL}^{-1}$; tetracycline, 3.3 $\mu\text{g mL}^{-1}$; and chloramphenicol, 10 $\mu\text{g mL}^{-1}$. DM800 derivatives carry the *acrA13* mutation (DiNardo *et al.*, 1982), rendering them more permeable to many antibiotics. For β -galactosidase assays and RNA extraction, the cells were first streaked on LB agar with appropriate antibiotics. Single colonies were used to inoculate liquid M9 minimal medium supplemented with glucose (0.2%), glycerol (0.2%), 40 $\mu\text{g mL}^{-1}$ L-cysteine, L-arginine,

Table 1. Bacterial strains used in this study

Strain	Description	References and source
DM800	$\Delta(\text{topA-cysB})204 \text{gyrB225}$	DiNardo <i>et al.</i> (1982)
EM205	DM800 $\text{topA}^+ \text{gyrB225}$	Masse <i>et al.</i> (1997)
EM209	DM800 $\text{topA}^+ \text{gyrB}^+$	Masse <i>et al.</i> (1997)
PH613	EM209 $\Delta(\text{argF-lac})\text{U169 zah735}::\text{Tn10 } \text{topA}^+ \text{gyrB}^+$	This work
PH615	EM205 $\Delta(\text{argF-lac})\text{U169 zah735}::\text{Tn10 } \text{topA}^+ \text{gyrB225}$	This work
PH617	DM800 $\Delta(\text{argF-lac})\text{U169 zah735}::\text{Tn10}\Delta\text{topA } \text{gyrB225}$	This work
MT008	PH613 $\lambda[\text{fooIB-lacZ}] \text{topA}^+ \text{gyrB}^+ +$	This work
MT009	PH615 $\lambda[\text{fooIB-lacZ}] \text{topA}^+ \text{gyrB225}$	This work
MT010	PH617 $\lambda[\text{fooIB-lacZ}]\Delta\text{topA } \text{gyrB225}$	This work
MT067	PH613 $\lambda[\text{fooIp-fooB-lacZ}] \text{topA}^+ \text{gyrB}^+$	This work
MT068	PH615 $\lambda[\text{fooIp-fooB-lacZ}] \text{topA}^+ \text{gyrB225}$	This work
MT069	PH617 $\lambda[\text{fooIp-fooB-lacZ}]\Delta\text{topA } \text{gyrB225}$	This work
MT070	PH613 $\lambda[\text{tetA-trcP-fooB-lacZ}] \text{topA}^+ \text{gyrB}^+$	This work
MT071	PH615 $\lambda[\text{tetA-trcP-fooB-lacZ}] \text{topA}^+ \text{gyrB225}$	This work
MT072	PH617 $\lambda[\text{tetA-trcP-fooB-lacZ}]\Delta\text{topA } \text{gyrB225}$	This work
MT073	PH613 $\lambda[\text{trcP-fooB-lacZ}] \text{topA}^+ \text{gyrB}^+$	This work
MT074	PH615 $\lambda[\text{trcP-fooB-lacZ}] \text{topA}^+ \text{gyrB225}$	This work
MT075	PH617 $\lambda[\text{trcP-fooB-lacZ}]\Delta\text{topA } \text{gyrB225}$	This work

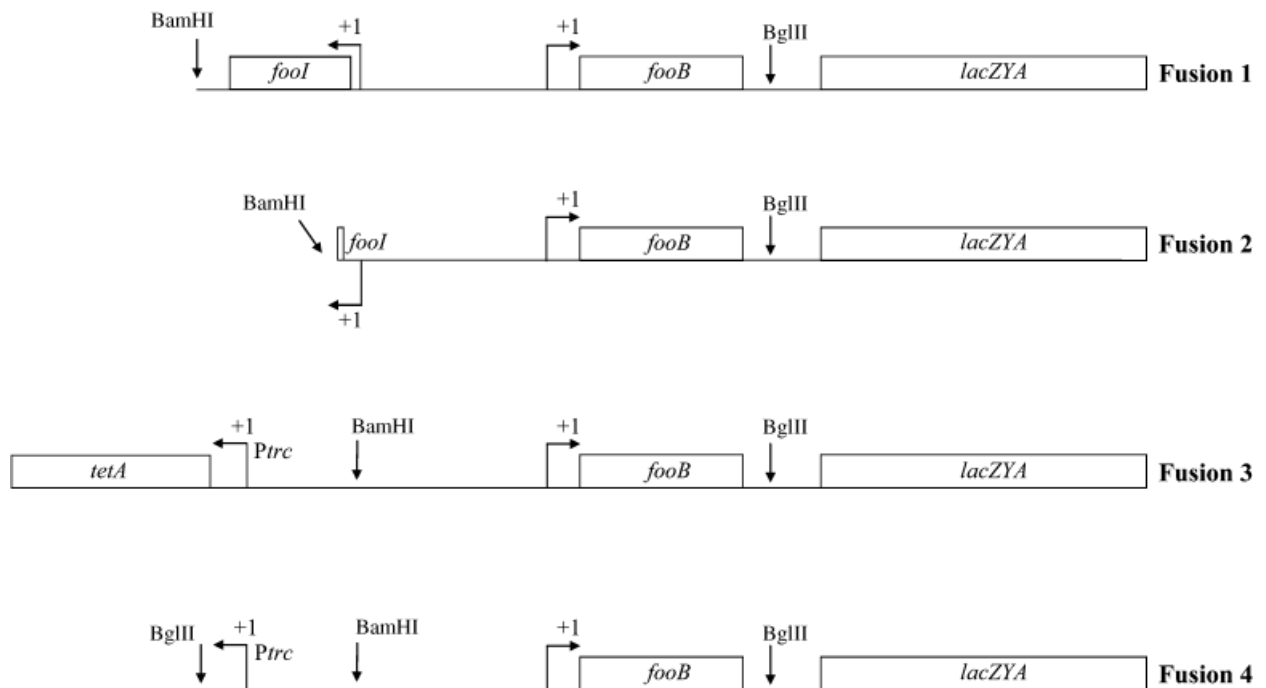


Fig. 1. Transcriptional fusions used in this study. Arrows indicate transcriptional start sites. Boxes represent ORFs *tetA* and *lacZYA* are not drawn to scale.

L-proline, L-glycine, L-isoleucine and L-valine, $20 \mu\text{g mL}^{-1}$ uracil and 5 g mL^{-1} thiamine. Overnight cultures were diluted (1/20) into M9 minimal medium supplemented with glycerol (0.4%), as the sole carbon source, and the other supplements. Cells were grown until an $\text{OD}_{600 \text{ nm}}$ of 0.8.

DNA supercoiling analysis

pBR322 $\Delta tet5'$, a pBR322 derivative with a 5' deletion in the *tetA* gene, was used to monitor the supercoiling state of the strains (Masse & Drolet, 1999). Plasmid DNA extraction and one-dimensional agarose gel electrophoresis in the presence of chloroquine at a concentration of 7.5 g mL^{-1} were performed as described in Masse & Drolet (1999).

β -Galactosidase assays

β -Galactosidase activity was measured according to Miller (1992). Each experiment was performed in duplicate at least three times. For statistical analysis, the Wilcoxon test was used to compare β -galactosidase activities between two independent samples.

RNA extraction and primer extension

RNA was extracted using the TRIZOL reagent (Invitrogen) as described (Wassarman & Storz, 2000). Primer extension analysis with 5'-GGCAACAGTACGCTCTCG-3' *fooB* pri-

mer and $10 \mu\text{g}$ of RNA was performed as described by Hraiky *et al.* (2000).

Calculation of the switch frequency

The switch frequency was calculated as described previously (Croft *et al.*, 2003). Briefly, strains were inoculated onto X-Gal M9 agar plates or LB plates. Colonies showing a uniform phenotype were suspended in M9 salts. Appropriate dilutions were spread onto M9 X-Gal glycerol agar supplemented with $40 \mu\text{g mL}^{-1}$ L-cysteine, L-arginine, L-proline, L-glycine, L-isoleucine, and L-valine, $20 \mu\text{g mL}^{-1}$ uracil, $5 \mu\text{g mL}^{-1}$ thiamine and $16.7 \mu\text{g mL}^{-1}$ kanamycin. After growth, the colonies were scored for a white or blue phenotype. The switch frequencies were calculated by the formula $(M/N)/g$, where M is the number of cells that underwent a phase variation, N the total number of cells evaluated and g the total number of generations that gave rise to the colony (for MT008 $g=25.70$; MT009 $g=22.82$; and MT010 $g=19.51$). The weighted average of frequencies was calculated based on the frequency switch of three colonies.

DNA methylation patterns analysed by DNA hybridization

In vivo GATC methylation states of the *foo* regulatory region in wild and mutant strains were analysed as described previously (Croft *et al.*, 2003). Single Lac⁺ colonies grown on M9-glycerol-X-Gal were used to inoculate M9-glycerol.

Genomic DNA from MT008, MT009 and MT010 strains was digested with restriction endonucleases PstI and HinfI. After precipitation of the restricted product, a second restriction was performed with DpnI, which only cuts methylated GATC sites, or MboI, which only cuts non-methylated GATC sites, and hybridized with a digoxigenin DNA probe that encompasses the two GATC sites. The digoxigenin-labeled DNA molecular weight marker VI (Roche) was used as a DNA migration reference. DNA restriction fragments were separated by electrophoresis on 1.5% agarose gel in TBE buffer and transferred by capillarity to a nylon positively charged membrane (Roche) in $20 \times$ SSC (0.3 M Na-citrate, 0.33 M NaCl, pH 7.0). DNA fragments were crosslinked to the filter with UV light using the GS Gene Linker (Biorad). The filter was probed with digoxigenin-labeled DNA (389 bp) generated by PCR using the PCR DIG Probe Synthesis kit (Roche) and primers DamF (5'-GAATTTGCTTTTGTGGCTTT-3') and DamR (5'-AGACATAGAGCCGGCAAC-3'). Hybridization was detected by chemiluminescence using CSDP (disodium 3-(4-methoxy spiro [1,2-dioxetane-3,2 {5'-chloro} tricyclo(3,31,1^{3,7}) decan[-4-yl]phenyl phosphate) (Roche) as a substrate for alkaline phosphatase. This analysis was carried out twice. Quantification was performed by densitometry analysis using the DIVERSITY DATABASE software from Bio-Rad.

Results and discussion

Maximal *fooB* expression requires DNA-negative supercoiling

To study the effect of DNA supercoiling and topoisomerases on *foo* operon expression, a set of isogenic strains carrying or not mutation(s) within genes encoding gyrase and/or DNA topoisomerase I was constructed. Strain DM800, a double- $\Delta topA$ *gyrB225* mutant (DiNardo *et al.*, 1982), was used to construct isogenic strains EM209 *topA*+*gyrB*+ and EM205 *topA*+*gyrB225*, such as any mutations apart from *gyrB225* that may have occurred in DM800 to compensate for the absence of *topA* will also be present in all strains. It was not possible to obtain *topA* mutants without revertants because they accumulate compensatory mutations in *gyr* (Stupina & Wang, 2005). Therefore difference in results obtained with mutant strains *gyrB225* is attributed to the *topA* mutation.

To monitor global supercoiling, pBR322 $\Delta tet5'$ was introduced in the strains. As shown in Fig. 2, the *gyrB225* mutation in EM205 reduced the global negative supercoiling of the reporter plasmid by about 25% compared with EM209 *topA*+*gyrB*+ isogenic strain, considered the wild-type strain. Because it is known that plasmid supercoiling generally reflects the chromosomal supercoiling level (Drli-

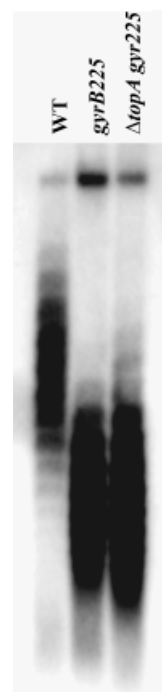


Fig. 2. DNA supercoiling in EM209 and its *gyrB225* (EM205) and $\Delta topA$ *gyrB225* (EM209) derivatives, pBR322 $\Delta tet5'$, a pBR322 derivative with a 5' deletion in the *tetA* gene (Masse & Drolet, 1999) was extracted and analyzed by agarose gel electrophoresis in the presence of $7.5 \mu\text{g mL}^{-1}$ chloroquine. In these conditions, the less negatively supercoiled DNA migrates faster.

ca, 1992), can be assumed that chromosomal negative supercoiling was also reduced by roughly 25% in the *gyrB225* mutant. The *topA* deletion in strain DM800 did not further reduce the negative supercoiling level as expected because the major function of topoisomerase I is not related to the regulation of global supercoiling. P1*vir* transduction (Miller, 1992) was used to introduce the $\Delta(argF-lac)U169$ zah735::Tn10 region into strains DM800 $\Delta topA$ *gyrB225*, EM205 *topA*+*gyrB225* and EM209 *topA*+*gyrB*+ to obtain, respectively, strains PH617, PH615 and PH613.

To measure *foo* promoter activity, isogenic strains harboring a single copy of transcriptional *fooB-lacZ* fusions were constructed (Fig. 1). As seen in Fig. 3, in a supercoiling-level background 25% lower than in the wild-type strain, expression of the *foo* operon measured using fusion 1, containing the entire regulatory region including the *fooI* and *fooB* genes, is decreased by about threefold (Fig. 3, compare fusion 1-1, MT008 *topA*+*gyrB*+ wild type with 2, MT009 *gyrB225*, $P < 0.05$). Additional deletion of the *topA* gene that does not alter global supercoiling further reduced *foo* expression by about threefold (Fig. 3, compare fusion 1-2, MT009 *gyrB225* with 3, MT010 $\Delta topA$ *gyrB225*, $P < 0.05$).

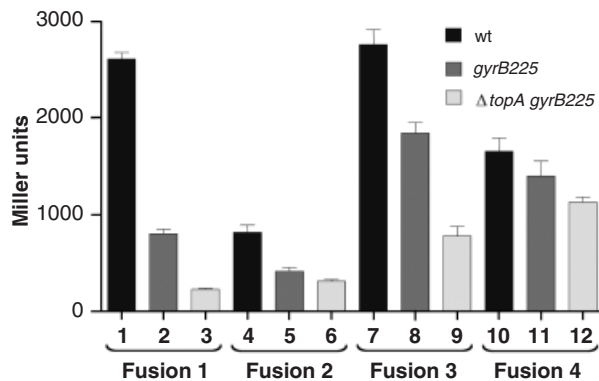


Fig. 3. Effect of DNA supercoiling on *foo* expression. β -galactosidase activities are indicated in Miller units. Values represent the mean of at least three independent experiments.

The *foo* promoter activity was also measured by primer extension. The 130 bases primer extension product indicated that the *fooB-lacZ* fusion transcriptional start site was 54-bp upstream of the ATG translational start codon (Fig. 4a). Figure 4b shows that *lacZ* transcription originated from the same promoter in MT009 and MT010 as MT008. The intensity of primer extension signals (Fig. 4b) matches the level of β -galactosidase activity (Fig. 3). Thus, differences in β -galactosidase activities are related to transcription. Altogether, these results demonstrated that both global and local supercoiling significantly affect *foo* promoter activity and that differences in expression observed between the *gyrB225* and the double- $\Delta topA$ *gyrB225* mutants are attributed to the major function of topoisomerase I that is related to transcription, most likely to relax transcription-induced negative supercoiling (Masse & Drolet, 1999).

Investigation on whether supercoiling also affected the phase variation control described previously was carried out (Crost *et al.*, 2003). ON to OFF (α) and OFF to ON (β) switches were scored in MT008, MT009 and MT010 by counting white and blue colonies on indicator plates. The wild-type and mutant strains clearly presented white Lac^- colonies and Lac^+ colonies with increased number of light blue colonies in MT009 and M010. The OFF to ON switch frequency was the highest in the three strains. As previously observed for wild-type and recombinant F165 strains (Harel *et al.*, 2000), in strain MT008, ON colonies were predominant, with more than 99% of Lac^+ cells in a Lac^+ colony. The β (OFF to ON) switch frequency of MT008 was around 4×10^{-2} cell generation $^{-1}$ whereas the α (ON to OFF) switch was less frequent ($\alpha = 7.37 \times 10^{-5}$ cell generation $^{-1}$). The ON to OFF transition was more frequent in mutants *gyrB225* MT009 ($\alpha = 6.82 \times 10^{-4}$ cell generation $^{-1}$) and $\Delta topA$ *gyrB225* MT010 ($\alpha = 9.23 \times 10^{-3}$ cell generation $^{-1}$) than in wild type. The single mutant *gyrB225* MT009 showed also a preference toward the ON state, but lower

numbers of Lac^+ colonies, 88.7%, in the OFF to ON switch. In the $\Delta topA$ *gyrB225* double-mutant MT010, lower percentages of ON colonies (Lac^+) were found in both transitions with 72.6% for OFF to ON and 85.4% for ON to OFF. Consistent with data on the repression of *foo* promoter activity by *gyrB225* and the *topA* null mutations, an increased number of light blue colonies in mutants was found and that although phase variation in the mutants still occurred its frequency was affected.

Supercoiling altered the methylation pattern of the *foo* regulatory region DNA

To determine whether the DNA supercoiling level could influence the methylation pattern of the two GATC sites of the *foo* regulatory region, Southern analyses were performed on chromosomal DNA of the lysogenic strains MT008 ($FooI^+$) and its derivative *gyrB225* and *gyrB225* $\Delta topA$ mutants as described previously (Crost *et al.*, 2003) (Fig. 5). Methylation pattern analysis of the *pap* regulatory region showed that when a GATC site is bound by Lrp, its methylation is prevented, and when a GATC site is methylated, Lrp binding is prevented. In the ON phase, Lrp binds to GATC_{dist} and GATC_{prox} is methylated, whereas in the OFF phase the opposite situation occurs (Braaten *et al.*, 1991). For *foo*, the correlation of methylation protection patterns with cells in the ON phase is less defined than for *pap* for which the affinity of Lrp is higher for the GATC_{prox} than for the GATC_{dist} region. As observed for recombinant *foo* strains (Crost *et al.*, in preparation), it was not possible to correlate the phase ON with a particular methylation status of the regulatory region because four different methylation patterns were observed in all cases, but their relative amounts varied depending on the strain (Table 2 and Fig. 5). However, the results showed that in topoisomerase mutants, the methylation status is different from that of MT008. The nonmethylated pattern of both GATC sites increased and the amount of methylated GATC_{dist} site decreased in both mutants *gyrB225* (MT009) and $\Delta topA$ *gyrB225* (MT010). Taken together, the results suggest that global and local supercoiling modulate the *foo* promoter activity by participating in either protection from methylation of the GATC sites and/or by influencing Lrp cooperative binding. Recently, global supercoiling was also shown to influence type 1 fimbrial phase variation (Kelly *et al.*, 2006) that, in contrast to Pap and Foo, requires the site-specific recombination of a short invertible element, *fimS*. The inversion is catalyzed by FimB or FimE and is influenced by auxiliary factors including Lrp (Lahooti *et al.*, 2005). Using novobiocin, a gyrase inhibitor, DNA relaxation was shown to be a prerequisite for ON-biased *fimS* inversion (Kelly *et al.*, 2006). In type 1 fimbriae, Lrp was demonstrated to collaborate with DNA supercoiling to determine the directionality of the DNA

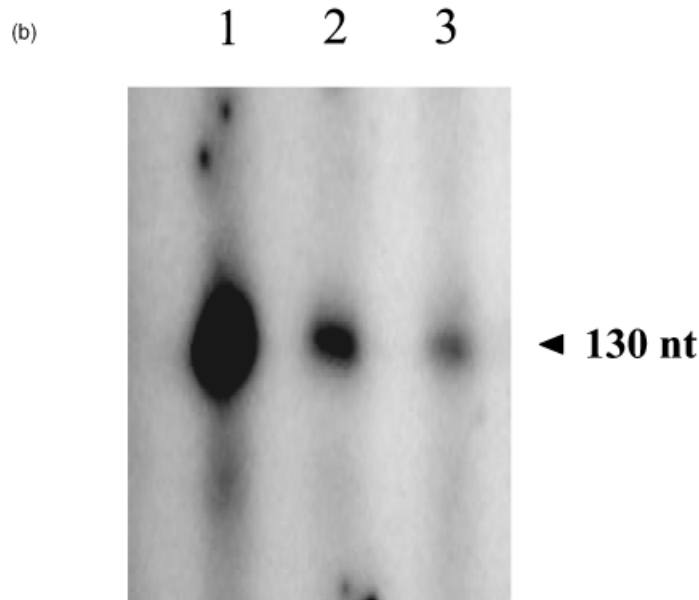
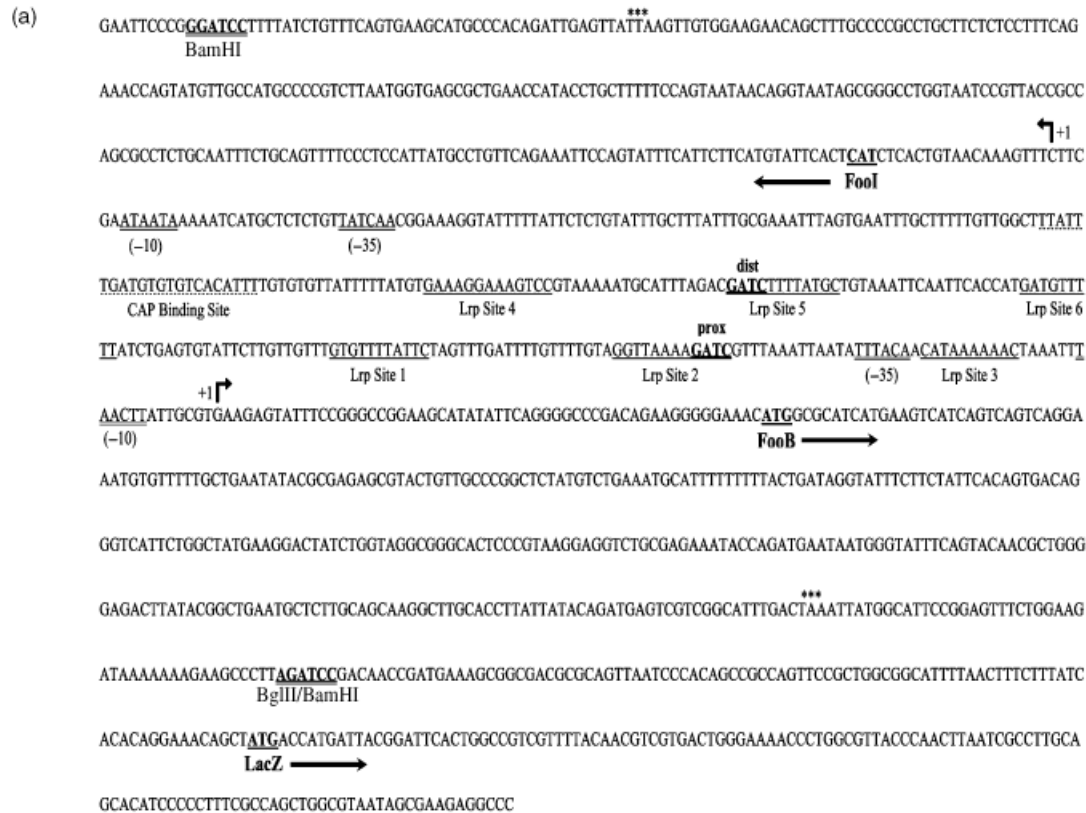


Fig. 4. Mapping of the transcriptional start site of the *foo* operon. (a) Nucleotide sequence of fragment carrying the *fooB* intergenic region within pRS551. The +1 initiation site of *fooB* has been mapped in this work. The *fooI* +1 initiation site has been determined by Desabrais, Harel & Drolet (unpublished results). Lrp, CAP and GATC indicate putative sequences for interaction with Lrp, CAP and DAM methylase, respectively. The identification of those regulatory sites is based on the strong sequence similarity of the *foo* regulatory region with the well-described *pap* operon (Hernday *et al.*, 2002). (b) Primer extension analysis using mRNA from strain MT008, MT009, MT010. The arrow indicates the length of the primer extension product.

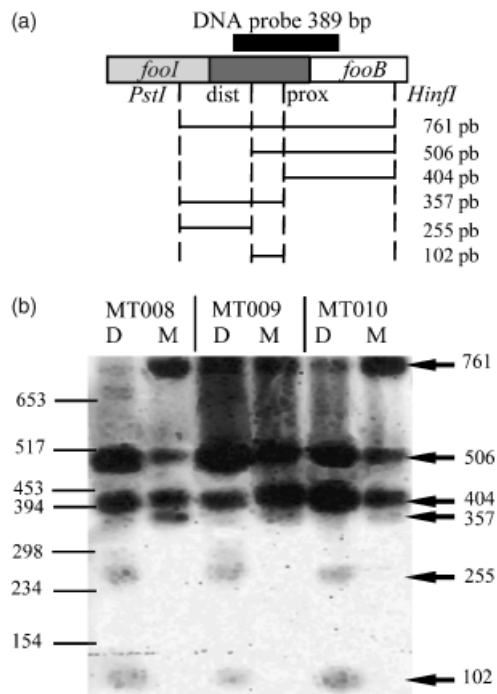


Fig. 5. Analysis of the *in vivo* GATC methylation states of the *foo* regulatory region. (a) Schematic representation of DNA fragments that may hybridize with the probe. dist corresponds to GATC_{dist} and prox to GATC_{prox}. (b) Southern blot of chromosomal DNA obtained from wild and mutant strains MT008 *fooI*⁺ *topA*⁺ *gyrB*⁺; MT009 *fooI*⁺ *topA*⁺ *gyrB*⁻; MT010 *fooI*⁺ *topA*⁻ *gyrB*⁻. The sizes of fragments of the digoxigenin-labeled DNA molecular weight marker V (Roche) are indicated on the left. Size of DNA fragments that hybridized with the probe are shown on the right. D, DpnI restriction of the PstI-Hinfl fragment. M, Mbol restriction of the PstI-Hinfl fragment.

inversion event, highlighting the importance of supercoiling in Lrp interaction with fimbrial DNA.

Fool function is sensitive to supercoiling

Fool is a positive coregulator of the *foo* operon (Berthiaume et al., 2004) that possibly acts by facilitating the interaction of Lrp with putative sites 4–6 to activate *foo* expression (Fig. 4a). To test if Fool function might be altered by changes in DNA supercoiling, a lysogen containing fusion 2 (*fooI*_P-*fooB*-*lac*) in which the coding region of *fooI*, but not its promoter, is absent (Fig. 1, fusion 2) were constructed. As shown in Fig. 3, the absence of *fooI* reduced *foo* expression threefold in the wild-type strain (Fig. 3, compare 1, MT008 Fool-positive with 4, MT067 Fool-negative, $P < 0.05$), as expected. In a strain in which global negative supercoiling is reduced by roughly 25% (*gyrB225*), the absence of Fool reduced by at most twofold *foo* expression (Fig. 3, compare 2, MT009 Fool-positive with 5, MT068 Fool-negative, $P < 0.05$) whereas it had no effect in the absence of DNA topoisomerase I (Fig. 3, compare lane 3, MT010 Fool-

positive with 6, MT069 Fool-negative, $P < 0.05$). Overall, these results suggest that DNA topology can impair Fool function.

Strong divergent transcription can fully compensate the absence of Fool in a wild-type strain

Because a mutation in *gyrB* considerably reduces *foo* expression, negative supercoiling must be sufficiently high for the correct expression of *foo*. Transcription-induced negative supercoiling has been shown to affect the promoter activity even in *topA*⁺ backgrounds. This phenomenon could possibly influence *foo* operon expression in wild-type cells as *fooI* and *fooB* promoters are divergently transcribed in the *foo* regulatory region. In addition, the fact that relative β -galactosidase production is similar in FoolI positive and negative *topA* null mutants, [Fig. 3, compare fusion 1, 3 (FoolI positive) and fusion 2, 6 (FoolI negative), $P < 0.05$] may suggest that *fooI* transcription-induced negative supercoiling, might contribute to *foo* expression in these mutants. To test this hypothesis, two additional lysogens [Fig. 1, fusions 3 (*tetA*-*trcP*-*fooB*-*lac*) and 4 (*trcP*-*fooB*-*lac*)] were constructed in which the coding region of *fooI* is absent and its promoter was substituted for the very strong *P_{trc}* promoter. In addition, in fusion 3 (*tetA*-*trcP*-*fooB*-*lac*), the *tetA* gene was placed under the control of *P_{trc}*. Transcription of this gene has been shown to generate very high levels of negative supercoiling because of the anchorage of the nascent TetA protein to the membrane that completely blocks the rotation of the transcribing RNA polymerase (Lynch & Wang, 1993). Transcription from *P_{trc}* is therefore expected to generate higher negative supercoiling in fusion 3 than in fusion 4 that carries only *P_{trc}*.

As observed in Fig. 3, divergent transcription of *tetA* from *P_{trc}* completely restored *foo* expression to its normal level in the wild-type strain (Fig. 3, compare lane 7, fusion 3 with lane 1, fusion 1). This result strongly suggests that transcription-induced negative supercoiling can fully compensate for the absence of Fool in wild-type strain (Fig. 3, lane 4, fusion 2). This is also supported by the fact that fusion 4, in which *P_{trc}* alone is expected to generate less negative supercoiling, did not activate *foo* expression to wild-type levels (Fig. 3, compare lane 7, fusion 3 and lane 10, fusion 4). Also predicted, if divergent transcription activates *foo* expression via transcription-induced negative supercoiling, is the fact that the stimulation of *foo* expression of fusion containing *P_{trc}*-*tetA* was higher than that of fusion *fooIB* in *gyrB225* strain in which global negative supercoiling is lower than in the wild-type strain (Fig. 3, compare lane 2, fusion 1 with lane 8, fusion 3; more than twofold stimulation). Transcription-induced negative supercoiling is expected to have a larger impact in the absence of DNA topoisomerase I,

Table 2. Methylation patterns of the *foo* regulatory region based on the results obtained by Southern analysis

Strain genotype	Methylation pattern	DNA molecules (%)	[Lac+] cells (%)
MT008 <i>topA</i> ⁺ , <i>gyrB</i> ⁺	—••	17	96
	—	5	
	•—	44	
	—•	31	
MT009 <i>topA</i> ⁺ , <i>gyrB</i> –	—••	20	70
	—	16	
	•—	33	
	—•	31	
MT010 <i>topA</i> ⁻ , <i>gyrB</i> –	—••	10	93
	—	19	
	•—	32	
	—•	39	

Dots indicate methylation of GATC sites (GATC_{dist} left dot; GATC_{prox} right dot).

because relaxation of such supercoiling is the major function of this enzyme (Masse & Drolet, 1999). This hypothesis is supported by the fact that in the $\Delta topA$ null mutant compared with the isogenic *topA*⁺ strain, the stimulation of *foo* expression by divergent transcription from *P_{trc}* was higher than from *P_{fool}* in fusion *foolB* (for example in Fig. 3, compare lane 3, fusion 1 with lane 12, fusion 4; there is almost fivefold stimulation). However, as opposed to the *topA*⁺ strains, divergent transcription from *P_{trc}* alone was slightly better than divergent transcription from *P_{trc}-tetA* in activating *foo* expression in the $\Delta topA$ null strain, suggesting that too much negative supercoiling can be detrimental to *foo* expression.

One interesting finding is that strong divergent transcription can bypass the *FooI* requirement for *foo* operon expression. This suggests that above a threshold level, transcription-induced negative supercoiling can directly activate the *foo* promoter. There is at least another precedent for promoter activation by transcription-induced negative supercoiling from the strong *P_{trc}* promoter. Indeed, it was shown that divergent transcription from *P_{trc}* in a *topA* null mutant can activate the supercoiling-sensitive *leu-500* promoter (Chen *et al.*, 1993). Such activation can take place in a *topA*⁺ background under specific growth conditions (Fang & Wu, 1998). Whether transcription from the divergent *fool* promoter can stimulate *foo* operon expression in its natural context remains unknown.

Both global and local supercoiling can significantly affect *fooB* expression. This is based on the fact that in the intergenic region within a DNA sequence where Lrp interacts, DNA methylation is reduced in the absence of DNA topoisomerase I. Furthermore, divergent transcription from a strong promoter enhanced *fooB* expression in the *topA* null mutant. The F165₁ fimbriae operon offers a good system to

study the interrelationship between both global and local supercoiling and the control of gene expression by global regulators.

Acknowledgements

Judith A. Kashul is thanked for revising the manuscript. Guy Beauchamp is also thanked for his help in statistical analysis. This work was supported by the Natural Sciences Engineering Research Council of Canada (NSERC) grants 225155 to M.D. and J.H.; 25120 to J.H. and by a NATO grant to C.M. M.-C.T. was funded by a scholarship from NSERC.

References

- Berthiaume F, Crost C, Labrie V, Martin C, Newman EB & Harel J (2004) Influence of L-leucine and L-alanine on Lrp regulation of *foo*, coding for F165₁, a Pap homologue. *J Bacteriol* **186**: 8537–8541.
- Braaten BA, Blyn LB, Skinner BS & Low DA (1991) Evidence for a methylation-blocking factor (*mbf*) locus involved in *pap* pilus expression and phase variation in *Escherichia coli*. *J Bacteriol* **173**: 1789–1800.
- Chen D, Bowater RP & Lilley DM (1993) Activation of the *leu-500* promoter: a topological domain generated by divergent transcription in a plasmid. *Biochemistry* **32**: 13162–13170.
- Crost C, Garrivier A, Harel J & Martin C (2003) Leucine-responsive regulatory protein-mediated repression of *clp* (encoding CS31A) expression by L-leucine and L-alanine in *Escherichia coli*. *J Bacteriol* **185**: 1886–1894.
- Crost C, Harel J, Berthiaume F, Garrivier A, Tessier MC, Rakotoarivonina H & Martin C (2004) Influence of environmental cues on transcriptional regulation of *foo* and *clp* coding for F165(1) and CS31A adhesins in *Escherichia coli*. *Res Microbiol* **155**: 475–482.
- Daigle F, Forget C, Martin C, Drolet M, Tessier MC, Dezfulian H & Harel J (2000) Effects of global regulatory proteins and environmental conditions on fimbrial gene expression of F165(1) and F165(2) produced by *Escherichia coli* causing septicaemia in pigs. *Res Microbiol* **151**: 563–574.
- DiNardo S, Voelkel KA, Sternglanz R, Reynolds AE & Wright A (1982) *Escherichia coli* DNA topoisomerase I mutants have compensatory mutations in DNA gyrase genes. *Cell* **31**: 43–51.
- Drlica K (1992) Control of bacterial DNA supercoiling. *Mol Microbiol* **6**: 425–433.
- Fang M & Wu HY (1998) Suppression of *leu-500* mutation in *topA*⁺ *Salmonella typhimurium* strains. The promoter relay at work. *J Biol Chem* **273**: 29929–29934.
- Harel J, Daigle F, Forget C, Tessier MC, Crost C & Martin C (2000) Phase variation of F165(1) (Prs-like) fimbriae from *Escherichia coli* causing septicaemia in animals. *Can J Microbiol* **46**: 1101–1107.

- Hernday A, Krabbe M, Braaten B & Low D (2002) Self-perpetuating epigenetic pili switches in bacteria. *Proc Natl Acad Sci USA* **99**: (suppl 4): 16470–16476.
- Hraiky C, Raymond MA & Drolet M (2000) RNAse H overproduction corrects a defect at the level of transcription elongation during rRNA synthesis in the absence of DNA topoisomerase I in *Escherichia coli*. *J Biol Chem* **275**: 11257–11263.
- Kelly A, Conway C, O Croinin T, Smith SG & Dorman CJ (2006) DNA supercoiling and the Lrp protein determine the directionality of *fim* switch DNA inversion in *Escherichia coli* K-12. *J Bacteriol* **188**: 5356–5363.
- Lahooti M, Roesch PL & Blomfield IC (2005) Modulation of the sensitivity of FimB recombination to branched-chain amino acids and alanine in *Escherichia coli* K-12. *J Bacteriol* **187**: 6273–6280.
- Liu LF & Wang JC (1987) Supercoiling of the DNA template during transcription. *Proc Natl Acad Sci USA* **84**: 7024–7027.
- Lynch AS & Wang JC (1993) Anchoring of DNA to the bacterial cytoplasmic membrane through cotranscriptional synthesis of polypeptides encoding membrane proteins or proteins for export: a mechanism of plasmid hypernegative supercoiling in mutants deficient in DNA topoisomerase I. *J Bacteriol* **175**: 1645–1655.
- Masse E & Drolet M (1999) Relaxation of transcription-induced negative supercoiling is an essential function of *Escherichia coli* DNA topoisomerase I. *J Biol Chem* **274**: 16654–16658.
- Masse E, Phoenix P & Drolet M (1997) DNA topoisomerases regulate R-loop formation during transcription of the *rrnB* operon in *Escherichia coli*. *J Biol Chem* **272**: 12816–12823.
- Miller JH (1992) *A Short Course in Bacterial Genetics*. Cold Spring Harbor, NY.
- Ngeleka M, Harel J, Jacques M & Fairbrother JM (1992) Characterization of a polysaccharide capsular antigen of septicemic *Escherichia coli* O115:K “V165”: F165 and evaluation of its role in pathogenicity. *Infect Immun* **60**: 5048–5056.
- Ngeleka M, Jacques M, Martineau-Doize B, Daigle F, Harel J & Fairbrother JM (1993) Pathogenicity of an *Escherichia coli* O115:K “V165” mutant negative for F165(1) fimbriae in septicemia of gnotobiotic pigs. *Infect Immun* **61**: 836–843.
- Opel ML, Arfin SM & Hatfield GW (2001) The effects of DNA supercoiling on the expression of operons of the *ilv* regulon of *Escherichia coli* suggest a physiological rationale for divergently transcribed operons. *Mol Microbiol* **39**: 1109–1115.
- Powell BS, Rivas MP, Court DL, Nakamura Y & Turnbough CL Jr (1994) Rapid confirmation of single copy lambda prophage integration by PCR. *Nucleic Acids Res* **22**: 5765–5766.
- Simons RW, Houman F & Kleckner N (1987) Improved single and multicopy *lac*-based cloning vectors for protein and operon fusions. *Gene* **53**: 85–96.
- Stupina VA & Wang JC (2005) Viability of *Escherichia coli topA* mutants lacking DNA topoisomerase I. *J Biol Chem* **280**: 355–360.
- Wassarman KM & Storz G (2000) 6S RNA regulates *E. coli* RNA polymerase activity. *Cell* **101**: 613–623.