N°d'ordre NNT : 2016LYSE1201

THESE de DOCTORAT DE L'UNIVERSITE DE LYON
opérée au sein de
l'Université Claude Bernard Lyon 1

Ecole Doctorale N° 341
Evolution, Ecosystèmes, Microbiologie, Modélisation

Spécialité de doctorat : Microbiologie

Soutenue publiquement le 03/11/2016, par :
Alexandre André Raymond Duprey

Régulation de la transcription des gènes de virulence bactériens : dynamique des complexes nucléoprotéïques

Devant le jury composé de :

Wisniewski-Dyé, Florence, Professeur Université Claude Bernard Lyon 1	Présidente
Matic, Ivan, Directeur de recherche CNRS	Rapporteur
Geiselmann, Hans, Professeur Université Joseph Fourier Grenoble	Rapporteur
Simonet, Pascal, Directeur de recherche CNRS	Examinateur
Nasser, William, Directeur de recherche CNRS	Directeur de thèse
Reverchon, Sylvie, Maître de conférences INSA Lyon	Co-directrice de thèse

Résumé

Les bactéries sont en permanence confrontées à des changements d'environnements. La régulation transcriptionnelle joue alors un rôle majeur dans l'adaptation des bactéries. En particulier, la bactérie phytopathogène *D. dadantii* s'est récemment adaptée à l'hôte végétal. Elle produit en particulier des pectate lyases (Pel) qui dégradent la pectine, ciment des parois végétales, et jouent un rôle majeur dans le développement de la maladie. Les gènes *pelD* et *pelE*, malgré la forte divergence dans leur expression, sont issus d'un transfert horizontal suivi d'une duplication récente. La question de l'intégration de ces gènes avec les régulations préexistantes s'est alors posée.

Dans un premier temps, les mécanismes moléculaires détaillés de la régulation de *pelD* ont été étudiés. Il a été montré que cette régulation s'appuie sur un promoteur divergent de forte affinité pour l'ARN polymérase mais de faible efficacité pour la transcription et sur un arrangement stratégique de quatre sites de fixation de répresseur FIS et deux sites de l'activateur CRP. Tous ces éléments interagissent entre eux pour produire une régulation fine de l'expression de *pelD*. L'origine de la divergence régulatrice entre les paralogues *pelD* et *pelE* a par la suite été explorée. De manière surprenante, la divergence entre ces deux gènes et leur sélection s'appuie presque exclusivement sur un décalage de la position du promoteur de *pelE* (« TSS turnover ») qui l'a transformé en initiateur de la dégradation de la pectine. Ce mécanisme très fréquent chez les eucaryotes pluricellulaires (homme, drosophile, souris…) n'avait jamais encore été décrit chez les bactéries.

A travers l'étude des promoteurs *pelD* et *pelE* de *D. dadantii*, de nouveaux mécanismes renforçant l'importance de la régulation transcriptionnelle dans les processus adaptatifs ont ainsi été découverts.

Mots clés

Pectate lyase, *Dickeya*, FIS, CRP, *pel*, régulation transcriptionnelle

Laboratoire

UMR5240 Microbiologie Adaptation Pathogénie, Université Lyon 1
Campus de la Doua
Bâtiment Lwoff - 10 rue Dubois
69622 Villeurbanne cedex

Abstract

Bacteria face frequent environmental changes. Transcriptional regulation plays a major role in the adaptation to these changes. In particular, the phytopathogen bacteria *Dickeya* have recently adapted to vegetal hosts. They produce Pecate lyases (Pel), among others, to degrade pectin in plant cell walls, which is necessary for disease development. The *pelD* and *pelE* genes, despite the strong divergence in their expression, originate from a horizontal gene transfer followed by a recent duplication. This raises the question of their integration into the preexisting regulatory networks.

Detailed molecular mechanisms of the transcriptional regulation of *pelD* were studied first. It was shown that this regulation relies on a high-affinity but low transcription efficiency divergent promoter and a strategic arrangement of four FIS repressor binding sites and two CRP activator binding sites. These elements interact together to fine-tune the expression of *pelD*. Next, the origin of the regulatory divergence between the paralogous genes *pelD* and *pelE* was explored. Surprisingly, their divergence and selection relies mostly on a TSS turnover which happened on the *pelE* regulatory region and transformed *pelE* into an initiator of pectin degradation. This widespread phenomenon in multicellular eukaryotes (human, fly, mouse…) had not yet been seen in bacteria.

To conclude, through the study of *D. dadantii pelD* and *pelE* promoters, new mechanisms highlighting the relevance of transcriptional regulation in adaptation were discovered in this work.

Keywords

Pectate lyase, *Dickeya*, FIS, CRP, *pel*, transcriptional regulation, TSS turnover

Remerciements

Je voudrais d'abord remercier les membres du jury, Hans Geiselman, Ivan Matic, Pascal Simonet et Florence Wisniewski-Dyé, pour avoir accepté d'évaluer ce travail.

Je tiens à remercier en particulier mes encadrants, William et Sylvie, pour leurs conseils et leurs directions tout en me laissant la liberté d'explorer les opportunités inattendues qui se sont présentées.

Une grande partie de ce travail n'aurait pas été possible sans nos collaborateurs : Céline Brochier pour la phylogénie, Georgi Muskhelishvili pour la mécanistique de l'action de FIS et Cendrine Moskalenko pour la microscopie à force atomique.

Je tiens à remercier spécialement les stagiaires qui ont contribué aux résultats présentés dans la thèse : Méganne qui a permis d'initier les projets OmpR et évolution et qui a construit l'hybride *pelD-pelE* sur lequel tant de choses reposent, Tiffany pour de nombreuses analyses sur le génome de *D. aquatica* et Guilhem pour m'avoir perdu deux plasmides et avoir tenté d'être plus têtu qu'OmpR (c'est OmpR qui a gagné).

Je remercie le personnel technique de l'unité pour leur soutien dans ce travail, Véronique et Jean-Michel pour la préparation des milieux et du matériel et Julien pour la purification de la majorité des protéines utilisées dans cette thèse. Je remercie les doctorants et chercheurs du laboratoire pour les discussions et leur soutien, en particulier Jiao pour l'analyse des données de RNAseq et de puces, Nicolas pour m'avoir appris R et Simon pour les analyses génomiques.

Je remercie spécialement le Tecan du 2ème étage pour toutes les données qu'il m'a fournies ainsi que, bien sûr, ses propriétaires (Agnès, Kévin et Kévin) pour avoir accepté de me laisser l'utiliser.

Enfin, j'aimerais remercier les différentes générations de stagiaires qui se sont succédés dans l'équipe (Rodrigue, Joanna, Romain, Marine, Clément, Pauline, Justine et probablement beaucoup d'autres dont j'ai oublié les noms) ainsi que les étudiants que j'ai eu en cours pour leurs perles dans les copies qui ont rendu les corrections moins tristes (« Albert Einstein mort au printemps dernier »…).

UNIVERSITE CLAUDE BERNARD - LYON 1

Président de l'Université	**M. le Professeur Frédéric FLEURY**
Président du Conseil Académique	M. le Professeur Hamda BEN HADID
Vice-président du Conseil d'Administration	M. le Professeur Didier REVEL
Vice-président du Conseil Formation et Vie Universitaire	M. le Professeur Philippe CHEVALIER
Vice-président de la Commission Recherche	M. Fabrice VALLÉE
Directeur Général des Services	M. Alain HELLEU

COMPOSANTES SANTE

Faculté de Médecine Lyon Est – Claude Bernard	Directeur : M. le Professeur J. ETIENNE
Faculté de Médecine et de Maïeutique Lyon Sud – Charles Mérieux	Directeur : Mme la Professeure C. BURILLON
Faculté d'Odontologie	Directeur : M. le Professeur D. BOURGEOIS
Institut des Sciences Pharmaceutiques et Biologiques	Directeur : Mme la Professeure C. VINCIGUERRA
Institut des Sciences et Techniques de la Réadaptation	Directeur : M. le Professeur Y. MATILLON
Département de formation et Centre de Recherche en Biologie Humaine	Directeur : Mme la Professeure A-M. SCHOTT

COMPOSANTES ET DEPARTEMENTS DE SCIENCES ET TECHNOLOGIE

Faculté des Sciences et Technologies	Directeur : M. F. DE MARCHI
Département Biologie	Directeur : M. le Professeur F. THEVENARD
Département Chimie Biochimie	Directeur : Mme C. FELIX
Département GEP	Directeur : M. Hassan HAMMOURI
Département Informatique	Directeur : M. le Professeur S. AKKOUCHE
Département Mathématiques	Directeur : M. le Professeur G. TOMANOV
Département Mécanique	Directeur : M. le Professeur H. BEN HADID
Département Physique	Directeur : M. le Professeur J-C PLENET
UFR Sciences et Techniques des Activités Physiques et Sportives	Directeur : M. Y.VANPOULLE
Observatoire des Sciences de l'Univers de Lyon	Directeur : M. B. GUIDERDONI
Polytech Lyon	Directeur : M. le Professeur E.PERRIN
Ecole Supérieure de Chimie Physique Electronique	Directeur : M. G. PIGNAULT
Institut Universitaire de Technologie de Lyon 1	Directeur : M. le Professeur C. VITON
Ecole Supérieure du Professorat et de l'Education	Directeur : M. le Professeur A. MOUGNIOTTE
Institut de Science Financière et d'Assurances	Directeur : M. N. LEBOISNE

Table des matières

Résumé **2**

Mots clés **2**

Laboratoire **2**

Abstract **3**

Keywords **3**

Remerciements **4**

Liste des figures **8**

Liste des tables **11**

Liste des abréviations **12**

I. Introduction bibliographique **13**

A. La régulation transcriptionnelle bactérienne : un système complexe et flexible 13

1. Les différents niveaux de la régulation de la synthèse des protéines chez les bactéries 13

2. La structure d'un promoteur bactérien 15

3. Les acteurs de la transcription 15

a. L'ARN polymérase bactérienne 15

b. Les facteurs σ 16

c. Les petits ligands 18

d. Les facteurs de transcription et leurs cofacteurs 19

e. Le rôle de la topologie de l'ADN 21

f. Les protéines associées au nucléoïde 24

4. Interactions entre des promoteurs proches 27

5. Evolution des réseaux de régulation transcriptionnelle 31

B. Les entérobactéries phytopathogènes : un mélange unique de gènes adaptés aux hôtes animaux et végétaux 36

1. Taxonomie et évolution 36

2. Les facteurs d'agression de *D. dadantii* 40

3. Les facteurs associés à la virulence de *D. dadantii* 41

C. Les régulateurs principaux de la virulence de *Dickeya dadantii* : fonctions et évolution 42

1. KdgR 44

2. CRP 44

3. FIS (1er article) 47

4. PecS 59
5. PecT 60
6. Le défi de *D. dadantii* : intégrer l'action des régulateurs anciens et nouveaux sur des gènes d'origine variée 61

II. Matériel et méthodes non publiés 61
A. Construction des plasmides supplémentaires 61
B. Surproduction et purification d'OmpR de *D. dadantii* 63
C. Retard de migration sur gel 63
D. Prédiction de la structure de la protéine OmpR de *D. dadantii* 63
E. Prédiction des sites de fixation OmpR dans le génome de *D. dadantii* 63
F. Tests phénotypiques chez *Dickeya* 63
1. Milieux solides 63
2. Milieux liquides 64
a. Activité Pel 64
b. Biofilms 64
3. Infections de plantes 64
G. Séquençage du génome de *D. aquatica* 64
H. Phylogénie des espèces *Dickeya*, *Serratia* et *Pectobacterium* 65

III. Résultats 65
A. *pelD* : une régulation transcriptionnelle complexe et optimisée (2ème article) 65
B. La nature de la connexion entre *pelD* et *div* : une proximité tridimensionnelle ? 91
C. Emergence et évolution des gènes *pelD* et *pelE* (3ème article) 93
D. L'acquisition de la sensibilité aux conditions environnementales de *pelE* (résultats non publiés) 115
E. *paeY* et *pemA* : une association récente avec *pelA/pelE/pelD* (résultats non publiés) 124
F. Deux nouveaux *Dickeya* pour explorer l'acquisition des gènes *pel* : *D. aquatica* et ATCC 39006 (résultats non publiés) 128

IV. Discussion générale 134

V. Références 136

Liste des figures

Figure 1 : Les différents niveaux de régulation génétique chez les bactéries. 14
Figure 2 : Structure typique d'un promoteur bactérien. 15
Figure 3 : Structure de l'ARN polymérase bactérienne holoenzyme en interaction avec l'ADN. ... 16
Figure 4 : Fonctionnement des anti-σ FlgM et RseA. 18
Figure 5 : Mise en place de la réponse stringente par RelA chez *E. coli.* 19
Figure 6 : Présentation schématique de la hiérarchie du réseau de régulation génétique chez *E. coli*. 20
Figure 7 : les principaux modes d'action des facteurs de transcription 21
Figure 8 : Interconversion entre torsion et vrillage. 22
Figure 9 : Mécanisme de suppression d'un supertour négatif par la topoisomérase I bactérienne. 23
Figure 10 : Modifications de surenroulement catalysées par différentes topoisomérases. 23
Figure 11 : régulation homéostatique du niveau de surenroulement bactérien par la topoisomérase I et la gyrase. 24
Figure 12 : Introduction de courbures dans l'ADN par HU 25
Figure 13 : Emballage de l'ADN par un octamère de LRP 26
Figure 14 : Pontage de l'ADN par H-NS 26
Figure 15 : Intégration des NAPs dans le contrôle de la topologie de l'ADN. 27
Figure 16 : Régulation transcriptionnelle du promoteur *nir* de *E. coli* par les NAPs 27
Figure 17 : Modulation de la transcription de *crp* par un promoteur divergent. 28
Figure 18 : Interférence par gêne stérique de deux molécules d'ARN polymérase fixées sur les promoteurs p_R et p_{RM} du phage λ. 28
Figure 19 : Interférences à distance de deux promoteurs convergents 29
Figure 20 : Modifications locales du surenroulement induites par la transcription et la réplication 30
Figure 21 : Structure de la région régulatrice de *fis* de *E. coli* 30
Figure 22 : Modification des réseaux de régulation lors d'événements de réduction de génomes 31
Figure 23 : Comparaison des réseaux de régulation associés au système à deux composants PhoP/PhoQ chez *Y. pestis* et *S. enterica*. 32
Figure 24 : Mécanisme de la répression xénogénique. 33
Figure 25 : Mécanismes de divergence de la régulation des gènes paralogues chez les bactéries. 35
Figure 26 : Place des phytopathogènes dans la phylogénie des entérobactéries. 37
Figure 27 : Phylogénie multi-locus détaillée des entérobactéries phytopathogènes. 38
Figure 28 : Evolution de la taxonomie des *Dickeya*. 39
Figure 29 : Dégradation extracellulaire et intracellulaire de la pectine. 41
Figure 30 : Présence et conservation des régulateurs principaux de la virulence de D. dadantii chez d'autres espèces bactériennes. 43
Figure 31 : Cascade de production de l'AMPc chez *E. coli*. 45
Figure 32 : Mécanismes régulateurs connus de CRP et KdgR sur les régions régulatrices *pelD* et *pelE*. 46
Figure 33: Fis coordinates the transition from the asymptomatic to the symptomatic phase in *D. dadantii*. 53

Figure 34: Control of *S. enterica* virulence by the DNA supercoiling state and Fis. 56
Figure 35: Comparative diagram of the functions regulated by Fis in *S. enterica* and *D. dadantii*. 57
Figure 36 : Alignement des régions régulatrices de *fis* chez *E. coli*, *S. enterica* et *D. dadantii* 58
Figure 37: Structure of the pelD regulatory region. 73
Figure 38: DNase I footprinting of FIS binding in the *pelD* regulatory region. 74
Figure 39: Comparison between the Gfp and Luc reporter systems in *D. dadantii* and the GUS system in *E. coli*. 75
Figure 40: Transcriptional activity of various *pelD* promoter constructs in *D. dadantii* wild type and *fis* mutant backgrounds. 76
Figure 41: Effect of FIS binding on *pelD* transcription in vitro. 77
Figure 42: Modification of the CRP2 binding by FIS. 78
Figure 43: Absence of modification of the CRP1 DNase I footprint by the presence of FIS. 79
Figure 44: Transcriptional regulation of the *div* promoter. 80
Figure 45: *in vivo* potassium permanganate footprinting of the *pelD* regulatory region. 81
Figure 46: Individual contributions of the CRP1 and CRP2 sites on *pelD* and *div* expression 82
Figure 47: Preferential binding of RNA polymerase at *div* does not correlate with stronger transcription. 83
Figure 48: Impact of *div* on *pelD* transcription. 84
Figure 49: Proposed model for the temporal control of *pelD* by FIS, CRP and the divergent promoter. 87
Figure 50: Predicted intrinsic and CRP-induced bending in the *pelD* regulatory region. 88
Figure 51: Regulatory structure of 3 different relevant promoters regulated by a divergent promoter. 89
Figure 52 : Effet des sites FIS sur la régulation de *div* 91
Figure 53 : Test de compétition entre les ARN polymérases de *pelD* et *div*. 92
Figure 54 : Effet d'insertions de 5 pb à différents endroits de la région régulatrice *pelD*. 93
Figure 55. Bayesian phylogeny of 53 PelAED proteins (317 amino acid positions). 98
Figure 56. Genomic contexts of the *pelA, pelD and pelED* genes in *Dickeya* and closely related species. 99
Figure 57. ML phylogeny of 55 *pelAED* genes (1137 nucleic acid positions). 100
Figure 58. Influence of the Pels of *D. dadantii* on *pel* gene transcription. 102
Figure 59. Transcriptional regulation of the *pelD/E* genes of *D. paradisiaca*, *D. solani*, *D. dianthicola* and *D. dadantii*. 103
Figure 60. Alignment of the regulatory regions of the *pelD/E* genes of *Dickeya* 104
Figure 61. Experimental characterization of the transcription start site of *D. paradisiaca* 105
Figure 62. Transcriptional behaviour of the hybrid *pelD-pelE* regulatory region. 107
Figure 63. Presence of TSS turnover in various bacterial species. 108
Figure 64. Model of evolution of the *pelD/E* regulatory regions in *Dickeya* 110
Figure 65 : Effet de OmpR et du relâchement de l'ADN sur l'expression de *pelD*, *pelE* et de l'hybride *pelD-pelE* (D21) 116
Figure 66 : Retard de migration sur gel d'acrylamide causé par OmpR sur les régions régulatrices de *pelE* et *pelD*. 117
Figure 67 : Empreinte à la DNase I causée par OmpR sur les régions régulatrices *pelD* et *pelE*. . 118
Figure 68 : Comparaison des structures des domaines C-terminaux d'OmpR chez *E. coli*, *S. enterica* et *D. dadantii*. 119
Figure 69 : Localisation des sites de fixation OmpR détectés par DNase I. 120
Figure 70 : Caractérisation phénotypique du mutant *ompR* 121

Figure 71 : Le rôle potentiel de OmpR et Fur dans la réponse des gènes de *D. dadantii* à l'acidité. 123
Figure 72 : Phylogénie bayésienne de 50 protéines PaeY (170 positions d'acides aminés)..... 125
Figure 73 : Phylogénie bayésienne de 63 protéines PemA (214 positions d'acides aminés) 126
Figure 74 : Cladogramme des 95 bactéries possédant au moins un gène homologue à *pelA*, *pelE*, *pelD*, *paeY* ou *pemA* présenté sur les arbres précédents. 128
Figure 75 : Photographies des plantes infectées par *D. dadantii* et *D. aquatica*..... 129
Figure 76 : Comparaison de l'activité pectate lyase de *D. dadantii* et *D. aquatica*. 129
Figure 77 : Comparaison de la résistance aux stress de *D. dadantii* et *D. aquatica* 130
Figure 78 : Quantification de la formation de biofilms chez *D. dadantii* et *D. aquatica* 131
Figure 79 : Phylogénie des Dickeya, Serratia et pectobacterium établie sur la base des séquences des gènes codant pour les protéines ribosomales concaténées (57 protéines ribosomales pour 47 espèces)..... 133
Figure 80 : Structures des régions régulatrices des gènes *pelE* et *pelD* incluant ATCC 39006 et *D. aquatica*..... 134

Liste des tables

Table 1 : Exemple de facteurs associés à la virulence chez *D. dadantii*..... 42
Table 2: Main virulence functions regulated by Fis in bacteria 50
Table 3 : Souches, plasmides et oligonucléotides non publiés 62
Table 4: Strains, plasmids and oligos used in this study 69
Table 5: Bacterial strains, plasmids and primers used in this study 113
Table 6 : Dosage β-glucuronidase de l'expression de différents gènes codant pour des facteurs de virulence ou associés à la virulence de *D. dadantii* dans une souche sauvage et un mutant *ompR* (G.Condemine communication personnelle). 115
Table 7 : Analyse fonctionnelle des cibles potentielles de OmpR sur le génome de *D. dadantii*. 121
Table 8 : Réponse des gènes du métabolisme du fer au stress acide..... 124
Table 9 : masses de pourriture induite par D. dadantii ou D. aquatica lors de l'infection de différents végétaux 128
Table 10 : Présence ou absence de gènes caractéristiques de *Dickeya* et *Serratia* dans la souche ATCC 39006 132

Liste des abréviations

ADN : Acide désoxyribonucléique
AMP : Adenosine monophosphate
AMPc : Adenosine monophosphate cyclique
ARN : Acide ribonucléique
ARNm : ARN messager
ARNP : ARN polymerase
ATP : Adenosine triphosphate
pb : Paire de base
BSA : Albumine de serum bovin
ChIP : Immunoprécipitation de la chromatine
CMC : Carboxymethylcellulose
CRP : Cyclic AMP receptor protein
CTD : Domaine C-terminal
DTT : Dithiothréitol
EPS : Exo-polysaccharides
FIS : Factor for inversion stimuation
FT : Facteur de transcription
Fur : Ferric uptake regulator
G : Glucose
h : Heure
H-NS : Heat-stable nucleoid-structuring protein
HU : Heat-unstable nucleoid-structuring protein
HTH : Hélice-tour-hélice
IHF : Integration host factor
KDG : 2-céto 3-deoxygluconate
L : litre
LB : Luria-Bertani
Lk : Nombre de liaison (linking number)
LRP : Leucine-responsive regulatory protein
M : Moles/litre
ME : Membrane externe
MI : Membrane interne
NAP : Protéine associée au nucléoide
NP40 : Nonidet P40
Pel : Pectate lyase
PGA : Acide polygalacturonique
ppGpp : Guanosine tetraphosphate
pppGpp : Guanosine pentaphosphate
RBS : Ribosome binding site
S : Saccharose
SPI : Ilot de pathogénie de *Salmonella*
Tw : Torsion (Twist)
Wr : Vrillage (Writhe)

I. Introduction bibliographique

A. La régulation transcriptionnelle bactérienne : un système complexe et flexible

1. Les différents niveaux de la régulation de la synthèse des protéines chez les bactéries

Les bactéries sont régulièrement confrontées à des changements de milieu, souvent brutaux. A l'inverse des cellules de nombreux eucaryotes pluricellulaires, les bactéries se trouvent directement confrontées aux changements physico-chimiques de leur environnement. La bactérie *Escherichia coli* par exemple alterne entre une vie dans un milieu à 37°C riche en nutriments à l'intérieur du mammifère hôte à un milieu tempéré pauvre en nutriments à l'extérieur. De plus, les bactéries peuvent alterner entre des styles de vie indépendants ou symbiotiques, c'est-à-dire associées avec un hôte, de manière parasitique ou mutualiste. Elles peuvent se trouver confrontées à d'autres micro-organismes et avoir à interagir avec eux de manière coopérative ou compétitive. Ces différents processus d'acclimatation impliquent souvent une modulation de l'expression génique. Les bactéries portent ainsi dans leur chromosome les gènes codant pour les protéines nécessaires à la survie dans ces environnements variés.

Les bactéries possèdent un premier ensemble de protéines dites « de ménage », qui assurent les fonctions vitales toujours requises quelles que soient les conditions de vie, telles que les acteurs de la réplication et de la réparation de l'ADN, de la synthèse d'ARN. Il existe également un second ensemble de protéines, bien plus nombreuses que les protéines de ménage (pour un génome « moyen », on estime qu'il y a environ 200 gènes de ménage sur 5000 gènes totaux [1]), nécessaires dans des conditions particulières. Une bactérie du sol peut par exemple posséder des protéines de biosynthèse d'hormones végétales qui ne sont utiles que lorsque la bactérie est associée à une plante [2]. Par souci d'économie d'énergie cellulaire, et, pour certaines bactéries mutualistes ou pathogènes qui sont confrontées aux défenses de l'hôte, de survie, il est donc nécessaire d'exprimer ces protéines accessoires au bon moment et dans les conditions adéquates. La régulation de la synthèse de protéines s'avère donc une nécessité.

Chez les bactéries, la synthèse de protéines s'effectue en 2 grandes étapes : transcription et traduction. L'information génétique contenue dans l'ADN est d'abord transcrite par l'ARN polymérase en ARN messager (ARNm), puis l'ARNm est traduit en protéine par les ribosomes. La régulation de la synthèse de protéines peut ainsi se faire à différents moments (Figure 1).

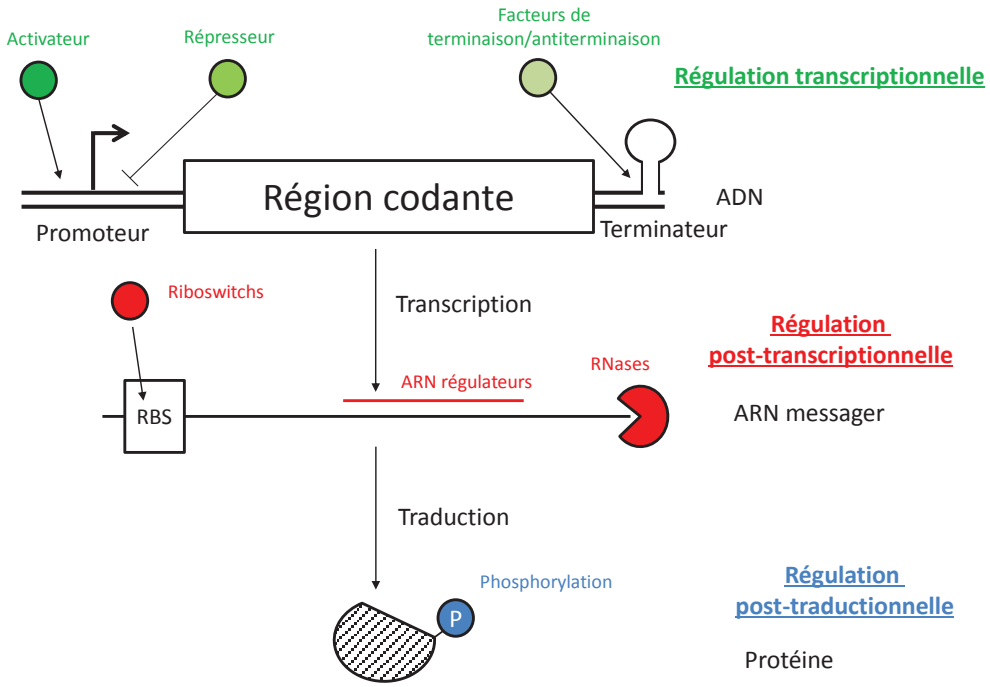


Figure 1 : Les différents niveaux de régulation génétique chez les bactéries.
Les acteurs de la régulation les plus fréquemment retrouvés sont représentés par des symboles colorés, chaque couleur correspondant à une étape de la régulation génétique (vert pour la transcription, rouge pour la traduction et bleu pour les modifications post-traductionnelle). Au niveau de la transcription, des protéines activatrices ou répressives modulent l'initiation de la transcription. Des facteurs de terminaison et d'antiterminaison peuvent moduler l'arrêt de la transcription à la fin du gène ou prématurément. Au niveau post-transcriptionnel, des petites molécules (ions, peptides…) peuvent modifier l'accessibilité du site de fixation du ribosome (RBS), des ARN régulateurs peuvent causer une dégradation du transcrit par une RNase ou, au contraire, le stabiliser ou moduler l'efficacité de sa traduction. Au niveau post-traductionnel, les protéines peuvent subir des modifications covalentes, la plus fréquente étant la phosphorylation impliquée dans de nombreux processus de signalisation.

Le premier niveau de régulation s'effectue au niveau transcriptionnel, l'objectif étant de moduler la quantité et la temporalité de production des ARNm. Les différents moyens permettant de réaliser cette régulation seront détaillés par la suite (voir partie 3-Les acteurs de la transcription).

La quantité d'ARN ainsi produite peut alors être modulée lors de la régulation post-transcriptionnelle. Le site de fixation du ribosome (RBS) peut être occlus ou au contraire exposé sous l'action d'une protéine ou d'un ligand, changeant alors son accessibilité pour les ribosomes, on parle alors de riboswitch. De plus, des ARN régulateurs peuvent se fixer sur les ARNm et moduler leur stabilité ou leur efficacité de traduction [3]. Il s'agit d'effets subtils qui ont pour but d'affiner l'efficacité de la prise en charge des ARNm par la machinerie de traduction [3].

Enfin, les bactéries sont capables d'effectuer des modifications post-traductionnelles sur les protéines, bien qu'elles soient moins courantes et moins diverses que chez les eucaryotes [4]. La modification la plus fréquente est la phosphorylation, qui est souvent retrouvée dans les cascades de signalisation bactériennes, et qui modifie l'affinité d'un régulateur pour l'ADN ou l'activité d'une enzyme [5]. Ce genre de systèmes présente l'avantage de moduler l'action d'une protéine rapidement, de manière réversible et à un coût énergétique relativement faible.

2. La structure d'un promoteur bactérien

La grande majorité des promoteurs bactériens sont organisés autour de deux séquences principales, appelées -35 et -10, et de séquences secondaires (UP, -10 étendu, discriminateur)[6]. Cette organisation, représentée en Figure 2 est commune à de nombreuses espèces bactériennes et facteurs σ différents.

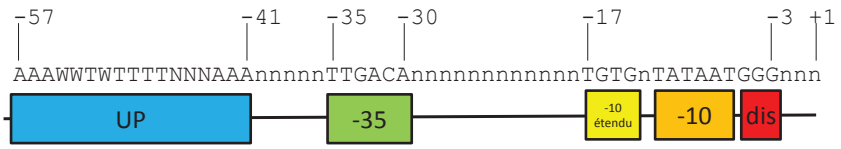


Figure 2 : Structure typique d'un promoteur bactérien.

Les différents constituants d'un promoteur bactérien usuel sont indiqués par des carrés de couleur représentés à l'échelle. Les séquences consensus pour le facteur σ^{70} de *E. coli* sont indiqués au-dessus [6]. W : A ou T. Dis : discriminateur.

La structure d'un promoteur peut être variable. D'abord, les séquences représentées sur la Figure 2 sont des consensus dégénérés. En particulier, la séquence -35 peut être sensiblement différente surtout dans le cas de promoteurs dépendant d'un activateur, au point que de nombreux promoteurs connus n'ont pas de séquence -35 identifiée. Ainsi le promoteur *pelD*, l'objet d'étude principal de la thèse, qui est activé par le régulateur du catabolisme des sucres CRP, possède une séquence -35 très dégénérée (CAAAAT). De même, l'élément UP, riche en bases A et T, favorise la transcription mais n'est pas indispensable au fonctionnement des promoteurs et peut donc être absent [7]. Le discriminateur peut potentiellement avoir n'importe quelle séquence, mais un discriminateur très riche en A/T ou en G/C rend le promoteur dépendant du (p)ppGpp (voir section 3-c-Les petits ligands). Ensuite, la distance entre les différents éléments est variable et une distance suboptimale réduit la transcription du promoteur. Par exemple, l'espacement entre le -35 et le -10 peut être de 16 à 19 pb, la distance optimale étant de 17 pb [8]. Enfin, chaque facteur σ reconnaît un ensemble de séquences particulières, qui ne sont pas toutes connues (voir section 3-b-Les facteurs σ).

3. Les acteurs de la transcription

a. L'ARN polymérase bactérienne

Chez les bactéries, la transcription des gènes est assurée par une seule enzyme, appelée ARN polymérase. Il s'agit d'un complexe formé de 6 sous-unités : αI, αII, β, β', ω, σ [9]. αI, αII, β, β'et ω sont toujours présentes dans le complexe et forment l'enzyme core, alors que le facteur σ ne s'associe que transitoirement avec l'enzyme core, formant ainsi l'holoenzyme [9,10] (Figure 3).

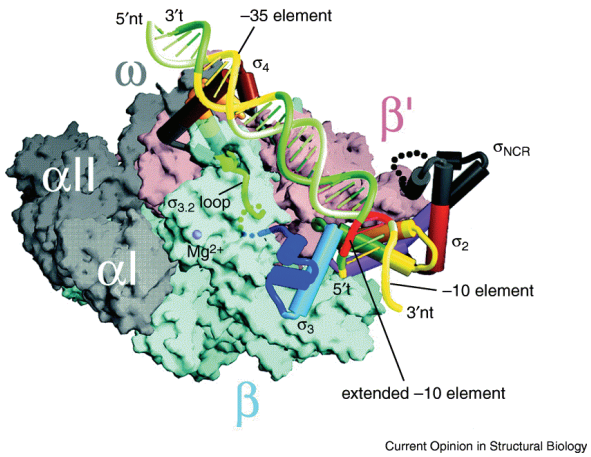


Figure 3 : Structure de l'ARN polymérase bactérienne holoenzyme en interaction avec l'ADN.
Chaque sous-unité de l'ARN polymérase est représentée par une couleur différente. Les sous-unités β et β' (cyan et rose) se placent autour de la double hélice de l'ADN et portent la sous-unité σ. σ effectue trois contacts majeurs avec l'ADN : -35 (jaune) avec σ_4 (rouge), -10 étendu (rouge) avec σ_3 (cyan) et -10 (jaune) avec σ_2 (rouge). Les autres sous-unités ne sont pas impliquées dans l'interaction avec cette région de l'ADN. Extrait de [9].

β et β' sont deux polypeptides homologues formant la structure centrale de l'ARN polymérase. Associés, ils forment une pince qui sert à guider l'ADN devant être transcrit à l'intérieur de l'ARN polymérase [9,11]. De plus, la pince et le site actif de l'enzyme ne peuvent être correctement formés qu'en présence d'un ion magnésium (Mg^{2+})[12].

αI et αII sont deux polypeptides identiques qui n'interagissent pas avec l'ADN présent à l'intérieur de l'ARN polymérase, mais qui sont essentiels au bon repliement des sous-unités β et β' [12]. Leurs domaines N-terminaux interagissent directement avec β et β', respectivement [12]. Leurs domaines C-terminaux, en revanche, n'interviennent pas dans le repliement de la protéine mais sont capables d'interagir avec des séquences d'ADN appelées UP situées en amont du site de fixation de l'ARN polymérase, permettant d'améliorer la transcription du promoteur associé [13].

ω est une petite sous-unité interagissant avec β' et qui jouerait un rôle de chaperonne pour le repliement de cette dernière [12].

σ se loge entre les deux sous-unités β et β' et, comme ces dernières, participe à l'interaction avec l'ADN [10]. σ joue un rôle clé dans l'initiation de la transcription, puisqu'il reconnaît les séquences promotrices sur l'ADN centrées sur les positions -35 et -10 par rapport à l'initiation de la transcription (Figure 2). Par la suite, la sous-unité σ se détache et l'ARN polymérase core peut initier la transcription en ouvrant la double hélice au niveau du -10 puis en démarrant la polymérisation de l'ARN.

b. Les facteurs σ

Contrairement à l'enzyme core, qui est unique, il existe plusieurs facteurs σ. Chaque facteur σ est produit dans des conditions particulières et reconnaît des séquences promotrices qui lui sont propres. Cela permet ainsi de réguler la transcription à un premier niveau : en fonction des facteurs σ disponibles dans la cellule, des ensembles de gènes différents seront transcrits.

En règle générale, les facteurs σ sont organisés en 4 domaines connectés par des régions non structurées [13], et les contacts avec l'ADN s'effectuent au niveau de plusieurs d'entre eux. Par

exemple, pour le σ^{70} de *Escherichia coli*, la séquence -35 est reconnue par le domaine 4, le -10 étendu par le domaine 3 et le -10 par le domaine 2 [10,14] (Figure 3).

Le nombre de facteurs σ varie selon les espèces, de 1 pour *Mycoplasma genitalium* à 63 pour *Streptomyces coelicolor.* La protéobactérie bactérie modèle *Escherichia coli*, possède 7 facteurs σ impliqués dans des processus divers.

σ^{70} (RpoD) est le facteur σ principal, dit « de ménage ». Il est majoritaire durant la phase exponentielle de croissance et permet une transcription efficace de la plupart des gènes requis pour une croissance rapide dans des conditions standard de *E. coli* [14].

σ^{32} (σ^{S}, RpoS) est produit lorsque la bactérie rencontre un stress environnemental (nutritionnel, oxydant, acide...), notamment lors de l'entrée en phase stationnaire de croissance (d'où le nom σ^{S}). Il permet la mise en place d'une réponse générale au stress en activant plus de 70 gènes [14].

Enfin, les 5 autres facteurs σ sont spécialisés dans une réponse spécifique à certaines conditions. σ^{32} (RpoH) répond à la présence de protéines cytoplasmiques non repliées (lors d'un choc thermique par exemple) et active la synthèse de chaperonnes. σ^{F} (FliA) active la transcription des gènes de flagelles. σ^{E} (RpoE) permet de combattre le stress membranaire et FecI met en place la réponse à la carence en fer [14]. σ^{N} (σ^{54}, RpoN) coordonne principalement le métabolisme de l'azote [15]. Il est intéressant de noter que les promoteurs reconnus par σ^{N} ont une structure unique : au lieu de séquences -35 et -10, les éléments de reconnaissance sont en -24 (TGGCACG) et -12 (TTGCW) [16].

La bactérie sur laquelle j'ai travaillé durant mon stage de master et en thèse, *Dickeya dadantii*, possède 9 facteurs σ. En plus des 7 facteurs présents chez *E. coli*, *D. dadantii* en possède deux autres : HrpL et Dda3937_02953 [17]. HrpL, en particulier, active la transcription des gènes liés au système de sécrétion de type III Hrp de cette bactérie, favorisant alors l'injection d'effecteurs dans les cellules de la plante afin de réduire les défenses de l'hôte et de contribuer à la destruction des cellules végétales [18]. Dda3937_02953 n'a pas de fonction connue. On voit ainsi l'intérêt du système modulaire des facteurs σ : ils peuvent facilement être transférés horizontalement avec les gènes qu'ils contrôlent et permettre une régulation appropriée de ces gènes dans le nouvel hôte au sein d'un îlot génomique tel que l'îlot *hrp/hrc* chez *D. dadantii* [19].

L'activité de ces facteurs σ peut elle-même être régulée par des anti- σ qui les séquestrent et les empêchent d'accéder à la core enzyme. Chez *E. coli*, par exemple, σ^{E} peut être complexé avec RseA et n'est relâché qu'en cas de stress membranaire (Figure 4).

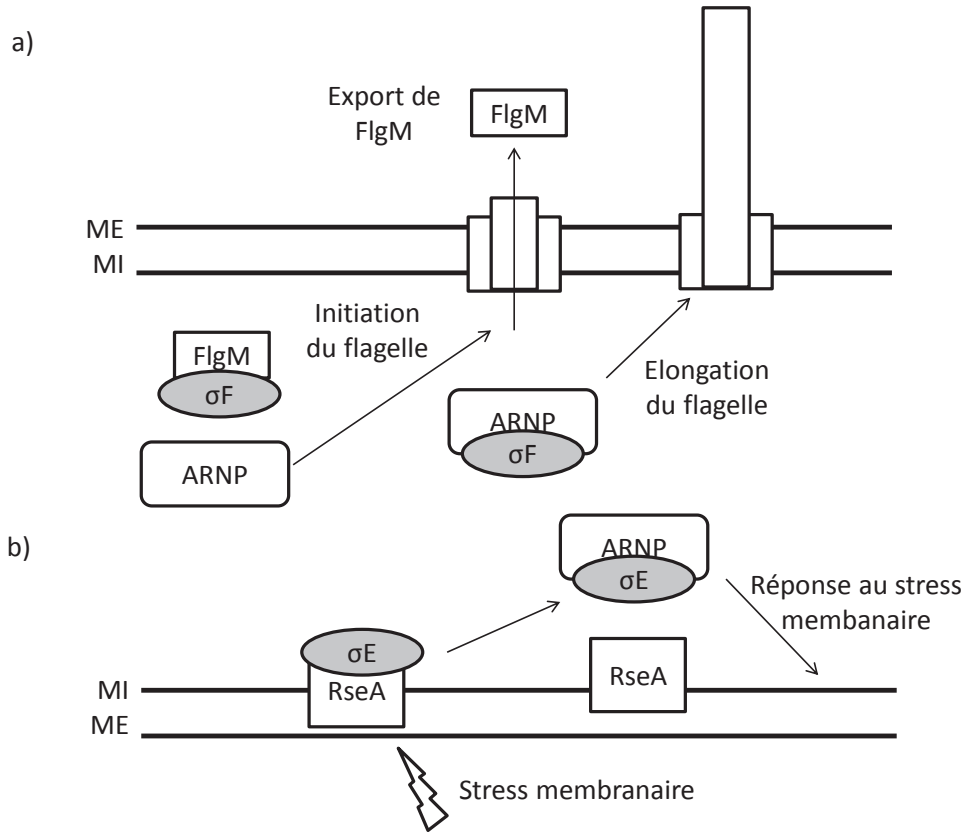


Figure 4 : Fonctionnement des anti-σ FlgM et RseA.

Le facteur sigma est représenté par un ovale gris et l'anti-sigma par un rectangle blanc. MI : Membrane interne. ME : Membrane externe. a) σ^F est complexé par FlgM. La formation du flagelle est initiée indépendamment de σ^F. Lorsque l'ancrage et le début du flagelle sont en place, FlgM est exporté, libérant σ^F. σ^F permet ensuite de produire les unités de structure du flagelle et donc son élongation. b) RseA est ancré dans la membrane et séquestre σ^E. Un stress membranaire cause une cascade de réponses encore mal comprise qui amène à la dissociation de σ^E, ce qui permet de transcrire les gènes liés à la réponse au stress membranaire [20].

c. Les petits ligands

L'activité de l'ARN polymérase peut être également modulée par la présence de petits ligands, le plus étudié et le plus répandu étant le (p)ppGpp (guanosine tetraphosphate et guanosine pentaphosphate)

Le ppGpp est synthétisé à partir de GDP et d'ATP et le pppGpp à partir de GTP et d'ATP. Cette synthèse est induite par différentes voies. L'une d'entre elles est bien conservée dans l'ensemble des bactéries et est induite lors d'une carence en acides aminés. Dans ces conditions, la présence d'ARNt non chargés en acides aminés entraine un arrêt de la traduction. Les ribosomes bloqués recrutent alors une (p)ppGpp synthase (appelée RelA chez *E. coli*) [21] (Figure 5).

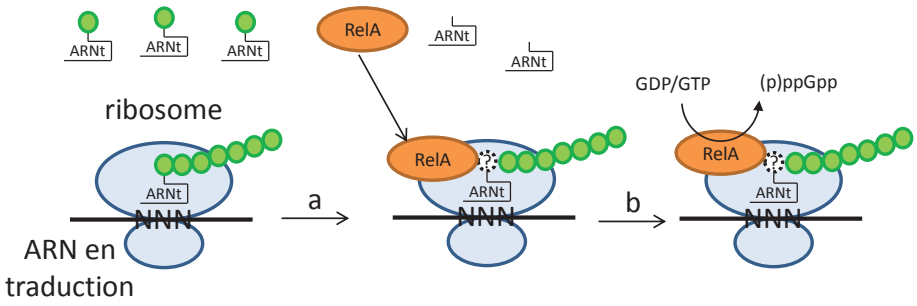


Figure 5 : Mise en place de la réponse stringente par RelA chez *E. coli.*
Le ribosome est représenté en bleu, les acides aminés en vert, l'ARNm en noir et RelA en orange. (a) La présence d'un ARNt non chargé en acide aminé provoque l'arrêt de la traduction et le recrutement de RelA. (b) RelA synthétise alors du (p)ppGpp. Adapté d'après [21].

Une deuxième voie de synthèse existe chez *E. coli* et la majorité des autres protéobactéries. D'autres carences telle que l'absence d'acides gras, de phosphate ou de fer induisent la synthèse de (p)ppGpp par SpoT [22]. Enfin, la dégradation du (p)ppGpp est assurée par SpoT chez les protéobactéries ou RelA chez les Gram + (qui ne disposent en général pas de SpoT) [22].

Au niveau transcriptionnel, le (p)ppGpp peut agir à deux niveaux. Au niveau direct, le (p)ppGpp peut déstabiliser les complexes instables ARNP-ADN (répression), ou augmenter l'efficacité des complexes stables en diminuant leur durée de vie pour favoriser la réinitiation de la transcritpion (activation) [22]. Le déterminant de l'effet activateur ou répresseur du (p)ppGpp est le discriminateur, une séquence située entre le -10 et le site d'initiation de la transcription interagissant avec le domaine 1 du facteur σ. Si elle est riche en G/C, le (p)ppGpp pourra avoir un effet répresseur, alors qu'il sera activateur avec un discriminateur riche en A/T. La protéine DskA favorise ces deux actions du (p)ppGpp [22]. Au niveau indirect, le (p)ppGpp déstabilise l'interaction enzyme core-σ^{70}, favorisant alors la fixation de facteurs σ alternatifs [23]. Les modes d'action du (p)ppGpp peuvent varier en fonction de l'espèce bactérienne considérée [22].

d. Les facteurs de transcription et leurs cofacteurs

Les facteurs de transcription (FT) sont des protéines capables d'interagir avec l'ADN et de moduler l'activité de l'ARN polymérase. Chez *E. coli,* on estime que 314 protéines (8% du génome) sont des facteurs de transcription [24]. Les FT sont hiérarchisés et peuvent eux-mêmes être régulés par des FT. En règle générale, un petit groupe de FT généraux (avec de nombreuses cibles) contrôle un grand groupe de FT spécifiques (avec peu de cibles) (voir Figure 6 pour *E. coli*).

Des tentatives ont été effectuées pour classer les FT en familles en se basant sur leurs structures, homologie et fonction. A l'origine, 13 familles ont été définies avec 37% de FT non associés à ces familles [24]. Depuis, certaines d'entre elles ont été divisées en sous-familles et de nouvelles familles ont été établies, pour un total de 58 chez la seule bactérie modèle *E. coli* [25].

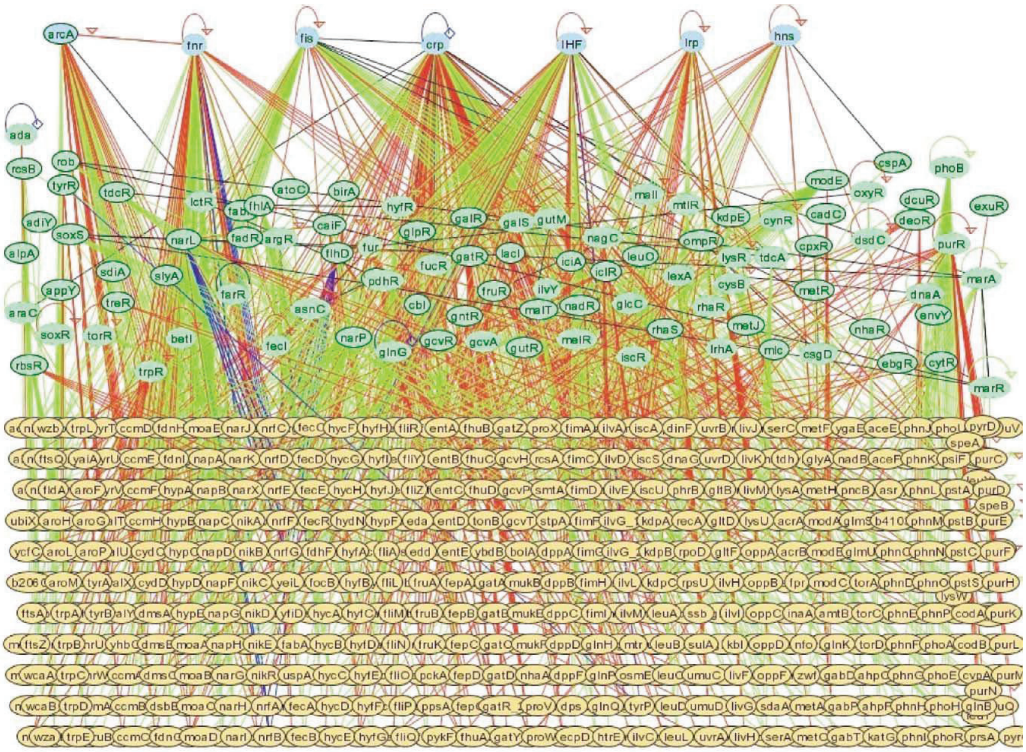


Figure 6 : Présentation schématique de la hiérarchie du réseau de régulation génétique chez *E. coli*.
Sept régulateurs généraux (CRP, FNR, IHF, FIS, ArcA, NarL et Lrp) contrôlent une majorité des effecteurs ainsi que de nombreux régulateurs spécifiques. Les lignes vertes indiquent une activation, les rouges une répression et les bleues une activité mixte [26].

Un facteur de transcription peut être activateur, répresseur ou les deux à la fois. Chez *E. coli,* ces 3 classes représentent 35%, 43% et 22% de tous les facteurs de transcription, respectivement [24]. En règle générale, un activateur se fixe d'abord sur une région régulatrice puis y favorise le recrutement de l'ARN polymérase [27] ou, plus rarement, l'ouverture de la double hélice [28]. La forme la plus courante de répression, quant à elle, est la gêne stérique : de par l'espace que le répresseur occupe sur l'ADN, il entre en compétition avec l'ARN polymérase et l'empêche de se fixer au promoteur. A nouveau, des variantes existent. Un répresseur peut ainsi inhiber l'ouverture de la double hélice, empêcher l'échappement de l'ARN polymérase ou faire barrage au passage de l'ARN polymérase en aval (Figure 7) [29].

Des mécanismes supplémentaires sont rendus possibles par l'interaction de plusieurs facteurs de transcription. Un répresseur peut agir en déplaçant un activateur et vice-versa, ou deux activateurs peuvent se fixer coopérativement sur leurs sites respectifs [27], on parle alors de co-activation. Enfin, un site de fixation peut initier une nucléation et attirer de nouvelles protéines sur la région régulatrice sans spécificité de séquence, ce qui permet à des sites situés à distance du promoteur d'avoir un impact sur la transcription (Figure 7). Fur [30] et H-NS [31], entre autres, sont capables d'utiliser ce mécanisme pour assurer une répression.

Des petits ligands peuvent moduler l'activité de ces facteurs de transcription. Ces petites molécules, de nature chimique diverse (nucléotides, peptides, phénols, métaux… [32–34]), se fixent sur un facteur de transcription et induisent des modifications conformationnelles qui favorisent ou défavorisent leur fixation sur l'ADN. L'exemple le mieux décrit est celui du répresseur LacI de *E. coli*. En l'absence d'allolactose, LacI se fixe sur la région promotrice de l'opéron *lac* et réprime les gènes associés. Cependant, en présence de lactose dans le milieu, ce dernier est transformé en allolactose qui se fixe sur LacI et diminue son affinité pour l'ADN [35].

Le domaine de fixation de l'allolactose et le domaine de fixation à l'ADN sont distincts, il ne s'agit donc pas, dans ce cas, d'une compétition directe mais d'une transition allostérique [35].

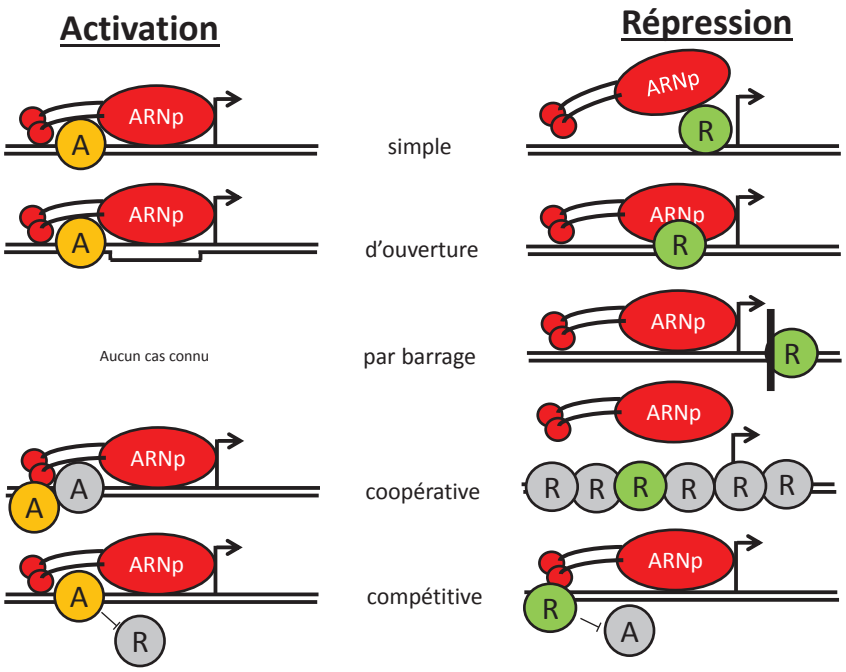


Figure 7 : les principaux modes d'action des facteurs de transcription

L'ADN est représenté en noir, l'ARN polymérase en rouge, l'activateur en orange et le répresseur en vert. Les facteurs de transcription peuvent agir directement sur l'ARN polymérase (simple), sur l'ouverture de la double hélice causée par l'ARN polymérase (d'ouverture) ou moduler l'action d'un activateur ou répresseur (coopérative et compétitive). De plus, les répresseurs peuvent aussi bloquer la progression de l'ARN polymérase (par barrage). Des activateurs ou répresseurs distincts peuvent être impliqués ensemble dans une régulation coopérative (voir par exemple FIS et IHF dans la Figure 16).

e. Le rôle de la topologie de l'ADN

Chez les bactéries, en l'absence de contraintes, la double hélice de l'ADN est une hélice droite qui effectue 1 tour tous les h_0=10,4 pb environ (la valeur exacte varie en fonction de la température, force ionique, séquence de l'ADN… [36]). Pour une molécule de longueur N (en pb) donnée, on peut alors définir le nombre de liaison (linking number) au repos Lk_0:

$$Lk_0 = \frac{N}{h_0}$$

Dans cet état, l'ADN est relâché. Si l'on prend un ADN linéaire que l'on circularise sans appliquer de contrainte, il prendra alors cet état relâché. En revanche, si l'on applique une torsion avant de le circulariser, l'ADN va devenir surenroulé positivement si la torsion est appliquée dans le même sens que l'hélice (sens droit) et négativement si elle est opposée à l'hélice (sens gauche) [37] (Figure 8 a-b). Il n'est ainsi pas possible de modifier fortement le surenroulement de l'ADN sans effectuer une cassure dans l'hélice [38]. Un tour cause un changement dans le nombre de liaison de +/- 1. Cette modification peut être quantifiée par la différence de liaison spécifique σ [36]:

$$\sigma = \frac{\mathrm{Lk} - \mathrm{Lk}_0}{\mathrm{Lk}_0}$$

En d'autres termes, la valeur de σ=-0,06 trouvée chez les bactéries signifie que, sur 100 tours de double hélice, 6 sont manquants et l'hélice n'en effectue donc que 94 [38].

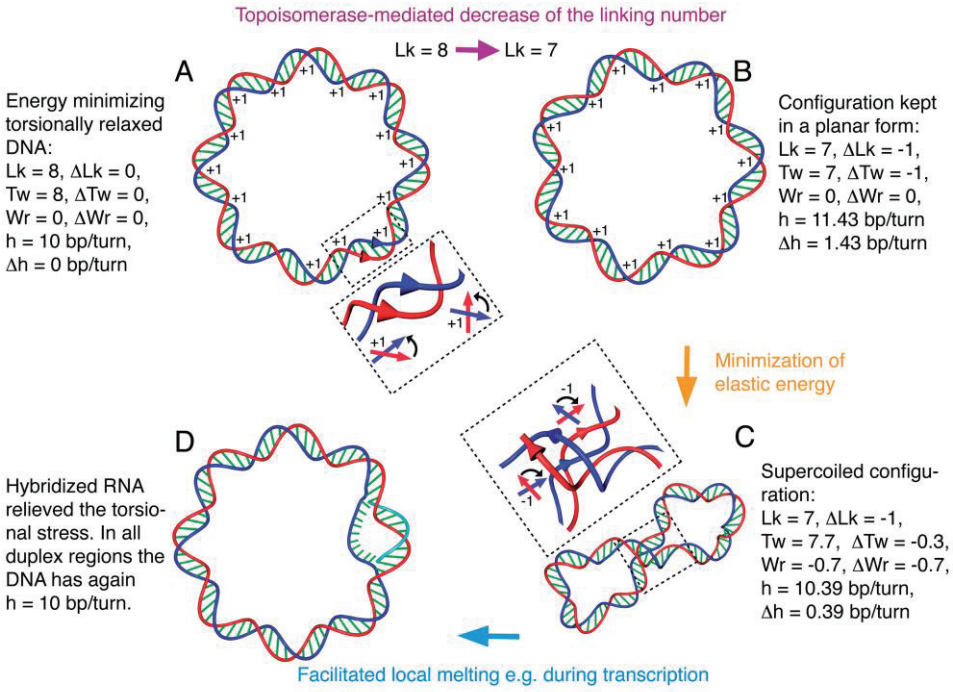


Figure 8 : Interconversion entre torsion et vrillage.
A : molécule l'ADN à l'équilibre. B : introduction d'un supertour négatif sous forme de torsion. C : conversion de la torsion en vrillage. D : ouverture de la double hélice facilitée par le surenroulement négatif. Lk : nombre de liason, Tw : Torsion, Wr : vrillage, h : périodicité de la double hélice. Extrait de [39].

La contrainte introduite par la modification du nombre de liaison peut être dissipée sous deux formes interconvertibles de surenroulement : la torsion (twist, Tw) et le vrillage (writhe, Wr). La torsion correspond au nombre de fois où les brins d'ADN se croisent et le vrillage au nombre de fois où la double hélice se croise. La somme de la torsion et du vrillage sera toujours égale au nombre de liaison [37,38] (Figure 8c).

Le niveau de surenroulement physiologique (σ=-0,06) n'est pas l'état d'équilibre (σ=0). Il est donc nécessaire d'assurer le maintien du surenroulement dans la cellule par une classe d'enzymes spécialisées : les topoisomérases.

Les topoisomérases sont divisées en 2 types, chacune subdivisée en plusieurs familles. La distinction entre les 2 types se fait sur la base du type de cassure introduite dans l'ADN pour modifier le niveau de surenroulement : simple brin pour le type I, double brin pour le type II [40]. Chez les bactéries, seules les familles IA et IIA sont représentées [40].

La famille IA comprend les topoisomérases I (aussi connue sous le nom de protéine ω ou TopA chez *E. coli*) et III [41]. La topoisomérase I permet de relâcher le surenroulement négatif : après avoir effectué la cassure simple brin de l'ADN caractéristique de leur type, elles permettent à la torsion stockée dans l'ADN (due au surenroulement négatif) de s'évacuer et ajoutent ainsi une unité au nombre de liaison [42] (Figure 9). Cette réaction ne requiert pas d'ATP. La topoisomérase III est une décaténase, permettant de ségréger les 2 chromosomes après une division cellulaire [42].

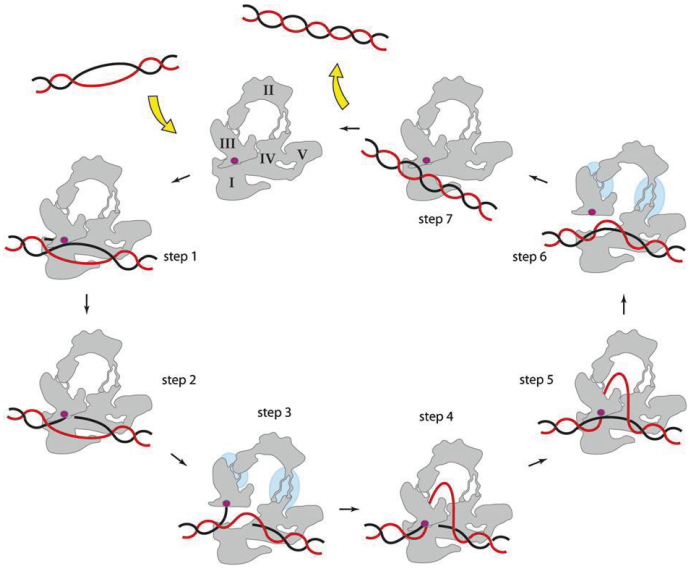


Figure 9 : Mécanisme de suppression d'un supertour négatif par la topoisomérase I bactérienne.
L'enzyme est représentée en gris et les brins d'ADN en couleur. La topoisomérase I introduit une cassure simple brin dans l'ADN (étape 1), supprime le tour négatif (étapes 2 à 5) et referme la cassure (étapes 6 et 7). Extrait de [42].

La famille IIA comprend les topoisomérases II (également appelées gyrases) et IV. L'ADN gyrase est particulièrement importante, puisqu'il s'agit de la seule topoisomérase capable d'introduire des supertours négatifs [40]. Elle agit en créant une cassure double brin dans l'ADN, en faisant passer un autre segment d'ADN à travers, puis en refermant la cassure. Ceci crée alors 2 vrillages négatifs (Figure 10), faisant alors diminuer le nombre de liaison de 2 [40]. Cette réaction requiert de l'ATP. Elle possède également d'autres fonctions partagées avec la topoisomérase IV (Figure 10). L'équilibre entre compaction et relâchement de l'ADN est ainsi assuré majoritairement par la topoisomérase I et la gyrase.

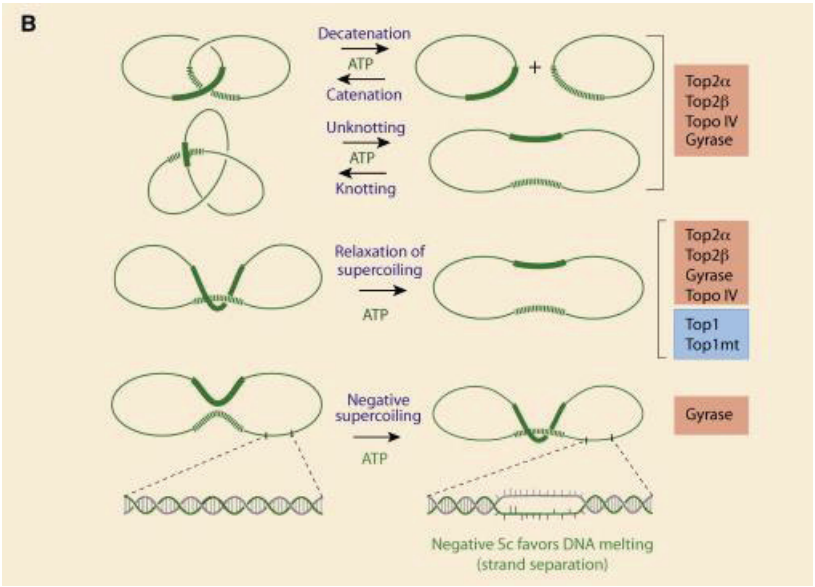


Figure 10 : Modifications de surenroulement catalysées par différentes topoisomérases.
Les réactions sont schématisées et associées aux différentes topoisomérases (en bleu pour le type IB, en rouge pour le type IIA). Top1 et Top2 sont des topoisomérases eucaryotes, alors que la Gyrase et Topo IV sont bactériennes. Extrait de [41].

En l'absence de tout autre facteur, le surenroulement négatif favorise la dissociation des brins d'ADN et donc l'initiation de la transcription [38] (Figure 8d). Il est ainsi possible de prédire

l'intensité de la transcription de certains gènes sur la seule base de la séquence de leurs régions promotrices en calculant la probabilité d'ouverture de la double hélice [43]. La transcription du gène *topA*, codant pour la topoisomérase I de *E. coli* suit cette règle générale [44]. Cependant, les modifications de surenroulement (σ) causent des changements dans la périodicité de la double hélice de l'ADN (h) et donc la géométrie de l'ADN, ainsi que dans les affinités de certains régulateurs (voir section suivante), ce qui peut avoir des effets variés en fonction de la région régulatrice considérée.

Les gènes *gyrA* et *gyrB* sont les cas les mieux décrits pour l'activation par la relaxation de l'ADN. Le phénomène est resté mystérieux pendant longtemps : des expériences de délétion [45] puis de mutagénèse dirigée [46] ont montré que la région -10 de *gyrA/B* était nécessaire et suffisante pour conférer cette propriété d'activation par la relaxation, mais que sa séquence ne montrait pas d'anomalie comparée à d'autres promoteurs réprimés par la relaxation. Par la suite, il a été montré que la région régulatrice de *gyrA/B* était particulièrement flexible et que les mutations réduisant ce phénomène diminuaient également l'activation du promoteur par la relaxation [47]. Deux mécanismes complémentaires ont alors été proposés : (i) la courbure permet un alignement optimum des séquences -10 et -35, cet alignement étant perdu en cas de compaction de l'ADN et (ii) la courbure attire certaines protéines associées au nucléoïde (voir section suivante, H-NS et HU ont été proposés), modulant alors la transcription du promoteur [47]. La régulation opposée de la transcription de *gyrA/B* et de *topA* par le surenroulement crée ainsi une boucle de rétroaction négative assurant l'homéostasie du surenroulement bactérien et permet d'éviter des niveaux de surenroulement extrêmes néfastes à la cellule (Figure 11).

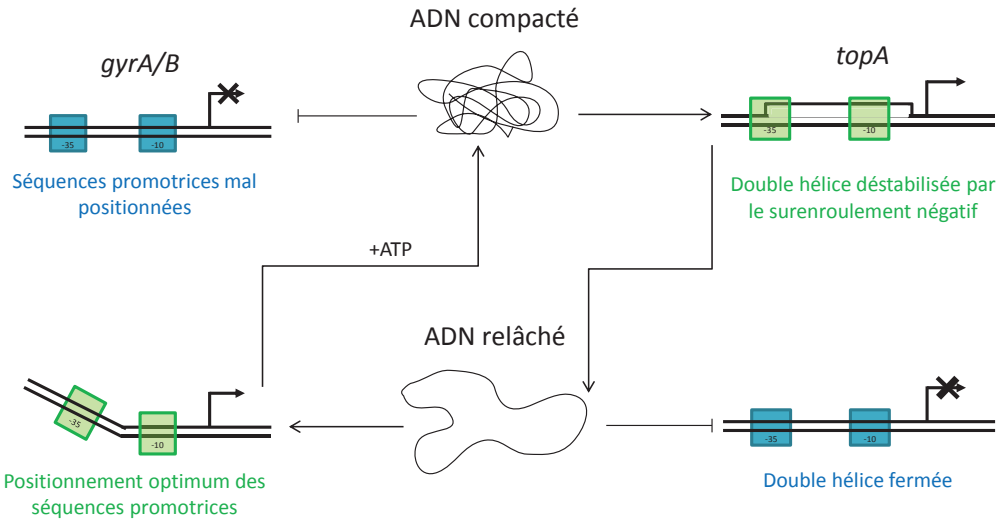


Figure 11 : régulation homéostatique du niveau de surenroulement bactérien par la topoisomérase I et la gyrase.

Un fort niveau de surenroulement de l'ADN (haut de la figure) permet la déstabilisation de la double hélice de l'ADN, favorisant la transcription de *topA*. La topoisomérase I cause alors un relâchement de l'ADN. Si ce relâchement devient trop important, la transcription de gyrA/B est activée par un mécanisme encore mal compris (bas de la figure). L'introduction de supertours négatifs par la gyrase permet alors de rétablir un ADN surenroulé à condition que de l'ATP soit présent.

f. Les protéines associées au nucléoïde

Les protéines associées au nucléoïde (NAPs) sont définies comme des « protéines qui altèrent la topologie de l'ADN pour le rendre plus compact et qui ont le potentiel d'influencer la transcription » [48]. Elles représentent l'équivalent fonctionnel des protéines histones des eucaryotes et de certaines archées, bien qu'elles ne partagent aucun point commun structurel

[48]. L'ADN de la bactérie *E. coli*, en l'absence des NAPs, représenterait en effet un volume de 200 μm^3, soit 400 fois le volume occupé en réalité (0.5 μm^3) [49]. Ce rôle structurant du nucléoïde bactérien par les NAPs est assuré via une concentration cellulaire forte (de l'ordre de 10000 à 50000 copies par cellule [50], avec un maximum de 178000 copies observé pour Dps chez *E. coli* en phase stationnaire de croissance), des fixations non spécifiques fréquentes et trois actions topologiques essentielles : courbure (bending), pontage (bridging) et emballage (wrapping).

La courbure consiste en l'introduction d'un angle dans l'ADN, la valeur de la courbure dépendant de la NAP considérée. HU (Heat-unstable nucleoid-structuring protein), une NAP présente dans presque toutes les bactéries et qui se fixe nonspécifiquement sur l'ADN, est un représentant de ce mode d'action, en étant capable d'induire des courbures d'environ 100°, ce qui peut ensuite donner lieu à des structures plus complexes et compactes (Figure 12). Des courbures plus importantes (180°) ont cependant été observées avec IHF [51].

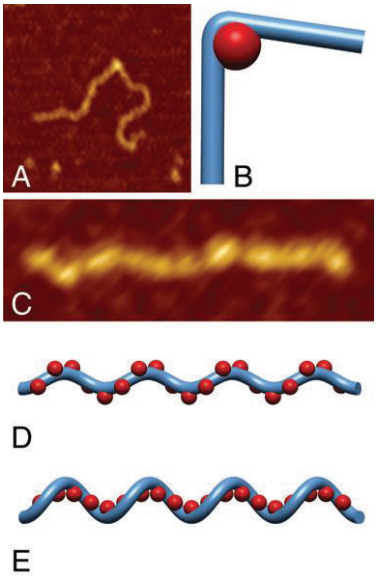


Figure 12 : Introduction de courbures dans l'ADN par HU

Des images de microscopie à force atomique de complexes HU-ADN (A, C) et leur interprétation (B, D, E) sont présentées. A, B : formation de courbures par HU. C, D, E : Formation de filaments nucléoprotéiques à haute concentration de HU. Extrait de [51]

L'emballage est une fonction partagée avec les histones eucaryotes, mais reste relativement rare chez les NAPs [48]. L'exemple le mieux décrit est LRP (Leucine-responsive regulatory protein), une NAP également très répandue chez les bactéries, qui peut former un octamère enroulant l'ADN sous une forme cylindrique (Figure 13).

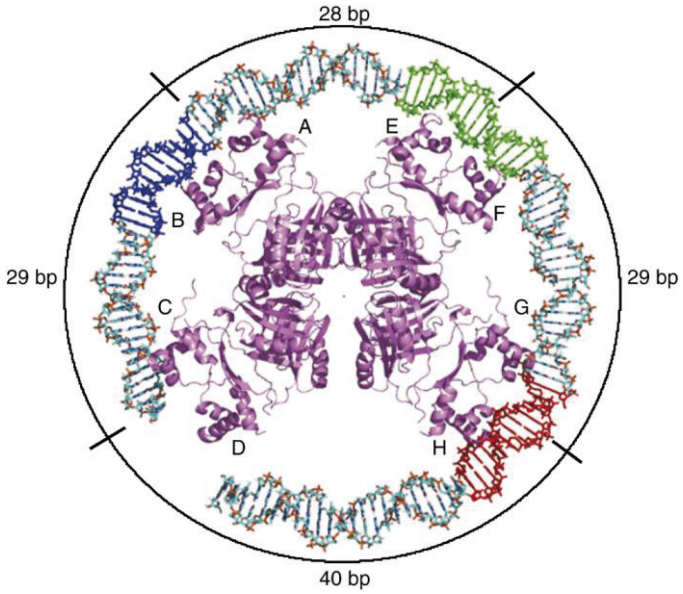


Figure 13 : Emballage de l'ADN par un octamère de LRP
Les protéines composant l'octamère sont indiquées par des lettres (de A à H). Extrait de [52].

Enfin, le pontage est un rapprochement de 2 doubles brins d'ADN suite à une multimérisation d'une NAP. Le phénomène est surtout étudié pour la NAP des γ-protéobactéries H-NS (Figure 14).

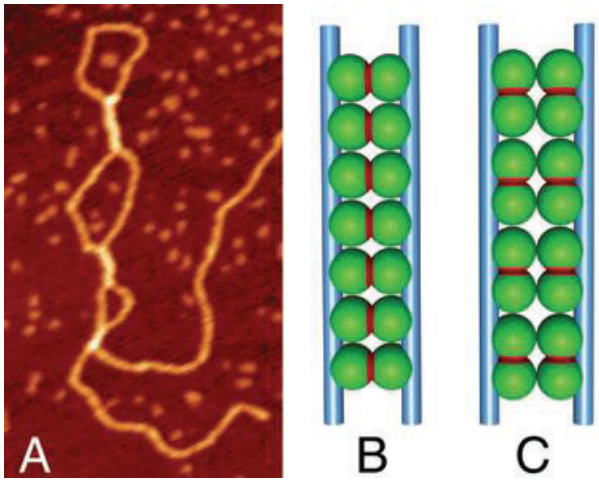


Figure 14 : Pontage de l'ADN par H-NS
A : image de microscopie à force atomique de complexes H-NS-ADN. B,C : hypothèses expliquant la formation des ponts. Les dimères sont représentées par deux boules vertes séparées par un trait rouge. L'orientation des dimères dans le pontage est débattue. Extrait de [51].

Au niveau structural, les NAPs permettent donc d'introduire des modifications locales dans la topologie globale de l'ADN définie par l'équilibre entre l'activité de la topoisomérase I et de la gyrase (Figure 15). Il est cependant fréquent que ces protéines soient également capables d'agir comme des FT, notamment FIS [53] et que, inversement, des FT interagissent avec la topologie de l'ADN (par exemple CRP [54]) au point que la distinction entre FT et NAP devienne floue. Par exemple, la régulation du promoteur du gène *nir* de *E. coli*, codant pour une nitrite reductase, fait appel à la fois à des propriétés structurantes et régulatrices des NAPs (Figure 16). Il semble plutôt y avoir un spectre continu de régulateurs ayant plus ou moins d'impact sur la topologie de l'ADN et sur l'expression des gènes.

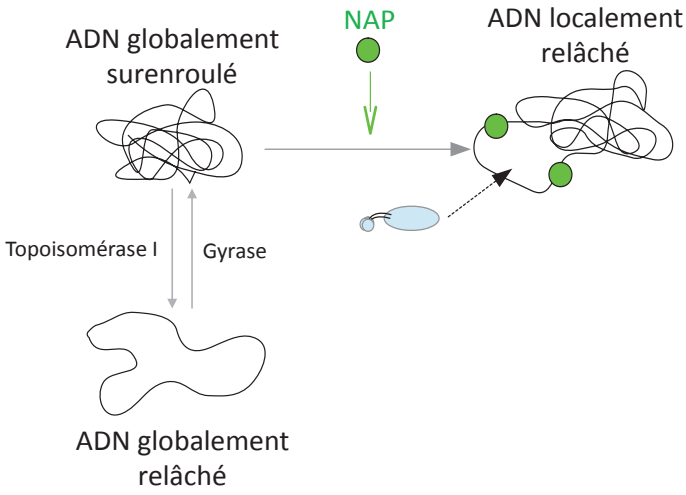


Figure 15 : Intégration des NAPs dans le contrôle de la topologie de l'ADN.
En plus de l'action de la topoisomérase I et de la gyrase qui définissent le niveau global de surenroulement (Figure 11), les NAPs (en vert) induisent des modifications locales qui peuvent attirer l'ARN polymérase (en bleu clair) ou d'autres régulateurs vers une région particulière du génome.

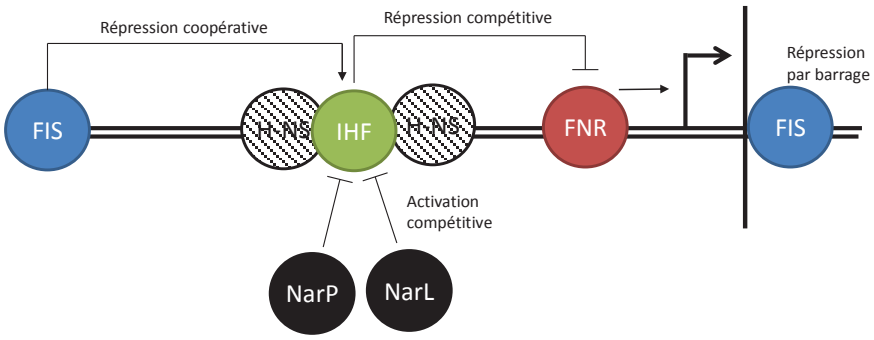


Figure 16 : Régulation transcriptionnelle du promoteur *nir* de *E. coli* par les NAPs
IHF joue ici le rôle d'un facteur de transcription répresseur classique agissant par compétition avec l'activateur FNR, l'action de INH étant elle-même régulée par compétition par NarP/NarL. FIS réprime *nir* à distance par le biais d'une action coopérative avec IHF via le site amont et probablement par barrage pour le site aval. H-NS réprime *nir* par un mécanisme inconnu indépendant de NarP/NarL [55].

4. Interactions entre des promoteurs proches

Les promoteurs des gènes ne servent pas que de site de reconnaissance et de fixation pour l'ARN polymérase. Plusieurs exemples où des promoteurs interagissent entre eux ont été décrits. Ces mécanismes peuvent être classés en 2 types : de contact et à distance.

Le cas le plus fréquent d'interaction de contact est la compétition directe entre promoteurs proches pour la fixation de l'ARN polymérase. Ce dernier est en effet un gros complexe capable d'occuper jusqu'à 60 pb sur l'ADN [11] (sans compter le CTD des sous-unités α), ce qui peut causer des compétitions entre ARN polymérases pour l'occupation de 2 promoteurs proches. Par exemple, la région régulatrice du gène *crp* porte 2 promoteurs divergents distants de seulement 2 nucléotides. Par défaut, le promoteur de *crp* est transcrit mais, en présence de fortes concentrations du cofacteur AMPc, le promoteur divergent est transcrit, empêchant alors l'ARN polymérase de se fixer sur *crp* [56]. Un tel mécanisme permet probablement de moduler l'action

de CRP, en compensant l'activité accrue de CRP du fait de la présence de son cofacteur par une diminution de la transcription de *crp* (Figure 17).

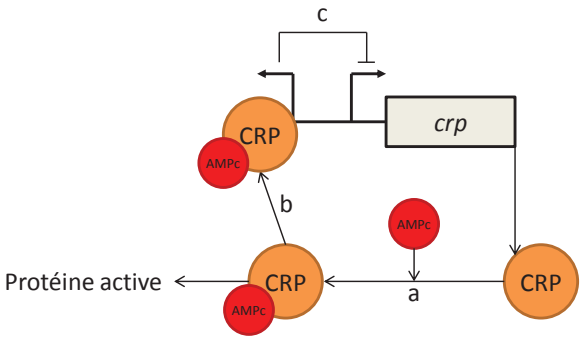


Figure 17 : Modulation de la transcription de *crp* par un promoteur divergent.
L'AMPc, cofacteur de CRP, permet de rendre la protéine active (a). Cette forme active est alors capable de promouvoir la transcription du promoteur divergent de *crp* (b) et ainsi d'empêcher la transcription de *crp* (c), créant une boucle de rétroaction négative.

Un cas artificiel [57] ou naturel [58] où les promoteurs divergents sont distants de 80 bp présente un comportement similaire : la fixation efficace de l'ARN polymérase sur les deux promoteurs divergents est mutuellement exclusive à cause de la gêne stérique occasionnée par les ARN polymérases (Figure 18).

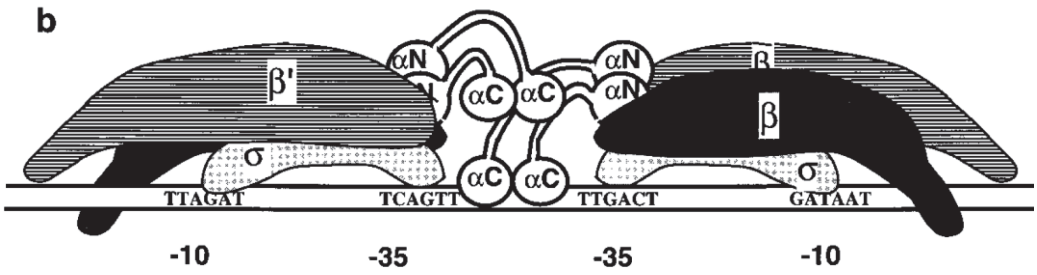


Figure 18 : Interférence par gêne stérique de deux molécules d'ARN polymérase fixées sur les promoteurs p_R et p_{RM} du phage λ.
Extrait de [58].

Dans le cas de promoteurs en tandem (dans la même orientation) ou convergents (face-à-face), l'influence peut se faire à plus longue distance. En effet, une polymérase transcrivant depuis le promoteur situé en amont (dans le cas d'un promoteur en tandem) ou depuis l'un des deux promoteurs convergents peut bloquer la transcription du second promoteur [59] (Figure 19). Ce blocage est favorisé par des séquences de pause sur l'ADN riches en bases A/T, qui arrêtent temporairement l'ARN polymérase [60].

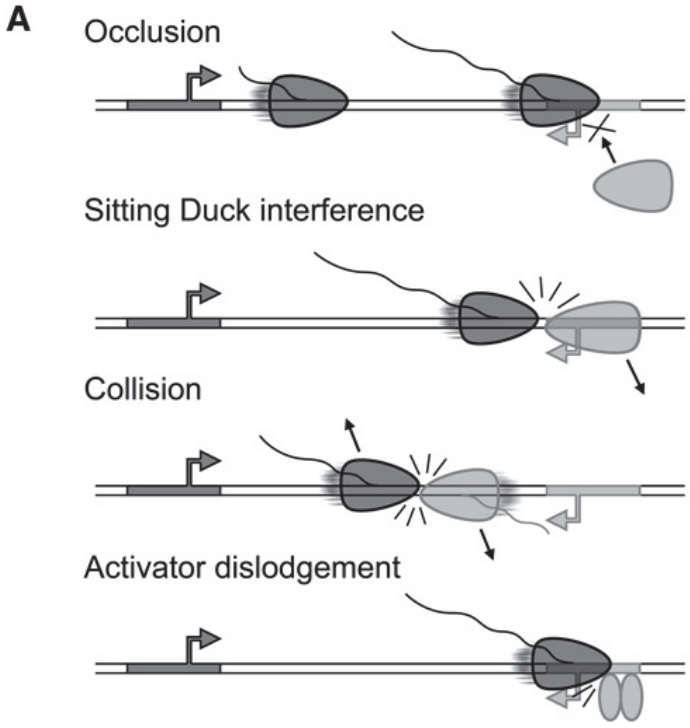


Figure 19 : Interférences à distance de deux promoteurs convergents

La polymérase lancée depuis le promoteur de gauche (gris foncé) peut réprimée la transcription du promoteur de droite (gris clair) en interagissant directement avec elle (occlusion, sitting duck et collision) ou en entrant en compétition avec l'un de ses activateurs (dislodgement). Extrait de [60].

Enfin, une seconde catégorie de mécanismes à distance, ne requérant pas de contact physique entre des molécules d'ARN polymérase, a été décrite. Ceux-ci se basent sur des transferts de surenroulement. Lors de la transcription, l'ARN polymérase ouvre la double hélice de l'ADN puis la referme. Cette opération ne modifie pas le nombre de liaison total mais crée un supertour positif en aval de l'ARN polymérase et un supertour négatif en amont environ tous les 10 paires de bases transcrites [37] (Figure 20). Les supertours ainsi générés sont capables d'influencer un promoteur proche. Par exemple, les supertours négatifs peuvent favoriser l'expression d'un promoteur divergent jusqu'à une distance de 4 kb [43]. Le nombre de supertours générés dépendant de la taille du transcrit, l'effet est plus prononcé lorsque le promoteur contrôle un transcrit long plutôt qu'un transcrit court [61].

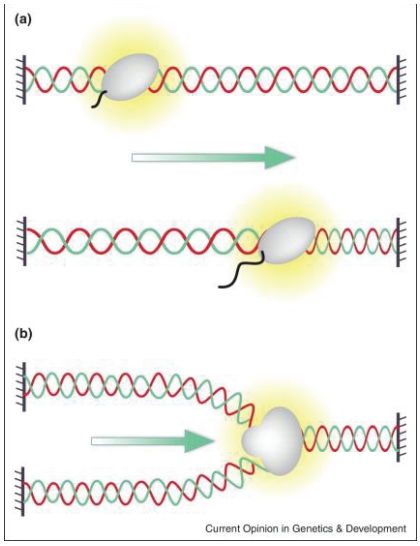


Figure 20 : Modifications locales du surenroulement induites par la transcription et la réplication
Lors de la polymérisation d'ARN (a) ou d'ADN (b), la double hélice est ouverte puis refermée, ce qui cause un surenroulement négatif en amont et positif en aval. Extrait de [37].

Tous ces mécanismes peuvent ensuite être combinés pour créer des systèmes encore plus complexes. La région régulatrice du gène *fis* chez *E. coli* est un bon exemple [62] (Figure 21). *fisP1* est le promoteur principal, organisé en tandem avec deux autres promoteurs plus faibles (*fisP2* et *fisP3*). Enfin, un promoteur divergent (*div*) situé 108 pb en amont de *fisP1* chevauche *fisP2*. Il en résulte un système très complexe avec de nombreuses interactions potentielles. Il a été montré que, malgré leur orientation différente, *fisP3* et *div* sont capables d'activer *fisP1* et que cet effet était perdu lors d'insertions de 5 pb (1/2 tour de double hélice) mais pas 10 pb (1 tour) entre les promoteurs [62]. Il a été proposé que les ARN polymérases fixées sur *div* et *fisP1* se dimérisent, stabilisant ainsi l'interaction ARNP-*fisP1*. Ce mécanisme requiert une géométrie particulière de la région, qui est perdue lors des insertions de 5 pb ou suite à un relâchement de l'ADN. Le mécanisme d'action de *fisP3* reste cependant encore inconnu.

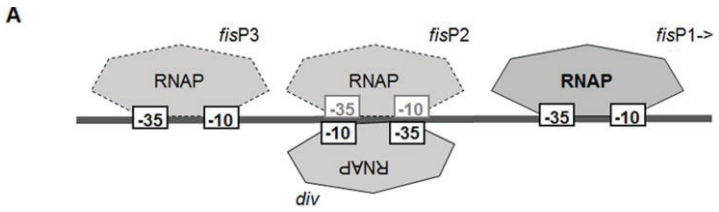


Figure 21 : Structure de la région régulatrice de *fis* de *E. coli*
Les quatre promoteurs de cette région régulatrice sont représentés par leurs boites -10 et -35, avec l'ARN polymérase fixée sur chacun d'eux. Extrait de [62].

Bien que les études détaillées sur l'importance de ces mécanismes *in vivo* soient encore rares, ces interactions entre promoteurs pourraient représenter une contribution non négligeable dans la régulation transcriptionnelle bactérienne. En effet, chez de nombreuses bactéries, les régions régulatrices des gènes portent souvent plus d'un promoteur [63] (92% chez *E. coli* portent au moins 2 promoteurs, avec en moyenne 3.1 promoteurs par région régulatrice), ce qui pose la question du rôle de ces promoteurs surnuméraires. De nombreux mécanismes restent donc probablement à découvrir dans ce domaine. L'un des défis les plus intéressants serait d'arriver à intégrer, dans l'espace et dans le temps, l'action des différents acteurs.

L'ensemble des mécanismes de régulation transcriptionnelle décrits jusqu'à présent forment

un réseau complexe optimisé pour la réponse aux signaux perçus dans un environnement donné, permettant alors aux bactéries de s'acclimater aux changements. Cependant, lorsque les bactéries changent durablement d'environnement, la flexibilité des réseaux de régulation intervient dans l'adaptation des bactéries.

5. Evolution des réseaux de régulation transcriptionnelle

L'évolution d'un réseau de régulation transcriptionnelle est un mécanisme complexe qui doit prendre en compte les gains ou pertes de régulateurs ou de gènes cibles. Les règles générales gouvernant cette évolution commencent seulement à être comprises.

Les bactéries possèdent un « stock » de facteurs de transcription dont la taille est corrélée à la taille du génome. En effet, la fraction de gènes dédiée à la régulation augmente avec la taille du génome : 12% pour *Streptomyces coelicolor* (8.7 Mbp), 6% pour *E. coli* (4.3 Mbp) et 0.8% pour *Mycoplasma genitalium* (0.6 Mbp) [64–66]. La régulation transcriptionnelle semble ainsi avoir un intérêt pour des génomes complexes correspondant à des bactéries en contact avec des environnements variables, plutôt que des bactéries symbiotiques à génome réduit confrontées à des environnements relativement stables. Seuls 5 régulateurs font partie du « génome minimal » : la NAP HU, le facteur d'antiterminaison NusG, les facteurs d'élongation NusA et GreA et la (p)ppGpp synthase RelA [65] mais aucun facteur de transcription ni facteur σ alternatif.

Au-delà de cette base minimale, le stock de facteurs de transcription peut varier en fonction des environnements auxquels sont confrontés les bactéries. Par exemple, chez les γ-protéobactéries, seuls 11 facteurs de transcription sont conservés (sur 141 pour *E. coli*), incluant 4 des 7 facteurs de transcription globaux identifiés chez *E. coli* (Figure 6) [67]. Il a été montré que les activateurs peuvent être facilement perdus, ce qui conduit à l'élimination postérieure de leurs gènes cibles non exprimés, alors que les répresseurs nécessitent la perte préalable de leurs cibles avant de pouvoir être eux-mêmes éliminés [67]. L'évolution des régulateurs est donc étroitement liée à celle de leurs gènes cibles (Figure 22).

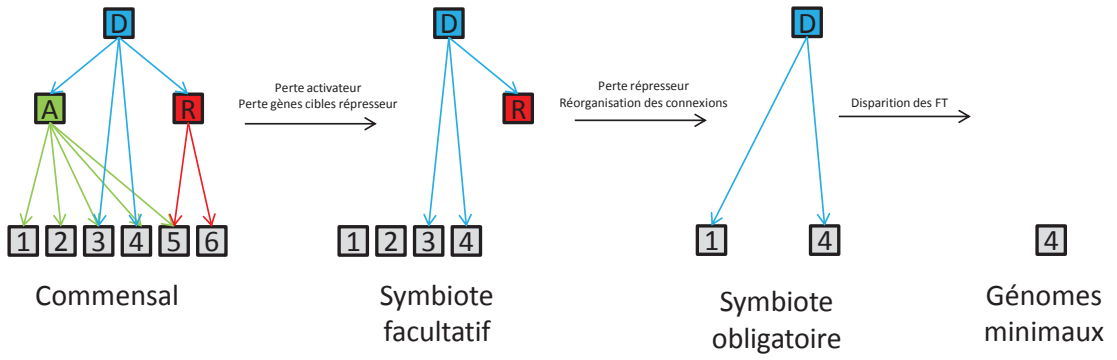


Figure 22 : Modification des réseaux de régulation lors d'événements de réduction de génomes

Les gènes cibles sont représentés en gris et numérotés, le régulateur double en bleu, l'activateur en vert et le répresseur en rouge. Lors d'une réduction de génome, les activateurs disparaissent en premier, laissant leurs gènes non régulés (gènes 1 et 2 par exemple). Les répresseurs suivent seulement une fois leurs gènes cibles perdus (gènes 5 et 6). Enfin, les régulateurs doubles ne disparaissent qu'à des étapes extrêmes de réduction. Adapté d'après [68].

En plus de la présence ou absence des régulateurs, un niveau supplémentaire de plasticité des réseaux de régulation est apporté par des modifications dans les connexions entre ces régulateurs et leurs gènes cibles ou entre eux. Il est en effet fréquent qu'un régulateur gagne ou

perde une cible [64]. PhoP est le régulateur de réponse du système à deux composants PhoP/Q , répondant à la concentration extracellulaire de Mg^{2+} et Ca^{2+} et également impliqué dans le contrôle de la virulence des bactéries pathogènes [69]. Bien que ce système soit présent dans de nombreuses entérobactéries, peu de cibles sont conservées dans le régulon *phoP* au sein des différentes espèces. Il s'agit exclusivement des gènes liés à l'homéostasie du magnésium, incluant une boucle de rétroaction négative agissant sur PhoP [70]. En revanche, bien que PhoP soit indispensable à la virulence de *Salmonella enterica* et *Yersinia pestis*, les gènes liés à la virulence régulés par PhoP chez ces deux espèces sont distincts [70]. Ce cas rappelle celui de FIS entre *S. enterica* et *D. dadantii* présenté avec le premier article [71], et de RpoE chez les entérobactéries [72]. Souvent, peu d'effet est observé lorsque les connexions sont modifiées, ce qui permet une plasticité importante. Par exemple, l'ajout d'un nombre important de connexions artificielles entre différents régulateurs chez *E. coli* n'a causé de modifications significatives de croissance que dans 15% des cas. De même, des modifications dans les réseaux de régulation impliquant des régulateurs globaux comme CRP ou RpoS ont été bien tolérées, indiquant une bonne robustesse des réseaux de régulation [73]. En revanche, d'autres exemples montrent une importance de ces connexions. Ainsi, l'intégration par *S. enterica* de PmrD, un régulateur intermédiaire associé au sytème à deux composants de réponse au Fe^{3+} PmrA/B, dans la réponse de PhoP permet une expression plus forte, plus tardive et plus persistante du gène *pbgP* conférant la résistance à l'antibiotique Polymyxine B, par rapport à *Y. pestis*. Ces modifications régulatrices améliorent ainsi la résistance à la Polymyxine B de *S. enterica* [64,74].

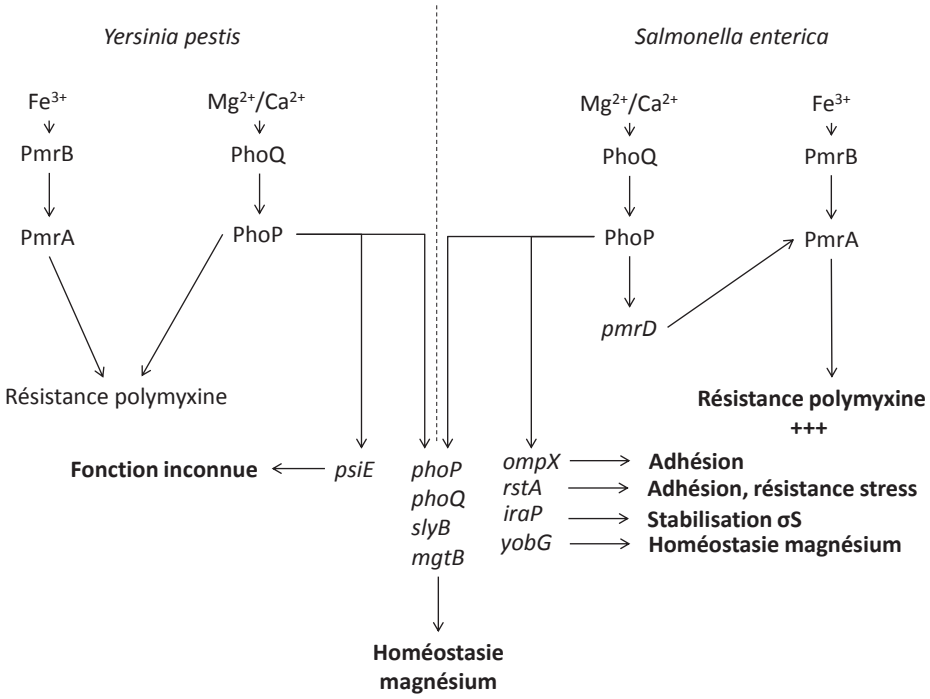


Figure 23 : Comparaison des réseaux de régulation associés au système à deux composants PhoP/PhoQ chez *Y. pestis* et *S. enterica*.

Un ensemble de gènes communs régulés par PhoP (*phoP*, *phoQ*, *slyB* et *mgtB*) assurent l'homéostasie du magnésium chez les deux espèces. Chez *Y. pestis*, la régulation de la réponse à la polymyxine est assurée directement par PmrA et PhoP, ce qui confère à cette bactérie une résistance limitée. En revanche, chez *S. enterica*, PhoP active la transcription de *pmrD*, ce qui améliore la résistance à la polymyxine. De plus, *S. enterica* possède des gènes cibles de PhoP uniques qui participent à sa virulence (*rstA* et *iraP*). Adapté d'après [64,70,74].

L'acquisition de nouveaux gènes pose des problèmes particuliers. Ces acquisitions peuvent se faire par transfert horizontal ou par duplication, le premier cas étant le plus fréquent chez les

bactéries [75].

Dans le cas des transferts horizontaux, les régulateurs des gènes ne sont pas forcément transférés avec les gènes cibles et, du fait de la faible conservation des régulateurs discutée plus haut, le gène nouvellement acquis se trouve dérégulé. Il en résulte une transcription souvent néfaste pour la bactérie, qui doit par conséquent réagir rapidement de manière à contrôler l'expression du gène nouvellement acquis. En règle générale, un régulateur peu spécifique, tel que la NAP H-NS chez les entérobactéries, va reconnaître le segment transféré via son taux de GC différent de celui du reste du génome et le réprimer [76]. Ce mécanisme, schématisé en Figure 24, s'appelle la répression xénogénique (xexogeneic silencing). Par la suite, une régulation se met en place via un régulateur spécifique (par exemple PhoP, VirB, LeuO ou Ler chez les entérobactéries [77,78]) pour lever la répression exercée par H-NS et adapter l'expression du gène transféré à son nouvel hôte. Bien que H-NS soit spécifique des entérobactéries, d'autres protéines telles que Rok chez *B. subtilis* [79] assurent la répression xénogénique dans les autres phyla, ce qui en fait un mécanisme répandu.

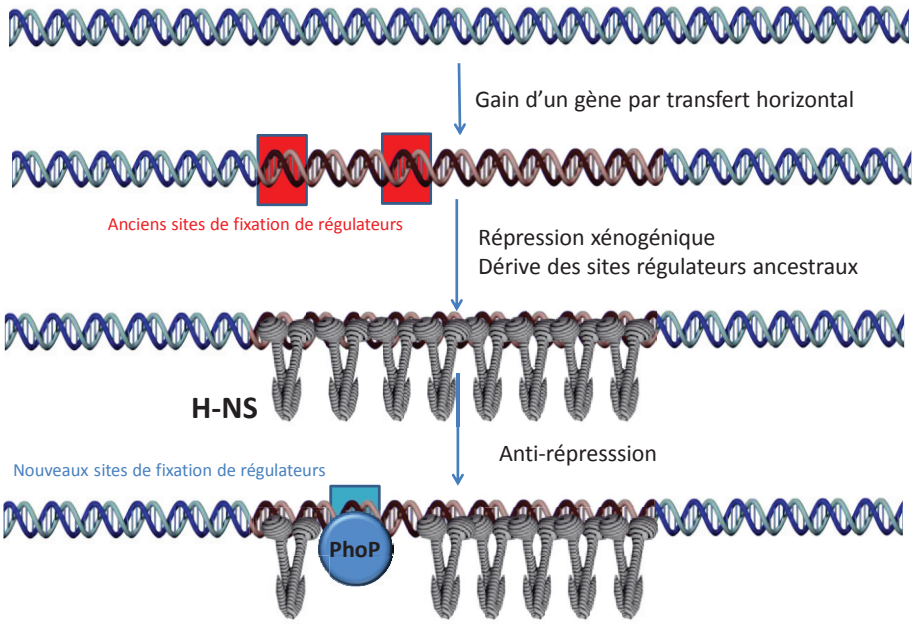


Figure 24 : Mécanisme de la répression xénogénique.
L'ADN ancestral est représenté en bleu et l'ADN transféré en rose. Les régulateurs de la bactérie donneuse sont rarement présents dans la bactérie receveuse, ce qui cause une expression anarchique des gènes transférés. Cette expression est supprimée par des répresseurs non spécifiques tels que H-NS, puis l'expression contrôlée des gènes est restaurée par apparition d'un site de fixation d'un régulateur de la bactérie receveuse (PhoP dans le schéma).

A l'inverse, il arrive que de grands groupes de gènes incluant leurs régulateurs soient transférés, et les mécanismes évolutifs sont alors différents. Les ilots de pathogénie SPI de *S. enterica* en représentent un exemple. Le SPI-2, un îlot de 25 kb nécessaire à la survie intracellulaire de *S. enterica*, code également pour ses propres régulateurs : le système à deux composants SsrA/SsrB [80]. Cependant, l'expression des gènes du SPI-2 est aussi contrôlée par des régulateurs chromosomiques (FIS, IHF, OmpR…) et même HilD, qui est codé par un autre îlot de pathogénie (SPI-1) acquis indépendamment du SPI-2 [80]. La régulation basique du SPI-2, importée au moment du transfert horizontal et assurée par SsrA/SsrB, a donc été affinée par l'intervention d'autres régulateurs déjà présents dans le génome de *S. enterica*.

L’intégration des gènes dupliqués pose des problèmes totalement différents. En effet, ces gènes ont initialement la même fonction et la même régulation. S’ils ne divergent pas et à moins de conférer un avantage par augmentation du nombre de copies du gène, une copie est rapidement perdue (à une vitesse d’environ 0.05 pertes/génome/génération [81]). Pour pouvoir se maintenir, les paralogues doivent acquérir des fonctions distinctes [82], ce qui se fait d’abord via des divergences dans leur régulation puis dans la fonction des protéines codées [83]. Les mécanismes de divergence de régulation connus sont nombreux et impliquent régulièrement des modifications dans les régulateurs des gènes dupliqués (Figure 25).

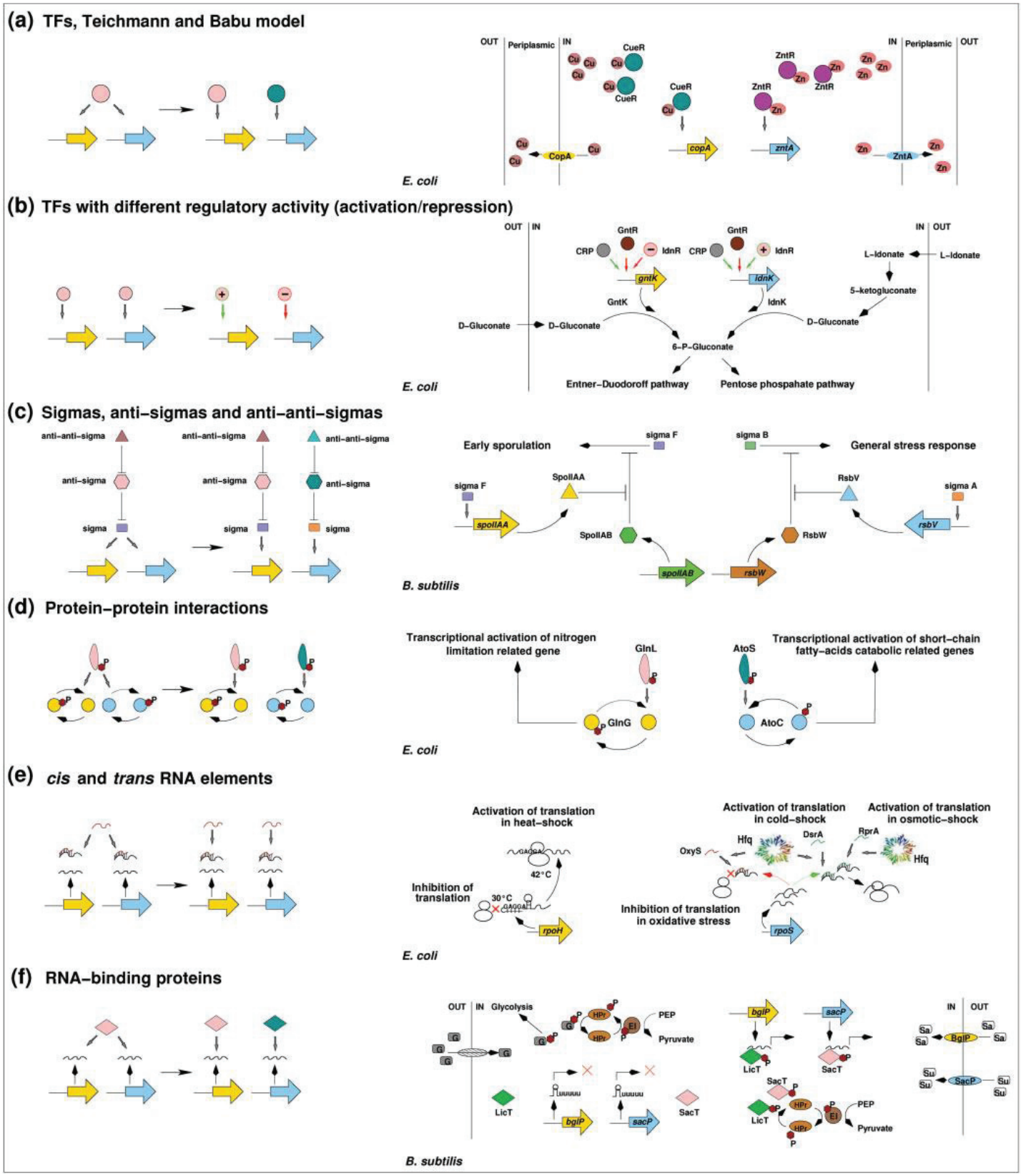


Figure 25 : Mécanismes de divergence de la régulation des gènes paralogues chez les bactéries.

Un nouveau facteur de transcription (a), facteur sigma (c), kinase (d), ARN régulateur (e) ou protéine liant l'ARN (f) est recruté pour réguler l'un des gènes dupliqués. Un régulateur conservé peut également changer son effet activateur ou répresseur (b). Pour chaque cas, un exemple est représenté sur la droite de la figure. Extrait de [84].

Ces modèles prédictifs établis à partir des espèces bactériennes actuelles se sont trouvés confirmés par les études d'évolution expérimentale. Dans l'expérience d'évolution à long terme de *E. coli*, plusieurs cultures indépendantes ont acquis une mutation sélectivement avantageuse

dans le gène *spoT*, modifiant alors l'expression de plusieurs autres gènes [85]. Ces souches ont également modifié le niveau de surenroulement à la fois global et local par des mutations dans *fis* et *topA* [86] et ont perdu certains gènes spécifiques tels que le régulateur du catabolisme du maltose *malT* et l'ensemble de l'opéron du catabolisme du ribose *rbs* [86]. De plus, les connexions des régulateurs ont aussi été modifiées puisque le régulon CRP s'est nettement enrichi, passant de 170 cibles à un maximum de 750 cibles dans l'une des cultures [87]. Bien que les autres régulateurs n'aient pas été étudiés, il est probable que les modifications dans les connexions soient aussi présentes pour d'autres régulons. Comme prédit, les modifications semblent donc s'effectuer principalement au niveau de la régulation des gènes et peu au niveau des fonctions des protéines.

B. Les entérobactéries phytopathogènes : un mélange unique de gènes adaptés aux hôtes animaux et végétaux

1. Taxonomie et évolution

Les entérobactéries les plus connue sont des pathogènes animaux (*Salmonella enterica*, *Yersinia pestis*, *Escherichia coli*…), mais elles comportent également plusieurs groupes de phytopathogènes. On distingue couramment 3 ensembles d'entérobactéries phytopathogènes (Figure 26 et Figure 27) :

-*Enterobacter*

-*Erwinia-Pantoea* (EP)

-*Pectobacterium-Brenneria-Samsonia-Lonsdalea-Dickeya* (PBSLD)

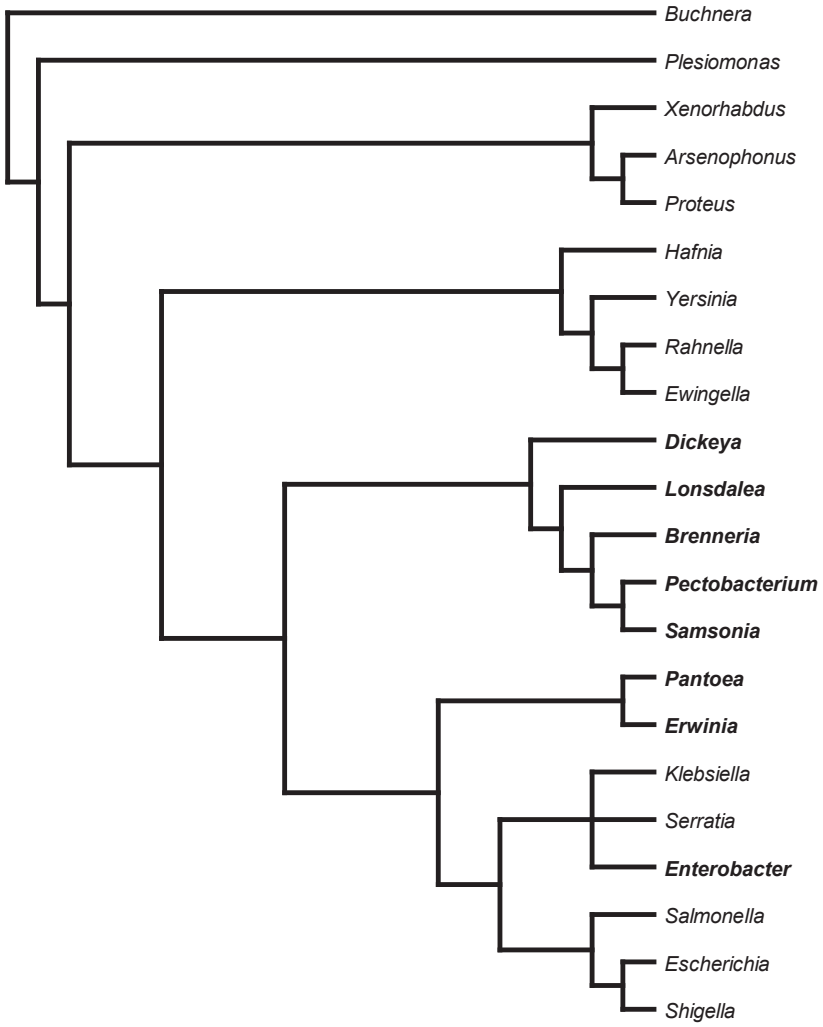


Figure 26 : Place des phytopathogènes dans la phylogénie des entérobactéries.

Les principaux genres d'entérobactéries sont représentés, ceux comprenant des phytopathogènes sont écrits en gras. Adapté d'après la phylogénie 16S présentée dans la littérature [88], avec les noms de genres actualisés pour correspondre à la taxonomie actuelle.

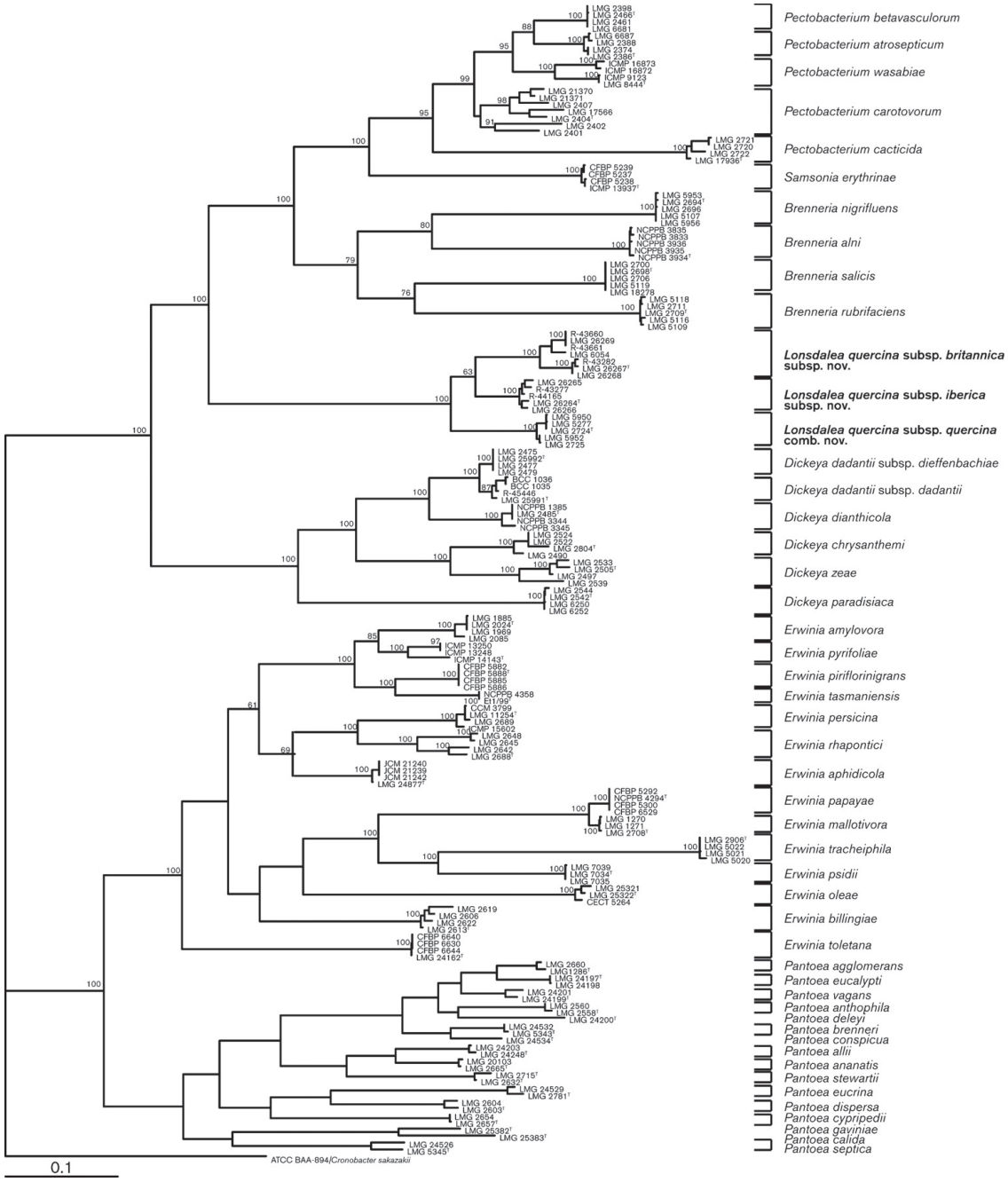


Figure 27 : Phylogénie multi-locus détaillée des entérobactéries phytopathogènes.

L'arbre a été construit par maximum de vraisemblance à partir des séquences des gènes *gyrB*, *rpoB*, *atpD* et *infB*. Les chiffres correspondent aux valeurs de bootstrap et l'échelle en bas à gauche à 0.1 substitutions/site. Extrait de [89].

La taxonomie de ces groupes a fortement évolué au cours du temps. A l'origine, toutes les entérobactéries phytopathogènes ont été groupées dans l'espèce *Erwinia* [90], puis elles ont été séparées en fonction des découvertes sur leurs hôtes, leurs mécanismes de virulence (Figure 28)… La taxonomie proposée en 1945 n'a pas été systématiquement utilisée par la communauté scientifique et, par exemple, jusqu'en 2005 les *Dickeya* étaient le plus souvent désignés, à tort, comme *E. chrysanthemi*, alors que la nomenclature officielle était *Pectobacterium chrysanthemi*.

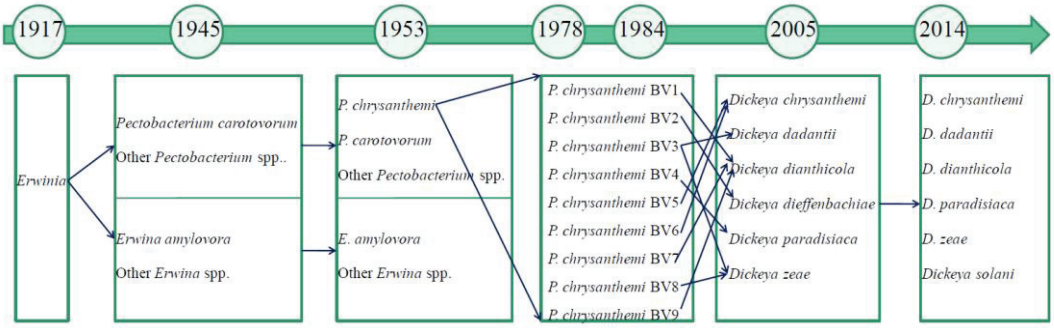


Figure 28 : Evolution de la taxonomie des *Dickeya*.
L'éclatement du genre *Erwinia* en différents groupes de bactéries phytopathogènes est représenté, avec les dates correspondantes en haut de la figure. Extrait de [91].

Enterobacter constitue un cas à part parmi les entérobactéries phytopathogènes. Le genre comportait en effet à l'origine uniquement des zoopathogènes proches de *Klebsiella* [92] mais certaines espèces phytopathogènes autrefois assignées au genre *Erwinia* ont été transférées vers le genre *Enterobacter* [93]. L'une des espèces phytopathogènes, *Enterobacter dissolvens*, a même ensuite été définie comme une sous-espèce d' *Enterobacter cloacae,* qui comprend aussi plusieurs souches d'importance clinique majeure [94,95]. La souche type d'*E. cloacae*, ATCC 13047, est par ailleurs capable d'infecter l'homme comme de causer la pourriture interne de l'oignon [96]. Les mécanismes exacts de la virulence de ces souches sur les végétaux restent encore inconnus. En conclusion, à l'inverse des genres qui vont être présentés par la suite, *Enterobacter* semble être versatile et comporter des souches adaptées à un hôte animal ou végétal, mais sans lien apparent avec les deux autres groupes d'entérobactéries phytopathogènes, que ce soit en termes de phylogénie (Figure 26) ou de symptômes.

Le groupe *Erwinia-Pantoea* se distingue par des mécanismes d'infection des végétaux relativement similaires. Pour les deux genres, les bactéries sont capables de se multiplier dans les tissus vasculaires, les bouchant alors et causant des symptômes variés selon la plante et l'espèce bactérienne [97,98]. Parmi les espèces les mieux décrites, *Pantoea stewartii* cause le flétrissement bactérien du maïs (maladie de Stewart) et *P. agglomerans* la galle des racines et du collet de la betterave [99,100]. *P. ananatis* possède quant à lui un spectre d'hôtes plus large : Eucalyptus, maïs, oignon, canne à sucre, melon… [97]. Certaines souches sont aussi des pathogènes opportunistes humains [99], comme *P. agglomerans* qui peut causer des septicémies potentiellement mortelles chez l'homme [101]. Du côté des *Erwinia*, *E. amylovora* et *E. pyrifoliae* causent le feu bactérien [98,102], symptôme considéré comme typique des infections à *Erwinia* bien que certaines autres espèces moins bien décrites causent des symptômes différents (*E. papayae*, par exemple, cause le chancre bactérien de la papaye [103]). A l'inverse de *Pantoea*, aucun *Erwinia* n'a de capacité connue à infecter les animaux. Ce groupe a donc développé des mécanismes de virulence reposant sur une capacité à survivre à l'intérieur du végétal hôte plutôt que l'utilisation de nombreux facteurs d'agression extracellulaires comme chez *Pectobacterium* et *Dickeya*.

Enfin, le groupe PBSLD comprend de nombreuses bactéries pectinolytiques. *Samsonia* (représenté par la seule espèce *S. erythrinae*) est faiblement pectinolytique et a été identifié comme pathogène de l'arbre erythrina, causant une défoliation et des lésions dans l'écorce [104]. Les mécanismes détaillés d'infection et les facteurs de virulence impliqués restent encore inconnus. *Lonsdalea* (représenté par la seule espèce *L. quercina)* cause le feu bactérien du chêne [100]. Malgré la similarité avec les *Erwinia*, le mécanisme d'action semble différent puisqu'il

dépend en partie d'une protéine à activité pectate lyase, HrpW [105]. *Brenneria* est surtout connu à travers l'étude de *B. salicis*, l'agent de la coloration bactérienne du saule. Cette espèce dégrade les parois des vaisseaux de l'arbre grâce à une activité cellulase, ce qui cause des fuites de sève [106]. Enfin, *Pectobacterium* et *Dickeya* sont très proches dans leurs modes d'infection. Les deux genres sont capables de causer la pourriture molle dans plusieurs végétaux grâce à la dégradation des parois végétales par une variété d'enzymes, majoritairement des pectinases [107]. Les différences entre espèces résident principalement dans les hôtes cibles et les conditions d'infection [108]. Par exemple, *D. dadantii* a un spectre large d'hôte (pomme de terre, tomate, endive…) alors que *P. atrosepticum* semble être optimisé pour l'infection de la pomme de terre. Cette différence mérite toutefois d'être nuancée car une nouvelle espèce de *Dickeya*, *D. solani*, s'attaque aussi majoritairement à la pomme de terre [109]. Un autre paramètre permettant de différencier les deux genres est la température à laquelle se développent les symptômes : 15-25° pour *Pectobacterium*, 30-35° pour *Dickeya* [110]. L'utilisation, entre autres, des capacités pectinolytiques ou cellulolytiques a donc permis aux bactéries de ce groupe de coloniser des plantes différentes avec des symptômes uniques et des facteurs d'agression dirigés contre les végétaux. En contrepartie et bien qu'elles gardent de nombreux gènes communs avec les entérobactéries zoopathogènes, elles ont totalement perdu leur capacité à s'associer avec des hôtes animaux comme *Enterobacter* ou *Pantoea*, ce qui en fait des pathogènes végétaux spécialisés.

Peu de données sont en général disponibles sur les mécanismes de virulence des entérobactéries phytopathogènes. *D. dadantii*, en revanche, a été particulièrement bien caractérisé et a permis de commencer à comprendre la manière dont ces bactéries se sont adaptées à l'hôte végétal en associant des gènes ancestraux adaptés aux hôtes animaux, des transferts horizontaux et des innovations évolutives adaptées à l'hôte végétal.

2. Les facteurs d'agression de *D. dadantii*

A l'inverse d'*Enterobacter*, *Pantoea* et *Erwinia*, les *Dickeya* possèdent un ensemble de facteurs d'agression sécrétés spécialisés dans la dégradation des parois végétales. Les parois végétales sont constituées majoritairement de fibres de cellulose reliées entre elles par l'hémicellulose et la pectine. *Dickeya* (comme *Pectobacterium*) s'attaque principalement à la pectine grâce à un ensemble d'enzymes agissant en synergie.

La pectine est principalement un polymère d'acide galacturonique portant des ramifications. Un premier ensemble d'enzymes comprenant les estérases PaeY (pectine acetylesterase) et PemA (pectine methylesterase) vise à enlever ces ramifications pour permettre l'action efficace des autres enzymes [111]. Le polymère linéaire est ensuite dégradé par plusieurs pectate lyases (9 extracellulaires chez *D. dadantii*) ainsi que des enzymes mineures comme des pectine lyases ou des pectate hydrolases. Ces enzymes extracellulaires sont sécrétées par deux systèmes de type 2 : Stt (qui sécrète uniquement la pectine lyase PnlH) et Out (qui sécrète toutes les autres enzymes) [107]. On estime que la diversité de pectate lyases est à l'origine du spectre d'hôtes large de *D. dadantii*, lui permettant de faire face à des structures de pectine différentes [112]. La dégradation se poursuit alors intracellulairement (Figure 29).

Les pectate lyases sont des acquisitions plus ou moins récentes par transfert horizontal. Par exemple, *pelB* et *pelC* sont présents chez *Pectobacterium* et *Dickeya* alors que *pelA*, *pelE* et *pelD* n'existent que chez *Dickeya* [113]. Les origines et l'évolution de ces gènes restaient cependant mal comprises et ont été au cœur du 3ème article.

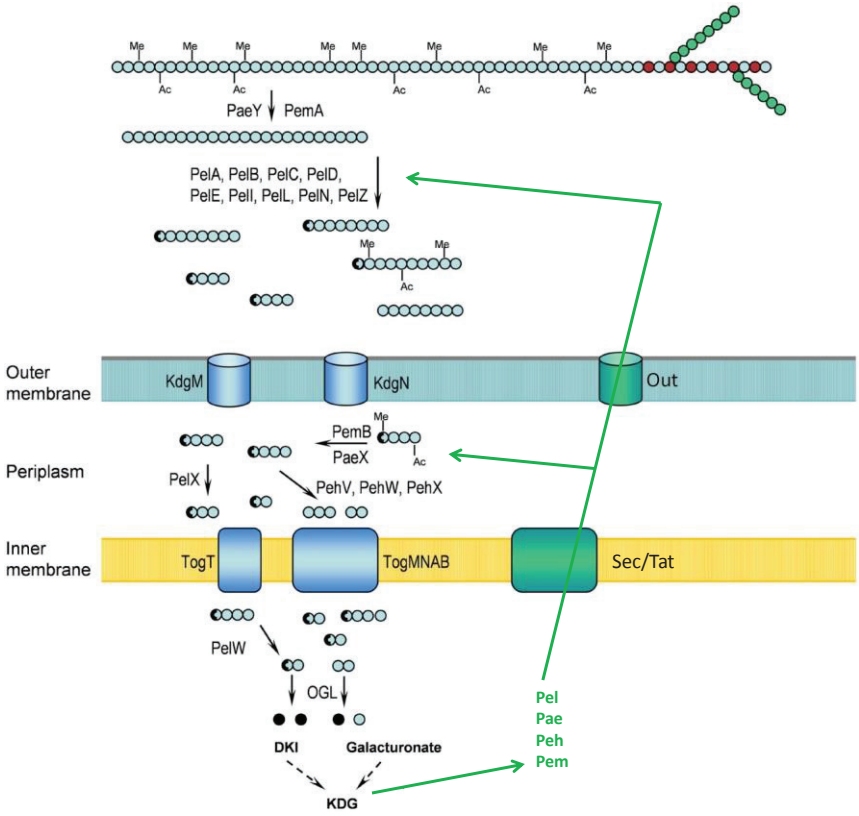


Figure 29 : Dégradation extracellulaire et intracellulaire de la pectine

Les groupements méthyl et acetyl sont clivés par les estérases PaeY et PemA, ce qui permet la dégradation du polymère par les pectate lyases extracellulaires (Pel). Les oligomères sont ensuite importés dans le périplasme par les porines KdgM et KdgN. Un deuxième ensemble d'enzymes comprenant notamment les pectine hydrolases (Peh) continue la dégradation. Enfin, Les oligomères sont importés dans le cytoplasme par le système Tog où la dégradation en monomère s'effectue et produit du KDG. Le KDG sert alors d'inducteur de la production de ces enzymes, créant une boucle de rétroaction positive (en vert sur la figure). D'autres voies non représentées ici permettent de dégrader le rhamnogalacturonane (chaîne bleue et rouge) et les chaînes latérales (vert). Adapté d'après [112].

La production des enzymes extracellulaires est finement régulée pour assurer la réussite de l'infection. La stratégie d'infection de *D. dadantii* est en effet de commencer par une phase asymptomatique, en se multipliant dans les espaces intercellulaires, jusqu'à ce que de bonnes conditions d'infection soient réunies [114]. Ensuite, *D. dadantii* induit soudainement et massivement la production des pectinases pour dégrader les cellules végétales avant qu'une réponse efficace de la planté hôte soit mise en place [115]. Cette régulation s'effectue grâce à de nombreux facteurs de transcription et NAPs dont certains seront détaillés dans la section C, ainsi que par des mécanismes post-transcriptionnels qui restent en majorité à comprendre ([107] et résultats non publiés).

Bien que les pectinases soient les plus nombreuses et les plus importantes dans la virulence de *D. dadantii*, deux autres types d'enzymes extracellulaires sont présents : 4 protéases et une cellulase [116]. Aucun rôle majeur dans la virulence de *D. dadantii* n'a pour l'instant été établi pour ces enzymes.

3. Les facteurs associés à la virulence de *D. dadantii*

D. dadantii possède de nombreux systèmes pour assurer sa survie dans la plante. Beaucoup sont des mécanismes généraux communs chez les entérobactéries (flagelle, chemotaxisme, enzymes antioxydantes, capture du fer… [116]) mais certains se distinguent par leur spécificité d'action ou leur évolution et vont être présentés ici.

Table 1 : Exemple de facteurs associés à la virulence chez *D. dadantii*

Fonction	Gènes	Présence dans les génomes	Ref
Formation de biofilms (cellulose)	*bcs*	Protéobactéries (mal conservé)	[117]
Flagelle	*fli*	Protéobactéries et Firmicutes (bien conservé)	[118]
Chemotaxisme	*che*	Protéobactéries et Firmicutes (bien conservé)	[119]
Compétition inter-espèces (système de sécrétion type VI)	*rhs*	*Dickeya* et *Pectobacterium*	[120]
Résistance au stress oxydant (pigment)	*indA*	*Dickeya*, *Serratia*, *Streptomyces*	[121]
Quorum sensing	*vfm*	Unique chez *Dickeya*	[122]
Inhibition des défenses végétales	*hrp*	β et γ-protéobactéries (phytopathogènes uniquement)	[123]

D. dadantii possède et utilise un système de sécrétion de type III pour inhiber les réactions de défense de la plante : Hrp [124]. Ce système, impliqué dans la virulence mais non essentiel, possède peu de molécules sécrétées connues, dont la protéine hairpin HrpN, qui contribue à l'induction de la réponse hypersensible de la plante [107,123,124]. Hrp est répandu chez les γ-protéobactéries phytopathogènes, même celles évolutivement éloignées, mais absent des autres γ-protéobactéries, ce qui indique des transferts horizontaux nombreux entre ces bactéries associées aux végétaux. Un système Hrp fonctionnel et impliqué dans la virulence a par exemple été décrit chez *Erwinia* [98], *Pseudomonas* [125] et *Xanthomonas* [126] et même chez les β-protéobactéries *Ralstonia* [127].

Comme de nombreuses bactéries, *D. dadantii* coordonne le processus infectieux à l'échelle de la population via l'utilisation de systèmes de communication interbactériennes (quorum sensing). Chez *Pectobacterium*, le contrôle de la virulence par quorum sensing est assuré par le système ExpI/ExpR [128], un système relativement répandu (sous d'autres noms) chez d'autres genres d'entérobactéries : *Yersinia*, *Pantoea* et *Erwinia*. Chez *D. dadantii*, ExpI/ExpR est fonctionnel mais n'intervient pas dans le contrôle de la virulence [122]. Le système de quorum sensing principal de *D. dadantii* est Vfm (codé par un ensemble de 24 gènes colocalisés organisés en 4 opérons), un système que seul les *Dickeya* possèdent et dont la molécule signal reste inconnue [122]. Ce système semble donc avoir été acquis par transfert horizontal mais son origine et l'intérêt d'avoir remplacé le système ancestral ExpI/ExpR par Vfm restent inconnus.

C. Les régulateurs principaux de la virulence de *Dickeya dadantii* : fonctions et évolution

Les régulateurs de la virulence de *Dickeya dadantii* sont nombreux et variés. Dix sont actuellement considérés comme majeurs (KdgR, PecS, PecT, CRP, H-NS, FIS, Fur, GacA, SlyA et MfbR [107]) mais de nouveaux sont régulièrement découverts. Des résultats non publiés de l'équipe indiquent par exemple que HFQ, ProQ et IHF sont indispensables à la virulence de *D. dadantii*. Je présenterai ici les trois (KdgR, CRP, FIS) qui ont été au centre de mes travaux ainsi que PecS et PecT, qui ont une histoire évolutive particulièrement intéressante. Ces cinq régulateurs présentent des fonctions et des histoires évolutives variées selon les espèces, qui sont résumées en Figure 30 et détaillées par la suite.



Figure 30 : Présence et conservation des régulateurs principaux de la virulence de D. dadantii chez d'autres espèces bactériennes.

Différents genres bactériens sont indiqués à gauche et positionnés sur un cladogramme établi à partir des données de la littérature. La présence de différentes fonctions associées à ces régulateurs est réprésenté par un code couleur dans le tableau. Un carré noir signifie que le régulateur est absent du génome considéré, vert que le régulateur est présent et a la fonction indiquée, jaune que la fonction indiquée est supposée par homologie, rouge que la fonction est absente et blanc que le régulateur présent mais qu'aucune donnée n'est disponible sur cette fonction.

1. KdgR

KdgR est un régulateur majeur de la virulence chez *Dickeya dadantii*. Il s'agit d'un répresseur qui contrôle directement environ 1% des gènes de cette bactérie. Les produits des gènes cibles de KdgR sont quasi-exclusivement liés à la dégradation de la pectine : enzymes extracellulaires (dont la majorité des Pel), composants du système de sécrétion Out, transporteurs de sucres [129]… En l'absence de pectine, KdgR se fixe sur l'ADN au niveau des séquences promotrices et empêche la transcription en général par occlusion du site de fixation de l'ARN polymérase [129,130], bien que *pelD*, qui a fait l'objet du deuxième article, soit un contre-exemple. La voie de dégradation de la pectine produit un intermédiaire, le 2-céto 3-deoxygluconate (KDG), qui se fixe sur KdgR et empêche son interaction avec l'ADN [111]. KdgR est donc l'acteur central de la réponse à la présence de pectine. Cependant, cette réponse n'est pas totalement abolie dans un mutant *kdgR* et d'autres voies pour le moment inconnues sont suspectées [111].

KdgR est présent essentiellement chez les γ-protéobactéries (*Salmonella*, *Escherichia*, *Xanthomonas*, quelques *Vibrio*… [129]). Chez *Pectobacterium,* KdgR remplit un rôle très similaire à celui de *Dickeya* : le contrôle de la production des enzymes extracellulaires en réponse à la présence de pectine dans le milieu, incluant même les cellulases et protéases (contrairement à *D. dadantii*). Le contrôle sur ces deux classes d'enzymes se fait majoritairement de manière indirecte et au niveau post-transcriptionnel via le système Rsm [131]. Chez *Xanthomonas oryzae*, un régulateur nommé KdgR contrôle le système de sécrétion de type III *hrp* mais pas les enzymes extracellulaires [132], contrairement à ce qui est observé chez *Dickeya* et *Pectobacterium*. Cependant, la protéine de *Xanthomonas* a une trop faible similarité avec celle de *D. dadantii* (20 % d'identité) et d'autres genres pourtant plus proches que *Dickeya* tels que *Pseudomonas* n'ont pas KdgR, ce qui met en doute son statut d'orthologue de KdgR. Il s'agit probablement d'un autre régulateur de la grande famille IclR à laquelle appartient KdgR, mais pas d'un orthologue de KdgR.

De manière plus surprenante, KdgR peut aussi être un régulateur important chez les entérobactéries zoopathogènes. Ces bactéries sont incapables de générer du KDG, permettant de lever la répression par KdgR. Pendant longtemps, il a été assumé que cette répression permanente exercée par KdgR constituait un « artefact » évolutif et n'apportait pas d'avantage sélectif à ces bactéries [129]. Cependant, il a été montré que *S. enterica* était capable de survivre efficacement dans la pourriture molle induite par *Dickeya* [133] ou *Pectobacterium* [134], ce qui laissait supposer un usage pour cette voie incomplète de dégradation de la pectine pour la vie saprophytique de *S. enterica.* KdgR a finalement été identifié comme l'un des régulateurs impliqués dans cette survie hors de l'hôte de *S. enterica* [135].

2. CRP

CRP (Cyclic Amp Receptor Protein) est, chez *D. dadantii*, un facteur de transcription clé dans la répression catabolique et l'activateur principal des gènes de virulence.

Au niveau de la répression catabolique, beaucoup de données sur ce régulateur ont été obtenues chez *E. coli,* mais les phénotypes observés et le fonctionnement de CRP est bien conservé entre *E. coli* et *D. dadantii* [111,136]. CRP est un dimère capable de fixer deux

molécules d'AMP cyclique (AMPc). En présence des deux molécules, la conformation de CRP change et la protéine acquiert la capacité de se fixer sur l'ADN [137]. La régulation de l'activité de CRP se fait donc majoritairement au niveau post-traductionnel, via le contrôle des niveaux d'AMPc.

Chez *E. coli*, l'AMPc est généré exclusivement par CyaA, une adenylate cyclase cytoplasmique. Cette dernière est activée en l'absence de glucose dans le milieu par une cascade de réactions présentées en Figure 31.

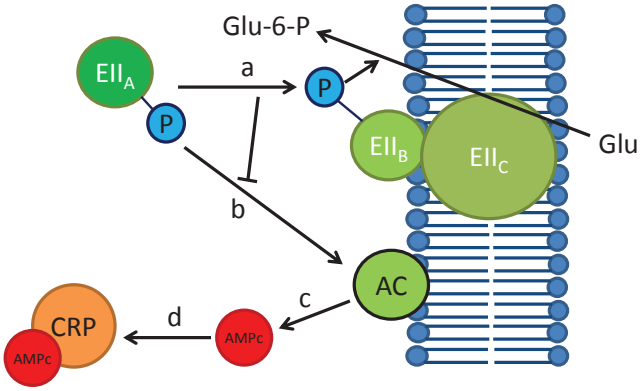


Figure 31 : Cascade de production de l'AMPc chez *E. coli*.
L'enzyme 2 (EII), chargée de l'importation et de la phosphorylation du glucose possède trois domaines A, B et C. Lorsque du glucose (Glu) est présent dans le milieu, le domaine A transfère son phosphate au domaine B, qui lui-même le transfère au glucose (a) pour produire du glucose-6-phosphate (Glu-6-P). En l'absence de glucose, le domaine A phosphorylé active l'adenylate cyclase CyaA (AC) (b) qui produit de l'AMPc (c). L'AMPc se lie alors à CRP (d), rendant la protéine active. Adapté de [138].

La dégradation de l'AMPc est assurée également par une seule phosphodiestérase, CpdA, qui transforme l'AMPc en AMP. Peu de choses sont connues sur la régulation de son activité, mis à part une activation par le fer [139], et il est communément supposé que la régulation de la concentration en AMPc s'effectue majoritairement au niveau de sa synthèse [138].

Une analyse transcriptomique du régulon CRP a été effectuée chez *E. coli.* Les cibles de CRP ainsi identifiées sont principalement des gènes liés au métabolisme. CRP active la production d'ARNt, des protéines ribosomales et de FIS, qui est lui-même un activateur de la synthèse de protéines (voir section 3-FIS (1er article)) ainsi que les gènes du métabolisme du carbone et des gènes de réponse au stress [140]. Le nombre de gènes cibles de CRP et l'existence de fixations peu spécifiques à l'ADN jouant un rôle principalement architectural ont amené certains auteurs à considérer CRP comme une NAP [141].

CRP a été initialement étudié pour sa fonction activatrice, avec deux mécanismes possibles. Le premier, de classe I, requiert un contact avec le domaine C-terminal de la sous-unité α de l'ARN polymérase. Cela requiert la présente du site CRP du même côté de la double hélice que l'ARN polymérase, c'est-à-dire avec un nombre de liaisons (Lk) entier entre les sites de fixation de CRP et le site d'initiation de la transcription [137]. En pratique, les distances fréquemment constatées sont d'environ 71, 82 ou 93 bp, avec un maximum connu de 161.5 bp [142]. Le second, de classe II, utilise un contact entre à la fois les domaines N et C terminaux de la sous-unité α de l'ARNP, et n'est possible qu'avec une distance d'environ 41.5 bp entre le site CRP et le

site d'initiation de la transcription [143]. CRP est également capable de courber l'ADN avec un angle de 90° à 130°[137], ce qui peut être utilisé dans des mécanismes d'activation comme celui de *malK* (impliqué dans le catabolisme du maltose) où la courbure induite par CRP permet de repositionner l'activateur MalT [144]. Sur les gènes *pel* étudiés, *pelD* et *pelE*, les deux mécanismes (classe I et II) sont présents (Figure 32).

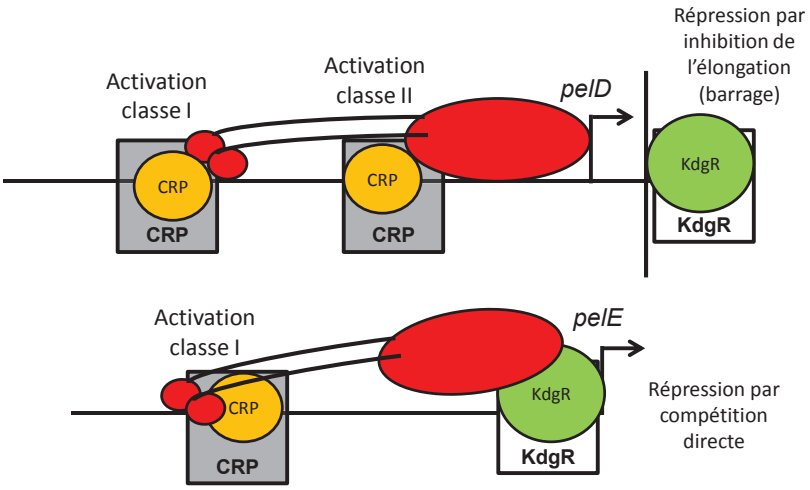


Figure 32 : Mécanismes régulateurs connus de CRP et KdgR sur les régions régulatrices *pelD* et *pelE*.
Sur *pelD*, deux sites CRP activent la transcription de *pelD*, l'un par un mécanisme de classe I et l'autre par un mécanisme de classe II. KdgR réprime l'élongation du transcrit *pelD* par barrage (Figure 7). Sur *pelE*, le seul site CRP agit par une activation de classe I et KdgR agit par compétition directe avec l'ARN polymérase.

Ainsi, chez *D. dadantii*, CRP est également un régulateur central de la virulence. L'augmentation des niveaux d'AMPc indique l'épuisement des ressources en sucres et ainsi la nécessité d'initier la dégradation de la pectine. CRP régule donc la majorité des gènes des voies de dégradation de la pectine, aussi bien extracellulaires (*pel*, *pemA*) qu'intracellulaires (*ogl*, *kdu*, *kdg*) [136]. Les mécanismes d'action de CRP décrits jusqu'à présent sont variés : régulation directe sans compétition connue (*pelB*, *pelC*, *pelE*), avec compétition avec KdgR (*ogl*, *kdgT*), avec compétition avec PecS (*pelD*) ou par coactivation avec un régulateur encore inconnu (*kdul*) [145,146].

Bien que les mécanismes d'activation de CRP soient les plus documentés, CRP peut aussi être un répresseur. Il réprime notamment le gène *crp*, avec un mécanisme complexe discuté précédemment dans la partie « Interactions entre des promoteurs proches », et *cyaA*, par une compétition directe avec l'ARN polymérase [137]. Ces gènes ne sont pas les seuls et de nombreuses enzymes sont aussi réprimés par CRP [140].

Bien que la répression catabolique soit un mécanisme presque universel chez les bactéries, CRP n'est ni toujours présent, ni forcément impliqué lorsqu'il est présent. CRP est présent dans la majorité des γ-protéobactéries mais ne semble pas exister dans les autres phyla. Chez *B. subtilis* et la majorité des Firmicutes, c'est CcpA qui assure un rôle similaire à CRP dans la répression catabolique [138]. Chez *Xanthomonas campestris*, CLP (orthologue de CRP) n'influence pas la répression catabolique mais contrôle la virulence [147]. Les *Pseudomonas* possèdent aussi un orthologue de CRP non impliqué dans la répression catabolique. Ils assurent leur répression catabolique par d'autres voies, notamment via Crc qui agit au niveau post-transcriptionnel et est

incapable de se lier à l'ADN [148]. Enfin, les *Vibrio* sont dans une situation intermédiaire où CRP contrôle la répression catabolique et la virulence mais n'a pas nécessairement besoin d'AMPc [149].

L'utilisation de CRP par les *Pseudomonas* est un cas très intéressant. Chez *Pseudomonas aeruginosa*, l'orthologue de CRP a été appelé Vfr (Virulence Factor Regulator), ce qui reflète les différences majeures dans le fonctionnement du régulateur. Vfr, contrairement à CRP, peut agir indépendamment de l'AMPc [150]. De plus, la production d'AMPc est assurée, en plus de CyaA, par une seconde adenylate cyclase, CyaB, qui est membranaire et qui est un senseur de calcium [151]. CyaB est même plus important que CyaA pour la virulence de *P. aeruginosa* [151]. Il en résulte une déconnexion entre Vfr/CRP et la répression catabolique, transformant Vfr en régulateur de la virulence dédié, capable de réguler la production de toxines, du système de sécrétion de type III et du quorum sensing [150]. Chez *Pseudomonas putida*, la situation est encore différente. Comme chez *E. coli*, CyaB est absent et l'AMPc est systématiquement requis pour l'action de CRP, mais ce dernier ne contrôle pas les gènes du métabolisme et seules quelques fonctions cibles mineures ont été identifiées [152].

En conclusion, la protéine CRP pourrait avoir d'abord été un régulateur de la virulence avant de devenir également un régulateur du métabolisme[153]. *D. dadantii*, comme de nombreuses autres entérobactéries [154,155], exploite cette régulation double virulence-métabolisme.

3. FIS (1er article)

Le rôle comparé de Fis dans la virulence des bactéries, se concentrant sur les modèles *Salmonella enterica* et *Dickeya dadantii*, a fait l'objet d'une revue parue en 2014 dans Trends in Microbiology [71]. L'article est retranscrit ci-dessous.

Bacterial virulence and Fis: adapting regulatory networks to host environment

Alexandre Duprey, Sylvie Reverchon and William Nasser

Université Lyon 1 F-69622 Villeurbanne; INSA de Lyon F-69621 Villeurbanne; CNRS UMR5240 Microbiologie, Adaptation et Pathogénie; France

Keywords: nucleoid-associated protein, virulence, Fis, DNA supercoiling, *Dickeya*, *Salmonella*

Running title: Fis and bacterial virulence regulation

Corresponding author: William Nasser

e-mail: William.nasser@insa-lyon.fr

Tel: (33) 472432695

Fax: (33) 472432686

Abstract

Pathogenic bacteria have to cope with adverse conditions, such as the host environment and the host defense reactions. To adapt quickly to environmental changes, pathogens have developed complex regulatory networks which ensure an adequate expression of their virulence genes. Recent evidence suggests that Fis, an abundant nucleoid-associated protein transiently produced during early exponential growth, plays a major role in these networks in several pathogenic bacteria. This review focuses on two enterobacteria, *Salmonella enterica* and *Dickeya dadantii,* which inhabit distinct ecological niches, to illustrate how Fis uses different strategies to coordinate virulence gene expression depending on the bacterial lifestyle.

Fis, a versatile regulator, optimizes the bacterial adaptive and virulence responses to a particular challenge.

Bacteria can colonize various environments and, consequently, adopt different lifestyles depending on their surroundings [156]. For example, they may live in planktonic form in liquid media or as biofilms on biotic or abiotic surfaces. In addition, bacteria often face sudden and brutal changes in these environmental conditions. An illustration of this is the situation encountered by the commensal gut bacterium *Escherichia coli* that, when excreted by a mammalian host, is confronted with a sharp decrease in temperature (37°C in the host and less than 25°C in the soil) and in nutrient availability [157]. Pathogens have to cope with many specific challenges, including acidic and oxidative stresses, evading or surviving the host defense reactions [158], accessing specific oligo-elements (notably iron [159]) and triggering a timely

attack on the host. The bacteria need to recognise these situations and react quickly and appropriately to adjust their genetic programs [160]. The production of virulence factors and the programs associated with virulence must be tightly coordinated to optimize the use of energy [161], the appropriate disease development and escape from the host defense system, culminating in the eventual dispersal of the bacteria. These considerations illustrate the need for genetic regulation to ensure that the appropriate virulence and acclimatization factors are produced at the right moment and at the level needed during infection. Most pathogenic bacteria have evolved sophisticated systems to detect hostile environments and trigger compensatory gene expression in order to survive within the host. A defect in any aspect of the fine-tuned molecular interactions leads to serious aberrant behaviour, or even cell death.

In bacteria, regulation of gene expression is exerted primarily at the level of transcription initiation using a large array of transcription factors (TFs) whose concentrations and activities change depending on specific environmental or metabolic signals. The most abundant regulators are the nucleoid-associated proteins (NAPs), coordinating the architecture and transcription of chromatin. Numerous TFs with more localised effects act upon the regulatory constraints imposed by a few highly abundant NAPs. In addition, bacterial DNA topology is highly responsive to environmental changes and, together with NAPs, it modulates the distribution of transcription machinery in the genome. This interplay of TFs and global constraints, imposed by DNA topology and NAPs, mediates the coordinated spatiotemporal expression of the different required virulence factors by sensing changes, in environmental conditions or the bacterial physiological states, at each stage of the infection process [160].

Recent findings have highlighted one of the NAPs, Factor for inversion stimulation (Fis), as a key regulator of virulence in several pathogenic bacteria such as pathogenic *E. coli* strains [162], *Shigella flexneri* [163], *Salmonella* [164,165], *Vibrio cholerae* [166] and *Dickeya dadantii* [167] (Table 2). Fis was initially characterized as a regulator of site-specific DNA recombination [168,169], but, since then, its role as a global regulator of growth, enhancing the synthesis of transfer and ribosomal RNAs, has become well established [170]. It has been found to influence the topological state of DNA in the cell by repressing DNA gyrase and stimulating topoisomerase I gene expression [171]. Fis can also affect DNA topology directly by binding to DNA. In particular, it acts to preserve the intermediate supercoiled forms of DNA, protecting them from extreme shifts towards the more relaxed or more negatively supercoiled ends of the topological spectrum. Thus, Fis is considered as being a local topological homeostat [38]. It is synthesized in large amounts during the early exponential phase of cells grown in rich medium but it becomes almost undetectable before the cells enter the stationary phase [172]. Hence, Fis is responsible for a sequential, or growth phase-dependent, regulation of its target genes. Considering its wide variety of mechanisms (direct [117] or indirect [173] interactions with RNA polymerase, interaction with other NAPs and TFs [165,174], local [175] or global [176] supercoiling modifications etc.), Fis appears to be a versatile regulator that is very efficient in controlling basic cellular processes and specific genetic programs such as virulence programs.

Table 2: Main virulence functions regulated by Fis in bacteria

Function	Involvement of Fis	Refs
Dickeya dadantii		
Adherence	Repression of cellulose synthesis during exponential growth phase	[117]
Motility	Activation of *fli* operon (flagella)	[167]
Survival to stress and manipulation of host defense response	Activation of *sap* (defense against antimicrobial peptides) and *hrpN* (effector of the type III secretion system)	[167]
Secretion	Activation of the Out type II secretion system	[167]
Production of plant cell wall degrading enzymes	Prevention of premature synthesis of plant cell wall degrading enzymes (pectate lyases and cellulases) secreted by the Out type II secretion system	[167]
Pathogenic *Escherichia coli*		
Adherence	Repression of curli synthesis during exponential growth phase	[177]
Multiple functions	Activation of *virF*, stimulating plasmid-related virulence	[163]
Production of toxin	Production of the plasmid-borne toxin Pet	[174]
Pasteurella multocida		
Survival	Activation of the production of capsule components, iron uptake and adhesion	[178]
Salmonella enterica		
Colonization of the host intestinal lumen	Activation of *hilA*, an activator of SPI-1-mediated adherence and penetration into gut epithelial cells	[48]
Survival to stress in macrophage vacuole	Activation of SPI-2-mediated intracellular survival in macrophages	[179]
Multiple functions	Activation of SPI-1 and SPI-2 type III secretion systems and motility-related genes (flagella and chemotaxis)	[53,164]
Shigella flexneri		
Multiple functions	Activation of *virF*, stimulating plasmid-related virulence	[180]
Vibrio cholerae		
Quorum sensing	Modulation of the production of HapR, a master regulator of quorum sensing	[166]

Considering the growing interest in understanding the concerted rearrangements of gene activities during the transition from saprophytic to pathogenic life style, this elucidation of the strategies used by the global regulators, NAPs, to coordinate virulence functions in species from different environmental niches could be a substantial contribution to a range of related fields concerned with the fundamental problems of bacterial gene regulation and its applied aspects (such as biotechnology and biomedicine).

The question we will explore in this review is that of how different pathogens, despite inhabiting distinct ecological niches and displaying different lifestyles, adapt their use of Fis to coordinate virulence programs. We will focus on the most frequently studied model organisms in this field, namely the extracellular plant pathogen *D. dadantii* and the facultative intracellular animal pathogen *Salmonella enterica*.

Fis plays a central role in the switch from the asymptomatic to the symptomatic phase in *D.*

dadantii.

D. dadantii (formerly *Erwinia chrysanthemi*) is a necrotrophic, Gram-negative plant pathogen that causes disease worldwide in a range of plant species, including many economically important vegetables such as potato, tomato and chicory [181]. It can also infect the model plant *Arabidopsis* [182]. It is also listed in the top 10 of the most important bacterial plant pathogens [183]. *Dickeya* was initially located in warm tropical regions, however, damage caused by *Dickeya* strains has recently increased in potato cultures in Europe and they are now considered emergent pathogens in this region [184]. Soft rot, the visible symptom, is mainly due to the production of degradative enzymes, mostly pectate lyases (Pels), that destroy pectin, an essential component of the plant cell wall [185,186]. However, an efficient colonization of the plant requires many additional factors including early factors (such as flagella, lipopolysaccharide, exopolysaccharide and cellulose synthesis), which permit adhesion of the bacteria to the plant, and acclimatization systems necessary for overcoming inhospitable conditions in the host (oxidative stress, acidic stress, iron starvation, and bactericidal compound) [117,119,187–189]. *D. dadantii* pathogenicity is clearly a multifactorial process.

The *D. dadantii* infectious cycle begins with its entry into the plant, via a wound or a natural opening (including the lenticel and stomata) [107]. Then, it settles and multiplies in the intercellular spaces, without causing any significant damage. In this early asymptomatic infection phase, the bacteria face several stresses, notably an acidic pH [190] but also oxidative stress [121] and low nutrient availability [107]. The production of aggression factors is low during this period in order to avoid induction of plant defenses until the bacterial density is sufficient for effective infection. The repression of energetically expensive Pel production also helps *D. dadantii* to devote energy instead to the acclimatization systems, which permit survival in hazardous conditions and, also, enable the bacteria to escape the host defense mechanisms. During the second stage, when the bacteria have successfully acclimatized to the intercellular plant environment and reached a high population density, *D. dadantii* produces and secretes massive amounts of aggression factors, mainly Pels [114]. The resulting destructuration of the plant cell wall causes the characteristic soft rot symptom of a *D. dadantii* infection. This complex host–pathogen interaction requires that the acclimatization systems, along with many additional virulence factors, are precisely controlled such that each is expressed at the appropriate time and place in the host. Fis plays a pivotal role in the transition from the asymptomatic to the symptomatic phase in *D. dadantii* by activating the early virulence factors and repressing the late virulence factors in a growth phase-dependent manner [116].

In the latent infection stage, Fis represses the production of cellulose, a major component of biofilms in *D. dadantii* [117] and activates flagella production to facilitate the planktonic life style and the colonization of the intercellular space of the host [117]. Fis also actives both the production of the Sap system, part of the *D. dadantii* defense mechanism against vegetal antimicrobial peptides that are encountered in plant intercellular spaces [191], and the Hrp type III secretion system, delivering effector molecules to the plant cells ([107], Lautier PhD thesis 2007, Université Claude Bernard Lyon, France). In addition, during this latent phase, Fis represses *pel* gene expression. These controls occur through a direct mechanism whereby Fis binds to the promoter regions of the operons mentioned [167,192].

During the transition to the symptomatic phase, a sudden and very high level expression of *pel* genes occurs. A complex regulatory system (reviewed in [107,116,185]) controls the production of Pels in *D. dadantii*. The growth phase and the presence of pectin have been

identified as the most important signals in this regulatory system [122,185,186]. Three main players are involved in these controls: (i) the KdgR protein that inhibits Pel production in the absence of pectin-breakdown products [146,186]; (ii) the NAP Fis that plays a key role in the growth phase-dependent regulation of *pel* genes by preventing their expression in the early stages of infection [167]; and (iii) the sugar catabolism regulator complex cAMP–CRP protein, which is the main activator of *pel* gene expression [145]. Fis collaborates with these two metabolite responding regulators, KdgR and CRP, to fine-tune the induction of disease by *D. dadantii* (Figure 33). The molecular mechanisms of this control involve a direct local action of these three regulators on transcription initiation, by RNA polymerase, at the *pel* promoters. During the early stages of infection, Fis, induced by the low cell density, binds to the *pel* promoter regions and hinders the CRP and RNA polymerase binding. On the contrary, the binding sites of Fis and KdgR are located in different areas so both proteins could bind simultaneously to the *pel* promoters. The effect of each individual protein, Fis or KdgR, on *pel* gene expression is lower than that observed in the presence of both proteins. Thus, Fis and KdgR cooperate to inhibit the activity of the CRP-RNA polymerase complex blocking *pel* gene expression [192]. As KdgR and Fis respond to different signals, this organization could allow for a gradual and coordinated de-repression of the *pel* gene during pathogenesis. In the initial steps of infection, the relatively low amount of pectin degradation products present in the plant intercellular spaces is not sufficient to alleviate KdgR repression and the *pel* genes are repressed by both Fis and KdgR. Progressively, the basal production of Pels leads to the initiation of pectin degradation, which in turn results in a more favorable environment for bacterial multiplication. As the cells enter in the transition from exponential to stationary growth phase, a strong decrease in Fis cellular content leads to a de-repression of *pel* gene expression. The concomitant increase in pectin-breakdown products completely removes the KdgR repression. These events result in a high-level production of Pels that leads to a strong attack on the host plant, and the consequent development of soft rot symptoms [167].

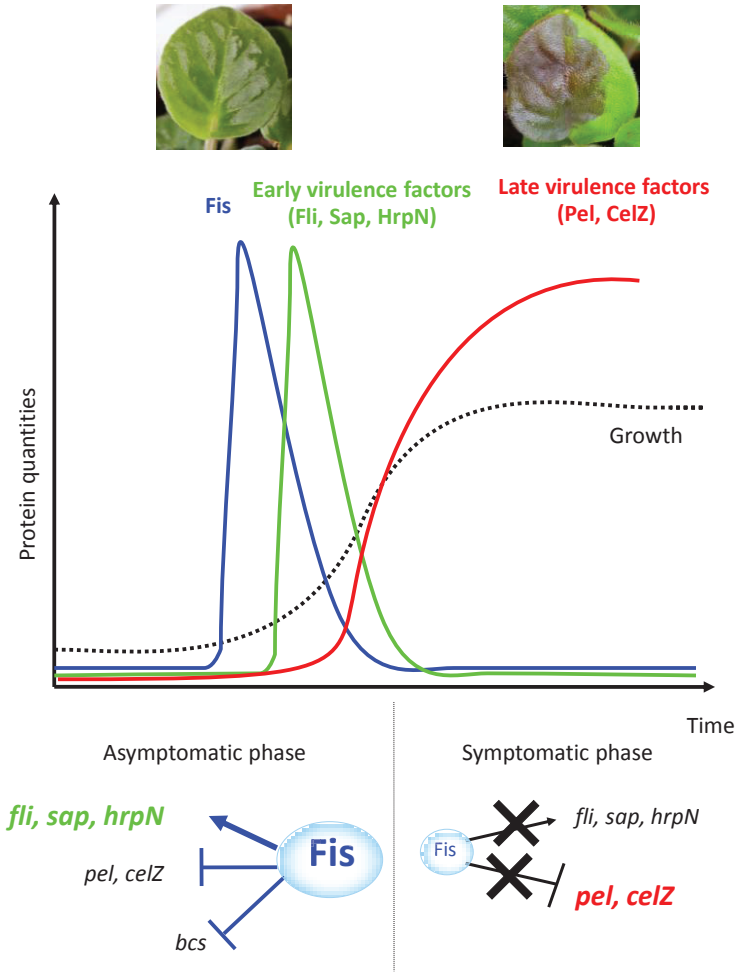


Figure 33: Fis coordinates the transition from the asymptomatic to the symptomatic phase in *D. dadantii*.
Upregulation is indicated by an arrow and downregulation by a line ending with a bar. Loss of regulation due to a decrease in Fis concentration is indicated by a cross on the corresponding arrow or line. In the asymptomatic phase, Fis, induced by the low bacterial density, mostly stimulates the production of motility machinery and additional factors required at this stage (Sap and type III secretion systems). At the same time, it represses the production of cellulose (encoded by the *bcs* operon), involved in the biofilm lifestyle, and the plant cell wall-degrading enzymes, mostly Pels and cellulase. The strong repression of *pel* genes prevent early detection by the host of pectin breakdown products (e.g., oligogalacturonides) that signal the presence of the pathogen before it reaches a population density appropriate for successful infection. As soon as the colonization of the host by *D. dadantii* is successful and bacteria enter into an active multiplication phase, the cellular concentration of Fis drops resulting in high-level production of the cell wall-degrading enzymes. This leads to a powerful attack on the host plant and the development of disease symptoms.

A second level of complexity is added by the intervention of DNA topology. Any relaxation of the DNA decreases the expression of *pel* genes [193]. In the *fis* mutant, repression of *pel* genes by DNA relaxation is increased two-fold during exponential growth, indicating a buffering role for Fis in *pel* promoter activity at a suboptimal superhelical density [192,193]. *D. dadantii* encounters both acidic and oxidative stresses, relaxing the DNA at an early stage in the course of infection. By buffering the impact of these two stresses, Fis might help *D. dadantii* to stabilize the basal level of Pel production, required to initiate the degradation of the pectin compounds. The impact of the topological buffering activity of Fis on the *pel* promoter is relatively limited as regards its direct repressive activity on *pel* gene expression [192,193]. In *D. dadantii*, the regulation of Pels

by Fis is, therefore, mostly influenced by physiological (growth phase) and metabolic stimuli rather than the environmental conditions (such as oxidative and acidic stresses). In the light of this finding, Fis can be considered as a global coordinator of virulence factors, establishing a balance between the two vital requirements of needing to remain hidden from the plant defense system during the asymptomatic phase and then efficiently initiating the degradation of plant cell walls in order to shift to the symptomatic phase.

S. enterica: overcoming the host defenses

S. enterica is a facultative intracellular enteric pathogen able to infect a wide variety of animals, including humans, mice and chickens. Many serovars have been described according to their host specificity, serovar Typhi (infects humans and other primates) and serovar Typhimurium (has a wide host range among mammals) are the most thoroughly studied due to their ability to infect humans. *S. enterica* enters the host by the oral route, survives the low pH of the stomach and, finally, reaches the intestines [194]. This bacterium colonizes the intestinal epithelia and then breaches the epithelial barrier by various means: either by uptake into epithelial cells, or traversing the M cells, or by phagocytosis [194]. This is the first important challenge encountered by *S. enterica*, since the intestinal epithelia are not meant to be crossed by bacteria.

Once the intestinal epithelia have been traversed, *S. enterica* is frequently taken up by immune cells, such as macrophages [194], which then attempt to degrade the bacteria. *S. enterica* uses several evasion mechanisms in order to survive intracellularly [195], multiply and, ultimately, escape by inducing apoptosis [196] or necrosis [197]. Surviving potential degradation by specialized cells, including macrophages, is the second key challenge *S. enterica* must overcome.

S. enterica virulence is multifactorial, like that of *D. dadantii*. However, it relies mostly on the horizontally-acquired *Salmonella* pathogenicity islands (SPI), more specifically the SPI-1 and SPI-2 [53,164]. The SPI-1 genes encode a type III secretion system, which is necessary for adhesion to epithelial cells and which secretes effectors causing cytoskeletal rearrangements in epithelial cells, facilitating subsequent engulfment and transport through the intestinal barrier [198]. HilA is a major regulator of the SPI-1 genes [199]. The SPI-2 genes encode another type III secretion system, which facilitates intracellular survival and multiplication by interfering with the host degradation pathways, as well as the two-component system SsrA-SsrB, which is one of the main regulatory system of SPI-2 [80].

A global transcriptomic study of Fis regulation [53,164]revealed that the most highly represented categories among Fis-regulated genes either belong to a SPI or are involved in motility, and that more than 90% of the genes in these categories were upregulated by Fis, independently of the infection stage they were involved in. This represents a striking contrast with *D. dadantii*, where upregulation and downregulation depends on whether the genes are required in the asymptomatic or the symptomatic phase of the disease.

The regulation of both SPI-1 and SPI-2 gene expression has been investigated in more detail and it has been shown that Fis activates their expression indirectly via the master regulators HilA and SsrB, respectively, and via the modulation of DNA supercoiling [179,200]. In the gut, a high supercoiling level in the nucleoid of *S. enterica* is induced by the anaerobic conditions and this

favors the expression of the SPI-1 genes. Fis stabilizes this high DNA supercoiling level in the gut and promotes SPI-1 gene expression. When entering the macrophages, the bacterial cells face oxidative stress, which relaxes their nucleoid and so facilitates the expression of the SPI-2 genes. During the initial period of macrophage infection, Fis enhances and accelerates this relaxation [48,179,200], allowing for the timely expression of the SPI-2 genes encoding essential intracellular survival factors. However, the Fis protein also prevents excessive DNA relaxation so that the *ssrA-ssrB* promoter remains active and allows SPI-2 gene expression to continue [48,179,200]. Hence, both SPI-1 and SPI-2 need the Fis protein for optimal gene expression in response to local changes in DNA topology [179]. In *S. enterica*, Fis participates in the transition from supercoiled to relaxed DNA levels when bacteria move from the gut lumen to the macrophage vacuoles, whereas in *D. dadantii* Fis mainly buffers any variations in DNA supercoiling.

In the light of these findings, Fis does not appear to act as a growth phase regulator in *S. enterica* but rather as an environmental regulator. The difference with *D. dadantii* may originate from the fact that bacterial density is not a key factor for the success of infection for *S. enterica.* Instead, the main challenges involve responding appropriately to different environments (epithelial cells, macrophages, and extracellular conditions) and their associated stresses. This is further supported by the presence of significant levels of Fis in the stationary phase in anaerobic conditions [56], reducing growth-phase dependent regulation in the gut.

As in *D. dadantii*, Fis regulates the virulence genes of *S. enterica* in interaction with other regulators, including NAPs (HU, H-NS and IHF [165,201]). Recently, it has been shown that Fis acts in concert with the master stress-response regulator, OmpR, to regulate the expression of SPI-1 and SPI-2 in response to DNA topology variations [176,202]. OmpR is an activator that preferentially binds to relaxed promoters while Fis binds preferentially to supercoiled DNA. During entry into the macrophages, Fis promotes DNA relaxation and increases the OmpR-mediated activation of *ssrA*. OmpR binding to relaxed DNA generates a locally supercoiled state that enhances Fis binding [176]. As a consequence, *ssrA* is synergistically activated by both Fis and OmpR, resulting in a further increase in SPI-2 gene expression. Conversely, OmpR and Fis display antagonist effects on the expression of the SPI-1 regulator HilD. Fis activates *hilD* expression and OmpR represses *hilD* expression [176,202]. The high supercoiling state of DNA in the gut decreases the binding of OmpR, relieves repression and facilitates activation by Fis [176,202]. This example shows the pivotal role played by Fis and DNA topology together in virulence regulation in *S. enterica* (Figure 34), in contrast to the observations made in *D. dadantii.*

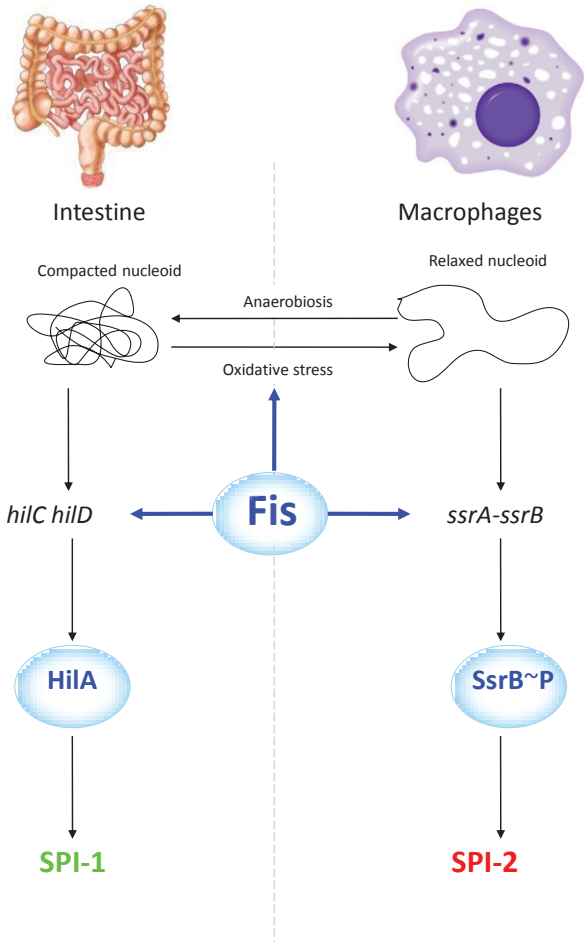


Figure 34: Control of *S. enterica* virulence by the DNA supercoiling state and Fis.

Up-regulation is indicated with an arrow. Only key regulators in the Fis control of SPI-1 and SPI-2 gene expression are represented. In the lumen of the gut, bacteria experience environmental conditions (anaerobiosis and high osmolarity) that increase DNA supercoiling. By stabilizing the compaction of the bacterial nucleoid in these early stages of infection, Fis induces the expression of the SPI-1 genes encoding lumen invasion effectors via a cascade reaction involving the SPI-1 activators (e.g. HilA, HilC and HilD). In contrast, when entering the macrophages, the cells face oxidative stress that relaxes DNA. By enhancing the relaxation of DNA whilst the bacteria are being engulfed by macrophages, Fis stimulates the production of the two component regulator SsrB which results, in turn, in the activation of the SPI-2 gene expression, encoding essential intracellular survival factors [176,179].

As the SPI-1 promoters are activated in high supercoiling conditions and the SPI-2 promoters in low supercoiling conditions, their expression should be mutually exclusive. However, there is evidence that several SPI-1 genes are activated in macrophages to mediate their apoptosis, thus liberating the bacteria from the infected macrophages [198]. Furthermore, some of the SPI-2 genes are primed whilst crossing the gut epithelia [201] and Fis and OmpR participate directly in this priming [201]. Consequently, Fis is involved not only in the adaptation to a given environment, but also in the anticipation of new environments, a widespread strategy among micro-organisms which allows them to adapt efficiently to sudden but predictable environmental changes [203].

Towards the evolutionary mechanisms of Fis regulation

There is now growing evidence that the regulation of bacterial virulence involves more than just a network of transcription factors. NAPs and DNA topology both have a role to play, to differing extents depending on the species, as illustrated by the example of Fis. This review, as summarized in Figure 35, shows that *S. enterica* and *D. dadantii* use Fis in different ways to regulate their main virulence functions. *D. dadantii* uses Fis primarily as a growth phase regulator, allowing a rapid transition between its asymptomatic phase and its symptomatic phase in order to overwhelm the plant defense systems. *S. enterica* uses Fis mainly as an environmental sensor, coupled with DNA topology, to detect the transition between extracellular and intracellular lifestyles and influence the binding and the expression of other regulators of its horizontally-acquired pathogenicity islands.

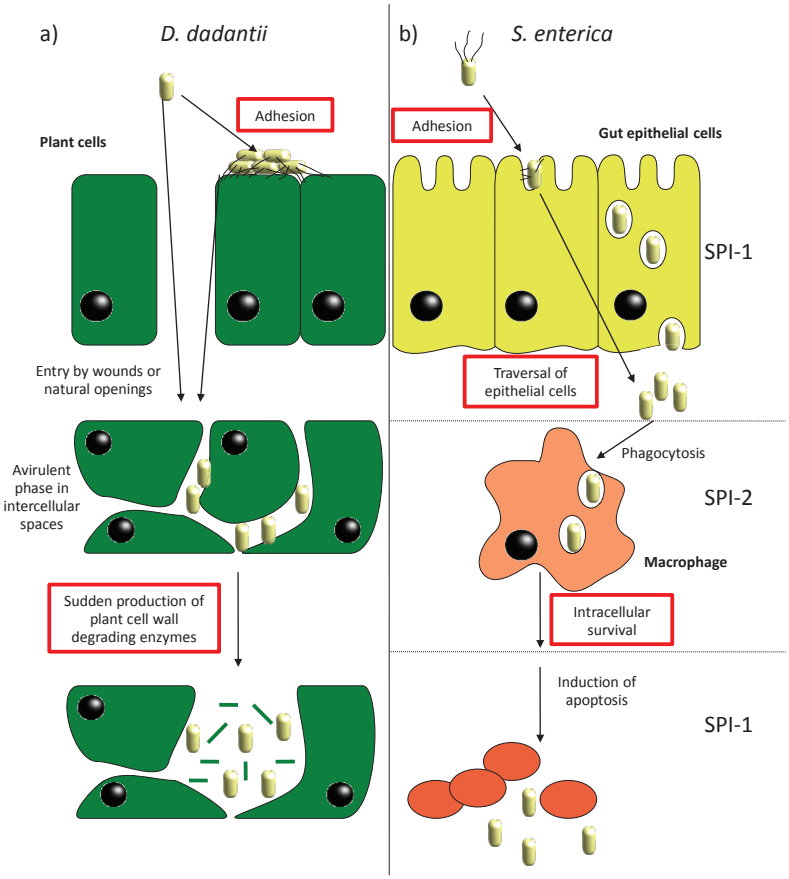


Figure 35: Comparative diagram of the functions regulated by Fis in *S. enterica* and *D. dadantii.*

Key stages of infection of both bacteria are indicated. Stages regulated by Fis are boxed in red. (a) The extracellular plant pathogen D. dadantii uses Fis to switch rapidly from the asymptomatic to the symptomatic phase in order to efficiently degrade the plant cell walls and invade the host plant. (b) The intracellular pathogen S. enterica uses Fis to cross the epithelial barrier and to escape from the host immune system, mainly by surviving in macrophage cells and infecting them.

This shows that each bacterium, although taxonomically related (both belong to the Enterobacteriaceae family), has adapted the use of Fis to their specific environmental challenges in order to control their virulence programs. Molecular mechanisms are increasingly well understood in a few model species *(S. enterica*, *D. dadantii* and, to a lesser extent, pathogenic *E. coli*). The virulence factors regulated by Fis differ from one species to another, implying an evolutionary process selecting the target genes for Fis regulation. Thus, the interspecies differences in Fis usage lead to questions about the evolutionary mechanisms that enabled Fis to fulfill such diverse roles (Box 1). The answers are not yet available, given the relatively small

number of organisms investigated. Extending current studies on the role of Fis in coordinating adaptive functions in commensal bacteria, or symbiosis programs in symbiotic species, from a range of environmental niches may provide valuable insights regarding the functional adaptability of Fis. An example of this is a recent report on the involvement of Fis in the colonization of plant roots by *Pseudomonas putida* [204]. Furthermore, and more generally, extending the functional characterization of the NAPs to more species with different lifestyles is a prerequisite to elucidating the precision with which these global regulators modulate the functions involved in host–bacteria interactions.

Box 1 : Outstanding questions

How did Fis come to be associated with such diverse virulence or metabolic functions ?

Does the central role of Fis in virulence regulatory networks extend to symbiosis or commensalism ?

Mise à jour bibliographique de l'article

Depuis la publication de l'article, les mécanismes permettant l'expression de FIS en anaérobiose en phase stationnaire chez *S. enterica* ont été élucidés. Alors que le promoteur principal (*fisP1*) est similaire à celui de *E. coli*, c'est-à-dire exprimé durant la phase exponentielle de croissance, un promoteur tandem situé en amont de *fisP1* est activé en conditions d'anaérobiose [205]. Ce promoteur possède une séquence -35 faible, ce qui est cohérent avec son statut de promoteur inductible, et n'existe ni chez *E. coli*, ni chez *D. dadantii* (Figure 36). Ceci suggère que le mécanisme d'activation de *fis* par l'anaérobiose en phase stationnaire est spécifique de *S. enterica* mais le ou les régulateurs impliqués sont encore inconnus.

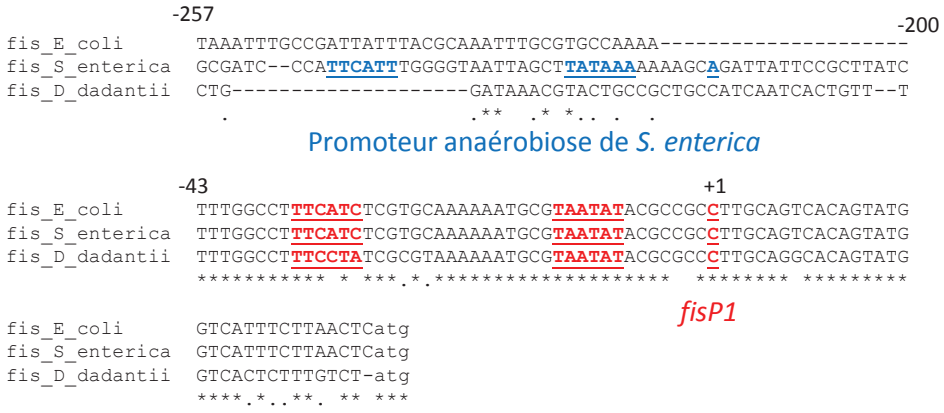


Figure 36 : Alignement des régions régulatrices de *fis* chez *E. coli*, *S. enterica* et *D. dadantii*

Les régions régulatrices ont été alignées par MAFFT puis les séquences -35, -10 et +1 présentées dans [205] ont été reportées pour le promoteur principal (fisP1, en rouge) et le promoteur activé en anaérobiose chez *S. enterica* (en bleu). Les positions des nucléotides sont indiquées par rapport à la région régulatrice de *S. enterica*. Par souci de simplicité, la région située entre -43 et -200 est omise de la figure.

Ces trois régulateurs (KdgR, CRP, FIS) mentionnés dans l'article et qui ont été au cœur de ce travail de thèse ont ainsi suivi un chemin évolutif assez similaire, avec une transmission verticale et une adaptation de leurs fonctions et cibles. Tous les régulateurs globaux de *D. dadantii* n'ont cependant pas suivi ce chemin, ce qui sera illustré au travers de 2 exemples : PecS et PecT.

4. PecS

PecS est un facteur de transcription dont le but est d'empêcher l'expression précoce de nombreuses fonctions de virulence [206]. Bien que ce rôle soit similaire à celui de FIS, un mutant *pecS* est hypervirulent, contrairement au mutant *fis*. Cette caractéristique pourrait être due au fait que, contrairement à FIS qui est un régulateur dual (activation des fonctions précoces et répression des fonctions tardives [71]), PecS agit essentiellement en tant que répresseur des fonctions tardives [187]. Les cibles de PecS sont variées. PecS réprime la production et l'export d'enzymes extracellulaires (gènes *out*, *pel*, *celZ*…), le quorum sensing (*vfm*), la mobilité (*fli*) et la producion d'indigoidine (*indA*) [187]. Au total, environ 140 gènes régulés (directement ou indirectement) par PecS ont été identifiés [187].

Le signal auquel répond PecS *in planta* est pour l'instant inconnu chez *Dickeya dadantii*. En revanche, en conditions de laboratoire, l'urate est capable de se lier au PecS de *Agrobacterium fabrum* et de le dissocier de l'ADN. Le domaine de fixation à l'urate est conservé chez l'orthologue de ce PecS rencontré dans différentes bactéries dont *Pectobacterium* et *Dickeya* [207]. L'urate est un antioxydant produit lors de l'infection à la fois par la plante et par certaines bactéries [208,209], ce qui laisse supposer que PecS sert de senseur du stress oxydant. Cette hypothèse est confortée par l'action répressive de PecS, dans des conditions de croissance standard, sur l'opéron *ind* codant pour le pigment bleu indigoidine impliqué dans la résistance au stress oxydant de *D. dadantii* (Table 1).

La distribution de PecS dans les génomes bactériens est assez surprenante (Figure 30). PecS est présent chez *Pectobacterium*, bien qu'il active la production des pectate lyases plutôt que de les réprimer [131], mais absent de nombreuses autres entérobactéries (*E. coli, S. enterica, Y. pestis, E. amylovora…*) et autres γ-protéobactéries (incluant les phytopathogènes *Xanthomonas*). *pecS* semble donc être une acquisition très récente par transfert horizontal, ce qui pose la question des mécanismes qui lui ont permis de prendre le contrôle de ses gènes cibles. Malgré cette absence quasi-généralisée de PecS chez les γ-protéobactéries, un xénologue est présent chez *Klebsiella pneumoniae.* Ce dernier régule la formation de biofilms par les fimbriae de type 1 (gènes *fim*) et répond également à la présence d'urate [210].

A plus longue distance, d'autres phytopathogènes possèdent un homologue de PecS. En plus de *Agrobacterium fabrum* (ATU0272, 45.7% identité) [211], on trouve également PecS chez *Ralstonia solanacearum* (RCFBP_21306, 37.5% identité). Il est intéressant de noter que, dans tous ces cas, *pecS* est colocalisé avec un gène codant pour un transporteur membranaire (*pecM* chez *D. dadantii*).

Enfin, PecS est également présent chez plusieurs *Streptomyces* où, à nouveau, il est impliqué dans la réponse au stress oxydant avec l'urate comme molécule signal [212]. De manière plus générale, PecS est très présente chez les actinobactéries. Par exemple, il existe également des homologues de PecS et PecM avec plus de 35% d'identité chez *Sanguibacter*, *Actinosynnema*, *Nocardiopsis*… En règle générale, la colocalisation *pecS-pecM* est toujours observée, ce qui laisse penser à une association ancestrale des deux gènes. De plus, chez les *Dickeya*, *pecS-pecM* sont

colocalisés avec les opérons *vfm* (quorum sensing) et *ind* (indigoidine), deux autres groupes de gènes acquis par transfert horizontal et régulés par PecS. Par ailleurs, PecM est impliqué dans l'export de l'indigoidine [213], ce qui laisse penser à un îlot génomique.

5. PecT

PecT est un facteur de transcription impliqué dans la thermorégulation de *D. dadantii* [214]. Les mutants *pecT* de *D. dadantii* et de *D. solani* sont hypervirulents du fait d'une production plus forte de pectate lyases [109]. Contrairement à KdgR et PecS, aucune étude globale des gènes cibles de PecT n'a été effectuée et peu de données sont donc disponibles sur ses cibles. Toutefois, des approches de génétique classique ont permis d'identifier un nombre relativement important de gènes régulés par PecT, incluant les gènes des enzymes extracellulaires (*pel*) [214], de la synthèse des EPS (*weaP*), de la mobilité (*fliC*), des cellulases (*celZ*) et des protéases (*prtC*) [215].

Pour les gènes *pel*, une étude plus poussée a été entreprise pour déterminer le mécanisme d'action de PecT. La fixation de PecT sur l'ADN dépend du niveau de surenroulement de l'ADN : PecT a une meilleure affinité pour l'ADN relâché. Or, à forte température, l'ADN se relâche et l'affinité de PecT pour les régions régulatrices *pel* augmente, ce qui permet de réprimer l'expression des *pel* à haute température [214]. De plus, l'expression de *pecT* est réprimée par H-NS [216], lui permettant de servir également d'intermédiaire dans la régulation des gènes *pel* par H-NS.

Le nom de la protéine PecT a été attribué suite à un crible génétique qui l'a identifiée comme répresseur de la synthèse des pectinases [217]. Cependant, ce nom est spécifique à *Dickeya* et PecT correspond à HexA chez *Pectobacterium* et LrhA chez la majorité des autres entérobactéries. PecT/LrhA a été caractérisé chez des bactéries ayant une variété d'hôtes.

En ce qui concerne les hôtes végétaux, chez *Pectobacterium*, PecT contrôle la production d'enzymes extracellulaires, le système de sécrétion de type III Hrp, le facteur σ alternatif RpoS et le quorum sensing (gène *ohlI*) [218]. Chez *Pantoea stewartii*, il joue un rôle dans la virulence mais la majorité des cibles de LrhA sont des gènes de fonction inconnue, dont des protéines fimbriales hypothétiques, ce qui laisse penser à un rôle dans la régulation de la mobilité [219].

Chez les pathogènes d'animaux et d'insectes (*E. coli, S. enterica*, *Xenorhabdus nematophila*), LrhA régule les fonctions associés à la mobilité : flagelles, formation de biofilm, attachement et chemotaxisme [220–223]. De plus, chez *X. nematophila*, LrhA contrôle également la production et la sécrétion d'enzymes, principalement la lipase extracellulaire XlpA [221].

Aucune étude ne s'est cependant intéressée à un rôle potentiel de LrhA dans la thermorégulation chez ces autres bactéries, ce qui ne permet pas de conclure quant à l'acquisition de ce rôle caractéristique de PecT. Il est cependant intéressant de noter qu'un lien entre température et mobilité est connu chez *E. coli* [224] et pourrait être lié à PecT/LhrA.

Quelques homologues potentiels avec environ 30% d'identité sont détectés chez différents genres plus éloignés (*Streptomyces*, *Bacillus*, *Xanthomonas*, *Agrobacterium*…). Cela reste faible et aucun de ces homologues n'a été caractérisé fonctionnellement. PecT/LrhA ne semble donc pas être présent en dehors des entérobactéries, ce qui indique une origine récente mais encore inconnue. PecT semble donc être à l'origine un régulateur de la mobilité et de la formation de biofilms, puisque ces fonctions sont conservées chez tous les exemples étudiés, puis être devenu un régulateur de certains gènes spécifiques de l'hôte (pectinases de *Dickeya* pour les plantes,

lipase de *Xenorhabdus* pour les insectes…).

6. Le défi de *D. dadantii* : intégrer l'action des régulateurs anciens et nouveaux sur des gènes d'origine variée

Comme on l'a vu jusqu'à présent, *D. dadantii* possède un assemblage hétérogène de facteurs de virulence. Certains sont anciens et partagés avec les entérobactéries zoopathogènes (flagelle) et ont été sélectionnés pour leur intérêt dans les hôtes végétaux. D'autres facteurs sont spécifiques de l'hôte végétal et sont issus d'une longue adaptation chez d'autres lignées bactériennes (système de sécrétion de type III Hrp), puis ont été acquis par *D. dadantii* par transfert horizontal. Enfin, *Dickeya* a lui-même mis en place des facteurs de virulence nouveaux, spécifiques de son mode d'infection (les trois enzymes paralogues PelA, PelE et PelD).

Les régulateurs qui contrôlent ces facteurs de virulence sont eux aussi d'origines diverses. Certains sont des régulateurs globaux de la virulence qui ont majoritairement continué de réguler les mêmes gènes (KdgR, PecT). D'autres sont des régulateurs ancestraux dont le régulon a subi des modifications majeures pour s'adapter aux contraintes de l'hôte végétal (FIS, CRP). Enfin, *Dickeya* a également intégré de nouveaux régulateurs acquis par transfert horizontal (PecS). Deux questions se sont alors posées.

Par quels mécanismes ces régulateurs d'origines diverses assurent-ils le contrôle fin des gènes de virulence ? Cette question a été explorée avec le second article, portant sur les mécanismes moléculaires de régulation de l'un des gènes codant pour une pectate lyase essentielle au pouvoir pathogène de *D. dadantii*, *pelD*.

Comment les nouveaux gènes de virulence créés par duplication dans cette bactérie ont-ils été sélectionnés ? Le troisième article répond à cette question en éclaircissant le déroulement de l'une des innovations évolutives majeures des *Dickeya*, à savoir l'acquisition de *pelA*, *pelE* et *pelD*.

II. Matériel et méthodes non publiés

Cette section couvre l'ensemble des techniques utilisées qui ne sont pas décrites dans les articles. Les autres se trouvent dans les articles de la partie III.

A. Construction des plasmides supplémentaires

pKD22 a été construit à partir de pKD0 par insertion de 5 bp entre les sites CRP1 et FIS1bis, via une mutagénèse dirigée avec les oligonucléotides 348 et 349. pKD23 a été construit à partir de pKD10 par mutagénèse avec les oligonucléotides 350 et 351.

Pour la construction de pCH2, deux produits PCR ont été assemblés dans pBC KS+ : le premier amplifié depuis pKD0 avec les oligonucléotides 380 et 381, le second amplifié depuis pKD0 avec les oligonucléotides 375 et 378.

Les plasmides correspondants des séries pI et pP ont alors été construits par clonage classique comme décrit dans le 2ème article.

Table 3 : Souches, plasmides et oligonucléotides non publiés

Souches	Description	Ref
E. coli		
BL21 pLys	F- *ompT hsdSB*(rB-mB-) *gal dcm* (DE3) pLysS (Cam^R)	Laboratoire
D. aquatica		
$174/2^T$	Souche isolée dans de l'eau de rivière en Angleterre	[225]
Plasmides	**Description**	**Ref**
pG-OmpR	pGEM-T portant, en insertion par T/A cloning : promoteur T7-RBS-séquence codante *ompR*-6XHis	Ce travail
pBC KS+	Vecteur de clonage, CmR	Stratagene
pKD22	pBluescript KS+ contenant une region régulatrice *pelD* entre les sites EcoRI et HindIII, avec une insertion de 5 pb entre CRP1 et FIS1bis	Ce travail
pCH2	pBluescript KS+ contenant une region régulatrice *pelD* entre les sites EcoRI et HindIII, avec une insertion de 6 pb entre CRP1 et CRP2	Ce travail
pPD22	pProbe AT' *gfp* contenant une region régulatrice *pelD* entre les sites EcoRI et HindIII, avec une insertion de 5 pb entre CRP1 et FIS1bis	Ce travail
pPH2	pProbe AT' *gfp* contenant une region régulatrice *pelD* entre les sites EcoRI et HindIII, avec une insertion de 6 pb entre CRP1 et CRP2	Ce travail
pID2	pProbe AT' *gfp* contenant une region régulatrice *pelD* entre les sites PstI et BamHI, avec FIS3 inactivé	Ce travail
pID3	pProbe AT' *gfp* contenant une region régulatrice *pelD* entre les sites PstI et BamHI, avec FIS2 et FIS3 inactivés	Ce travail
pID10	pProbe AT' *gfp* contenant une region régulatrice *pelD* entre les sites KpnI et BamHI, avec FIS1 et FIS1BIS inactivés	Ce travail
pID22	pProbe AT' *gfp* contenant une region régulatrice *pelD* entre les sites KpnI et BamHI, avec une insertion de 5 pb entre CRP1 et FIS1bis	Ce travail
pID23	pProbe AT' *gfp* contenant une region régulatrice *pelD* entre les sites KpnI et BamHI, avec FIS1, FIS1BIS inactivés et la séquence -10 de *div* restaurée à un état sauvage.	Ce travail
pIH2	pProbe AT' *gfp* contenant une region régulatrice *pelD* entre les sites KpnI et BamHI, avec une insertion de 6 pb entre CRP1 et CRP2	Ce travail
Oligonucléotides	**Séquence**	**Ref**
338	AAGAATTCTAATACGACTCACTATAGGGAAGGAGGTATATAATGC	Ce travail
339	TTAAGCTTTCAGTGATGATGGTGGTGATGTGCTTTACTGCCGTCC	Ce travail
348	GAAAAATTAATTCAACATTCATAACTAATTTCTAAGTTACCGGTCACGATCACACTTTAG	Ce travail
349	CTAAAGTGTGATCGTGACCGGTAACTTAGAAATTAGTTATGAATGTTGAATTAATTTTTC	Ce travail
350	CAGATAAAAAGtCTTATACATATAGCTGAATTTAAAAtAAAAATTAATTC	Ce travail
351	GAATTAATTTTTaTTTTAAATTCAGCTATATGTATAAGaCTTTTTATCTG	Ce travail
375	AAAAGCTTTACGGAAGACACTCGTGTG	Ce travail
378	AAAGATCTCCTCATAAAAAAAACGAGATTTTG	Ce travail
380	AAAGATCTCTAATTAATTTTATCTAAAGTGTG	Ce travail
381	AAGAATTCTCAACTGTTTGGGTTATTTCAG	Ce travail
426	AGTAGCGTGAGTGCGCGC	Ce travail
427	ATTACCGCCGCGATGCATG	Ce travail

B. Surproduction et purification d'OmpR de *D. dadantii*

La séquence codante d'OmpR a été amplifiée par PCR à partir du génome de *D. dadantii* 3937 par les oligonucléotides 338 et 339. L'oligonucléotide 338 contient un promoteur T7 et un RBS fort, 339 contient un tag 6XHis et un codon stop. Le produit ainsi amplifié a été inséré dans le plasmide pGEM-T par T/A cloning, donnant le plasmide pG-OmpR.

Ce plasmide a alors été transformé dans la souche de *E. coli* BL21 pLys. Les cellules ont été cultivés en LB + ampicilline + chloramphénicol à 30°C jusqu'à DO_{600} = 0.4. La surproduction d'OmpR a été induite par ajout d'1 mM d'IPTG suivie d'une croissance pendant 2 h. Les cellules ont par la suite été concentrées par une centrifugation de 10 min à 6000 rpm, puis le culot obtenu a été suspendu dans un tampon de lyse (50 mM NaH_2PO_4, 300 mM NaCl, 1 mM imidazole, pH 8) et les cellules ont alors été cassées à l'aide de la presse de French. La protéine OmpR contenue dans le lysat soluble a été purifiée sur résine Ni-NTA (His-select nickel affinity gel, Sigma) selon les instructions du fournisseur puis dialysée deux fois contre 500 mL de tampon de conservation (50 mM NaH_2PO_4, 300 mM NaCl, 50% v/v glycérol, pH 8) avant d'être conservée à -20°C.

C. Retard de migration sur gel

La réaction a été réalisée dans un volume final de 20 µL de composition Tris-HCl 10 mM, KCl 70 mM, DTT 1 mM, BSA 200 µg/mL, poly(dI-dC) 10 µg/mL, $MgCl_2$ 10 mM, $CaCl_2$ 5 mM, NP40 0,1% (m/v), glycérol 5% (m/v), acetyl phosphate 50 mM et des concentrations variables de OmpR purifé. Le mélange a été incubé 30 min à 30°C pour phosphoryler OmpR, puis 100 ng d'ADN marqué par fluorescence ont été ajoutés et les mélanges ont été incubés 25 min à 30° C. Les complexes ADN-OmpR ont alors été séparés sur gel de polyacrylamide 4 % non dénaturant avec un tampon Tris-Acécate (Tris-HCl 6,75 mM, Acétate de sodium 3,3 mM, pH 7,9) puis visualisés par fluorescence avec l'appareil ChemiDoc MP System (Bio-rad, Etats-Unis).

D. Prédiction de la structure de la protéine OmpR de *D. dadantii*

Puisque la structure obtenue par diffraction aux rayons X de la protéine OmpR disponible sur les bases de données ne contient que le domaine C-terminal, la séquence de ce domaine de la protéine d'OmpR de *D. dadantii* a été récupérée depuis UniProt (E0SJT3). La prédiction de la structure de la protéine a alors été réalisée avec l'outil geno3D (https://geno3d-prabi.ibcp.fr/) en utilisant la structure du domaine C-terminal de OmpR de *E. coli* comme modèle (pdb 10DD).

E. Prédiction des sites de fixation OmpR dans le génome de *D. dadantii*

La séquence consensus de fixation d'OmpR a été déterminé par MEME (http://meme-suite.org/tools/meme) à partir des sites identifiés par ChIP chez *S. enterica* [226]. Ce consensus a alors été utilisé pour chercher des sites de fixation par FIMO (http://meme-suite.org/tools/fimo) dans la base de données « Upstream sequences : Prokaryotic, *Dickeya dadantii* 3937 uid52537 » avec une p-value limite de 10^{-4}. Les gènes dont la région amont porte au moins un site OmpR ont alors été regroupés selon leur classification COG.

F. Tests phénotypiques chez *Dickeya*

1. Milieux solides

5 µL d'une suspension bactérienne à DO_{600} de 0.2 a été déposée sur différents milieux gélosés : M63G + PGA 0.2% (pectinase), GL lait 0.63% (protéase), M63G CMC 0.2% (cellulase), LB + agar 0.4% (mobilité). La dégradation du PGA a été visualisée par la formation de halos blancs lors de l'ajout d'acétate de cuivre en solution aqueuse saturante (10% m/v). La dégradation du

CMC a été visualisée par la formation de halos jaunes lors de l'ajout de rouge Congo 0.5 % m/v suivi d'un rinçage au NaCl 1M.

Pour les tests de résistance aux stress, la suspension initiale a été dilué de 10 en 10 jusqu'à une DO_{600} de 2.10^{-6}. 5 µl de chacune des 6 dilutions ont alors été déposés sur les milieux suivants : M63G, M63G + acide malique pH 5, M63G + H_2O_2 100 µM, M63G + NaCl 0.3 M. Les boîtes ont alors été incubées pendant 24 h à 28°C.

2. Milieux liquides

a. Activité Pel

Des cultures bactériennes ont été effectuées en milieu M63S + CaCl2 0,1 mM + PGA 0,2% m/v. A chaque prélèvement, 1 mL de culture a été mélangé à 100 µl de toluène pour perméabiliser les membranes et libérer les enzymes intracellulaires. L'activité pectate lyase a ensuite été dosée selon la méthode décrite dans [227].

b. Biofilms

Des cultures bactériennes d'un volume de 2 mL ont été effectuées en milieu M63S à partir d'une DO initiale de 0.003 en plaques 24 puits pendant 48 h à 30 °C. Le surnageant a été prélevé puis le biofilm a été rincé par 1 mL de M63 afin d'éliminer les bactéries planctoniques sédimentées. Ces deux premières fractions ont alors été réunies et la DO_{600} mesurée (cette DO sera désignée S par la suite). Le biofilm a ensuite été resuspendu dans 1 mL de M63 et la DO_{600} mesurée (cette DO sera désignée B).

$$DO_{totale} = B + 3S$$

$$\%_{adhérence} = \frac{B}{B + 3S}$$

3. Infections de plantes

Des précultures bactériennes ont été effectuées en M63S et diluées en milieu M63S à une DO_{600} dépendant de la plante à infecter : 0.2 (endive) ou 1 (pomme de terre et tomate). Pour les endives, 5 µL de la suspension ont été infiltrées dans une coupure effectuée au milieu de la feuille et longue de 2 cm. Pour les pommes de terre, 5 µL ont été injectés avec un cône de micropipette à quelques mm de profondeur. Pour les tomates, 200 µL ont été injectés avec un cône de micropipette à environ 3 cm de profondeur. Les plantes ont alors été mises dans des bacs humides et scellés pendant 18 h (endives) ou 42 h (pomme de terre et tomate).

G. Séquençage du génome de *D. aquatica*

L'ADN génomique de *D. aquatica* 174/2 a été extrait avec le kit Wizard Genomic DNA purification (Promega) sur une culture en phase exponentielle tardive de croissance (DO_{600} = 0.9). La préparation de la banque SMRTbell suivie du séquençage sur PacBio RSII ont été effectués par Eurofins (Allemagne). Les reads ainsi obtenus ont alors été assemblés par CANU 1.3 (http://canu.readthedocs.io/en/stable/), produisant deux contigs. La fin du premier contig chevauchait le début du deuxième sur une longueur de 16 kb, permettant de joindre les contigs. Le gap restant (entre la fin du deuxième contig et le début du premier) a été comblé par PCR sur l'ADN génomique de *D. aquatica* avec les oligonucléotides 426 et 427, produisant alors un génome complet. Une annotation automatique de ce génome a alors été effectuée sur RAST (http://rast.nmpdr.org/)

H. Phylogénie des espèces *Dickeya*, *Serratia* et *Pectobacterium*

Les séquences des protéines ribosomales des espèces de *Dickeya*, *Pectobacterium* et *Serratia* présentes dans RiboDB (https://ribodb.univ-lyon1.fr/ribodb/ribodb-in.cgi) ont été récupérées. Ces séquences ont servi à extraire les séquences des protéines ribosomales des 24 autres espèces de *Dickeya* présentes sur GenBank et du protéome de *D. aquatica* traduit depuis le génome séquencé. Les séquences nucléiques des gènes associés ont alors été alignées au niveau protéique par MUSCLE (implémenté dans Seaview) puis nettoyées par BMGE. Les séquences ont alors été concaténées et l'arbre construit par IqTree 1.4.2 avec le modèle GTR.

III. Résultats

A. *pelD* : une régulation transcriptionnelle complexe et optimisée (2ème article)

Le second article se concentre sur la mécanistique de l'action de FIS et CRP sur *pelD* et la découverte d'un promoteur divergent impliqué dans la régulation transcriptionnelle de *pelD*. L'article a été publié en 2016 dans le journal BBA-gene regulatory mechanisms [228].

Temporal Control of *Dickeya dadantii* Main Virulence Gene Expression by Growth Phase-Dependent Alteration of Regulatory Nucleoprotein Complexes

Short title: ***pelD* Regulation by Nucleoprotein Complexes Alteration**

Alexandre Duprey[1,2,3], Georgi Muskhelishvili[1,2,3], Sylvie Reverchon[1,2,3] and William Nasser[1,2,3]*.

[1] Université Lyon 1, F-69622 Villeurbanne, France

[2] INSA de Lyon, F-69621 Villeurbanne, France

[3] CNRS UMR5240 Microbiologie, Adaptation et Pathogénie, Villeurbanne, France

* To whom correspondence should be addressed. Tel: +33 (0)4 72 43 26 95; Fax: +33 (0)4 72 43 26 86; Email: william.nasser@insa-lyon.fr

Abstract

In bacteria, important genes are often controlled at the transcriptional level by several factors, forming a complex and intertwined web of interactions. Yet, transcriptional regulators are often studied separately and little information is available concerning their interactions. In this work, we dissect the regulation of the major virulence gene *pelD* in *D. dadantii* by taking into account the effects of individual binding sites for regulatory proteins FIS and CRP, and the impact of a newly discovered divergent promoter *div*. Using a combination of biochemistry and genetics approaches we provide an unprecedented level of detail on the multifactorial regulation of bacterial transcription. We show that the growth phase dependent regulation of *pelD* is under the control of changing composition of higher-order nucleoprotein complexes between FIS, CRP, *div* and *pelD* during the growth cycle that allow sequential expression of *div* and *pelD* in the early and late exponential growth phases, respectively. This work highlights the importance of "orphan" promoters in gene regulation and that the individual binding sites for a regulator can serve several purposes and have different effects on transcription, adding a new level of complexity to bacterial transcriptional regulation.

Highlights

- The transcriptional regulation of the *pelD* virulence gene of *D. dadantii* is complex
- *pelD* is regulated by a divergent orphan promoter *div*
- Individual FIS binding sites contribute differentially to *pelD* regulation

- CRP activates *pelD* and *div* asymmetrically
- Nucleoprotein complexes change to create a sequential expression of *pelD* and *div*

Keywords

pel gene, Nucleoid-associated protein, FIS, CRP, Transcriptional regulation

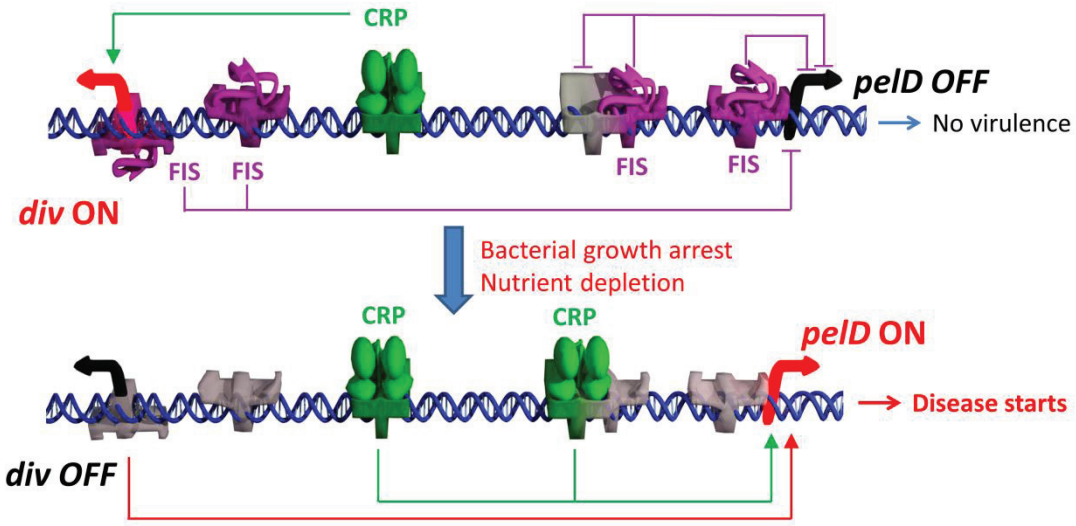
Abbreviations

TF: Transcription Factor

NAP: Nucleoid-Associated Protein

PGA: Poly-Galacturonic Acid

Graphical abstract



1. Introduction

During the bacterial life cycle, the cells can be confronted with a variety of environments differing in nutrient composition/availability, temperature, pH, salinity and biotic interactions. Successful adaptation to these variable environmental conditions requires appropriate phenotypic modifications. For this purpose, bacteria modify the gene expression primarily at the level of transcriptional regulation, mediated by transcription factors (TFs) binding at regulatory site(s) in the gene promoter regions. Whereas repression by direct steric hindrance leading to promoter occlusion, and activation by recruitment of RNA polymerase by a TF, are the most frequently described regulation mechanisms [27], most genes are controlled by several TFs often competing for binding sites in the gene regulatory region, or for interaction with RNA polymerase. Furthermore, a small class of highly abundant TFs, known as nucleoid-associated

proteins (NAPs), regulate transcription via a variety of DNA topology-related mechanisms such as looping [229], modulation of the spatial proximity of DNA-binding factors [193], and alteration of DNA configuration [230]. This potential of utilising topological changes in DNA entails complex regulatory mechanisms, which are not well understood.

FIS (factor for inversion stimulation) is one of the major NAPs in γ-proteobacteria, capable of bending and looping the DNA and acting as a global regulator of transcription and growth [48,171,231]. In adition to direct binding effects, FIS can influence global DNA topology via regulation of DNA topoisomerases [232]. In particular, FIS acts to preserve the moderately supercoiled forms, protecting the DNA from shifts towards the strongly relaxed or highly negatively supercoiled extremes of the topological spectrum [233]. In *E. coli*, FIS is highly abundant in early to mid-exponential growth phase when the bacterial growth rate is high, whereas afterwards its level rapidly decreases becoming undetectable in stationary phase [50]. FIS can activate transcription initiation by various mechanisms such as recruitment of RNA polymerase at the promoter, induction of a conformational change in the promoter DNA, or both [234]. However, FIS is a dual regulator and can also act as a repressor counteracting the expression of inessential gene products during rapid growth [234]. Thus, the changing concentration of FIS during the growth cycle can both activate and repress the production of growth-related proteins and modulate the expression of virulence factors in a growth phase-dependent manner [48,71] leading to a sequential production of specific proteins during infection.

Dickeya dadantii is a phytopahogenic bacterium originating from tropical areas, which has been recently identified as a causal agent of crop damage in temperate regions (PQR - EPPO database on quarantine pests, available from http://www.eppo.int/DATABASES/pqr/pqr.htm). *D. dadantii* has a broad host spectrum, including economically important crops such as corn (*Zea mays*), potato (*Solanum tuberosum*) and chicory (*Cichorium endivia*) [108]. The main infection symptom is soft rot, which results from the action of extracellular enzymes, mainly pectate lyases (Pels) degrading pectin, an essential component of the plant cell wall. *D. dadantii* produces more than ten Pels with different biochemical properties and expression patterns [227,235]. The synthesis of *D. dadantii* Pels is fine-tuned according to environmental or metabolic signals [116], such as the presence of pectic compounds (via KdgR repressor) [130,145,236], and the oxidative and acidic stresses [188,193]. It was shown that the sugar catabolism regulator complex cAMP-CRP is the main direct activator of the *pel* genes [145,188]. Furthermore, the abundant NAP FIS modulates the expression of the *pel* genes in a growth phase-dependent manner [167]. Both these factors bind to the *pel* gene regulatory regions making these genes ideal for exploring how bacteria integrate various mechanisms to control transcription initiation. Among the Pels, PelD production is most strongly affected by pectin derivatives and growth-phase, which makes it a key element in the induction of *D. dadantii* virulence [237,238]. The *pelD* gene is inactive during exponential growth when the bacterial density is low, and strongly expressed in early stationary phase when the bacterial density is sufficient to overcome the plant defence mechanisms. Previously, it was demonstrated that FIS binds two regions in *pelD* [192], but their impact in modulating *pelD* expression remains unclear. Furthemore, it was observed that RNA polymerase binds at the *pelD* upstream AT-rich region more tightly than at the *pelD* core promoter itself [146,192]. Notably, binding of CRP and FIS induces overlapping protection patterns in the region between the upstream RNA polymerase binding site and the *pelD* promoter. However, neither the role of these overlapping binding sites, nor the impact of upstream RNA polymerase binding site on *pelD* transcription have been elucidated.

In this work, by combining *in vivo* transcriptional activity assays using reporter genes with *in vitro* biochemical studies, we describe the complexity of the *pelD* promoter control showing that the key component in the temporal pattern of *pelD* expression is the dynamic balance between the binding of RNA polymerases at either *pelD* or a newly discovered upstream divergent promoter modulated by FIS, CRP and DNA topology. By integrating the individual effects of regulators resolved at a single binding site scale with the data on polymerase binding site occupation, we propose a model of *pelD* transcriptional regulation during the growth cycle.

2. Materials and methods

2.1. Bacterial strains, plasmids, media, culture conditions and molecular biology methods

Escherichia coli and *Dickeya dadantii* were cultured at 37 °C in LB (Luria-Bertani, [239]) and 30 °C in M63S (M63 [239] supplemented with sucrose (S) 0.2% w/v), respectively. Antibiotics were used at the following concentrations: Ampcillin 100 µg/mL, Chloramphenicol 20 µg/mL or 4 µg/mL for *E. coli* and *D. dadantii*, respectively. Liquid cultures were grown in a shaking incubator (150 r.p.m.). Media were solidified by the addition of 1.5% agar (w/v).

Plasmids from the pKD series were obtained either by direct synthesis (pKD0, pKD1, pKD2, pKD3, pKD4 and pKD5) or by QuickChange mutagenesis. pKD7 was mutated from pKD0 using primers pelDfis1bismut FW and pelDfis1bismut REV. pKD9 and pKD10 were mutated from pKD5 and pKD7, respectively, using primers pelDFis1mutFW and pelDFis1mutREV. pKD24 was mutated from pKD0 using primers 35rev v2 FW and 35rev v2 REV (Table 4). pKD26 and pKD27 were mutated from pKD0 using primers 395/396 and 397/398, respectively. pKD28 was mutated from pKD26 using primers 397/398.

Plasmid pJV was constructed by amplifying the double terminator B0014 (http://parts.igem.org/Part:BBa_B0014) from the plasmid pIG50 [240] using primers B0014 FW and B0014 REV. The EcoRI-PvuII PCR product was then cloned into pJCD0 [241], yielding a plasmid with terminators on both sides suitable for *in vitro* transcription. Like all pUC-derived plasmids, pJV bears the noncoding RNA-I inside its replication origin which was used as an internal standard.

Plasmids from the pPD series, pID series, pJD series, pLD0 and pND0 were obtained by restriction assembly of the corresponding pKD insert into pProbe AT', pProbe AT' (with different restriction sites for the pPD or pID series, see Table 4, such that *pelD* controls *gfp* transcription in the pPD series, and *div* controls *gfp* transcription in the pID series), pJV, pUCTer-Luc-Cm or pNB4, respectively, using standard molecular biology protocols [242].

Table 4: Strains, plasmids and oligos used in this study

Strains	Description	Ref
D. dadantii		
3937 (A4922)	Wild-type strain isolated from *Saintpaulia ionantha*	[243]
A4474	3937 *fis*::Cm^R	[167]
E. coli		
CSH50	F- *araBAD-0 M(pro-lac) λ- rpsL*-(strR) *thi- fimE1::IS1-*	[239]

I3506	CSH50 *fis::*Cm^R		[244]

Plasmids	Description	Inactivated FIS sites	Ref
pNB4	Reporter vector with *uidA*, AmpR		[245]
pProbe-AT '	Reporter vector with *gfp*, AmpR		[246]
pBluescript SK+	Cloning vector, AmpR		Stratagene
pBluescript KS+	Cloning vector, AmpR		Stratagene
pUCter-Luc-Cm	Reporter vector with *luc*, AmpR CmR		[247]
pJV	*In vitro* transcription vector AmpR		This work
pWN2481 (pKD0)	pBluescript KS+ containing the wild-type *pelD* regulatory region between sites EcoRI and HindIII	none	[146]
pKD1	pBluescript SK+ containing *a* mutated *pelD* regulatory region at site SmaI (T7 promoter oriented)	2	This work
pKD2	pBluescript SK+ containing *a* mutated *pelD* regulatory region at site SmaI (T7 promoter oriented)	3	This work
pKD3	pBluescript SK+ containing *a* mutated *pelD* regulatory region at site SmaI (T7 promoter oriented)	2,3	This work
pKD4	pBluescript SK+ containing *a* mutated *pelD* regulatory region at site SmaI (T7 promoter oriented)	1, 2, 3	This work
pKD5	pBluescript SK+ containing *a* mutated *pelD* regulatory region at site SmaI (T7 promoter oriented)	1bis, 2, 3	This work
pKD7	pBluescript KS+ containing a mutated *pelD* regulatory region between sites EcoRI and HindIII	1bis	This work
pKD9	pBluescript SK+ containing *a* mutated *pelD* regulatory region at site SmaI (T7 promoter oriented)	1, 1bis, 2, 3	This work
pKD10	pBluescript KS+ containing a mutated *pelD* regulatory region between sites EcoRI and HindIII	1, 1bis	This work
pKD24	pBluescript KS+ containing a mutated *pelD* regulatory region between sites EcoRI and HindIII, with *div* inactivated	none	This work
pKD26	pBluescript KS+ containing a mutated *pelD* regulatory region between sites EcoRI and HindIII, with CRP2 inactivated	none	This work
pKD27	pBluescript KS+ containing a mutated *pelD* regulatory region between sites EcoRI and HindIII, with CRP1 inactivated	none	This work
pKD28	pBluescript KS+ containing a mutated *pelD* regulatory region between sites EcoRI and HindIII, with CRP1 and CRP2 inactivated	none	This work
pND0	pNB4 containing the wild-type *pelD* regulatory region between sites EcoRI and HindIII	none	This work
pLD0	pUCter-Luc-Cm containing the wild-type *pelD* regulatory region between sites NheI and XhoI	none	This work
pPD0	pProbe-AT ' containing the wild-type *pelD* regulatory region between sites EcoRI and HindIII	none	This work
pPD1	pProbe-AT ' containing a mutated *pelD* regulatory region between sites EcoRI and HindIII	2	This work
pPD2	pProbe-AT ' containing a mutated *pelD* regulatory region between sites EcoRI and HindIII	3	This work
pPD3	pProbe-AT ' containing a mutated *pelD* regulatory region between sites EcoRI and HindIII	2,3	This work
pPD4	pProbe-AT ' containing a mutated *pelD* regulatory region between sites EcoRI and HindIII	1, 2, 3	This work
pPD5	pProbe-AT ' containing a mutated *pelD* regulatory region between sites EcoRI and HindIII	1bis, 2, 3	This work
pPD7	pProbe-AT ' containing a mutated *pelD* regulatory region between sites EcoRI and HindIII	1bis	This work
pPD9	pProbe-AT ' containing a mutated *pelD* regulatory region between sites EcoRI and HindIII	1, 1bis, 2, 3	This work
pPD24	pProbe-AT ' containing a mutated *pelD* regulatory region between sites EcoRI and HindIIi, with *div* inactivated	none	This work
pPD26	pProbe-AT ' containing a mutated *pelD* regulatory region between sites EcoRI and HindIII, with CRP2 inactivated	none	This work
pPD27	pProbe-AT ' containing a mutated *pelD* regulatory region between sites EcoRI and HindIII, with CRP1 inactivated	none	This work
pPD28	pProbe-AT ' containing a mutated *pelD* regulatory region between sites EcoRI and HindIII, with CRP1 and CRP2 inactivated	none	This work
pJD0	pJV containing the wild-type peID regulatory region between sites EcoRI and SalI	none	This work
pJD1	pJV containing a mutated *pelD* regulatory region between sites EcoRI and SalI	2	This work
pJD2	pJV containing a mutated *pelD* regulatory region between sites EcoRI and SalI	3	This work
pJD3	pJV containing a mutated *pelD* regulatory region between sites EcoRI and SalI	2,3	This work

pJD9	pJV containing a mutated *pelD* regulatory region between sites EcoRI and SalI	1, 1bis, 2, 3	This work
pJD10	pJV containing a mutated *pelD* regulatory region between sites EcoRI and SalI	1, 1bis	This work
pJD24	pJV containing a mutated *pelD* regulatory region between sites EcoRI and SalI, with *div* inactivated	none	This work
pID0	pProbe-AT ' containing a wild-type *pelD* regulatory region between sites KpnI and BamHI	none	This work
pID24	pProbe-AT ' containing a mutated *pelD* regulatory region between sites KpnI and BamHI, with *div* inactivated	none	This work
pID26	pProbe-AT ' containing a mutated *pelD* regulatory region between sites KpnI and BamHI, with CRP2 inactivated	none	This work
pID27	pProbe-AT ' containing a mutated *pelD* regulatory region between sites KpnI and BamHI, with CRP1 inactivated	none	This work
pID28	pProbe-AT ' containing a mutated *pelD* regulatory region between sites KpnI and BamHI, with CRP1 and CRP2 inactivated	none	This work

Primers		Ref
DNase pel FW	HEX-ACTATAGGGCGAATTGG	This work
DNase pel REV	FAM-AATTAACCCTCACTAAAGG	This work
B0014 FW	AACAGCTGTCACACTGGCTCACCTTC	This work
B0014 REV	AAGAATTCAATAATAAAAAAGCCGGATTAATAATC	This work
bla 3B4 fluo	HEX-CAGGAAGGCAAAATGCCGC	This work
pelDfis1bismut FW	GCTGAATTTAAAAtAAAAATTAATTCAAtATTCATAACTAAAAG	This work
pelDfis1bismut REV	CTTTTAGTTATGAATaTTGAATTAATTTTTaTTTTAAATTCAGC	This work
pelDFis1mutFW	TTTCAGATAAAAACtCTTATACATATAGtTGAATTTAAAA	This work
pelDFis1mutREV	TTTTAAATTCAaCTATATGTATAAGaGTTTTTATCTGAAA	This work
35rev v2 FW	GAATTTAAAAGAAAAATTAATTCttCATTCATAACTAAAAGTTACC	This work
35rev v2 REV	GGTAACTTTTAGTTATGAATGaaGAATTAATTTTTCTTTTAAATTC	This work
300	ACTATAGGGCGAATTGG	This work
395	CATAAAAAAAACGAGATTTTGAagtCAAAATAAACAATCGAAAACGC	This work
396	GCGTTTTCGATTGTTTATTTTGactTCAAAATCTCGTTTTTTTTATG	This work
397	CTAAAAGTTACCGGTCACGAagtCACTTTAGATAAAATTAATTAGC	This work
398	GCTAATTAATTTTATCTAAAGTGactTCGTGACCGGTAACTTTTAG	This work

2.2. *In vivo* transcriptional activity measurements

Overnight bacterial precultures in M63S were diluted to an OD_{600} of 0.03 in M63S + $CaCl_2$ 0.1 mM + Polygalacturonic acid (PGA) 0.2 % w/v (*Dickeya dadantii*) or LB (*E. coli*). The medium was supplemented with D-luciferin (Perkin-Elmer, USA) at a final concentration of 450 µg/mL when working with Luc. When required, novobiocin was added at the start of the culture, at a final concentration of 25 µg/mL.

For fluorescence and luminescence, 200 µL of this dilution were then distributed in a Cellstar 96-well plate (655090, Greiner Bio-one, Germany). The OD_{600}, fluorescence and luminescence were monitored for 48h using an Infinite 200 pro plate reader (Tecan, Switzerland). For fluorescence, the excitation and emission wavelengths were 485 and 530 nm, with a bandwidth of 9 and 20 nm, respectively.

The normalised fluorescence is defined as the ratio of the fluorescence of the considered

promoter in early stationary phase (which corresponds to 20 h growth for a WT strain) by the OD_{600}, normalised to the ratio calculated for the WT promoter in a WT strain (arbitrarily set as 100).

mRNA levels were computed from fluorescence data using the model described in [248]. Due to the stability of the Gfp used (> 20 days half-life, [249]), degradation was neglected. Because *pelD* and *div* are not expressed at the same time (see Figure 44a), the maximum expression of both promoters is usually represented in the figures, which means early log growth phase for *div* and late log for *pelD.*

For GUS, the specific beta-glucuronidase activity of the whole culture was measured at a comparable growth stage using the method described in [245].

2.3. DNase I footprinting

For quantitative fluorescent DNase I footprinting, the protocol published in [145] was modified as follows. Fluorescently labelled primers DNase pel FW and DNase pel REV (Table 4) were used to amplify the *pelD* regulatory region from plasmids of the pKD series by PCR. A clean-up was then performed using the Nucleospin Gel and PCR Clean-up kit (Macherey-Nagel, the Netherlands). 50 ng of the labelled DNA was used per reaction, which represents a 2-fold increase in DNA quantity compared to radioactivity. Protein concentrations were adapted accordingly. In the end, the samples were analysed on an AB 3730 XL capillary sequencer (Applied biosystems, USA). Peak areas were normalised to the total fluorescence of each sample.

2.4. Potassium permanganate footprinting

Potassium permanganate reactions were performed on supercoiled plasmid, as described in [192]. The reaction products were divided into 2 equal parts. 15 cycles of primer extension were performed with Phusion polymerase (New England Biolabs, USA) with the fluorescently labeled primers DNase pel REV, for one part, and Bla3B4 fluo for the other part (which was used as a standard). After the primer extensions, both reactions were mixed together and analysed on an AB 3730 XL capillary sequencer (Applied biosystems, USA).

For *in vivo* reactivity assays, *D. dadantii* wild type cells harbouring plasmid pKD0, containing the full *pelD* and *div* promoters, were grown in minimal M63 medium containing sucrose (0.2%, W/V) and polygalacturonate (0.2%, W/V) until the required stages. Then KMnO4 was added at 10 mM (final concentration) for 1 min as described [250] to the batch cultures growing at 30°C under vigourous shaking. The cells were collected by centrifugation and the plasmid was extracted with Machery Nalgel Spin Miniprep Kit. The isolated plasmids (125 ng) were used as template for 5 cycles of primer extension with Taq polymerase (New England Biolabs, USA) and ^{32}P-labeled primer 300 (Table 4). Reaction products were then loaded onto a sequencing gel.

2.5. *In vitro* transcription

Supercoiled pJD plasmids were incubated with CRP (purified from *D. dadantii*, used at 100 nM concentration unless indicated otherwise) and RNA polymerase (commercial holoenzyme from *E. coli* purchased from Epicentre*,* used at 40 nM concentration unless indicated otherwise) for 5 min at 30 °C in a reaction buffer containing 10 mM Tris-HCl pH 8, 0.5 mM dithiothreitol, 150

mM KCl, 200 µM cAMP and 0.1% (v/v) Nonidet P-40. Transcription was initiated by the addition of 0.2 mM ATP, GTP and CTP, 0.01 mM cold UTP and 0.1 µCi/µL [α-P^{32}]-UTP and allowed to proceed for 15 min. Reactions were stopped by the addition of an equal volume of 20 mM EDTA, 0.025% (w/v) bromophenol blue and 0.025% (w/v) Xylene cyanol in formamide and loaded on a sequencing gel.

3. Results

3.1. Identification and inactivation of the FIS binding sites

The cooperative binding of CRP at two distinct sites in the *pelD* regulatory region was established in a previous study [192]. Furthermore, it was found that binding of FIS resulted in protection of two DNA regions (from position -150 to -119 and -48 to +7 relative to the transcription initiation site at +1), one of which overlapped the *pelD* promoter (Figure 37). However, the exact number of FIS binding sites and their functional role remained unknown.

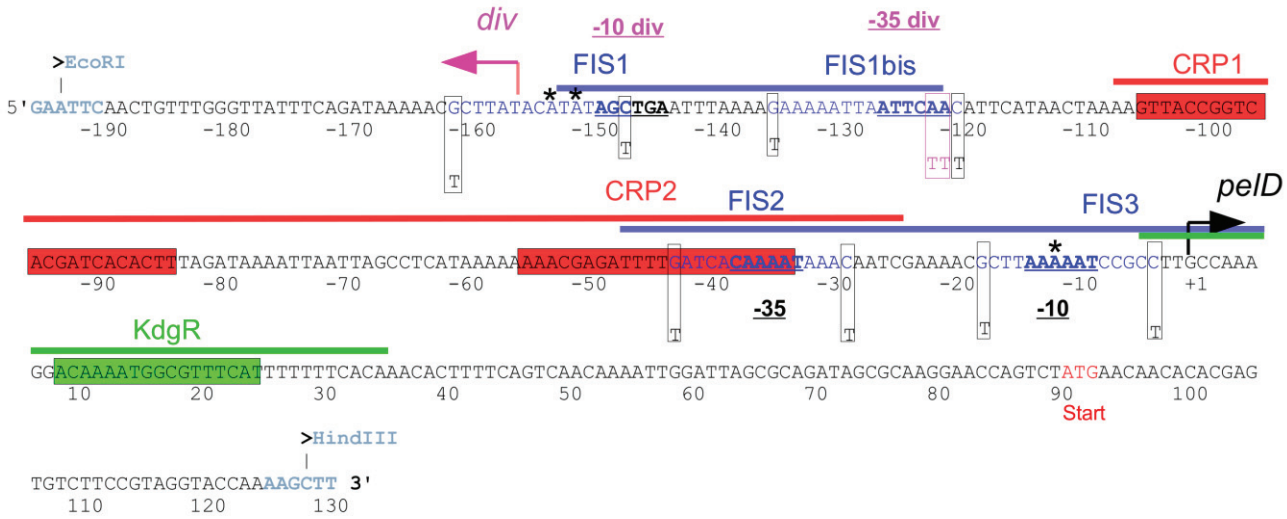


Figure 37: Structure of the pelD regulatory region.

The sequence of the *pelD* regulatory region is annotated with the experimentally verified binding sites of CRP (red box) and KdgR (green box) [145,146]. The -10 and -35 elements are in bold and underlined, the translational start codon is in red letters, and the transcription initiation site is indicated by a broken arrow [146]. The positions of the -35, -10 elements and initiation site of the div promoter are indicated in pink. The two regions protected by FIS [192] and harbouring the four identified FIS binding sites are indicated by blue solid lines. Mutated positions are enclosed in a rectangle (black and pink for inactivation of FIS sites and *div* promoter, respectively). The region protected by CRP [192] and harbouring two CRP binding sites is indicated by red solid line. The gray boxes indicate A/T tracts centered at +9, +18 and +26 downstream of *div* initiation start site. Stars indicate the main Ts targeted by potassium permanganate in Figure 47, which are situated on the reverse strand.

Consequently, a bioinformatic search for the FIS binding consensus on the *pelD* regulatory region was performed *via* Virtual Footprint (http://prodoric.tu-bs.de/vfp/vfp_promoter.php [251]), using the weight matrix established for *E. coli*. Only the sites with a score greater than 3.00 were considered (the mean score for referenced FIS binding sites being 2.87). This highlighted three binding sites, which were designated FIS1, FIS2 and FIS3 (Figure 37). FIS2 and FIS3 matched the previously determined FIS I experimental footprint [22], whereas FIS1 was

located upstream partially overlapping the FIS II experimental footprint [192]. Manual examination of the sequence revealed a 4th potential site, named FIS1bis, that was not detected by Virtual Footprint due to its increased length (16 bp) compared to the consensus FIS site (15 bp). The position of FIS1bis matches the previously identified experimental FISII footprint [192]. Among these four putative FIS binding sites, both FIS1 and FIS1bis are located far upstream, whereas FIS2 and FIS3 overlap the *pelD* core promoter. Importantly, FIS2 also overlaps the CRP2 binding site. These four putative sites were inactivated separately, or in various combinations, by substituting the conserved G at the position 1 and the conserved C at the position 15 by Ts [252], as described in Figure 37.

The success of inactivation was confirmed by DNase I footprinting demonstrating a disappearance of protection by FIS for all of these sites (Figure 38), except for FIS1bis which showed barely detectable FIS binding on the WT template. Inactivation of FIS1, FIS2 or FIS3 did not noticeably modify the FIS footprints at the other sites (Figure 38a), which implies that binding of FIS at these sites is largely independent. In addition, major differences in affinity were observed, with FIS2 and FIS3 having a higher affinity than FIS1 and Fis1bis (Figure 38b).

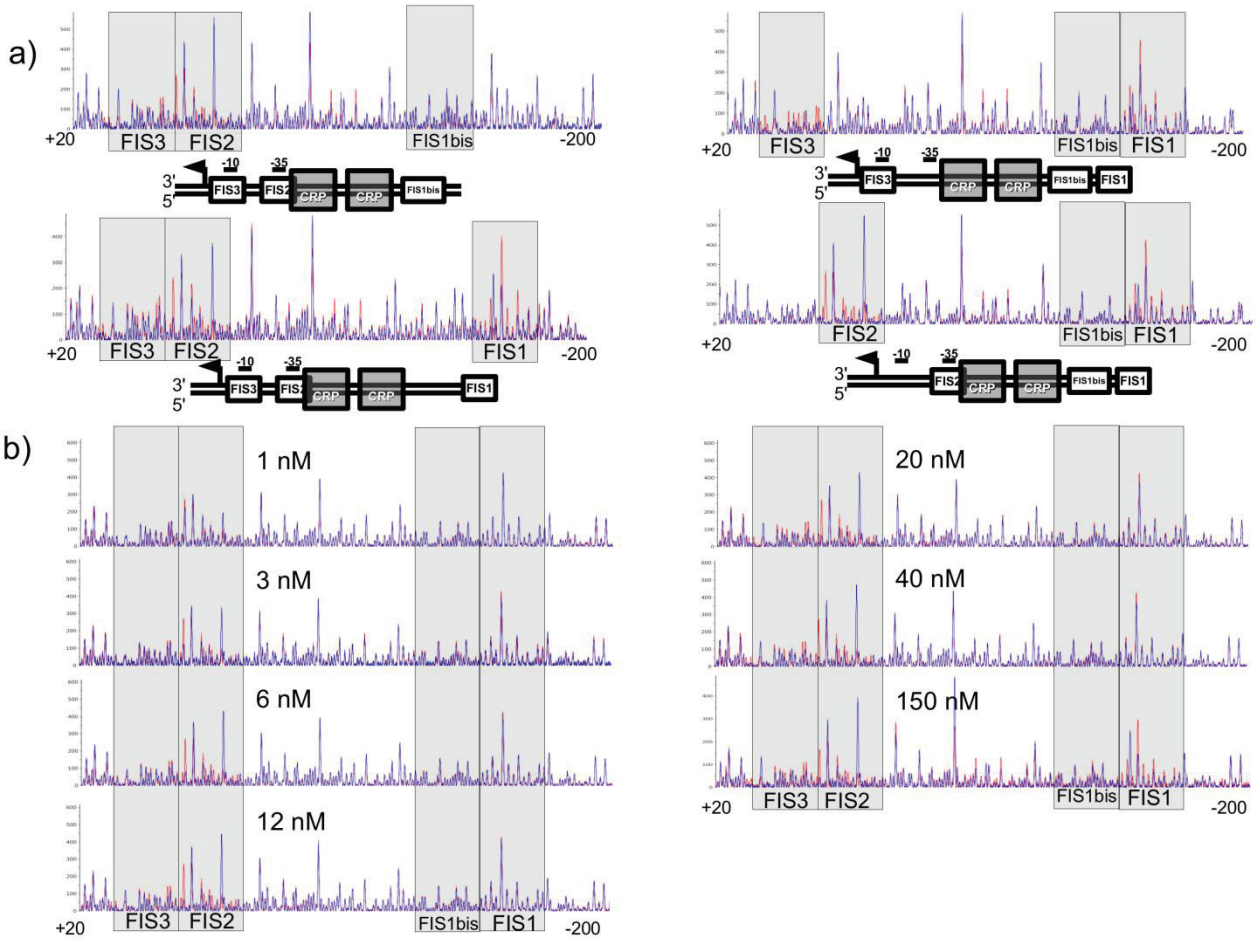


Figure 38: DNase I footprinting of FIS binding in the *pelD* regulatory region.

DNase I footprinting was performed using fluorescently labelled linear *pelD* DNA. The profiles represent the pattern obtained for the template strand. The red and the blue profiles correspond respectively to the samples without and with FIS. a) Profiles for *pelD* DNA samples with inactivated FIS1, FIS1bis, FIS2 or FIS3 sites at 150 nM FIS concentration. b) Profiles of the wild type *pelD* DNA incubated with increasing concentrations of FIS (from 1 to 150 nM)

3.2. The pProbe-Gfp reporter system allows accurate quantification of promoter activity in *D. dadantii*

Previously, most of the quantitative transcriptional activity measurements in *D. dadantii* have been performed by using beta-glucuronidase (GUS) genomic insertions [167], or by quantitative RT-PCR (q-RT-PCR) [122,193]. Both methods are relatively time-consuming, due to the need to recombine the numerous transcriptional fusions into the chromosome or to extract RNA, respectively. Alternatively, heterologous GUS fusions on plasmids have been successfully used in *E. coli* [253]. We explored two reporter genes, Green fluorescent protein (Gfp) and Firefly luciferase (Luc), which would allow to perform easy, plasmid-based, real-time and automated measurements of gene expression in *D. dadantii*. Gfp is a highly stable protein (> 20 days half-life, [249]), while GUS is moderately stable (4h half-life, [39]) and Luc is an unstable protein (10 min half-life, [254]). The expression of a *pelD* transcriptional fusion with each reporter (Figure 39) in *D.dadantii* wild type (WT) and its *fis* derivative backgrounds was measured. We found that the amplitude of the differences was negatively correlated with the stability of the proteins (ranging from a fold change of 1.4 for Gfp to 7.0 for Luc). However this increased sensitivity comes at the cost of decreased accuracy, especially in the *fis* mutant, with the coefficient of variation being 20% for Luc, but only 7% for Gfp. This results in weaker statistical tests for the Luc system e.g. a Welch test comparing WT to *fis*- conditions gives $p<10^{-9}$ for Gfp but only $p<10^{-3}$ for Luc (n=12 per condition for each reporter system). Consequently, to obtain reproducible results, the Gfp system was used for further measurements, since despite lower sensitivity even small (but statistically significant) differences in fluorescence could be assumed to be biologically relevant.

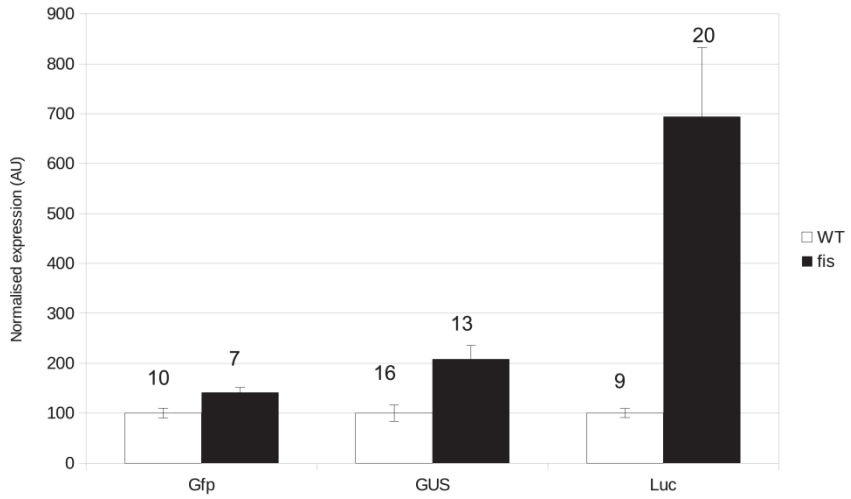


Figure 39: Comparison between the Gfp and Luc reporter systems in *D. dadantii* and the GUS system in *E. coli*.
Bacteria transformed with the wild-type *pelD* regulatory region in a reporter plasmid were cultured in equivalent conditions (see materials and methods). The normalised fluorescence and GUS specific activity were taken in early stationary phase and the luminescence at the maximum. Data represents the mean of 3 (GUS) to 12 replicates (Gfp and Luc) and error bars represent the standard deviation. Numbers above the error bars indicate the coefficient of variation (in %).

3.3. The *pelD* FIS binding sites are partially redundant

We next investigated the role of the different FIS binding sites in the modulation of *pelD* expression. To achieve this, plasmid-borne *gfp* transcriptional fusions were constructed with the

mutated *pelD* promoters and transformed in *D. dadantii* wild type (WT) and *fis* mutant strains. The fluorescence of the cells was measured in early stationary phase, when the transcription of *pel* genes is maximal [186]. The results are shown in Figure 40.

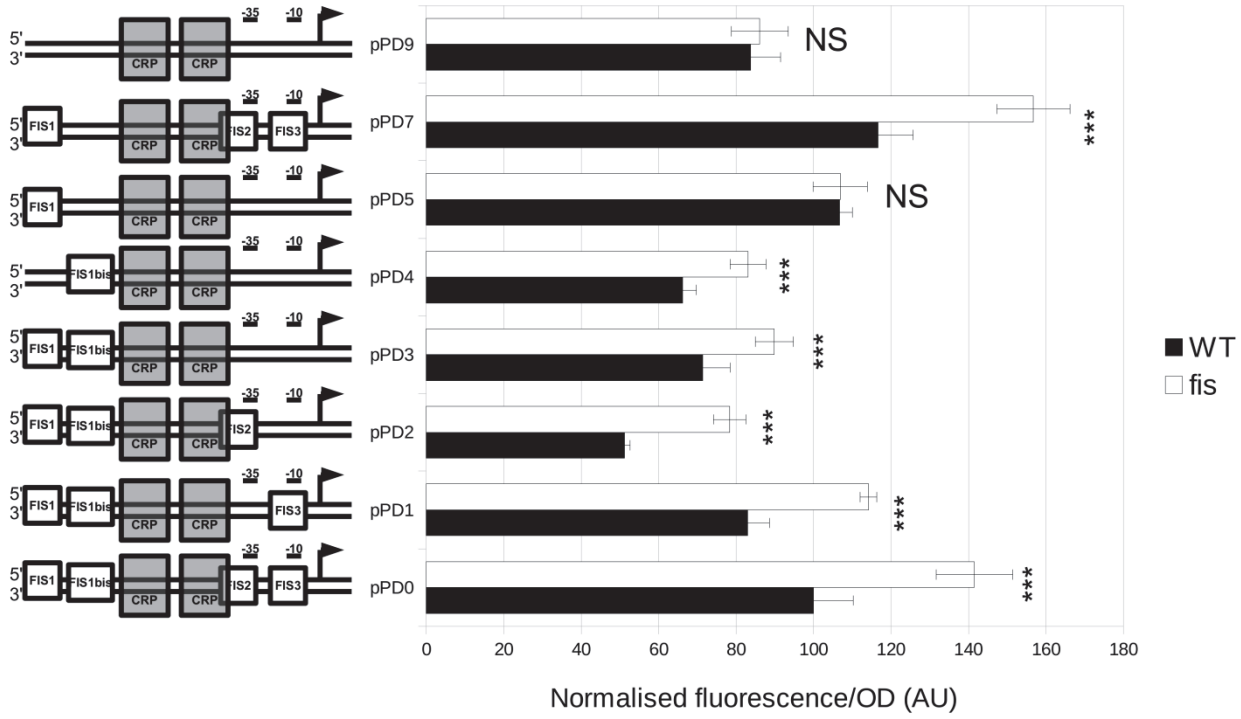


Figure 40: Transcriptional activity of various *pelD* promoter constructs in *D. dadantii* wild type and *fis* mutant backgrounds.

D. dadantii wild type and *fis* strains transformed with various *pelD* constructs (pProbe-Gfp plasmids D0 to D9 represented schematically on the left) were cultured in 96 well plates and the normalised fluorescence was determined in early stationary phase. Data represents the mean of 7 to 34 replicates and error bars represent the standard deviation. NS: not significant; ***: p<0.001 (Tukey's HSD test). Note that some of the introduced mutations affect the basal *pelD* promoter activity due to the overlap between the FIS binding sites and the core promoter sequence.

We found that as expected, FIS acted as a repressor [167], since the *pelD* activity was significantly increased in the *fis* mutant (see D0 in Figure 40). Inactivating either FIS1bis, FIS2 or FIS3 separately (compare D0 with D1, D2 and D7) had no significant impact on the repression by FIS (the fold-change remained constant at 1.40). Simultaneous inactivation of FIS2 and FIS3 (D3) alleviated the repression by FIS from 1.40-fold to 1.25-fold ($p<10^{-4}$, t-test for contrasts), and this residual repression can be attributed to the intact upstream sites. Furthermore, inactivation of FIS1, FIS2 and FIS3 had a similar effect as inactivation of FIS2 and FIS3 (compare D4 and D3), suggesting that FIS1 does not play any role in *pelD* repression *in vivo*. However, only the simultaneous inactivation of FIS1bis, FIS2 and FIS3 resulted in a complete elimination of repression by FIS, comparable to the construct in which all FIS binding sites were inactivated (compare D5 and D9). It thus appears that FIS2, FIS3 and FIS1bis are required for full repression of the *pelD* promoter by FIS *in vivo*. *In vitro* transcription reactions performed with *pelD* plasmid constructs yielded similar results (Figure 41). Inactivation of both FIS2 and FIS3 together was necessary to fully remove the repression of *pelD* by FIS, but, unlike *in vivo*, FIS1bis inactivation showed no effect, in keeping with the barely detectable occupancy of FIS1bis site *in vitro* (Figure 38).

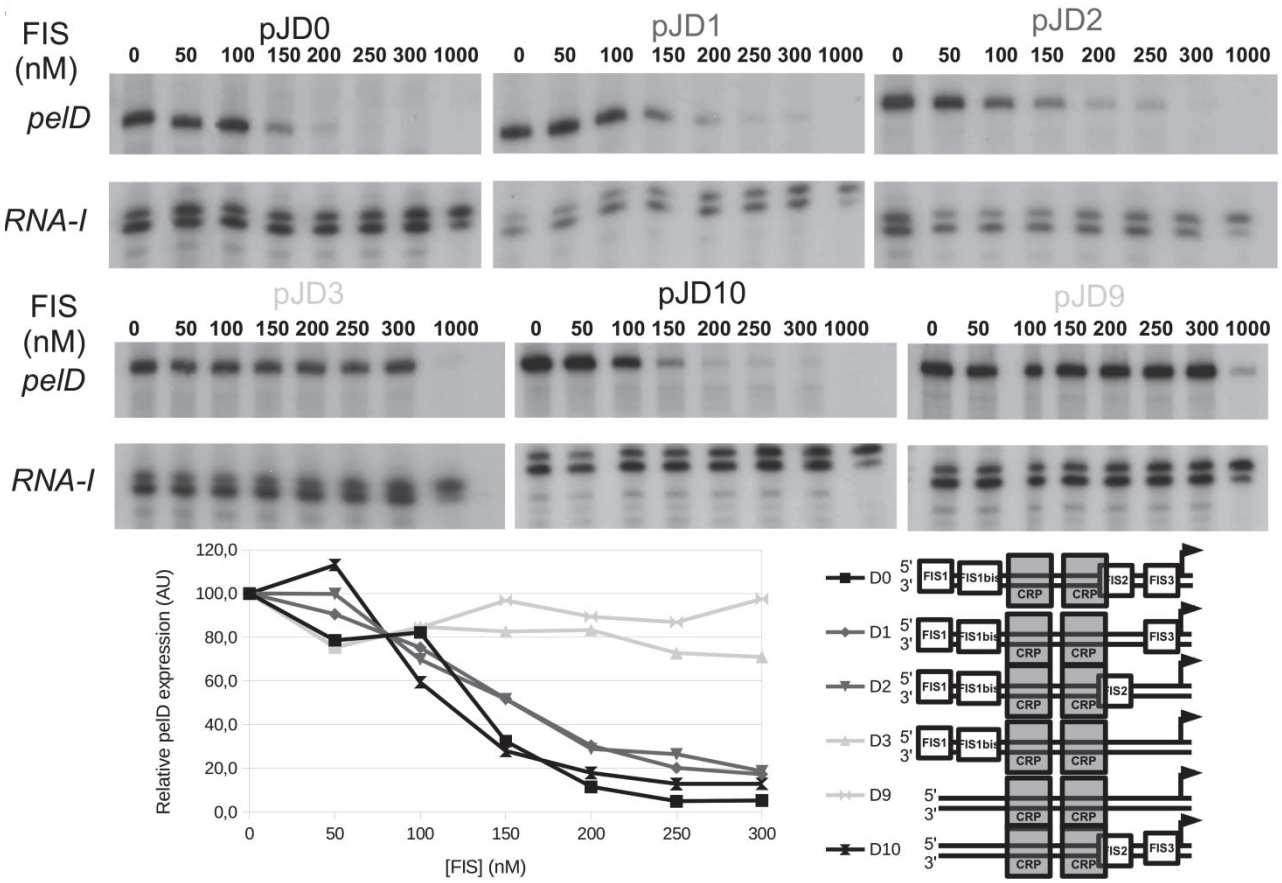


Figure 41: Effect of FIS binding on *pelD* transcription in vitro.

In vitro transcription was performed using several *pelD* constructs (plasmids D0, D1, D2, D3, D9 and D10 indicated in the bottom right panel) in presence of 100 nM CRP, 40 nM RNA polymerase and varying concentrations of FIS. The RNA-I internal control was used as a standard for quantification of the *pelD* transcript (bottom left panel).

3.4. FIS competes with binding of CRP only at the CRP2 site

Whereas the occupancy of either of the strong FIS2 and FIS3 sites overlapping the RNA polymerase (RNAP) recognition elements appears sufficient to prevent RNAP binding at the *pelD* promoter, FIS2 also overlaps the CRP2 binding site (see Figure 37), suggesting a potential competition between FIS and CRP for promoter binding. Therefore, quantitative DNase I footprinting was performed to examine whether the presence of FIS had an impact on the CRP footprint at the CRP2 site (Figure 42). On the wild-type regulatory region, a footprint by CRP at the CRP2 site was visible only in the absence of FIS, whereas on templates with inactivated FIS2 (or both FIS2 and FIS3) site(s), the protection of CRP2 site by CRP was observed despite the presence of FIS. In contrast, the protection of CRP1 site was independent of the presence of FIS on both the wild type and mutated templates (Figure 43). This suggests that binding of FIS at FIS2 specifically interferes with occupation of the CRP2 site, whereas FIS does not interfere with CRP binding at CRP1. We infer that binding of FIS represses the *pelD* promoter activity by competing not only with RNA polymerase binding, but also with binding of the activator CRP.

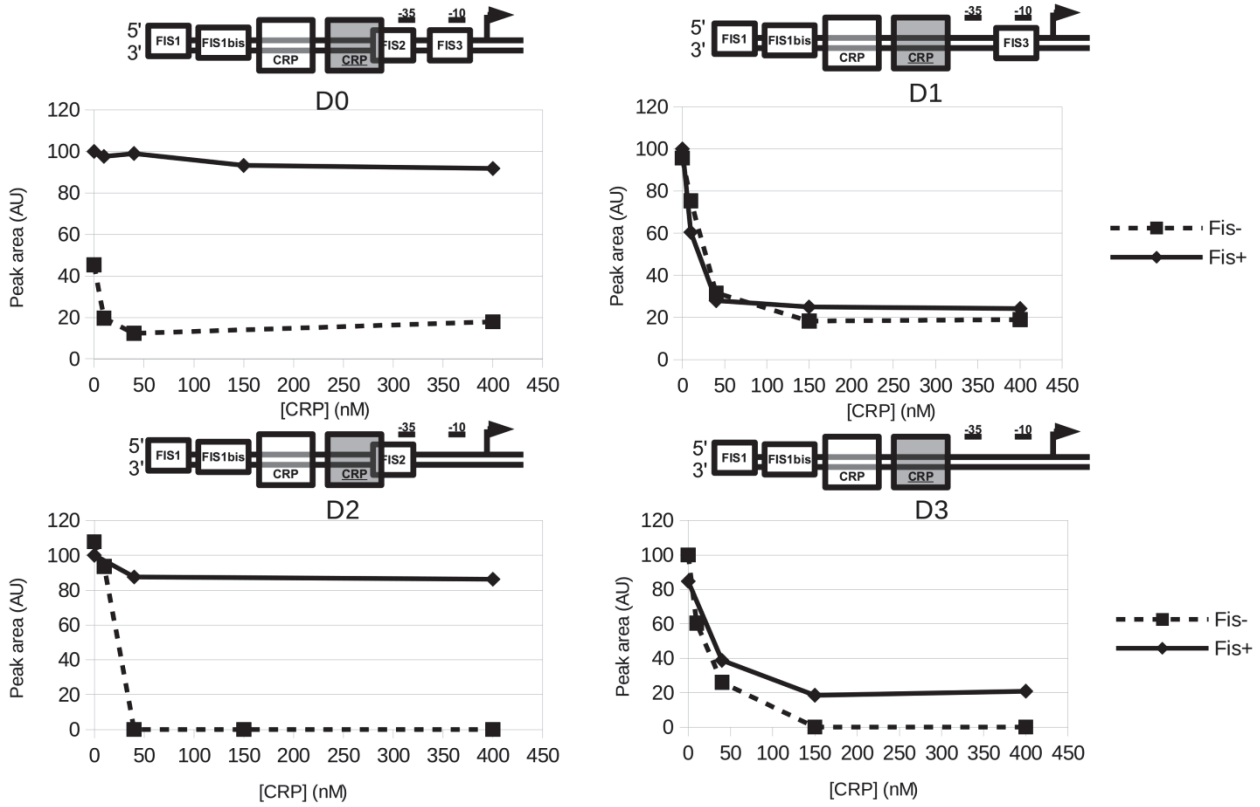


Figure 42: Modification of the CRP2 binding by FIS.

Quantitative DNase I footprinting using fluorescently labelled linear *pelD* DNA fragments lacking one or both of the FIS binding sites overlapping the *pelD* promoter. The reactions were performed in the presence or absence of FIS (150 nM) and with increasing concentrations (from 0 to 400 nM) of CRP. The area of the peaks at one position within the CRP2 binding site (at position -29 relatively to the transcription start site) is represented for different CRP concentrations. This position is protected by CRP and becomes hypersensitive in presence of FIS. Absence of variation in this area (as observed with D0 and D2 in presence of FIS) means that CRP cannot bind the corresponding site. Consistent results were obtained also with two other positions analysed within the CRP2 footprint (-47 and -21).

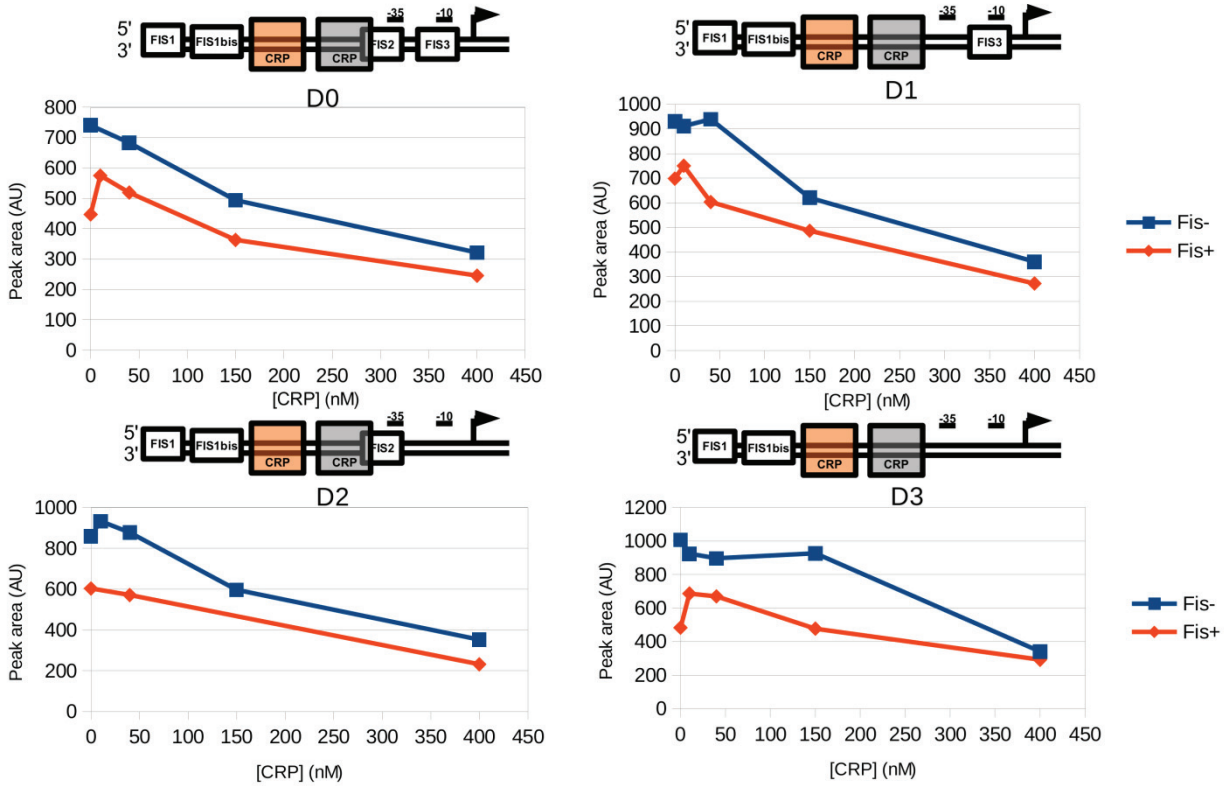


Figure 43: Absence of modification of the CRP1 DNase I footprint by the presence of FIS.

A quantitative DNase I footprint was performed on fluorescently labelled pelD regulatory regions, in identical conditions as in Figure 42. The area of the peaks at one position inside CRP1 (at position -85 relatively to the transcription start site) for different CRP concentrations is represented.

3.5. The upstream divergent promoter is activated by supercoiling and repressed by FIS

The upstream FIS binding sites (FIS1 and FIS1bis) are located too far upstream of *pelD* core promoter to be directly involved in competition with RNA polymerase for binding. However, previous DNAse I footprinting experiments have revealed that this upstream region contains a high affinity binding site for RNA polymerase (extending from pos. -142 to -104) in divergent orientation to *pelD* [146]. Indeed, recent RNAseq data analysis (*Jiang et al.* in preparation) confirmed by our *in vitro* transcription experiments identified a divergent transcript initiating from *pelD* upstream region and devoid of any meaningful ORF. We propose that this region comprises a divergent promoter (denoted hereafter as *div*), which potentially could influence the expression of *pelD* and also be regulated by FIS and CRP *via* FIS1/FIS1bis and CRP1 sites, respectively. Indeed FIS1/FIS1b overlap the -10 and -35 elements of the *div* promoter, whereas CRP1 is centred at -60.5 with respect to the *div* transcription initiation site (Figure 37). This latter organisational feature is characteristic of class I CRP-dependent promoters, suggesting that CRP1 site might be used for activation of both the *pelD* [26] and *div* promoters by CRP.

To examine the relationship between the *pelD* and *div* promoters we measured the amount of mRNA produced from both promoters on the plasmid constructs *in vivo*. We found that the peak of *pelD* expression observed as expected, on transition to stationary phase, was preceded by a peak of *div* expression, observed early during the exponential growth phase (Figure 44a), whereby the *div* promoter showed significantly lower expression levels than *pelD*. Measurements of the untwisting of the *div* and *pelD* promoters using *in vivo* $KMnO_4$ reactivity assay (see

Materials and methods) confirmed that the peak of *div* promoter untwisting during exponential phase preceded that of *pelD* occurring on transition to stationary phase (Figure 45), albeit with relatively less precision and temporal resolution compared to the *in vivo* expression data. Interestingly, whereas these two promoters were sequentially activated during growth of *D. dadantii,* the decline of *div* activity clearly coincided with activation of *pelD.* We asked whether this correlation – given the close proximity of the promoters - indicated a possible link in their regulation mediated by binding proteins. Indeed, not only was the expression of both *pelD* and *div* increased in the *fis* mutant, but also the peak of *div* expression was shifted to the right along the growth curve (Figure 44a).

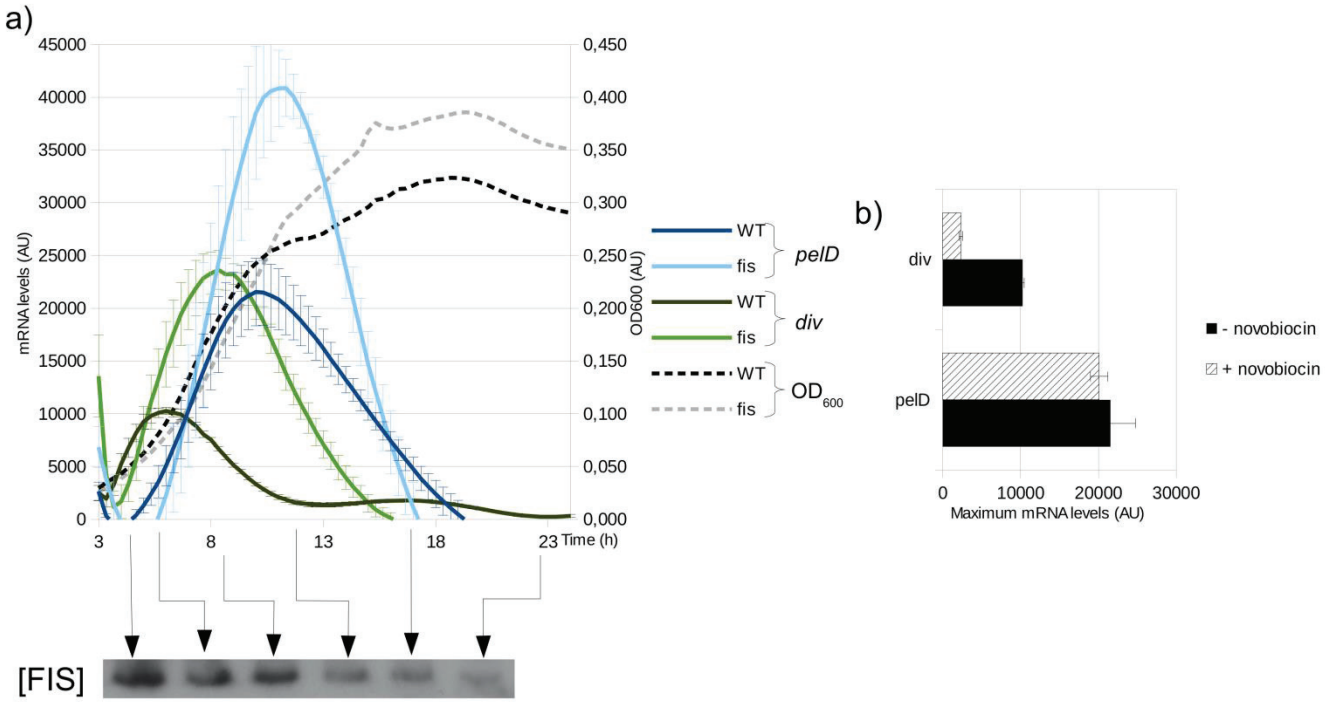


Figure 44: Transcriptional regulation of the *div* promoter.

Dickeya dadantii wild type and *fis* mutant cells were transformed with constructs carrying *pelD* promoter-*gfp* or *div* promoter-*gfp* fusions and cultured in 96 well plates. Sublethal concentrations of novobiocin (25 µg/mL) were added at the start of the culture. a) The mRNA levels produced from different constructs in wild type and *fis* mutant cells were computed from fluorescence data as previously described [248]. Western blots indicating the levels of FIS at the indicated time points during the growth cycle. b) Maximum calculated mRNA levels obtained during the experiment in the presence or absence of novobiocin.

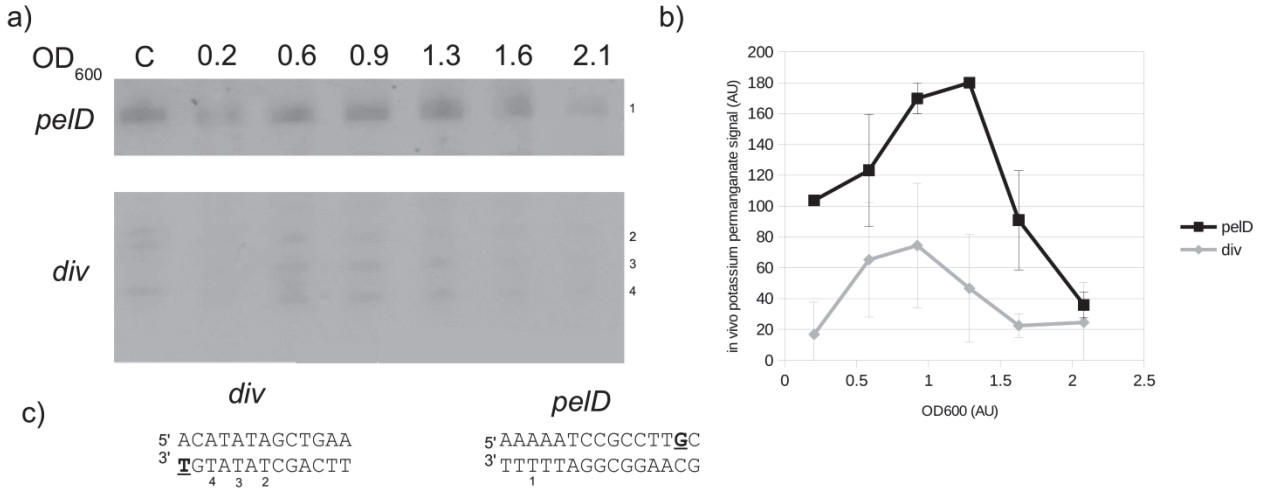


Figure 45: *in vivo* potassium permanganate footprinting of the *pelD* regulatory region.

pKD0 (wild-type *pelD*) was transformed into *D. dadantii.* Bacteria were then cultured in M63S + PGA and an *in vivo* potassium permanganate footprinting was performed as described in materials and methods. (a) image of one replicate of the experiment. Lane C corresponds to an *in vitro* potassium permanganate footprinting control performed on the same plasmid with 100 nM CRP and 40 nM RNAP. Bands are numbered and localised in part c of the figure. (b) quantification of the intensity of the bands, represented as mean +/- SD of 2 independent biological replicates. (c) Position of the attacked nucleotides of part a. Both strands are shown and transcription start sites are underlined.

It is noteworthy, that despite the de-repression of *div* promoter in the *fis* mutant, its activity sharply declined during the transition of cells to stationary phase, which is characterised by decrease in global DNA superhelicity (Figure 44a). Since unlike *pelD*, the peak of *div* activity was observed in early exponential growth phase characterised by high levels of negative DNA superhelicity, we compared the response of the *div* and *pelD* promoters to DNA relaxation (Figure 44b). We found that in contrast to *pelD* promoter, the *div* promoter activity was substantially decreased on DNA relaxation by novobiocin, a coumarin drug inhibiting the generation of negative superhelicity by DNA gyrase [255]**.** Therefore, whereas FIS clearly represses *div* promoter activity (Figure 44a), the observed reduction of *div* transcription later during the growth cycle (observed in both the wild type and *fis* mutant) could be explained by relaxation of DNA on transition of the cells to stationary phase [256].

3.6. The two CRP binding sites synergistically activate both *pelD* and *div*

To investigate the impact of the CRP1 and CRP2 sites in *pelD* and *div* promoter expression we inactivated these sites either separately or together (see legend to Figure 46). The resulting constructs were used to quantify the promoter expression in the wild type strain. The data are shown in Figure 46. We found, as expected, that inactivation of either CRP1 or CRP2 separately strongly impairs the *pelD* promoter activity. However, this effect is more pronounced in the absence of CRP2. This observation is consistent with observed higher affinity of CRP2 compared to CRP1 (compare Figure 42 and Figure 43) and its optimal position (centred at -42.5 bp) for activating the *pelD* promoter. It thus appears that whereas CRP2 is the main activator site, both CRP binding sites synergistically activate the *pelD* promoter. At the same time, the CRP1 and CRP2 sites also act synergistically on the *div* promoter but in this case the impact of CRP1 site is predominant, in keeping with its optimal location (-60.5 bp) with regard to the *div* transcription initiation site (see Figure 37). Thus, binding of CRP to CRP1 and CRP2 sites synergistically modulates both *pelD* and *div* promoter expression, although the impact of individual sites at each promoter is different.

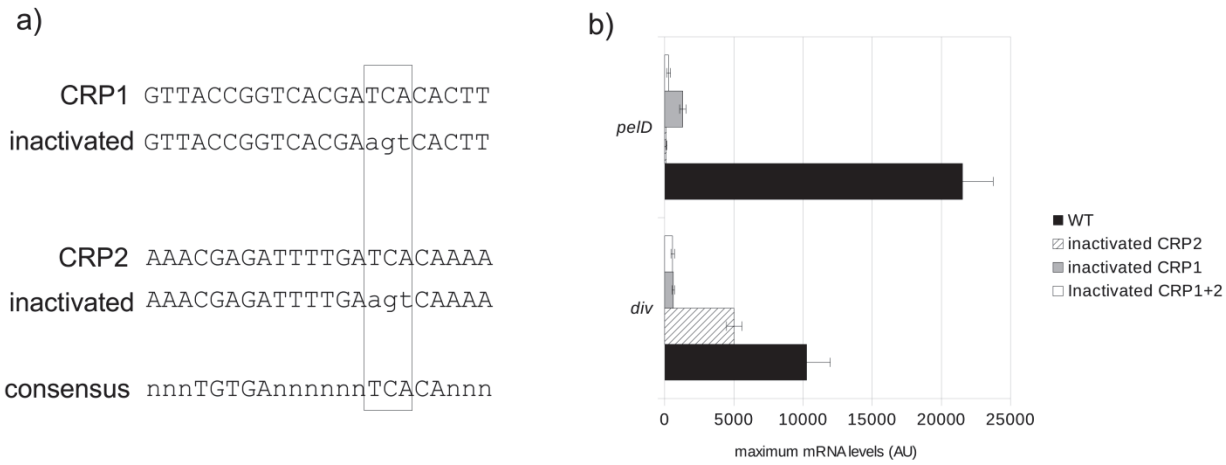


Figure 46: Individual contributions of the CRP1 and CRP2 sites on *pelD* and *div* expression.

(a) CRP sites were inactivated by site-directed mutagenesis, substituting several of the most conserved nucleotides in the consensus, using the same approach as for FIS sites (see methods). (b) Expression of the mutated promoters was measured as in figure 6b. Data represents the mean of 10 replicates and error bars indicate standard deviation.

3.7. The divergent promoter potentially competes with *pelD* for RNA polymerase binding *in vitro*

To gain more insight into the relationship between the *div* and *pelD* promoters we investigated *in vitro* the promoter opening using potassium permanganate reactivity assay. In these experiments, conducted with various concentrations of RNA polymerase and CRP, we observed that under conditions of low transcriptional activity (low concentrations of both CRP and RNAP), the opening of both promoters was similar, whereas under conditions favouring high levels of transcription (high concentrations of both RNAP and CRP), an up to 6-fold stronger reactivity to $KMnO_4$ is observed at the *div* promoter compared to the *pelD* promoter (Figure 47). These results suggest that under conditions supporting active transcription *in vitro*, the *div* promoter can gain an advantage over the *pelD* promoter in binding RNA polymerase.

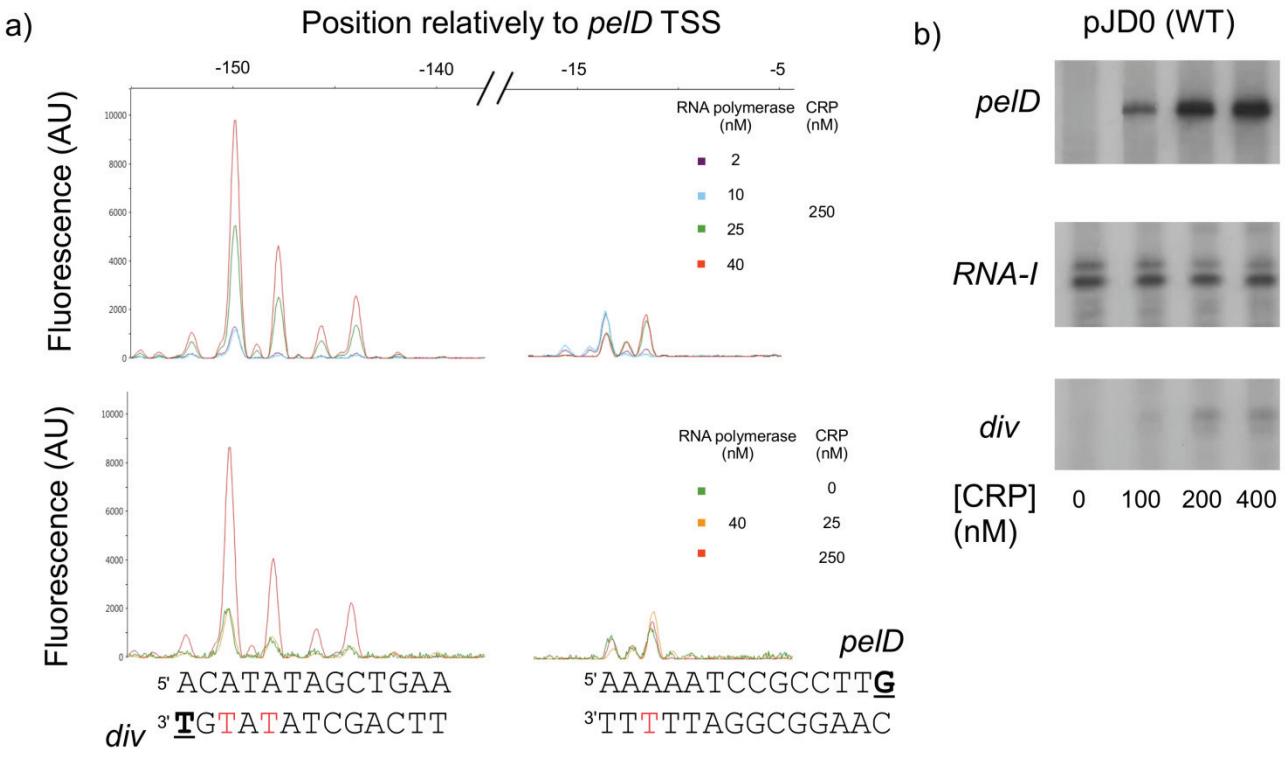


Figure 47: Preferential binding of RNA polymerase at *div* does not correlate with stronger transcription.

a) Potassium permanganate reactivity of the supercoiled *pelD* regulatory region in the presence of varying concentrations of CRP and RNA polymerase as indicated on the right of the figure. The *pelD* and *div* transcription initiation sites are underlined in the corresponding sequences shown below the reactivity profiles. Both starnds are shown and attacked Ts are indicated in red. The capillary electrophoresis fluorescence profiles were normalised to the *pelD* peaks and overlaid with the position of each peak relatively to the *pelD* transcription start site. b) *In vitro* transcription using the supercoiled wild type *pelD* promoter was performed in the presence of 40 nM RNA polymerase and varying concentrations of CRP. The *pelD*, RNA-I (internal standard) and *div* transcripts are shown.

In the *in vitro* transcription reactions performed in parallel, we found that although both the *pelD* and *div* promoters were activated by CRP, this effect was much more pronounced for the former (Figure 47b), probably due to the strong activating effect of the high-affinity binding CRP2 site centred at classical -42.5 bp with respect to the *pelD* promoter transcription startpoint. The discrepancy between the efficient opening and low transcriptional activity of the *div* promoter suggests that this promoter may act as a trap for RNA polymerase molecules readily forming stable open complexes but only rarely initiating transcription. One plausible explanation for this behavior is the strong -35 hexamer (TTGAAT) of *div*, which combined with further stabilization of polymerase binding by CRP (especially via CRP1 site) could be detrimental to promoter clearance [29].

3.8. Inactivation of *div* reduces the *pelD* activity *in vitro*

To directly assess the effect of the *div* promoter on *pelD* activity we compared the promoter function of wild type *pelD* construct and a construct with a mutation (D24) inactivating the -35 element of the *div* promoter and demonstrating a strong reduction of *div* opening by $KMnO_4$ reactivity assay (Figure 48a).

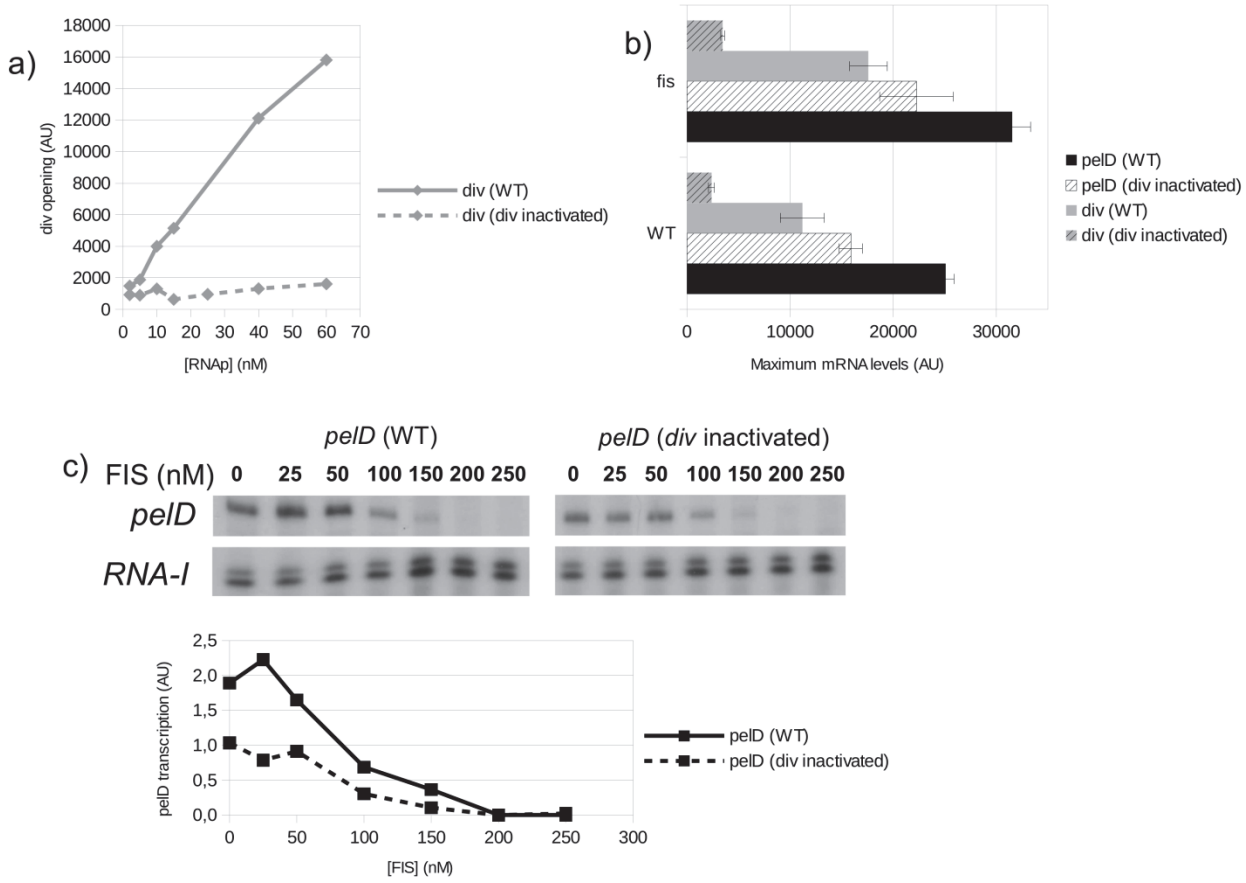


Figure 48: Impact of *div* on *pelD* transcription.

The *div* promoter was inactivated by mutating the -35 box (see Figure 37) and the impact of this mutation was assessed. a) Potassium permanganate reactivity (opening) of the wild type and inactivated *div* promoter in the presence of 100 nM CRP and varying concentrations of RNA polymerase. The promoter opening was quantified as described in the legend to Figure 44. b) *In vivo* transcription of *pelD* and *div* constructs with intact and inactivated *div* promoter carried out under the same conditions as in Figure 44a. *pelD* and *div* expression was measured in real-time for 48 hours, like as in Figure 44a. Because *pelD* and *div* are not expressed at the same time (see Figure 44a), the maximum expression of both promoters is taken for this graph, which means early log for div and late log for *pelD*. c) *In vitro* transcription of *pelD* and *div* constructs with intact and inactivated *div* promoter in the presence of 100 nM CRP, 40 nM RNA polymerase and varying concentrations of FIS.

Using both these constructs we compared the *pelD* mRNA levels produced in wild type and *fis* mutant cells. We observed that inactivation of *div* decreased *pelD* transcription in both the *D. dadantii* wild type and *fis* mutant cells (Figure 48b), indicating that *div* exerts an activating effect on *pelD* transcription and that activation is largely independent of FIS (this effect was slightly more pronounced in wild type cells showing an 1.56-fold increase in *pelD* expression in presence of an intact *div* in WT cells, and an 1.41-fold in the *fis* cells).

Using both these constructs we also performed *in vitro* transcription experiments (Figure 48c). With the construct carrying inactivated *div* we observed a 2-fold decrease of *pelD* transcription compared to that with intact *div*, confirming in a purified system that *div* activates *pelD*.

4. Discussion

4.1. The *pelD* regulatory region structure

The purpose of this work was to provide an integrated picture of the temporal regulation of the *pelD* promoter driving the expression of a major virulence factor of the plant pathogen *D. dadantii*. Upstream of the *pelD* promoter we identified a divergent promoter (*div*) which is not associated with any meaningful ORF and the activity of which is strongly reduced by DNA relaxation *in vivo*. This sensitivity to supercoiling is consistent with maximum *div* activity during early exponential phase, when the DNA superhelical density is maximal, and its sharp reduction during transition of cells to stationary phase and associated DNA relaxation [256]. In contrast, the *pelD* promoter is activated on transition to stationary phase concomitantly with the decline of *div* promoter activity. We found that the relationship between the closely spaced *div* and *pelD* promoters is not trivial, as under conditions supporting high rates of transcription *in vitro*, the *div* promoter binds RNAP and forms open complexes much more efficiently than *pelD* and could thus potentially compete with the latter (Figure 47a). However, inactivation of the *div* promoter significantly decreases *pelD* promoter activity (Figure 48bc). Furthermore, despite the facile formation of open complexes by RNAP, the *div* promoter is considerably less active than *pelD in vivo* (Figure 44a) and is barely transcribed *in vitro* (Figure 47b). The reason for this discrepancy is not entirely clear, but the structural organisation of *div* strongly resembles that of a divergent promoter identified in the upstream region of the *E. coli fis* operon [62]. Both these weakly transcribed promoters demonstrate similar intrinsic properties such as overly strong untwisting in conjunction with reasonably close-to-consensus -35 hexamer sequences (TTGCAT for the *fis* divergent promoter, TTGAAT for the *pelD* divergent promoter).

In addition to studying the effects of two previously described CRP binding sites on *pelD* and *div*, we identified four FIS binding sites, two of which (FIS2/3) overlap the *pelD* promoter sequence, and two are located upstream (FIS1/1bis) and overlap with the *div* promoter. The FIS2 and FIS3 sites act as repressor sites inhibiting the *pelD* expression during exponential growth phase (Figure Figure 44a), whereas the FIS1 and FIS1bis sites, although weak, can repress *div* transcription during early exponential growth *in vivo*, when the FIS concentration is high (Figure Figure 44a). However, it is noteworthy that inactivation of FIS1bis site affected *pelD* activity only *in vivo* (compare pPD4 and pPD5 in Figure 40) but not *in vitro* (compare pJD3 and pJD9 in Figure 41). This could be attributed to differences in the *in vitro* and *in vivo* conditions and indicate more complex relationships, perhaps involving changes of other regulators not known at this time.

Repression of *pelD* by FIS appears to be mediated by two mechanisms. The FIS2 site is the main repressor site preventing the binding of both the RNA polymerase and the activator CRP at CRP2 site, whereas FIS3 site prevents the binding of RNA polymerase, but does not interfere with CRP binding. The second CRP binding site (CRP1) is suitably positioned to activate both the *div* and *pelD* promoters by providing interaction surfaces for the C-terminal domains of the α subunits (α-CTDs) of RNAP [257]. Indeed, our data suggest that binding of CRP at CRP1 and CRP2 synergistically activate both the *pelD* and *div* promoters. However, the relative contribution of these two sites to the activity of the *pelD* and *div* promoters is different. The pertinent question is whether the identified organisation of the DNA binding sites and the observed relationships between the binding molecules of RNAP and global regulators can be rationalised with regard to the critical function of *pelD* gene in bacterial virulence.

4.2. Mechanism of growth phase dependent regulation of *pelD*

For *D. dadantii* it is vital to completely repress the expression of *pelD* (as well as other *pel* genes) in exponential phase, especially since any leaky *pel* expression would trigger a massive

production of Pels due to the KdgR positive feedback loop [236]. If the bacterial population density is low, an early activation of *pel* genes results in an early activation of the plant defense reactions and ultimately in bacterial death [258] as observed, for example, in *D. dadantii* mutant lacking the *fis* gene [167]. This justifies an establishment of redundant mechanisms for efficient *pelD* repression during exponential growth.

Our data suggest that *D. dadantii* uses three different mechanisms to prevent aberrant expression of *pelD* during exponential growth phase: i) direct hindrance of the RNA polymerase binding *via* FIS2 and FIS3, ii) competition between FIS and activator CRP for binding of the overlapping FIS2 and the CRP2 sites and, iii) potential competition of the *div* promoter with *pelD* promoter for RNA polymerase binding. Our observations allow us to propose a model for regulation of the expression of *pelD* by taking into account all these factors and the expression patterns of FIS [50] and the *div* promoter (Figure 49). CRP is obviously central for this regulation, as it is absolutely required for the activity of both the *pelD* and *div* promoters. The DNA region with two CRP binding sites organised at optimal positions for activating two divergent (*pelD* and *div*) promoters would serve as a trap for RNA polymerase directing transcription in one or another direction depending on the transcriptional environment. The expression peak of the *div* promoter in *D. dadantii* is observed at 6h during early exponential phase, i.e. when FIS is near its maximum concentration (Figure 44a, lower panel). Given that FIS1/1bis are low affinity sites, high FIS concentrations perhaps weakly interfere with binding of RNAP to *div* but fully repress *pelD* via FIS2/3. In addition, in early exponential phase the DNA is highly supercoiled [233], allowing maximal *div* promoter activity mainly supported by CRP binding at CRP1 site centred at -60.5 with respect to the *div* transcription initiation site. Indeed, whereas FIS precludes the binding of CRP at CRP2 site and activation of *pelD*, it does not interfere with binding of CRP at CRP1 site and activation of *div*, consistent with different impact of individual CRP binding sites on the *pelD* and *div* promoter activity (Figure 46b). At this growth stage the high-affinity *div* promoter can serve as an efficient trap withdrawing the RNA polymerase molecules that could otherwise bind at the *pelD* promoter. During mid-exponential phase the concentration of FIS is still sufficient to repress the *pelD* promoter via high affinity FIS2/3sites, but on transition to stationary phase (12h), the FIS concentration subsides (see Figure 44a, lower panel) enabling transcription of *pelD.* Our data suggest that the burst of *pelD* expression on transition of cells to stationary phase is likely facilitated by two additional mechanisms. First, the reduced transcription of *div* (e.g. due to the global relaxation of DNA), would abolish competition with *pelD*. Second, due to the increased cAMP levels at this stage [256], the cooperative binding of cAMP-CRP to CRP1 and CRP2 activator sites [146] would stabilise RNAP binding at *pelD* leading to its maximal expression. We propose that growth phase dependent competition between FIS and CRP binding in conjunction with alterations of DNA topology provide an efficient temporal mechanism enabling both tight repression and strong activation of the *pelD* promoter. It is likely that binding and bending of the DNA by CRP [48] at the CRP1 and CRP2 sites generally facilitates the recruitment of the RNAP molecules, whereas the changing pattern of CRP site occupation (binding of CRP at CRP1 initially, and then at both CRP1 and CRP2), determines the distribution of polymerase between *div* and *pelD*. Interestingly, the *pelD* regulatory region is predicted to be intrinsically bent by an angle of around 50° (Figure 50a) and the two helically phased CRP sites could potentially create a full 180° bend (Figure 50b), bringing *pelD* and *div* in close spatial proximity and promoting RNAP transfer between both. This mechanism implies that inactivation of the *div* promoter, despite its potential to compete with the *pelD* promoter for RNAP binding, could lead to decreased *pelD* activity.

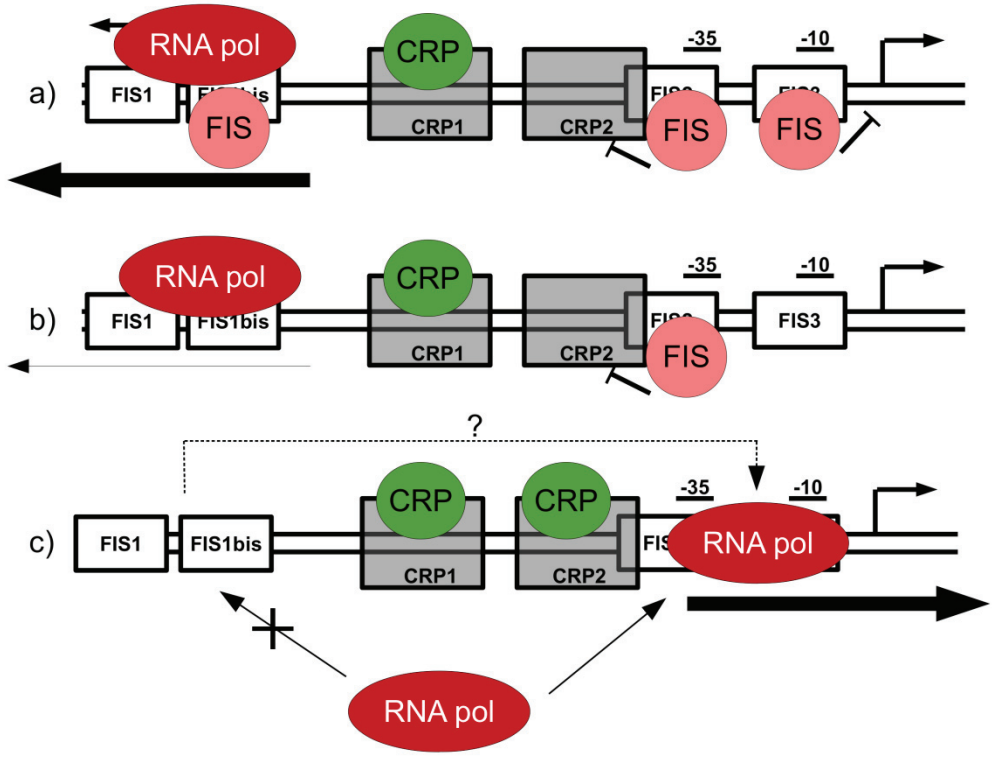


Figure 49: Proposed model for the temporal control of *pelD* by FIS, CRP and the divergent promoter.

For clarity, the *pelD* regulatory region is represented in a linear fashion. Binding sites are represented by boxes (white for FIS, grey for CRP). Proteins are represented by ellipses: red for RNA polymerase, green for CRP, pink for FIS. Large horizontal arrows indicate transcription. Lines ending with a bar indicate inhibition. a) During early exponential phase, FIS sequentially occupies most of its binding sites and both the *pelD* and *div* promoters are repressed, albeit to different extent, since at this stage *div* is activated by high negative superhelicity. The *div* promoter efficiently binds RNA polymerase due to its high affinity for RNAP and the low affinity of the competitor FIS to the FIS1 and FIS1bis sites. Binding of CRP at CRP1 site also favours binding of polymerase at *div* and so, the RNAP is "parking" at *div.* b) During the late exponential phase, DNA relaxation inactivates *div* and presumably destabilises the binding of RNAP. At this stage, *pelD* is still repressed due to the overlapping high-affinity FIS sites. c) During transition to stationary phase FIS levels subside and *pelD* is de-repressed. RNA polymerase now preferentially binds *pelD*. The activating effect of *div* on *pelD* at this stage is most likely due to a transfer of idling RNAP molecules previously "parking" at *div* towards *pelD.* This results in very strong transcription of *pelD*.

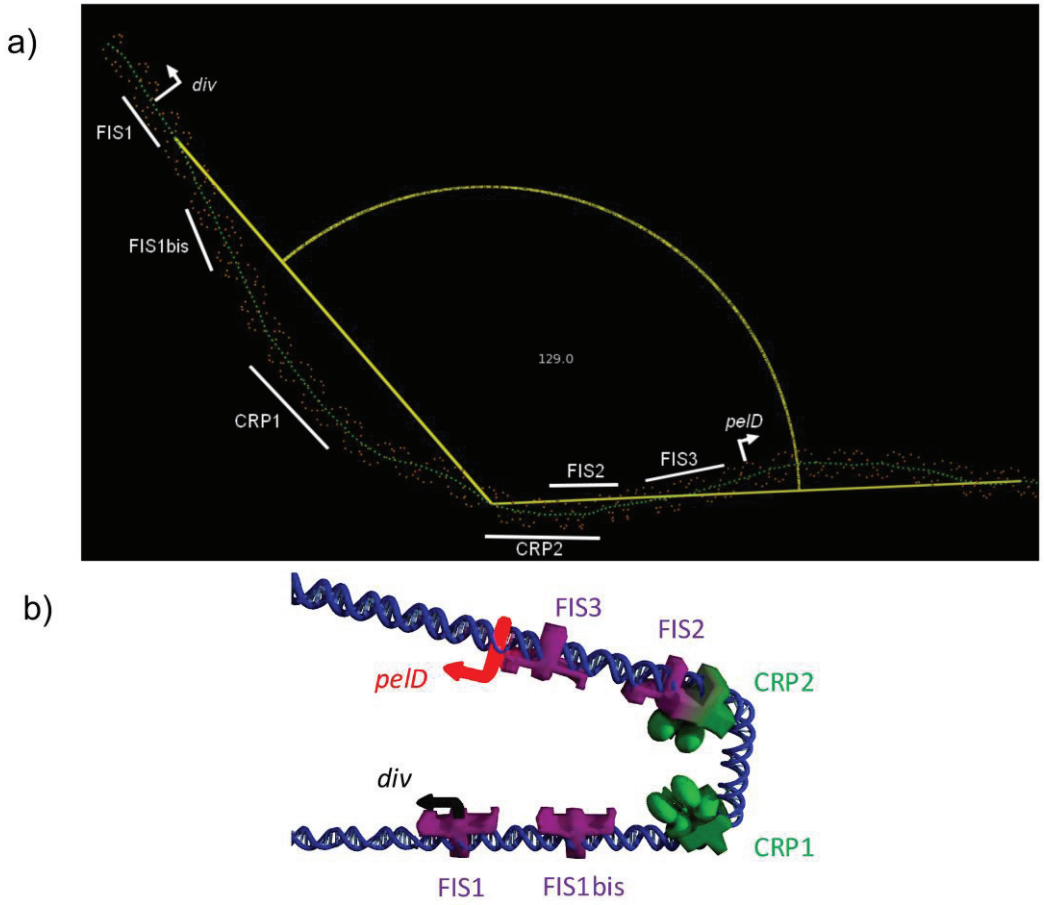


Figure 50: Predicted intrinsic and CRP-induced bending in the *pelD* regulatory region.

a) The intrinsic curvature of the *pelD* was modeled using DNAcurve with the Bolshoi-Trifonov model (http://www.lfd.uci.edu/~gohlke/dnacurve/) and the angle (in yellow) was measured using PyMOL, which corresponds to a bending angle of 51°. The DNA molecule is represented by orange dots and its trajectory in green dots. The main features of the *pelD* regulatory region are shown in white. b) Schematic representation of the CRP-induced bend in the *pelD* regulatory region. Considering that CRP can bend DNA to an angle of around 90°, the presence of both CRP sites could bring *pelD* and *div* in close proximity, facilitating RNAP transfer between both.

In stationary growth phase, *pelD* expression decreases below the detection limit independently of FIS (Figure 44a), showing that FIS is involved only in the initiation of *pelD* expression. This decrease is likely due to other regulators, among which KdgR which has already been identified [186]. Indeed, in stationary growth phase, PGA is totally depleted which makes KdgR repress the *pel* genes, including *pelD* [186].

4.3. Biological implications

Our data strongly suggest that the growth phase dependent regulation of *pelD* is under the control of changing composition of higher-order nucleoprotein complexes formed in the *pelD* regulatory region during the growth cycle. Such “alteration” of the composition of regulatory nucleoprotein complexes during the growth cycle [259] seems to be a common design principle employed for the growth phase-dependent gene regulation. Similar regulatory mechanisms can be found not only at different promoters in single species but also in different species, whereby involvement of several FIS binding sites in promoter repression is not unusual. A well-known example is the negative autoregulation of the *fis* promoter involving numerous FIS binding sites [260], and there are other examples, such as the *mom* [252] and the *gyr* promoters of *E. coli* [171] and *Salmonella* [261], to name just a few. However, multiple FIS binding sites can also be used to exert either an activating or repressing effect [262].

On the other hand, there are only a few reports of a direct competition of FIS with an activator [263,264]. The case of the *crp2* promoter is particularly relevant [264] as it shares many similarities with the *pelD* promoter such as two high-affinity FIS binding sites overlapping the -10 and -35 elements, whereby the FIS site overlapping the -35 element also competes with binding of the activator CRP. Moreover, the *crp* regulatory region harbours also a FIS-regulated divergent promoter competing with the *crp1* promoter, as well as an upstream CRP binding site, which is little affected by FIS. Yet, there is no sequence homology between the *D. dadantii pelD* and *E. coli crp* promoters, suggesting a convergent evolution towards similar structural organisation of gene regulatory regions.

Also the *E. coli fis* promoter is subject to complex regulation involving FIS, CRP and a supercoiling-dependent divergent promoter [62]. In this latter case the RNAP molecule binding an upstream divergent site appears to directly interact with the RNAP bound at the *fis* promoter and this interaction is implicated in sensing the changing DNA superhelicity during the growth cycle. Similarly, in the *pelD* regulatory region it is the *div* promoter which strongly depends on supercoiling and thus may serve as a sensor of DNA superhelicity transmitting its changes to the *pelD* promoter. Whether the two RNAP molecules binding in the *pelD* regulatory region can interact, is an open question that merits further investigation. Theoretically, the 155 bp separation between the *div* and *pelD* transcription start sites (given that RNAP footprints are typically 70 bp long) could suffice to avoid steric hindrance between the RNAP molecules, but assuming that polymerase on average wraps about 90 bp of on binding the promoter DNA [265,266] there could be competition, as is for example, in the case of the *crp* promoter ([56,264] and Figure 51). It is noteworthy that also in the *crp* regulatory region, only one of the promoters (*crp2*) appears to be strongly sensitive to supercoiling, suggesting that the topological state of the DNA is one major variable involved in modulating the growth phase-dependent impacts of regulatory nucleoprotein complexes assembled on divergent promoter modules [56,264].

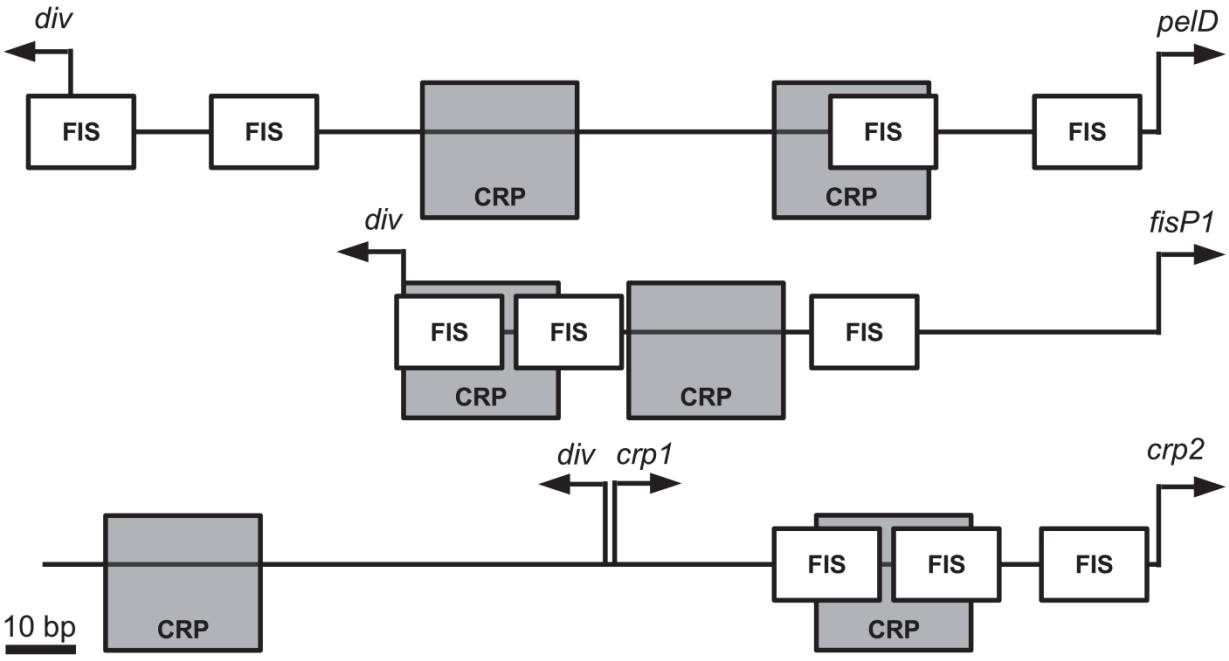


Figure 51: Regulatory structure of 3 different relevant promoters regulated by a divergent promoter.

The *pelD* (*D. dadantii*), *fis* (*E. coli*) and *crp* (*E. coli*) regulatory regions are represented. The known promoters are represented with an arrow, FIS sites by white boxes and CRP sites by gray boxes, as described in the litterature and this work. The promoter which is coupled to *div* is drawn on the right. All regulatory regions and binding sites are represented to scale. The scale bar in the bottom left corner represents 10 base pairs.

We infer that the organisation of *pelD* promoter region reveals a strategic design of protein binding sites evolved for reconciliation of strong expression with regulatability. The *pelD* gene is one among the many bacterial virulence genes known to be under the control of several TFs and be also tightly environmentally regulated [267]. Indeed, it is not uncommon to see more than five different TFs acting on a single virulence gene [116,268,269]. By taking into account the binding of CRP, FIS and RNA polymerase at multiple sites in *pelD*, as well as many regulators unconsidered in this study, such as H-NS and PecT, which are key components in the thermoregulation of the *pel* genes [214], or PecS, which is known to interact with CRP [146], a highly dynamic regulatory mechanism has to be envisioned. Comprehensive understanding of the strategic design underlying the spatial organisation of regulatory binding sites in *pelD* and the topological constraints imposed on the cooperative and antagonistic interactions between various TFs remains an important focus for further studies.

5. Funding information

This work was supported by the Centre National de la Recherche Scientifique; Université de Lyon; and Institut National des Sciences Appliquées [BQR 2013 « Topologie de l'ADN et organisation du nucléoïde » to S.R.]. The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

6. Acknowledgement

The authors would like to thank the teams of Agnès Rodrigue and Xavier Charpentier for providing access to their Tecan plate reader. We are grateful to Julien Wawrzyniak for the purification of FIS and CRP proteins and Xuejiao Jiang for her assistance with RNAseq data analysis that first allowed the detection of *div* transcription.

B. La nature de la connexion entre *pelD* et *div* : une proximité tridimensionnelle ?

Le deuxième article a montré que *div* contribue à la régulation de *pelD* et permet notamment l'intégration des sites FIS1 et FIS1bis dans la régulation de *pelD*. A l'inverse, les sites FIS2 et FIS3, malgré leur éloignement de *div*, contribuent fortement à sa régulation puisqu'une région régulatrice inactivée pour ces deux sites ne présente pas de répression de *div* par FIS (voir pID3 sur la Figure 52). Le rôle des sites FIS1 et FIS1bis sur div est ambigu. En effet, l'inactivation de FIS1 dans la région régulatrice D10 affecte la séquence -10 de *div* et cause une augmentation majeure de l'expression basale de *div* (facteur 5). Sur cette région, la régulation de *div* par FIS est diminuée par rapport à la région sauvage. La réversion de la mutation dans le -10 de *div* (D23) a permis de ramener l'expression basale du promoteur à des niveaux plus proches du promoteur sauvage. En revanche, sur cette région, FIS régule *div* de manière comparable à la souche sauvage. Le site FIS1 partiellement inactivé suffit-il à assurer cette répression ? Il est ainsi difficile de conclure sur la contribution de FIS1/1bis à la régulation de *div*. Pour clarifier ce point, il serait intéressant de comparer ces effets *in vivo* à des expériences *in vitro* afin d'écarter l'effet d'éventuels régulateurs non identifiés et d'explorer la fixation de FIS sur FIS1 dans la région D23.

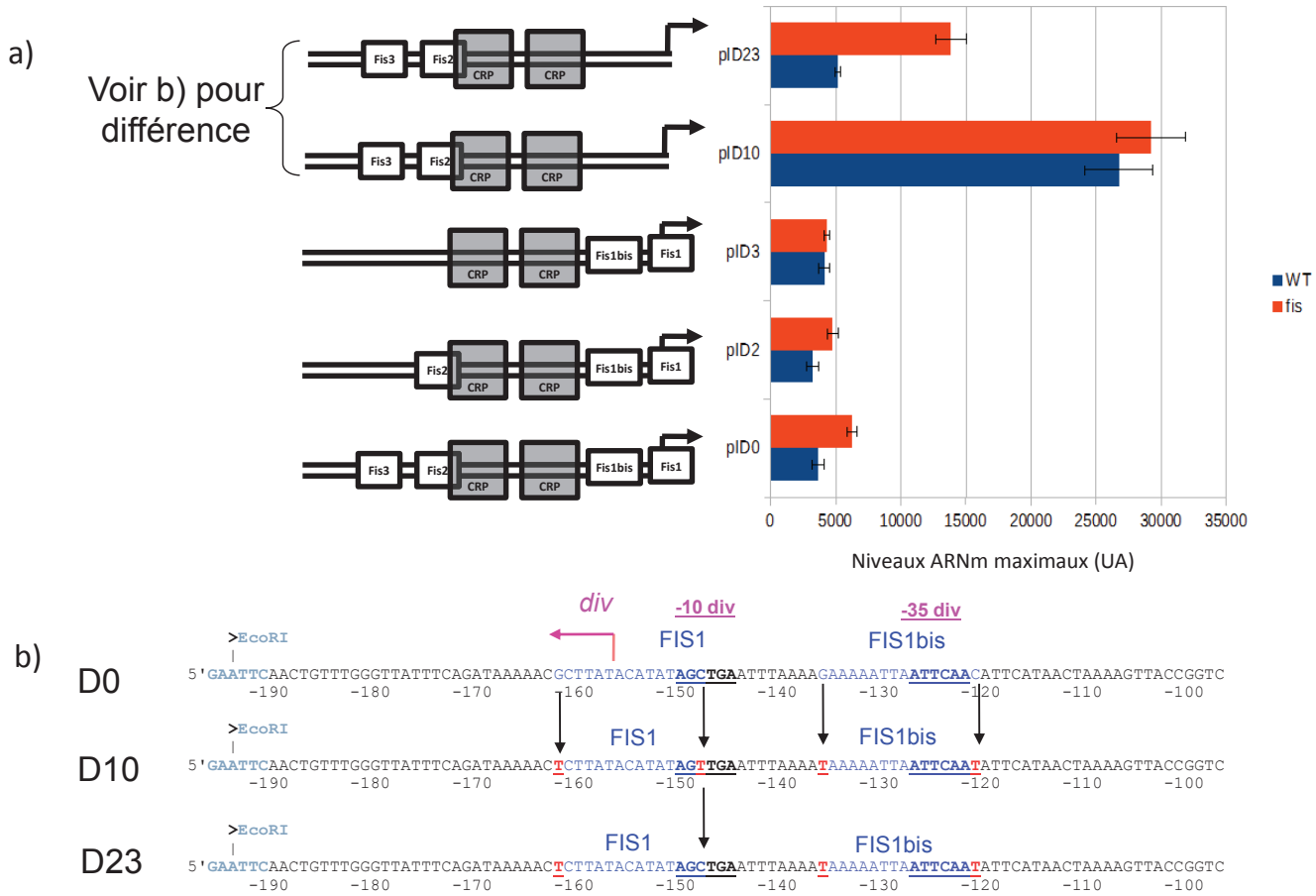


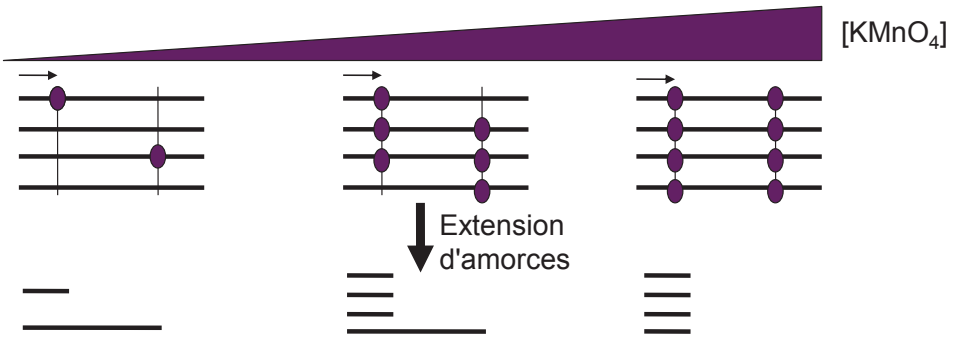
Figure 52 : Effet des sites FIS sur la régulation de *div*

a) Les sites FIS ont été inactivés comme précédemment et l'expression de *div* a été mesurée en utilisant des fusions transcriptionnelles *gfp*. Les moyennes des quantités maximales d'ARNm calculées sont représentées et les barres d'erreur représentent l'écart-type de 9 à 12 répliques. b) Construction de la région régulatrice D23. Div est représenté en rose, ses régions -10 et -35 soulignées et les sites FIS en bleu. Les nucléotides mutés lors d'une inactivation de site FIS sont en rouge. Les flèches indiquent les nucléotides mutés lors des deux étapes de mutagénèse successive ayant produit D10 puis D23 à partir de D0.

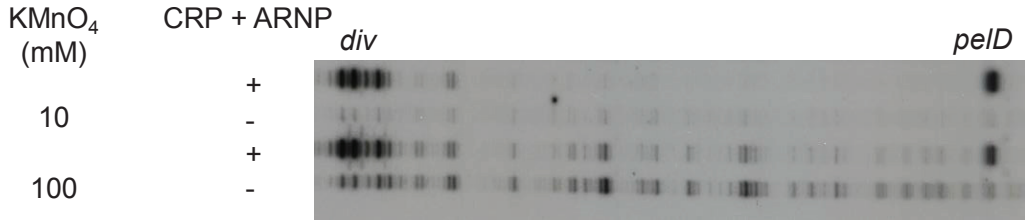
pelD et *div* semblent ainsi être en interaction, s'influençant réciproquement et permettant l'action de sites FIS éloignés des promoteurs. La nature de cette forte connexion entre les deux promoteurs reste inconnue et peut avoir de multiples origines (voir section I-A-4).

Un mécanisme simple fréquemment rencontré est la compétition entre ARN polymérases. Les promoteurs *pelD* et *div* sont distants de 155 pb, ce qui est en théorie suffisant pour permettre une fixation simultanée de deux molécules d'ARN polymérase. Ceci a été testé par empreinte au $KMnO_4$ (Figure 53). L'extension d'amorces effectuée lors de cette expérience s'arrête au premier nucléotide modifié, ce qui crée un biais en faveur des petits fragments lorsqu'un brin d'ADN a subi plusieurs modifications comme lors de la fixation de deux ARNP. Ce phénomène s'amplifie à concentration élevée de $KMnO_4$. Lorsque *div* correspond au plus petit fragment, l'augmentation de concentration de $KMnO_4$ crée un biais en faveur de *div*. Ce biais s'inverse en faveur de *pelD* lorsque le brin complémentaire est examiné. Ces observations indiquent que la fixation simultanée de deux molécules d'ARN polymérase est possible, ce qui exclut un mécanisme de compétition directe. De même, l'absence de réponse de *pelD* au relâchement induit par la novobiocine (voir 2ème et 3ème article) exclut une connexion topologique entre les deux promoteurs par transfert de surenroulement.

a)



b)



c)

	Radioactivité		Fluorescence		Fluorescence (2ème brin)	
$KMnO_4$	10 mM	100 mM	10 mM	100 mM	10 mM	100 mM
Rapport *div/pelD*	3,5	5,7	1,4	1,8	3,5	2,5
Variation	+ 62%		+ 29 %		- 29 %	

Figure 53 : Test de compétition entre les ARN polymérases de *pelD* et *div*.

a) Principe du test [57]. Les lignes horizontales représentent une population de molécules d'ADN, les lignes verticales la position des promoteurs et les cercles violets les nucléotides modifiés par le $KMnO_4$. Lors de l'extension d'amorces réalisée sur les plasmides modifiés au $KMnO_4$, l'ADN polymérase s'arrête à la première modification rencontrée. A concentration élevée de $KMnO_4$ et s'il existe des ADN modifiés aux deux emplacements (donc des fixations simultanées sur les deux promoteurs), un biais se crée vers les petits fragments. A l'inverse, si la fixation de l'ARN polymérase sur les deux promoteurs est indépendante, ce biais n'est pas observé. b) Image du test réalisé avec un marquage radioactif avec deux concentrations de $KMnO_4$ en présence ou absence de (25 nM CRP + 50 nM ARNP). *div* était le plus petit fragment. c) Quantification des essais réalisés avec marquages radioactifs et fluorescents. Sur le 2ème brin, *pelD* est le plus petit fragment.

En revanche, les deux sites CRP sont distants de 50 pb, soit environ 5 tours de double hélice. CRP est capable de courber l'ADN à 90° [54], ce qui pourrait créer un « coude » dans la région régulatrice *pelD* et engendrer une proximité spatiale des promoteurs *pelD* et *div*. Pour tester l'importance de cette géométrie particulière, des insertions de 5 pb ont été effectuées entre les sites FIS1bis et CRP1 d'une part, et CRP1 et CRP2 d'autre part. Dans le premier cas (D22), le coude est conservé mais les promoteurs *pelD* et *div* se trouvent sur des faces opposées de l'ADN (Figure 54). Ceci cause une diminution modérée (facteur 2) de l'expression de *pelD* sans affecter celle de *div*. Le coude est perdu dans le deuxième cas (H2), ce qui cause une chute d'expression importante à la fois pour *pelD* et *div*. Ces deux observations semblent indiquer que l'activation de *pelD* par *div* se ferait par transfert d'ARN polymérases fixées sur *div* vers *pelD* grâce à la proximité spatiale des deux promoteurs induite par CRP. Cette proximité spatiale semble aussi être importante pour l'expression de *div*, ce qui indique une forme de réciprocité entre *pelD* et *div*.

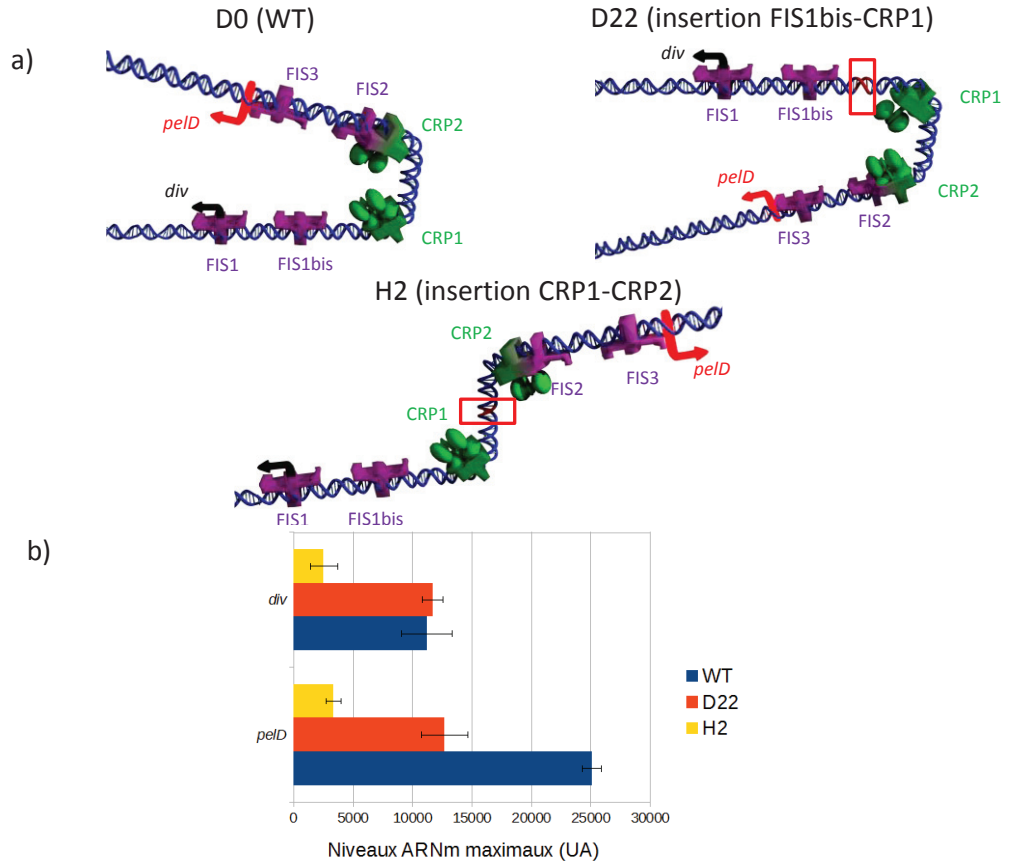


Figure 54 : Effet d'insertions de 5 pb à différents endroits de la région régulatrice *pelD*.

a) Effet des insertions sur la géométrie de la région régulatrice *pelD* en supposant que CRP courbe l'ADN d'environ 90°. L'ADN de la région régulatrice sauvage est représenté en bleu, celui de l'insertion de 5 bp en rouge, *pelD* en rouge, *div* en noir, les sites de fixation FIS en violet, CRP et ses sites de fixation en vert. b) Effet des insertions sur la transcription de *pelD* et *div*. Les quantités d'ARN maximales produites ont été mesurées comme précédemment. Les données sont représentées comme moyenne +/- écart-type de 8 à 12 répliques.

C. Emergence et évolution des gènes *pelD* et *pelE* (3ème article)

Le 3ème article explore la question de l'origine des gènes *pelD* et *pelE* et de leur maintien dans le génome des *Dickeya*, impliquant un mécanisme évolutif commun chez les eucaryotes pluricellulaires mais inconnu jusqu'à présent chez les bactéries : le TSS turnover. L'article est actuellement en révision pour FEBS journal.

Relevance of Transcriptional Start Site Turnover in the Evolution of Bacterial Paralogous Genes: the *pelE-pelD* Virulence Genes in *Dickeya*.

Short title: **Bacterial Evolution of Paralogs: the *pelE-pelD* Case**

Alexandre Duprey[a], William Nasser[a], Simon Léonard[a], Céline Brochier-Armanet[b], Sylvie Reverchon[a]*

[a] Univ Lyon, Université Claude Bernard Lyon 1, INSA-Lyon, CNRS, UMR5240, Microbiologie, Adaptation et Pathogénie, 10 rue Raphaël Dubois, F-69622, Villeurbanne, France

[b] Univ Lyon, Université Claude Bernard Lyon 1, CNRS, UMR5558, Laboratoire de Biométrie et Biologie Évolutive, 43 bd du 11 novembre 1918, F-69622, Villeurbanne, France

*Corresponding author

E-mail: sylvie.reverchon-pescheux@insa-lyon.fr

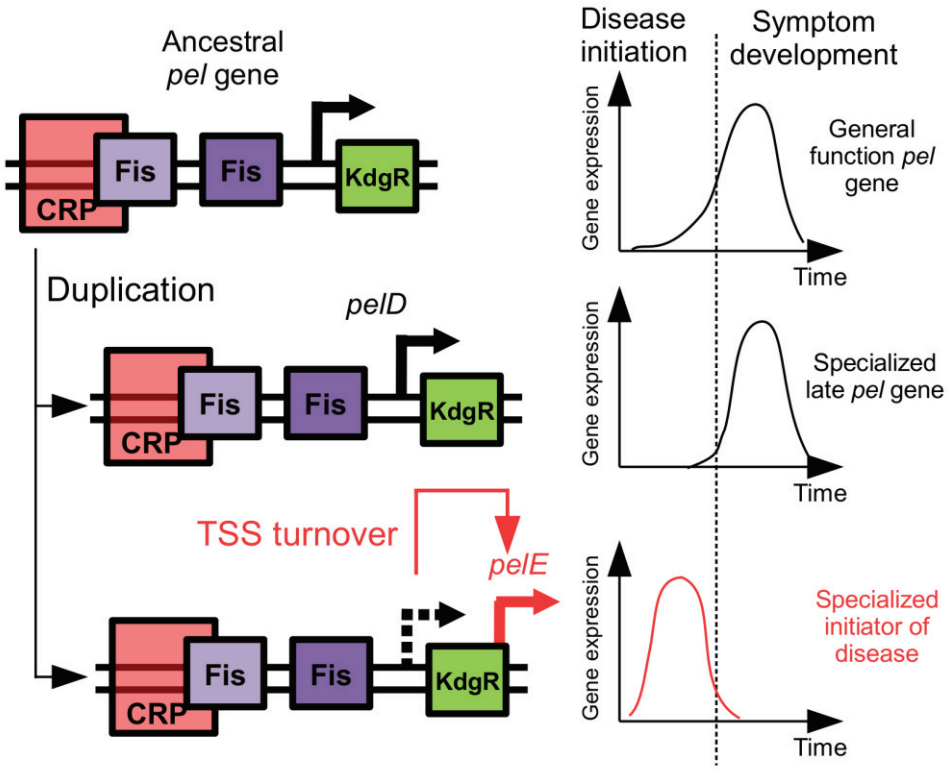
Keywords

Adaptation, promoter, transcription, pectate lyases, gene duplication

Abstract

After a gene duplication event, the resulting paralogous genes frequently acquire distinct expression profiles, roles and/or functions but the underlying mechanisms are poorly understood. While transcription start site (TSS) turnover, i.e. the repositioning of the TSS during evolution, is widespread in eukaryotes, it is less documented in *Bacteria*. Using *pelD* and *pelE*, two closely related paralogous genes encoding key virulence factors in *Dickeya,* a gamma proteobacterial genus of phytopathogens, we show that *pelE* has been selected as an initiator of bacterial aggression, while *pelD* acts at a later stage, thanks to modifications in the transcriptional regulation of these two genes. This expression change is linked to a few mutations that caused a shift in the position of the *pelE* TSS and the rapid divergence in the regulation of these genes after their duplication. Genomic surveys detected additional examples of putative turnovers in other bacteria. This first report of TSS shifting in *Bacteria* suggests that this mechanism could play a major role in paralogous genes fixation in prokaryotes.

Graphical abstract



Introduction

The successful adaptation of living organisms to new environments is a complex and dynamic process involving the modification and optimization of existing functions, the acquisition of new functions, and the loss of inappropriate ones, but also changes in regulatory networks [270,271]. Gene duplication plays key roles in organismal evolution. Duplicate genes initially have identical sequences but diverge in coding and regulatory regions during subsequent evolution. Changes in coding regions may lead to the acquisition of new functions whereas divergence in regulatory regions could result in changes in expression profiles.

In bacteria, several mechanisms contributing to the divergence in gene expression have been identified. They include the duplication and divergence of transcription factors and modifications in regulatory region architectures [84,272]. In eukaryotes, TSS turnover, i.e. the repositioning of the transcription start site, is the main mechanism responsible for paralogous gene expression divergence [273,274]. This affects about 85% of human duplicated gene pairs [273], and in *Drosophila*, paralogous genes having undergone a TSS turnover display higher expression divergence than other duplicated genes, highlighting the importance of this mechanism in evolution [274]. This interesting mechanism, however, has never been documented in bacteria.

Using *Dickeya* as a case study, we analyze in detail the structure and the evolution of the regulatory regions of *pelD* and *pel*E, two closely related paralogous genes. These phytopathogenic gammaproteobacteria are responsible for soft rot disease in a wide range of plant species [108]. The soft rot symptom mainly results from the degradation of pectin, an essential component of the plant cell wall, by a variety of pectinolytic enzymes [112]. An effective infection process requires the correct spatiotemporal expression of different virulence factors in response to environmental conditions, resulting from the ongoing host-pathogen interactions. More precisely, the recognition of a variety of signals from the host induces the sequential synthesis of pectinases, mainly pectate lyases (Pels) to initiate the degradation of plant pectin[107,185]. Pectin is then cleaved extra and intra-cellularly to produce 5-keto-4-deoxyuronate (DKI), 2,5-diketo-3-deoxygluconate (DKII) and 2-keto-3-deoxygluconate (KDG) [111]. These inducers bind to the repressor KdgR, thereby suppressing its ability to repress pectinase synthesis and creating a positive feedback loop [236].

Each *Dickeya* species produces about a dozen different Pels with various biochemical properties and roles during infection [227,275]. In *D. dadantii*, PelD and PelE are among the most important for virulence [238]. Interestingly, many of the *pel* genes are assumed to be recent paralogs, including *pelE* and *pelD* [190], which are also paralogous to *pelA*. The corresponding enzymes PelD and PelE have indeed highly similar sequences (78% amino acid identity, 89% amino acid similarity) and are biochemically nearly identical [227]. In contrast, these two paralogs have very different expression patterns. In fact, *pelD* is most strongly inducible by pectin but has a very low pectin-independent expression while *pelE,* although inducible by pectin, has a high expression level even in the absence of pectin [193]. In addition to this pectin inducibility, *pel* genes are regulated by more than a dozen known regulators, coordinating their expression with that of other virulence genes in a complex regulatory and effector cascade during disease development [116]. Unlike other previously reported examples of regulation of bacterial duplicated genes[84], it was found that the expression of *pelD* and *pelE* is regulated by the same set of transcription factors (TFs), the main ones being the activator CRP, the repressors FIS and KdgR [130,145,192,193,236]. The regulatory mechanisms enabling these common TFs to differentially modulate *pelD* and *pelE* gene expression are unknown. We demonstrate here that the functional divergence of *pelD* and *pelE* in *Dickeya* originates from a TSS turnover event, establishing for the first time the existence of such a mechanism in bacteria. Genome-wide analyses allowed the detection of TSS turnovers in other important pathogens such as *Xanthomonas* and *Staphylococcus,* suggesting it could be a general mechanism for the fixation of paralogous genes in both prokaryotes and eukaryotes.

Results

The evolutionary history of the *Dickeya pel* locus

To clarify the relationships among the *pelE* and *pelD* sequences in *Dickeya*, we performed the phylogenetic analysis of these genes and their closest homologues. A preliminary analysis confirmed the close proximity between *pelD* and *pelE* and identified *pelA* from *Dickeya* and *pel* sequences from *Lonsdalea* and *Serratia* ATCC 39006 as their closest relatives (Figure 55). Interestingly, in *Dickeya*, *pelA*, *pelD* and *pelE* are located at the same locus (Figure 56). An accurate analysis focused on these genes with a larger taxonomic sampling showed that *pelD* and *pelE* cluster robustly together while *pelA* diverges first (Bootstrap Values (BV) = 96% and 71%, Figure 57), suggesting that two tandem duplication events occurred during the diversification of this genus, the first one leading to *pelA* while the second corresponds to the emergence of *pelD* and *pelE*. The branching position of *pel* sequences from *Lonsdalea* and *Serratia* ATC 39006 is not significantly supported (BV=79% and BV < 50%, respectively, Figure 57), meaning that their precise relationship with *Dickeya* sequences remains unresolved. The three *Lonsdalea* strains harbour two closely related *pel* genes, suggesting that an independent gene duplication event occurred in the common ancestor of this genus (D3, BV = 68% and 100%, Figure 57). Interestingly, *D. paradisiaca*, a basal branching *Dickeya* species[276] contains only one *pelD/pelE* homologue, indicating that the second duplication event occurred after the divergence of this species. This is further supported by the lack of any trace of gene loss (e.g. pseudogene, deletion marks) at the *pel* locus in the genome of *D. paradisiaca* (Figure 56). In contrast, in some species, one of the *pel* genes has become a pseudogene, e.g. *pelA* in *D. dianthicol*a and *pelE* in *D. chrysanthemi*.

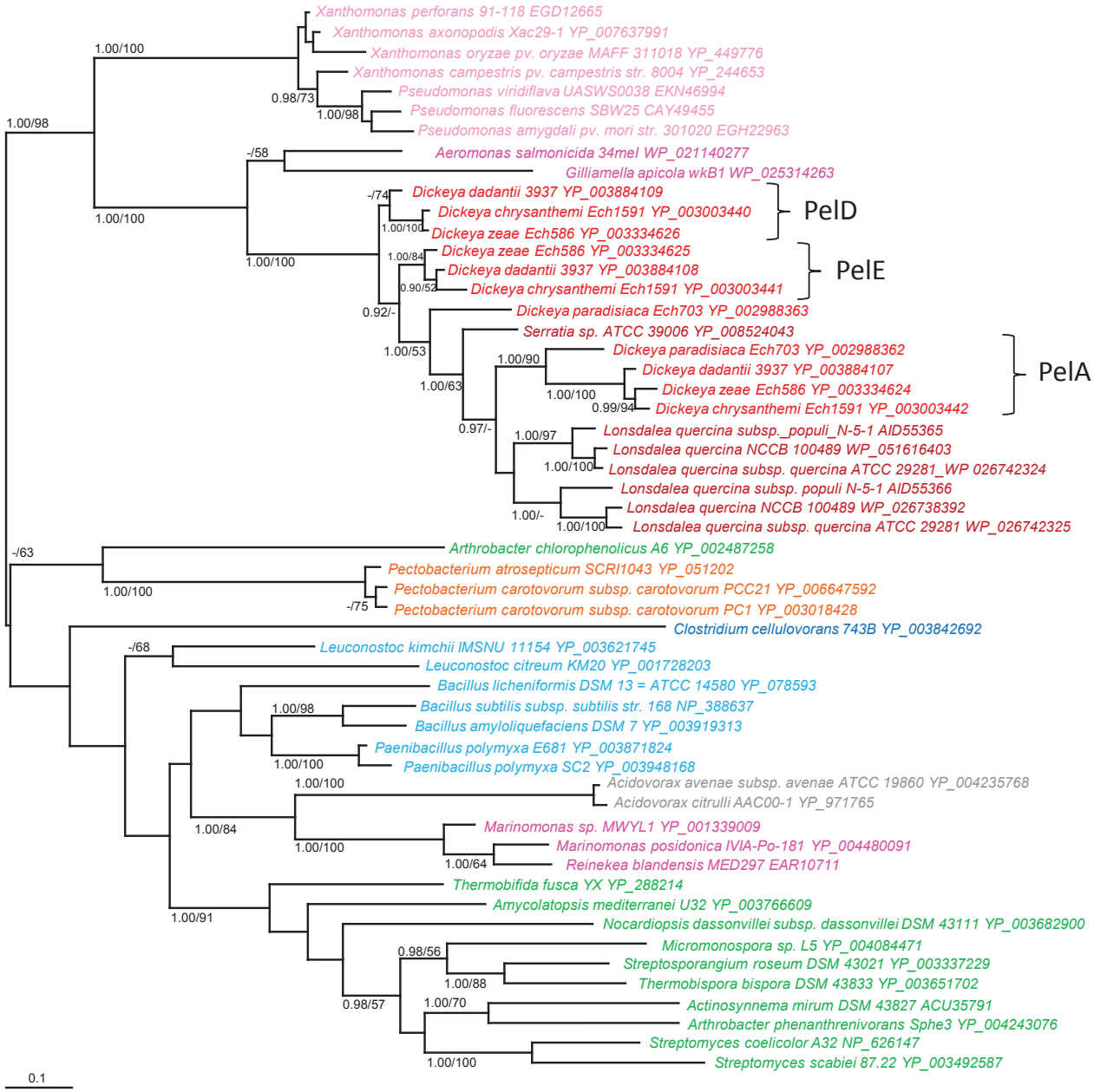


Figure 55. Bayesian phylogeny of 53 PelAED proteins (317 amino acid positions).

The tree was inferred with MrBayes using a mixed evolution model and a gamma distribution with 4 categories (Γ4). The scale bar represents the average number of substitutions per site. The first number (between 0 and 1) next to the branches corresponds to Bayesian posterior probabilities and the second number (between 0 and 100) corresponds to bootstrap values of the ML phylogeny inferred using PhyML with the LG+ Γ4 model on the same data set (100 replicates of the original data set). A - is indicated if the probability is lower than 0.9, or if the bootstrap value is lower than 50. For clarity, nothing is indicated if the branches have both a bootstrap value lower than 50 and a probability lower than 0.9. *Pectobacterium* is in orange, *Dickeya* in red, other *Enterobacteriaceae* in dark red, *Xanthomonadales, Pseudomonadales* and related orders in light pink, other *Gammaproteobacteria* in pink, *Alphaproteobacteria* in purple, *Deltaproteobacteria* and *Betaproteobacteria* in grey, *Bacteroidetes* in light green, *Actinobacteria* in green, *Bacilli* in light blue, *Clostridia* in dark blue and other *Bacteria* in black.

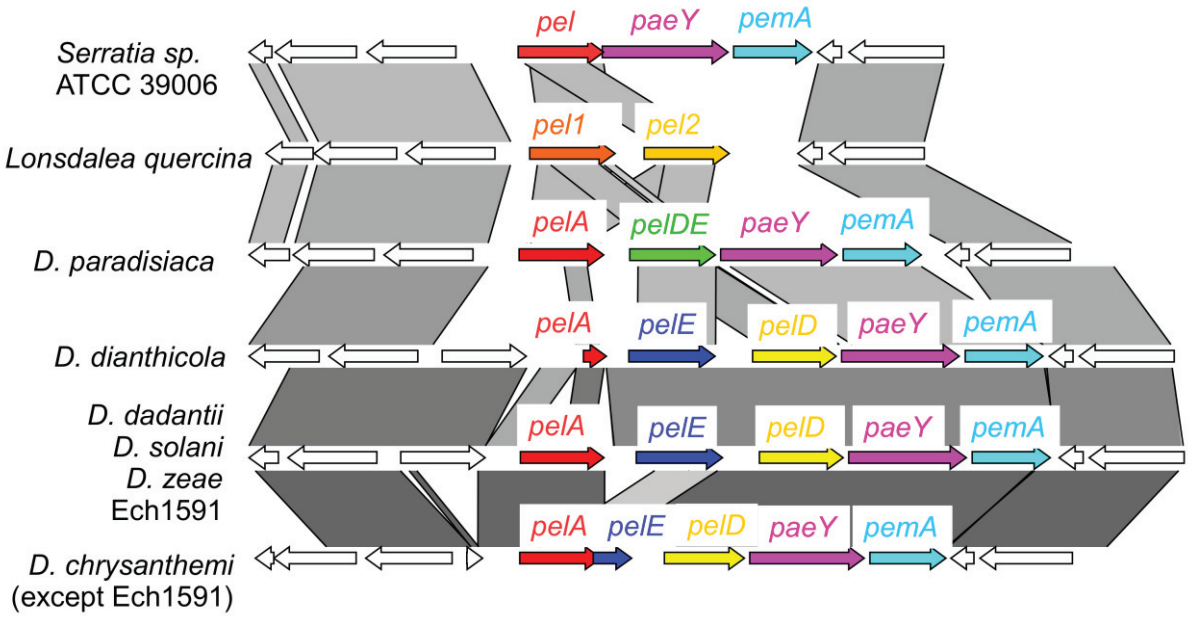


Figure 56. Genomic contexts of the *pelA, pelD and pelED* genes in *Dickeya* and closely related species.

The loci containing the *pelA, pelD and pelE* were compared using EasyFig. Gray areas indicate similarity between the corresponding regions. *pelA, pelD, pelE, paeY* and *pemA* genes are colored. The names of these genes were attributed according to their position in the phylogenetic tree of Figure 55, using the nomenclature of *D. dadantii* 3937 as a reference.

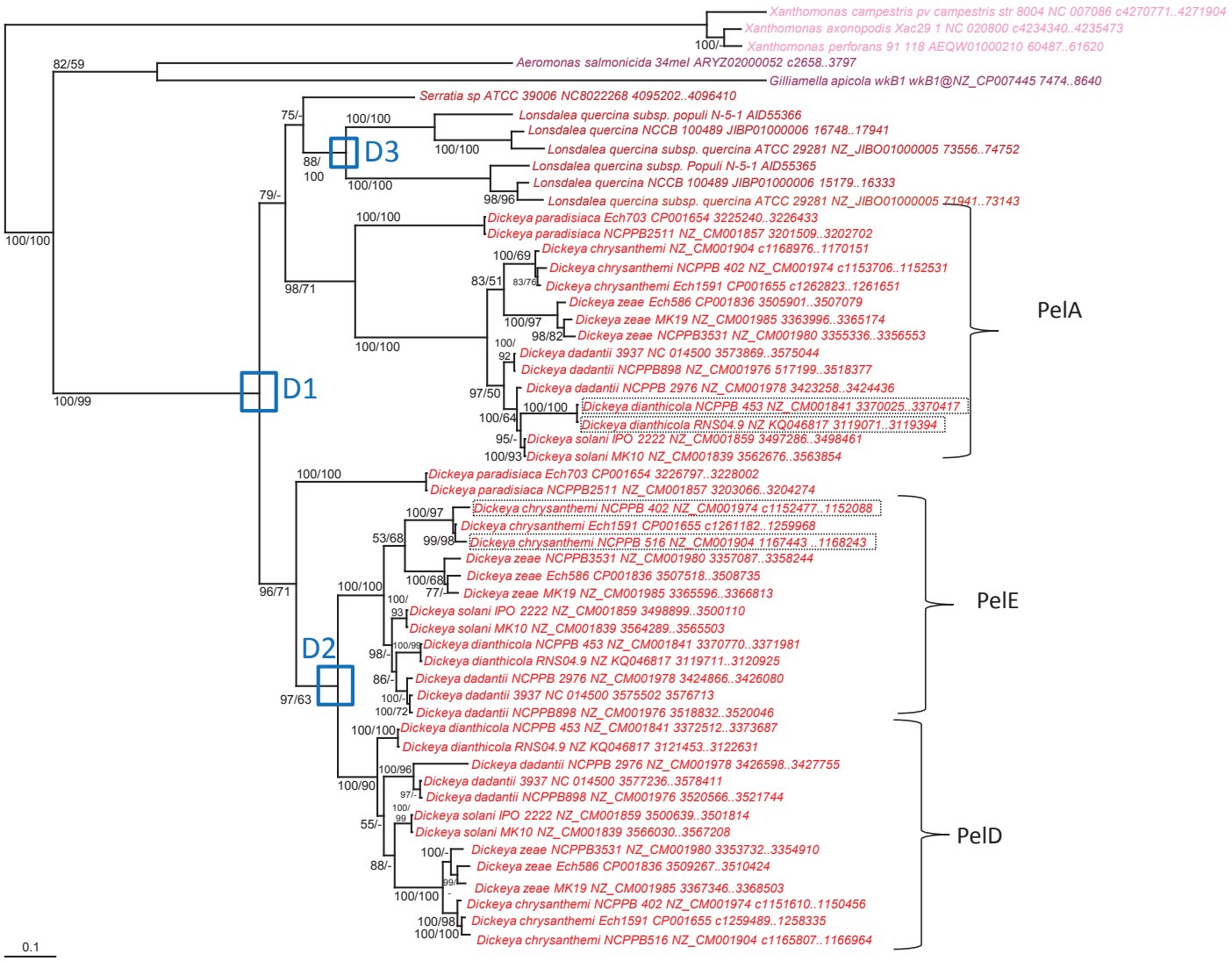


Figure 57. ML phylogeny of 55 *pelAED* genes (1137 nucleic acid positions).

The tree was inferred using nhPhyML with the GG model. The scale bar represents the average number of substitutions per site. The first number next to the branches corresponds to bootstrap values computed with the nucleotide multiple alignment and the second number corresponds to bootstrap values of a ML phylogeny computed with PhyML using the LG+ Γ4 model on the protein sequences (379 amino acid positions). A - is indicated if the bootstrap value is lower than 50 or the branch does not exist in the protein ML phylogeny. For clarity, nothing is indicated if both branches have a bootstrap value lower than 50. *Dickeya* is in red, other *Enterobacteriaceae* in dark red, *Xanthomonadales, Pseudomonadales* and related orders in light pink and other *Gammaproteobacteria* in pink (consistently with Figure 55). Truncated (i.e. likely nonfunctional) proteins are boxed in black dotted lines. Blue boxes accompanied by a "D" indicate duplications.

The relationships among *Dickeya* sequences (Figure 57) are overall consistent with the species phylogeny: *D. dadantii, D. solani* and *D. dianthicola* sequences generally group together while *D. zeae* and *D. chrysanthemi* sequences form a separate cluster [276]. This suggests that the *pel* genes were transmitted vertically during the diversification of *Dickeya*. There are some exceptions, like the *Dickeya dianthicola pelD* which does not group with *D. dadantii* and *D. solani*, but the associated BV are very low.

To conclude, two unique duplication events in *Dickeya* have led to the selection of up to three *pel* paralogous genes. However, since PelD and PelE are biochemically nearly identical, the

selective pressure that allowed the fixation of *pelD* and *pelE* could be related to their differential expression.

PelE is an initiator of pectin degradation in *Dickeya*

Previous works have suggested that either PelE [193] or PelD [190,275] could be initiators of pectin degradation in *Dickeya*: *pelE* has high expression level even in the absence of pectin while *pelD* has weak expression level in the absence of pectin. However, *pelD* was described as the first transcribed *pel* gene in synthetic cultures and during the infection [190,275]. These reports have used chromosomal *pel::uidA-Kan* fusions disrupting the corresponding *pel* open reading frame [190,275] and thus, likely created a bias due to inactivation of the *pel* gene. Consequently, the temporal expression of *pelE* and *pelD* genes and their role as initiators of pectin degradation remain a conundrum.

To solve this enigma, the relationship between *pelD* and *pelE* was evaluated by measuring the expression of *pelD* in a *pelE* mutant and *vice-versa* (Figure 58a) and comparing their expression to wild type (WT) *D. dadantii*. We used low-copy plasmid-borne transcriptional *pel-gfp* fusions that do not require the inactivation of any gene. In the WT, the *pelD* expression starts after 5.5 hours and reached its maximum level at 11 hours, whereas the *pelE* expression starts 4 hours post inoculation and reached its maximum at 9.3 hours. Thus, *pelE* is expressed earlier than *pelD*. We found that in the *pelE* mutant, the expression of *pelD* is severely delayed by about 14 hours indicating that *pelE* is needed for the timely activation of *pelD*, while the expression of *pelD* was not modified in a *pelD* mutant (Figure 58a). The expression of *pelE* was also modified in a *pelE* mutant, showing however only a 2-3 hours delay of the peak (compared to 14 hours for *pelD* expression) and a 2-fold decrease in expression but not in a *pelD* mutant (Figure 58b). Interestingly, in a *pelE* mutant, *pelD* expression becomes biphasic, with a minor peak taking place at the same time as the WT strain. This could be due to the presence of another, less efficient, initiator Pel or to the presence of pectin debris among the pectin added into the media.

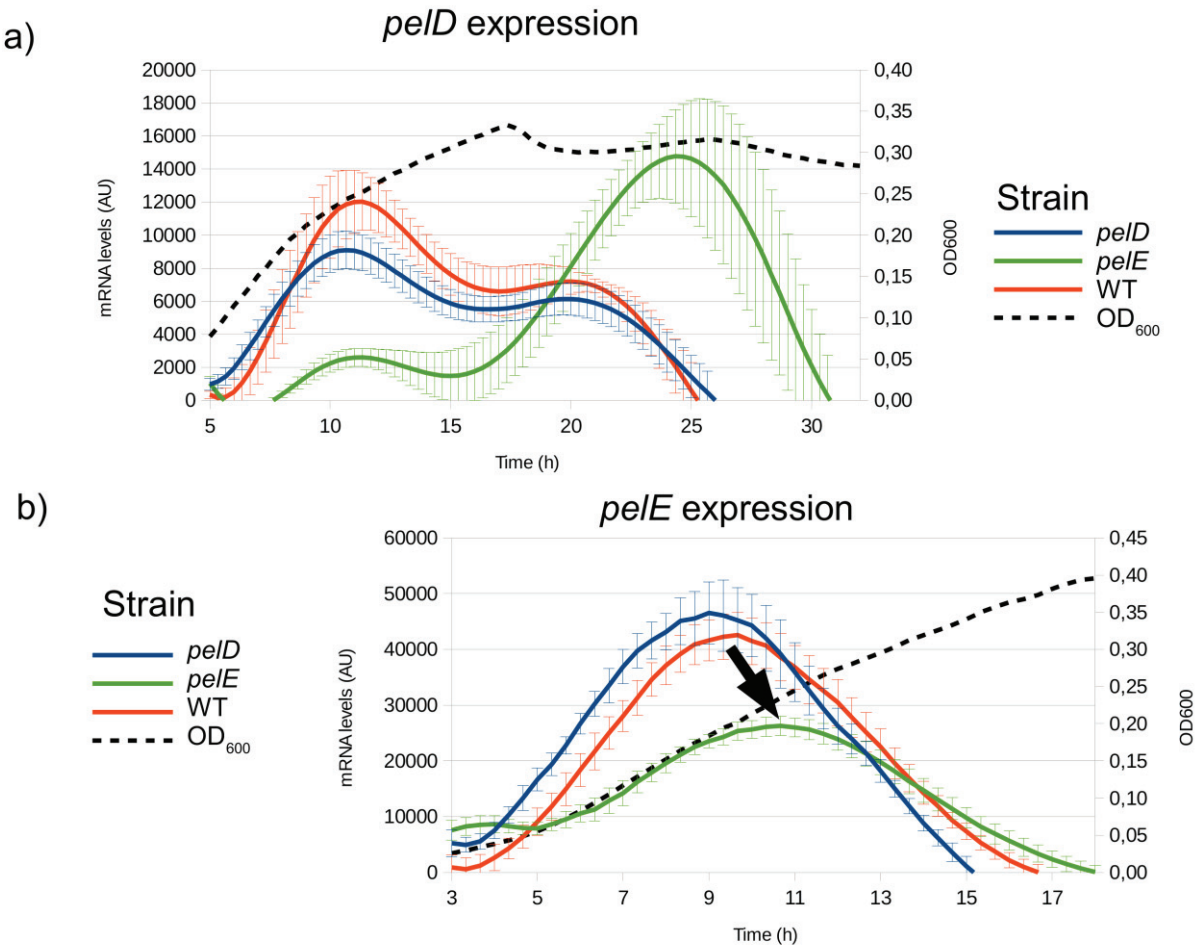


Figure 58. Influence of the Pels of *D. dadantii* on *pel* gene transcription.

Dickeya dadantii strains (*pelD*, *pelE* and WT backgrounds) transformed with a pProbe-Gfp controlled by either (a) the *pelD* promoter or (b) the *pelE* promoter were cultured in 96 well plates in M63S + Pectin. mRNA levels were computed from fluorescence data as described in materials and methods. Curves represent the mean of 9 to 12 replicates and the error bars represent the standard deviation. The black arrow highlights the shift in the position of *pelE* expression peak between the WT and *pelE* strains. Note that, due to the greater shift in *pelD* expression compared the *pelE* shift, the x-axis scale is different for both graphs.

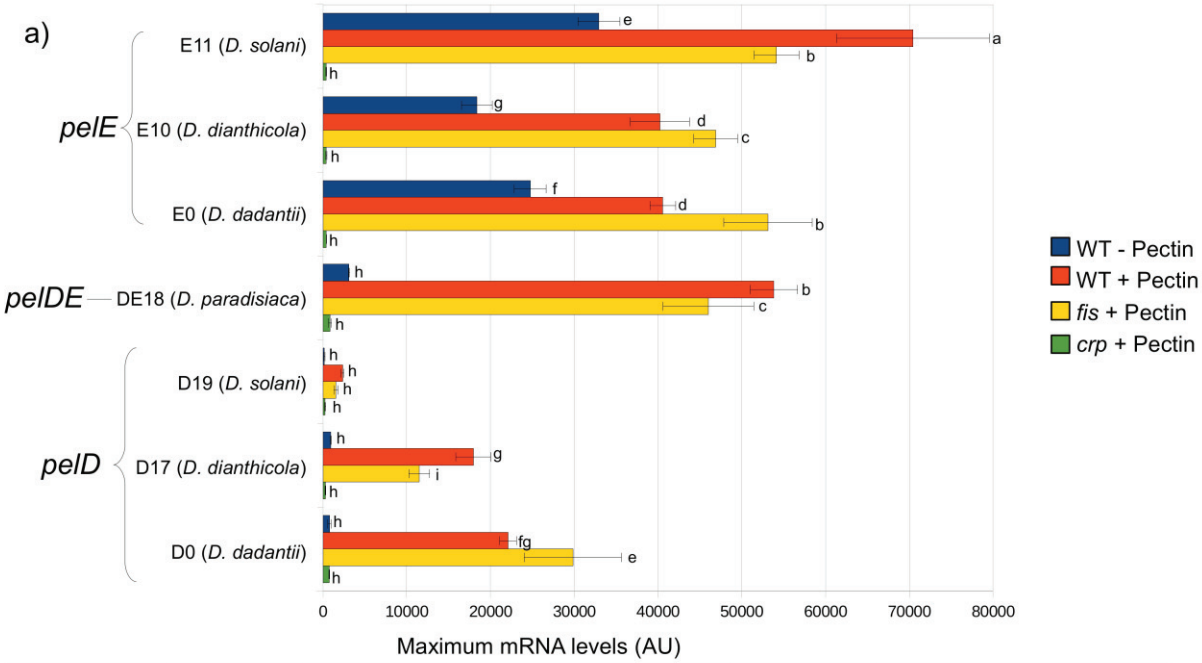
Thus, PelE is a powerful activator for *pelD* expression and weaker for *pelE* expression. In contrast, PelD does not affect *pelE* or *pelD* expression. Overall, these results suggest that PelE acts as an initiator of pectin degradation generating the initial inducers of *pel* gene expression (DKI, DKII, KDG) required to prevent the binding of the repressor KdgR to its operators [236].

Expression of *pelD* and *pelE* from various *Dickeya* species

pelD/E promoter expression from different *Dickeya* species such as *D. dianthicola*, *D. solani* and *D. paradisiaca* was examined and compared to the *pelD/E* expression pattern in *D. dadantii*. *D. chrysanthemi* and *D. zeae* were not considered because their *pelD/E* regulatory regions display similar organisation compared to *D. dianthicola*.

The plasmid-borne transcriptional *pel-gfp* fusions from different *Dickeya* species were expressed in *D. dadantii,* which was chosen as a reference genetic background to eliminate the potential effects owing to differences in either TF concentrations or optimal culture conditions. Several conditions known to affect *pel* transcription were tested: the inactivation of CRP, the

global regulator of sugar metabolism which is the main activator of the *pel* genes [145], the presence or absence of pectin and inactivation of the growth phase regulator FIS [167], known to repress *pel* genes in *D. dadantii* [192]. The produced *pel-gfp* mRNA levels were calculated in real-time from *gfp* reporter fusions, and their maxima are represented in Figure 59. Regarding the pectin-independent *pel* expression, *D. solani* and *D. dianthicola* are similar to *D. dadantii*: *pelE* has a high pectin-independent expression while *pelD* has a low pectin-independent expression. In addition, all *pel* genes are strongly activated by CRP (Figure 59). The effect of FIS is less obvious, activating or repressing *pelE* and *pelD* depending on the species.



b)

Fold change relative to WT + Pectin	D0 (*D. dadantii*)	D17 (*D. dianthicola*)	D19 (*D. solani*)	DE18 (*D. paradisiaca*)	E0 (D. dadantii)	E10 (D.dianthicola)	E11 (D. solani)
WT - Pectin	-29	-20	-15	-17	-1,6	-2,2	-2,1
fis	1,4	-1,6	-1,5	-1,2	1,3	1,2	-1,3
crp	-30	-66	-10	-65	-102	-106	-186

Figure 59. Transcriptional regulation of the *pelD/E* genes of *D. paradisiaca*, *D. solani*, *D. dianthicola* and *D. dadantii*.

The plasmid-borne transcriptional *pelD/E-gfp* fusions from different *Dickeya* species [*D. paradisiaca* (strain NCPPB 2511); *D. dianthicola* (strain NCPPB 453); *D. dadantii* (strain 3937); *D. solani* (strain IPO 2222)] were expressed in *D. dadantii* 3937 wild-type, *fis* and *crp* cells. a) The maximum mRNA quantity produced by various *pelD/E* regulatory regions was measured as described in materials and methods. Data represents the mean of 8 to 12 replicates and error bars represents the standard deviation. Significant differences are indicated by different lowercase letters on the right of the bar (Tukey's HSD, $p<0.05$). b) Corresponding fold-changes relatively to the wild-type strain in presence of pectin for each promoter.

Furthermore, differences exist among these species. For instance, the *pelD* promoter from *D. solani* is 10-fold weaker than that of *D. dadantii*, its closest relative, under inducing conditions (with pectin) and is less inducible by CRP (10-fold decrease in the *crp* mutant, compared to 30-fold for *D. dadantii*), which indicates a defect in CRP activation.

Finally, *D. paradisiaca* displays intermediate features: its unique *pelD/E* gene, alike the *pelD*

genes found in other *Dickeya* species, is strongly inducible by pectin (17-fold), but has a stronger pectin-independent expression, which nevertheless does not approach the level of a *pelE* gene.

To conclude, while the activation by CRP is shared by both *pelD* and *pelE*, their response to pectin is different, the response to FIS shows no clear pattern between species or types of genes and *D. paradisiaca* unique *pelD/E* gene shows features intermediate between *pelD* and *pelE*.

Structural modifications of the *pelD* and *pelE* regulatory regions

To decipher the observed differences in the expression of *pelD* and *pelE*, we studied the structure of the corresponding regulatory regions in all *Dickeya* species (except *D. chrysanthemi*'s *pelE* pseudogene). The regulatory regions were aligned and the binding sites for RNA polymerase, FIS, CRP and KdgR were identified according to literature data, homology with experimentally validated binding sites and bioinformatic predictions (see materials and methods) (Figure 60). For *D. dadantii*'s *pelD,* the relevance of all these sites has been elucidated recently [228]. For *D. paradisiaca*, the bioinformatic search revealed no satisfactory transcription start site. Consequently, it was determined experimentally by reverse transcription on *D. paradisiaca*'s *pelD/E* mRNA (Figure 60). The position of the transcription start site revealed a very weak putative -35 hexamer, explaining why the promoter prediction algorithms failed to detect it.

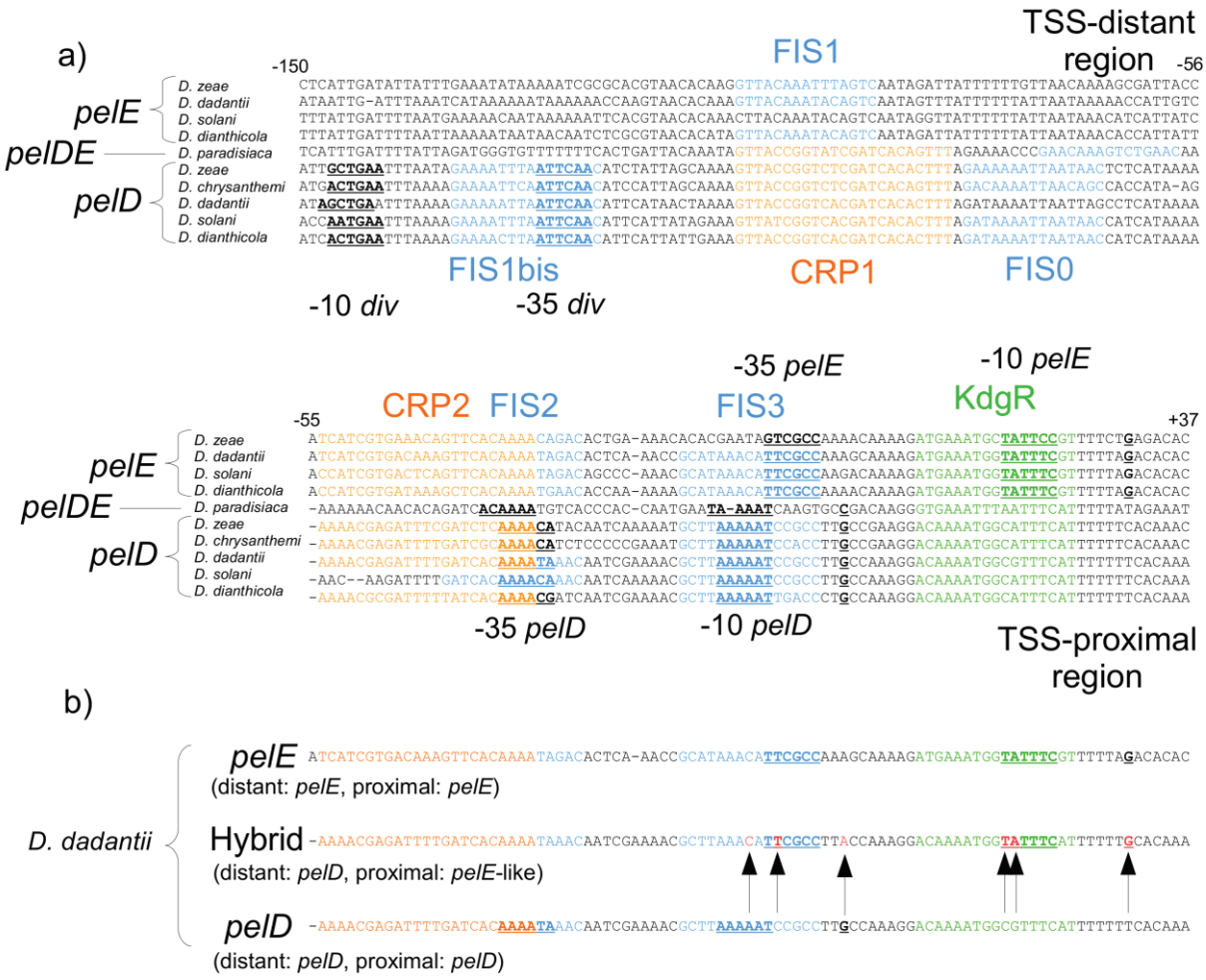


Figure 60. Alignment of the regulatory regions of the *pelD/E* genes of *Dickeya*.

TSS, -10 and -35 sequences of *pelE, pelD* and *div* promoters are underlined. Fis sites are indicated in blue, CRP in orange and KdgR in green. a) Alignment of the variable (TSS-distant) sequence and conserved (TSS-proximal)

sequences. Strains used are the following: NCPPB 2511 (*D. paradisiaca*), Ech1591 (*D. chrysanthemi*), Ech586 (*D. zeae*), NCPPB 453 (*D. dianthicola*), 3937 (*D. dadantii*), IPO 2222 (*D. solani*). Little to no intra-specific variability was observed. Nucleotides are numbered with respect to *D. dadantii pelD* TSS b) Construction of the hybrid *pelD*/*pelE* regulatory region by introducing point mutations (indicated in red on the hybrid regulatory region).

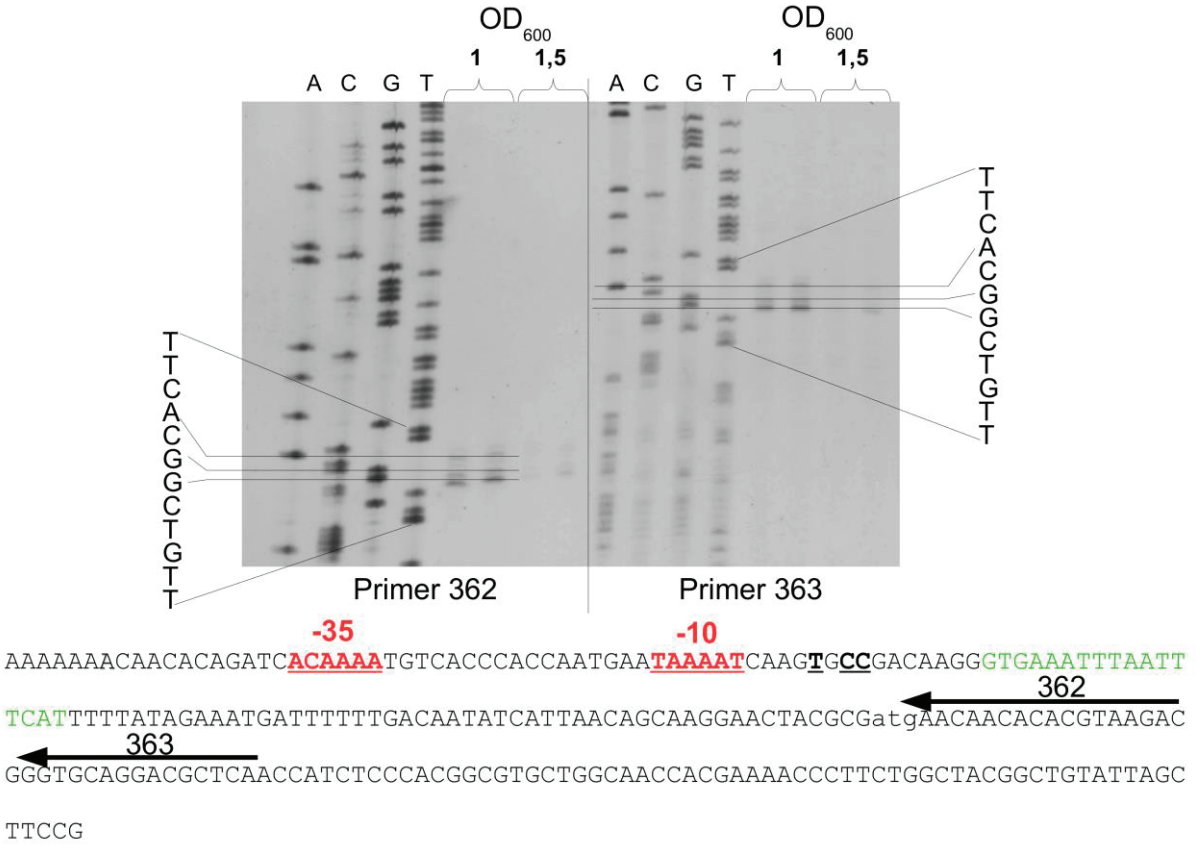


Figure 61. Experimental characterization of the transcription start site of *D. paradisiaca*.

Total RNA was extracted from *E. coli* DH5α expressing a plasmid-borne *pelDE-gfp* fusion at the OD600 indicated at the top of the Fig. A reverse transcription was then performed using radiolabeled primers 362 (left) or 363 (right). Sequencing reactions using the same primers were used to position the observed TSS. The identified TSS are underlined in black, and associated -10 and -35 sequences are underlined. Primer positions are indicated by an arrow. The KdgR binding site, which can also be seen on , is marked in green.

The alignment revealed the presence of a highly conserved “proximal” region (Figure 60a), comprising conserved CRP2, FIS2, FIS3 and KdgR sites for both *pelD* and *pelE* regulatory regions. The most notable difference in this region is the modification of the position of the RNA polymerase binding site (-10 and -35 sequences), which is shifted 24 bp downstream on *pelE* compared to *pelD*. This shift changes the relative binding position of all regulators with respect to RNA polymerase binding site. While being divergent in sequence, the *pelD/E* promoter of *D. paradisiaca*'s is aligned with that of *pelD* genes. *D. solani*'s *pelD* lacks the high-affinity CRP2 binding site, which could explain its relatively low expression and defect in CRP activation observed in our experiments (Figure 59).

Further upstream, the *pelD* and *pelE* regulatory regions differ. All the *pelD* promoter regions contain the recently described divergent promoter *div* [228], FIS1bis and CRP1 sites. Interestingly, all but *D. dadantii* have an additional predicted FIS binding site between CRP1 and CRP2, which we designate FIS0. The CRP1 site is also conserved in *D. paradisiaca*, as well as a FIS

site downstream of CRP1, which due to its similar location we also designate as FIS0. However, the *div* promoter and FIS1bis are not conserved in *D. paradisiaca*. The *pelE* distant regulatory regions demonstrate only a relatively well conserved single FIS binding site (FIS1) of unknown function, described previously [192].

A promoter shift explains many aspects of the divergence between *pelD* and *pelE*

The two major observations we made by comparing the *pelD* and *pelE* regulatory regions are a downstream shift (24 bp) of the promoter, and a variable distant region. We tried to evaluate whether any of these two modifications could lead to divergence between the expression of *pelD* and *pelE* in the ancestor of the studied *Dickeya* by shifting the *D. dadantii*'s *pelD* regulatory region at a position similar to that of *D. dadantii*'s *pelE*. This was achieved by introducing six point mutations in the *pelD* regulatory region (Figure 60b). The "opposite" hybrid regulatory region (i.e. a *pelE* region with a 24bp upstream shifted promoter) yielded no detectable expression. This absence of expression is probably linked to the absence of CRP1 binding site in the hybrid regulatory region. Indeed, we demonstrated that CRP1 is necessary for expression of a *pelD*-like promoter [228].

We next measured real-time mRNA levels produced by the *pel* promoter constructs as described above. As shown on Figure 62, the hybrid regulatory region behaves like a *pelE* promoter, with a high pectin-independent expression level, low induction by pectin (2-fold) and an earlier expression than *pelD*. This strongly suggests that a promoter shift could account for the expression difference between *pelD* and *pelE*, which can be readily achieved by six base substitutions (Figure 60b) that are sufficient to turn a *pelD*-like promoter into a *pelE*-like promoter. PelE could then become an initiator of pectin degradation and one of the most important aggression genes in most *Dickeya* species. Taken together, our data strongly suggest that the patterns of *pel* gene expression are primarily associated with divergence of *pel* promoter positions.

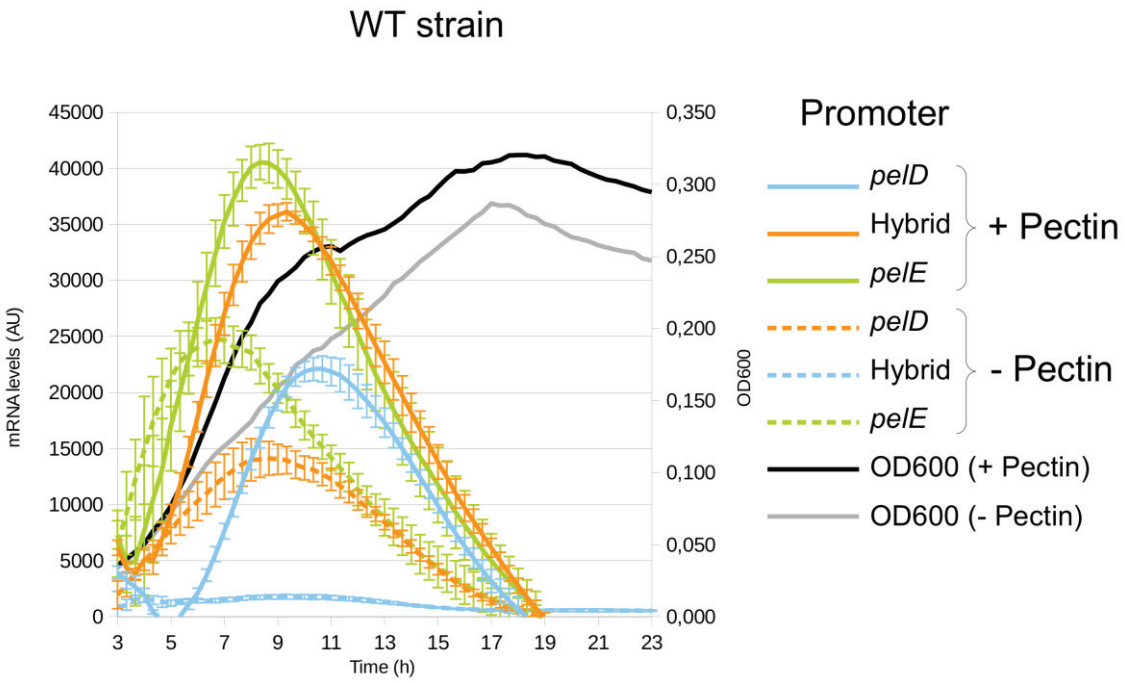


Figure 62. Transcriptional behaviour of the hybrid *pelD-pelE* regulatory region.

mRNA quantities transcribed from *D. dadantii*'s *pelD*, *pelE* or hybrid regulatory region (*pelD* with a 24bp downstream-shifted promoter) were measured in real time as described in material and methods. Data represents the mean of 9 to 12 replicates and error bars indicate the standard deviation. OD_{600} are represented in black and gray.

TSS turnover is widespread in bacteria

Having established the existence of TSS turnovers in *Dickeya*, we attempted to see evidence of this mechanism in other bacterial species. We considered the model enterobacterium *Escherichia coli*, the phytopathogen gammaproteobacterium *Xanthomonas campestris* and the human firmicute pathogen *Staphylococcus aureus* (Figure 63).

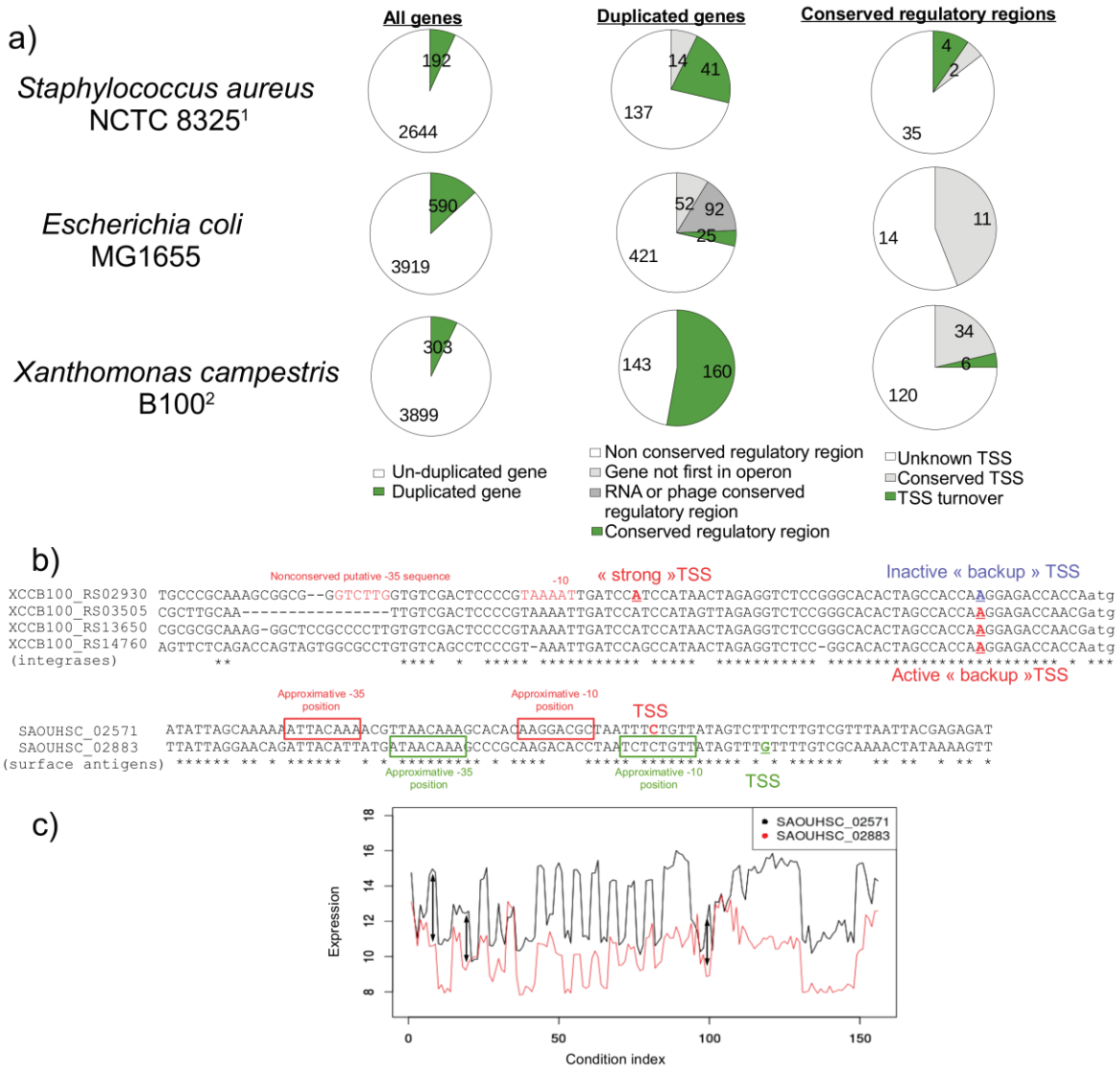


Figure 63. Presence of TSS turnover in various bacterial species.

a) Search for TSS turnovers in bacterial genomes coupled with TSS mapping data. A TSS turnover was detected if the proteins and the corresponding regulatory regions were homologous and TSS were known for both genes but not located at the same position. Numbers correspond to the number of genes in each category. 1: stable RNAs and phage genes were not present in the original data set. 2: because the *Xanthomonas* data set had no functional or operon information, we could not exclude operons, RNA and phage genes from the analysis and these categories are included inside "conserved regulatory regions". b) Examples of TSS turnovers detected with the analysis. In *Xanthomonas*, a strong upstream TSS is present in only one case. In the other genes, the -35 sequence of the « strong » TSS is scrambled and a « backup » TSS is used instead. No expression data is available. In *Staphylococcus*, mutations around the hypothetical -10 and -35 sequences may have relocated the TSS. c) Expression levels of the two paralogous surface antigen genes in *S. aureus*, plotted for each of the 157 conditions tested. Black vertical arrows highlight example conditions where strong differential regulation of the genes is observed.

We searched for paralogous genes with conserved regulatory regions in these genomes and examined the conservation of the experimentally determined TSS (see materials and methods). The analysis was limited by the high number of unknown TSS: 85, 56 and 75% of duplicated regulatory regions could not be compared in *S. aureus*, *E. coli* and *X. campestris* respectively because at least one paralogous gene had no known TSS. Among the duplicated regulatory regions with known TSS, 10 out of 57 (18%) showed evidence of a putative TSS turnover. This proportion is smaller than in human for example (85% [273]) but represents a significant fraction of duplicated regulatory regions and can be detected in different bacterial phyla. Interestingly,

the TSS turnover in the regulatory region of *S. aureus* surface antigens coincides with differential expression in some experimental conditions (Figure 63b and c). Consequently, TSS turnover could be widespread in bacteria.

Discussion

Despite their key role in the virulence of *D. dadantii*, the benefit to maintain the closely related paralogs *pelE* and *pelD* remained elusive. In this study, we demonstrate that, following a gene duplication event, PelE was selected as an initiator of pectin degradation while PelD acted in later stages. This distinction between PelD and PelE is well conserved in *Dickeya* as all tested *pelE* promoters have a high pectin-independent expression (Figure 59), although some strains lack a dedicated *pelE* (i.e. most *D. chrysanthemi)* and may have recruited other *pel* genes to initiate pectin degradation.

Our transcriptional (Figure 59) and structural (Figure 60) data suggest that the promoter of the ancestor of *pelD* and *pelE* was initially located at a position similar to that of *pelD* (Figure 64) and was composed of CRP1, FIS0 and KdgR binding sites, as suggested by finding of such organization in *D. paradisiaca*, which diverged before the gene duplication at the origin of *pelD* and *pelE*. Accordingly, the CRP1 binding site, which appeared earlier than CRP2, would be too distant to activate a *pelE*-like promoter [142] and support an efficient expression of the *pelD/E* ancestral gene. This *pel* promoter likely had a simple regulation, as the one observed in *D. paradisiaca* (Figure 59). The emergence of CRP2, FIS2 and FIS3 binding sites in the lineage leading to other *Dickeya* species may have improved the control by FIS and CRP, rendering the CRP1 site at least partially redundant.

Subsequently, the gene duplication yielded *pelD* and *pelE*, whereby the *pelE* promoter was shifted by 24 bp downstream, acquiring its main features: a strong pectin-independent expression and an earlier induction than *pelD* (Figure 59 and Figure 62). Sensitivity to pectin is under the control of KdgR [130,236]. Since KdgR overlaps the TSS on *pelD* and the -10 hexamer sequence on *pelE*, it is conceivable that KdgR is a more efficient repressor when occluding the TSS (acting as a roadblock instead of acting as a competitor [29]). This does not exclude the implication of other unknown regulators involved in pectin inducibility of *pel* genes. Opposite to pectin sensitivity, CRP and FIS do not appear to be involved at this stage of the divergence between *pelD* and *pelE* as the regulation by CRP seems ancestral and the response to FIS is highly variable (Figure 59). From this point on, PelE could already act as an initiator of pectin degradation. Further modifications affected both promoters (Figure 64), the most notable being that *pelD* "recycled" the apparently redundant CRP1 site to control its divergent promoter *div* and establish a tighter control of its expression in the absence of pectin as well as better inducibility, relying on the balance between *pelD* and *div* [228]. *pelE*, in contrast, acquired a totally new TSS-distant region of yet unknown function

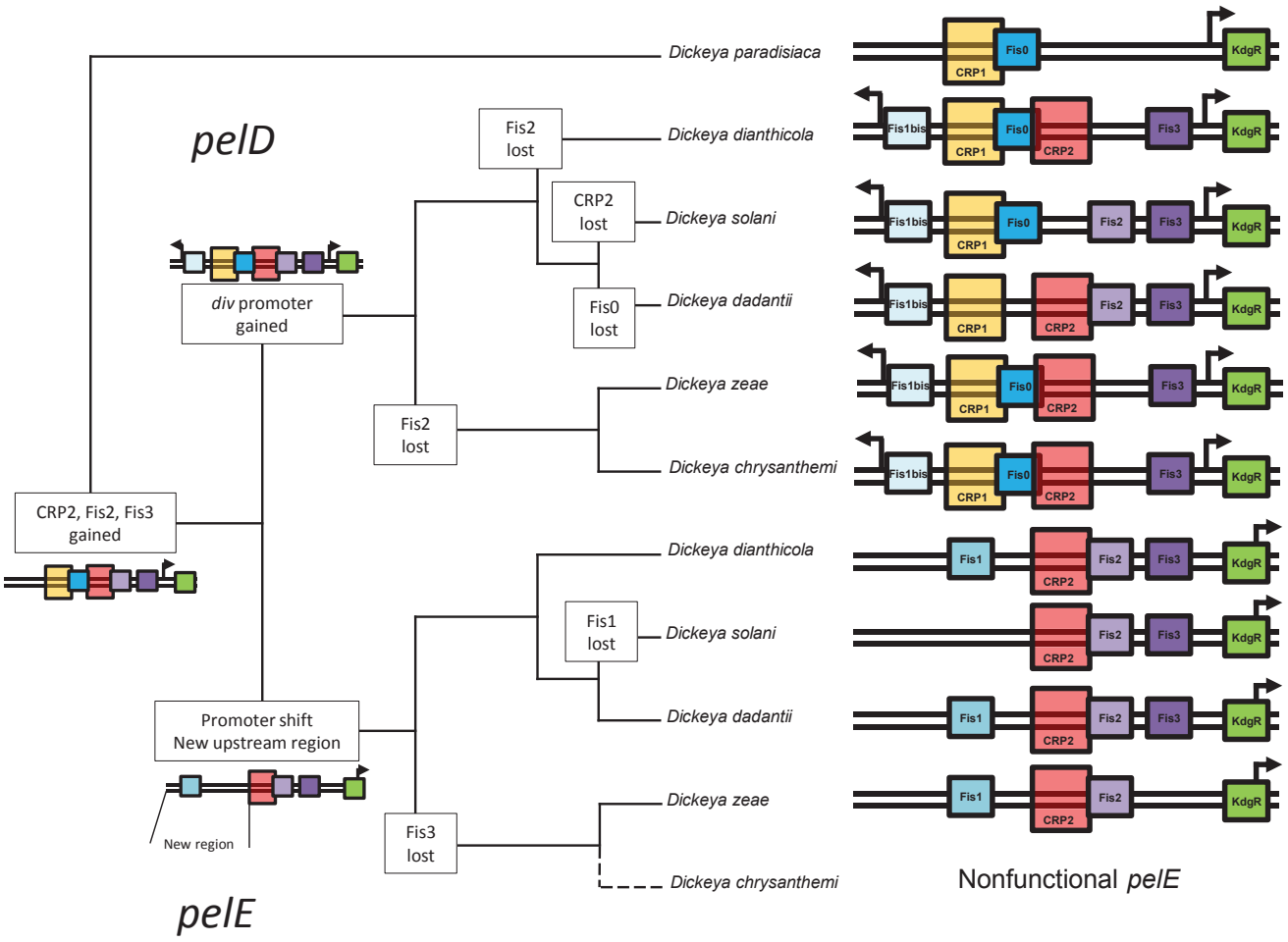


Figure 64. Model of evolution of the *pelD/E* regulatory regions in *Dickeya*.
Regulatory regions of all *pelD/E* genes from *Dickeya* species are reported on a *Dickeya* phylogeny adapted from the most recently published [276]. Most parcimonious events leading to the observed promoter structures are indicated on the branches. The potential structures of the regulatory regions before and after the *pelD/E* duplication are indicated next to the corresponding branches.

While PelE is an initiator of pectin degradation, and a *D. dadantii pelE* mutant has reduced virulence on *Saintpaulia*, no reduction of virulence is observed on potato tuber or chicory leaves [237]. This suggests the existence of other pectin degradation initiation pathways, independent of *pelE*, that are used depending on the host.

Due to their high variability, bacterial regulatory regions are key players in the evolution of paralogous genes [84]. Several examples of post-duplication regulatory modifications are documented [84]. This is the case for instance of the acquisition of an aerobic citrate degradation pathway in the long-term experimental evolution experiment of *E. coli* [277]. In this experiment, the *citT* gene, coding for an anaerobic citrate transporter, was duplicated in tandem, which resulted in the second copy of *citT* being under control of a neighbouring aerobic promoter. Consequently, the mutants acquired an initial aerobic Cit+ phenotype which was later improved by gene amplification, allowing faster growth on citrate-containing media [81,277].

Here, we propose a new, more subtle mechanism sufficient by itself to explain the

divergence between the expression and function of paralogous genes. It arises from a few point mutations in the target regulatory regions resulting in the shift in the promoter positions, i.e. a TSS turnover. This shift changes the relative binding position of all regulators with respect to RNA polymerase binding site and causes profound changes in the regulation of these genes. To our knowledge, this is the first report that turnover of transcription start sites contributes to divergence in expression between duplicate genes in bacteria. Unfortunately, the lack of a comprehensive knowledge about TSS positions in bacteria does not allow to perform reliable genome-wide searches for this phenomenon yet. Nevertheless, we have been able to detect a few other cases of putative TSS turnover (Figure 63). Alongside these putative new TSS turnovers, the *pelE-pelD* system could be used to explore other aspects in duplication-divergence mechanisms relevant to evolutionary and systems biology. To conclude, despite differences in genetic regulation, this mechanism is now reported in two kingdoms of life (*Bacteria* and *Eucarya*) suggesting that it is likely very general.

Experimental Procedures

Phylogenetic analyses

A search for bidirectional best blast hit homologous proteins to *D. dadantii*'s PelA, PelE and PelD was performed using Rsat 2015 [278]. Redundant sequences showing little to no difference were discarded, except for *Dickeya*.

Unrooted protein Maximum likelihood (ML) phylogenies of Pel proteins (106 sequences) were then built as described below. According to these preliminary trees, *Dickeya* sequences and their closest relatives were selected and used for an in-depth analysis in *Dickeya* with a broader taxonomic sampling. More precisely, *s*equences from additional *Dickeya* strains were added to this set of sequences and all were aligned using the MUSCLE program implemented in SEAVIEW version 4.6 [279]. The nucleotide sequences were translated, aligned at the protein level with MUSCLE and reverted to original nucleotide sequences in order to preserve the codon structure. The resulting multiple alignments were trimmed with GBLOCKS v. 091b implemented in SEAVIEW [280].

Protein ML phylogenies were inferred using PhyML v. 3.1 [281] with the Le and Gascuel (LG) model [282], a gamma distribution with 4 discrete classes of sites (Γ4) and an estimated alpha parameter. Nucleotide ML trees were inferred with nhPhyML [283], with the Galtier and Gouy (GG) model [284], in combination with a Γ4 distribution. The discrete version of the model was considered with three values of G+C equilibrium content (0.25, 0.5 and 0.75). ML tree branch supports were computed with the non-parametric bootstrap procedure implemented in PhyML (100 replicates of the original alignments).

Protein Bayesian phylogenies were inferred using MrBayes v3.2.2 [285] with a mixed evolution model+Γ4. Four chains were run in parallel for 10^6 generations. The first 2000 generations were discarded as burn-in. The remaining trees were sampled every 100 generations to build a majority consensus tree and compute posterior probabilities.

Prediction of regulator binding sites

All sites which were similar to experimentally validated binding sites (in *Dickeya dadantii* [145,192,228]), with no mutation in a highly conserved position, were taken into account. Further sites were detected with Virtual Footprint (http://www.prodoric.de/vfp/). Only the sites that had a score greater than at least one of the previously determined sites were conserved. Promoters were detected using Bprom (http://linux1.softberry.com/berry.phtml?topic=bprom&group=programs&subgroup=gfindb) and manually verified to be compatible with the observed CRP activation (i.e. with at least one CRP site situated around -41.5, -61.5, -93.5 or -103.5 relatively to the *pel* TSS [286]).

Genome-wide prediction of TSS turnovers

Nucleotide sequences and genome annotation of *Escherichia coli* MG1655, *Staphylococcus aureus* NCTC 8325 and *Xanthomonas campestris* B100 were downloaded from NCBI. For each organism, gene sequences were translated with MACSE [287] and were compared using BLASTP from the Blast+ application [288] with default parameters. Genes were clustered into the same family of paralogs if (i) protein sequence identity was greater than 50 % and (ii) the alignment length was at least 80% of the size of the longest protein. Genes coding for proteins having at least one paralog were considered “duplicated genes”.

For each duplicated gene, the 250 nt upstream region was considered to be the regulatory region. Regulatory regions were compared using Blastn with a word size of 7 and otherwise default parameters. When two regions corresponding to duplicated genes produced an alignment of at least 40 nt (gaps not included), they were considered “conserved regulatory regions”.

TSS positions were retrieved from Biocyc and the literature [289,290] and were placed on the alignments of the regulatory regions. A TSS turnover was detected if the TSS of two homologous regulatory regions were distant of more than 5 nt on the alignments. All automatically detected TSS turnovers were manually verified. When this comparison was not possible because either or both TSS were unknown, genes were classified in “unknown TSS”.

Bacterial strains, plasmids, media, culture conditions and molecular biology methods

D. dadantii was cultured using standard conditions (M63S, i.e. M63 minimal medium supplemented with sucrose 0.2 % and, when needed, pectin 0.2 %, at 30°C) as previously described [228].

The *pelD/E* regulatory regions from *D. dianthicola* and *D. solani* (D17, D19, E10 and E11) were amplified by PCR from genomic DNA using the corresponding primers (see Table 5). The *pelDE* regulatory region from *D. paradisiaca* was obtained by direct synthesis from Genecust Europe (Luxembourg) from the sequence published on Genbank for strain NCPPB 2511 (NZ_CM001857). The hybrid *pelD-pelE* regulatory region was obtained with 2 rounds of site-directed mutagenesis from *D. dadantii's pelD* regulatory region (D0). The various regulatory regions were then cloned into pProbe-AT'.

Table 5: Bacterial strains, plasmids and primers used in this study

Strains	Description	Ref
D. dadantii		
3937 (A4922)	Wild-type strain isolated from *Saintpaulia ionantha*	[243]
A1798	3937 pelD::*uidA*-Kan^R	[275]
A1828	3937 pelE::*uidA*-Kan^R	[275]
A2507	3937 *crp::*Cm^R	[136]
A4474	3937 *fis::*Cm^R	[167]
D. dianthicola		
NCPPB 453^T	Wild-type strain isolated from *Dianthus*	[276]
D. solani		
IPO 2222^T	Wild-type strain isolated from *Solanum tuberosum*	[276]

Plasmids	Description	Ref
pProbe-AT '	Reporter vector with *gfp*, AmpR	[246]
pPD0	pProbe-AT ' containing *D. dadantii*'s *pelD* regulatory region between sites EcoRI and HindIII	[228]
pPD17	pProbe-AT ' containing *D. dianthicola*'s *pelD* regulatory region between sites EcoRI and HindIII	This work
pPD18	pProbe-AT ' containing *D. paradisiaca*'s *pelD* regulatory region between sites EcoRI and HindIII	This work
pPD19	pProbe-AT ' containing *D. solani*'s *pelD* regulatory region between sites EcoRI and	This work
pPD21	pProbe-AT ' containing *D. dadantii*'s hybrid *pelD-pelE* regulatory region between sites EcoRI and HindIII	This work
pPE0	pProbe-AT ' containing *D. dadantii*'s *pelE* regulatory region between sites EcoRI and HindIII	This work
pPE10	pProbe-AT ' containing *D. dianthicola*'s *pelE* regulatory region between sites EcoRI and HindIII	This work
pPE11	pProbe-AT ' containing *D. solani*'s *pelE* regulatory region between sites EcoRI and HindIII	This work

Primers		Target	Ref
312	AAAGAATTCACTTTAGTCATTTCAAATGAAAATAG	*pelD, D. dianthicola*	This work
313	AAAAAGCTTTTGGTTCCTGCAGAAGACAC		This work
314	AAAGAATTCTAATTTATTGATTTTAATTAAAAATAATAAC	*pelE, D. dianthicola*	This work
315	AAAAAGCTTTGGGTTCCCAGGGTAG		This work
318	AAAGAATTCTCATTTGGGTTATTTCAGATAAAAAC	*pelD, D. solani*	This work
319	AAAAAGCTTTTTGTTCCTGCAGAAG		This work
320	AAAGAATTCTAATTTATTGATTTTAATGAAAAAC	*pelE, D. solani*	This work
321	AAAAAGCTTTTGGTTCCCAGGGCAG		This work
326	CAATCGAAAACGCTTAAAcATtCGCCTTaCCAAAGGACAAAATGG	Hybrid *pelD-pelE*, 1st round	This work
327	CCATTTTGTCCTTTGGtAAGGCGaATgTTTAAGCGTTTTCGATTG		This work
328	CCAAAGGACAAAATGGtaTTTCATTTTTTgCACAAACACTTTTCAGTC	Hybrid *pelD-pelE*, 2nd round	This work
329	GACTGAAAAGTGTTTGTGcAAAAAATGAAAtaCCATTTTGTCCTTTG		This work

362	GTCTTACGTGTGTTGTTC	*D. paradisiaca pelD/E*	This work
363	TGAGCGTCCTGCACC	TSS	This work

In vivo transcriptional activity measurements

D. dadantii was transformed with various pProbe plasmids and was cultured in 96 well plates as described previously [228]. Real-time mRNA levels were computed from fluorescence data using the model described in [248]. Because different *pel* genes or mutant strains resulted in irrelevant changes in the temporality of gene expression, the maximum mRNA levels were considered in Figure 59.

TSS determination by reverse transcription

Total RNA was extracted from *E. coli* DH5α expressing a plasmid-borne *D. paradisiaca pelDE-gfp* fusion (pPD18) using the SV Total RNA isolation system (Promega, USA). A reverse transcription was performed on 1 µg total RNA using the primers 362 or 363 (see Table 5), using the Revertaid first strand cDNA synthesis kit (Fermentas, USA) according to the manufacturer's instructions. Sanger sequencing reactions were performed using the same primers on the pPD18 plasmid using the T7 sequencing kit (USB, USA) according to the manufacturer's instructions.

Conflict of interest statement

The authors declare that no conflicting interests exist.

Author contributions

Conceptualization, A.D., W.N., C.B.-A. and S.R.; Formal Analysis, A.D., C.B.-A. and S.L.; Investigation, A.D.; Writing – Original Draft, A.D.; Writing – Review & Editing, A.D., W.N., C.B.-A., S.L. and S.R.; Visualization, A.D. and C.B.-A.; Supervision, W.N., C.B.-A. and S.R.

Acknowledgements

We would like to thank Méganne Billet for building the hybrid *pelD-pelE* regulatory region, as well as the cloning of the *D. solani* and *D. dianthicola* regulatory regions.

We are grateful to Agnès Rodrigue, Xavier Charpentier and their teams for allowing us to use their Tecan plate reader, Georgi Muskhelishvili and Florence Hommais for their critical reading of the manuscript.

This work was supported by the Agence Nationale de la Recherche combicontrol (grant number ANR-15-CE21-0003-01) and Investissement d'Avenir (grant number ANR-10-BINF-01-01 "Ancestrome" to C.B-A). C.B-A is member of the Institut Universitaire de France

D. L'acquisition de la sensibilité aux conditions environnementales de *pelE* (résultats non publiés)

Le troisième article a montré que le décalage du promoteur entre *pelD* et *pelE* est à l'origine des caractéristiques d'expression uniques de *pelE* (plus précoce et moins sensible à la présence de composés pectiques). Cependant, il existe d'autres différences connues entre l'expression de *pelD* et *pelE*, notamment une plus grande sensibilité de l'expression de *pelE* aux conditions environnementales [193]. Plus précisément, *pelE* est plus sensible au relâchement induit artificiellement par la novobiocine (inhibiteur de l'ADN gyrase) ou naturellement par le stress acide ou oxydant. De plus, des résultats non publiés indiquent également que *pelE* serait régulé plus fortement par OmpR (Table 6).

Table 6 : Dosage β-glucuronidase de l'expression de différents gènes codant pour des facteurs de virulence ou associés à la virulence de *D. dadantii* dans une souche sauvage et un mutant *ompR* (G.Condemine communication personnelle).

	Gène	WT	*ompR*	Rapport	Régulation
Facteurs d'agression	pelA	4,9	3,4		
	pelB	36	32		
	pelC	28	23,5		
	pelD	8,1	25	-3,1	Répression
	pelE	59,5	1360	-23	Répression
	pelI	66	57		
	pemA	89,5	102		
	paeX	50	71		
	paeY	94	89		
	pehN	136	116		
	pehV	41	12,8	3,2	Activation
	pehW	33,5	4,8	7,0	Activation
	pehX	64	6,3	10	Activation
Facteurs associés à la virulence	kdgM	233	882	-3,8	Répression
	kdgN	710	82	8,7	Activation
	outC	407	449		
	indA	743	595		
	hrpN	5,2	4,9		
	opgB	2360	1814		
	togM	152	120		
Régulateurs	pecS	4742	8316	-1.8	Répression
	pecT	3753	4067		
	kdgR	1100	1194		

Les gènes activés ou réprimés d'un facteur au moins égal à 1.5 sont indiqués en noir, les autres en gris. Les activités sont exprimées en µmol de produit libéré/minute/mg PSB.

OmpR est le régulateur de réponse du système à deux composants OmpR/EnvZ, qui a été initialement caractérisé pour son implication dans la perception de l'osmolarité et/ou de l'acidité chez de nombreuses bactéries [291]. Des travaux récents chez *S. enterica* indiquent également un lien entre OmpR et le surenroulement de l'ADN [176], ce qui ferait de OmpR un régulateur

environnemental global ne se limitant pas à la perception des stress acides et osmotiques. En particulier, chez les entérobactéries zoopathogènes telles que *Yersinia* et *Salmonella*, OmpR est un régulateur central de la virulence [291,292]. Or, lors de l'infection, *D. dadantii* est confronté séquentiellement à un stress acide puis osmotique [247], mais le seul rôle publié d'OmpR est l'activation de la transcription des porines impliquées dans le catabolisme de la pectine (KdgM et KdgN) par un mécanisme encore inconnu [293]. En particulier, aucune information sur la régulation des gènes *pel* par OmpR n'est disponible. Puisque PelE est un initiateur de la dégradation de la pectine, la régulation très forte de l'expression de *pelE* par OmpR pourrait se répercuter sur le reste du processus infectieux. La question de l'acquisition de cette sensibilité et de ses déterminants se pose alors.

La régulation de *pelD*, *pelE* et de l'hybride *pelD-pelE* (Figure 60b) par OmpR et la novobiocine a donc été explorée en utilisant le système de fusions transcriptionnelles *gfp* couplé au calcul en temps réel des quantités d'ARN messagers détaillé dans le 2ème article (Figure 65). Comme observé dans le 3ème article et dans la Table 6, *pelD* est relativement peu sensible au relâchement de l'ADN et à OmpR. En particulier, l'expression maximale de *pelD* n'est pas modifiée mais devient uniquement légèrement plus précoce dans le mutant *ompR*, malgré le retard de croissance de la souche. De même, *pelE* est réprimé par le relâchement induit par la novobiocine (facteur 2) et par OmpR (facteur 1.5). L'hybride *pelD-pelE* conserve des caractéristiques plutôt proches de *pelD* sur ces aspects. En effet, l'hybride est très peu sensible à la novobiocine et à OmpR (facteur 1.2). Il faut noter que, comme rapporté dans le second article, les effets observés avec le système Gfp sont moins marqués qu'avec le système GUS de la Table 6, qui ont de plus été effectués dans un milieu différent (M63 glycérol dans la Table 6, M63S + PGA pour la Figure 65)

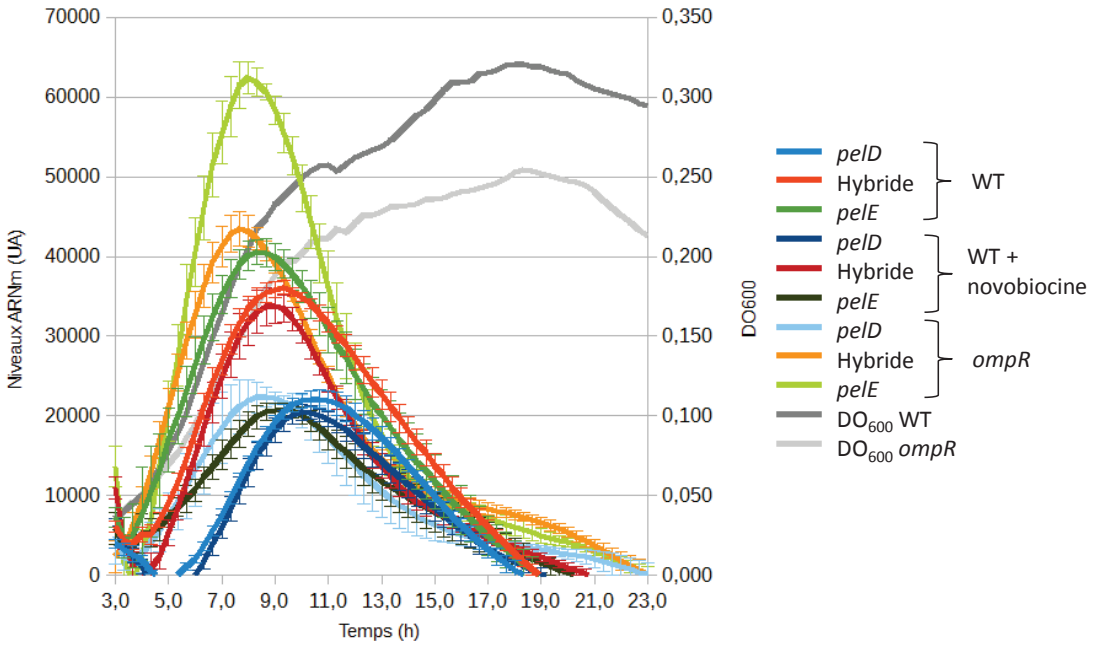


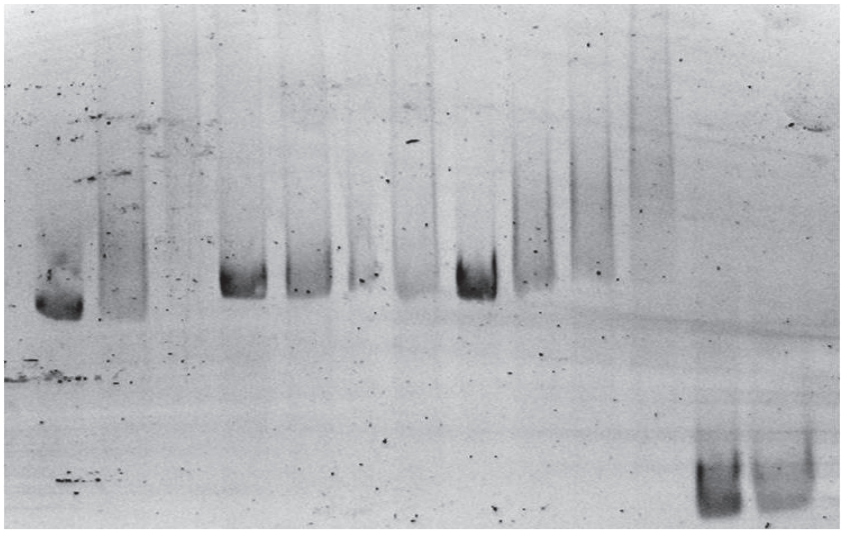
Figure 65 : Effet de OmpR et du relâchement de l'ADN sur l'expression de *pelD*, *pelE* et de l'hybride *pelD-pelE* (D21).

Les souches sauvages ou du mutant *ompR* de *D. dadantii* portant des fusions transcriptionnelles plasmidiques *pel-gfp* (pPD0, pPD21 ou pPE0) ont été cultivées en plaque 96 puits et les quantités d'ARNm ont été calculées en temps

réel comme décrit dans le matériel et méthodes de la section A. Des concentrations sublétales de novobiocine (25 µg/mL) ont été ajoutées dès le début de la culture pour induire un relâchement de l'ADN. Les niveaux d'ARNm sont indiqués comme la moyenne de 9 répliques et les barres d'erreur représentent l'écart-type.

La conservation de la réponse au surenroulement suite au décalage du promoteur de l'emplacement *pelD* vers *pelE* était attendue. En effet, la réponse d'un promoteur au surenroulement de l'ADN dépend des propriétés globales de la région régulatrice (voir section I-A-3-e). Lors de la construction de l'hybride, les quelques substitutions effectuées sont probablement insuffisantes pour avoir un effet fort sur les propriétés physiques de l'ensemble de la région régulatrice *pel*.

Pour clarifier l'observation concernant l'impact de OmpR sur l'expression des gènes *pelE* et *pelD*, une étude plus approfondie visant à déterminer la position et l'affinité des sites de fixation OmpR a été entreprise. Tout d'abord, l'affinité globale d'OmpR pour les deux promoteurs a été comparée par retard de migration sur gel (Figure 66). Pour *pelD* et *pelE*, un complexe net ADN-OmpR n'a pas pu être observé. Cependant, d'après la vitesse de disparition de la forme libre, *pelE* semble avoir une meilleure affinité pour OmpR que *pelD*, ce qui contribue à expliquer l'effet plus marqué d'OmpR sur *pelE* que *pelD*. Les profils peuvent résulter soit de la formation d'agrégats, soit d'une déphosphorylation de la protéine OmpR en cours de migration, phénomène précédemment observé dans les expériences similaires effectuées avec les formes natives de OmpR [294].



ADN	*ompF* (Témoin +)			*pelD*				*pelE*				Témoin -	
OmpR (µM)	0	0,3	2	0	0,3	1	2	0	0,3	1	2	0	2

Figure 66 : Retard de migration sur gel d'acrylamide causé par OmpR sur les régions régulatrices de *pelE* et *pelD*.
Les régions régulatrices testées ont subi un marquage fluorescent identique à celui effectué pour les empreintes à la DNase I (voir matériel et méthodes de la section A) puis ont été mis en contact avec différentes concentrations d'OmpR phosphorylé à l'acetyl phosphate. Les complexes OmpR-ADN ont ensuite été séparés sur gel d'acrylamide non dénaturant et révélés par fluorescence.

En plus de l'affinité de la protéine pour la région régulatrice, le nombre et la position des sites de fixation d'OmpR peut contribuer à expliquer les différences observées entre *pelD* et *pelE*. Afin de déterminer la position exacte des sites de fixation, des empreintes à la DNase I ont été

effectuées (Figure 67). Pour *pelE*, une empreinte nette au niveau de la base 280, en partant du début de l'oligo marqué, se forme dès 300 nM d'OmpR (site OmpR2), ainsi que deux empreintes mineures à 600 nM (tailles 140 et 340). Les empreintes sont moins claires pour *pelD* puisque des signaux semblent apparaitre tout le long de la région régulatrice. Une zone de taille 180 semble cependant être un site de fixation spécifique et porte un très fort pic d'hypersensibilité apparaissant dès 300 nM. La meilleure affinité d'OmpR pour *pelE* observée en retard sur gel résulte donc à la fois de la présence d'un site de forte affinité (OmpR2) mais également de deux autres sites semblant inexistants sur *pelD*.

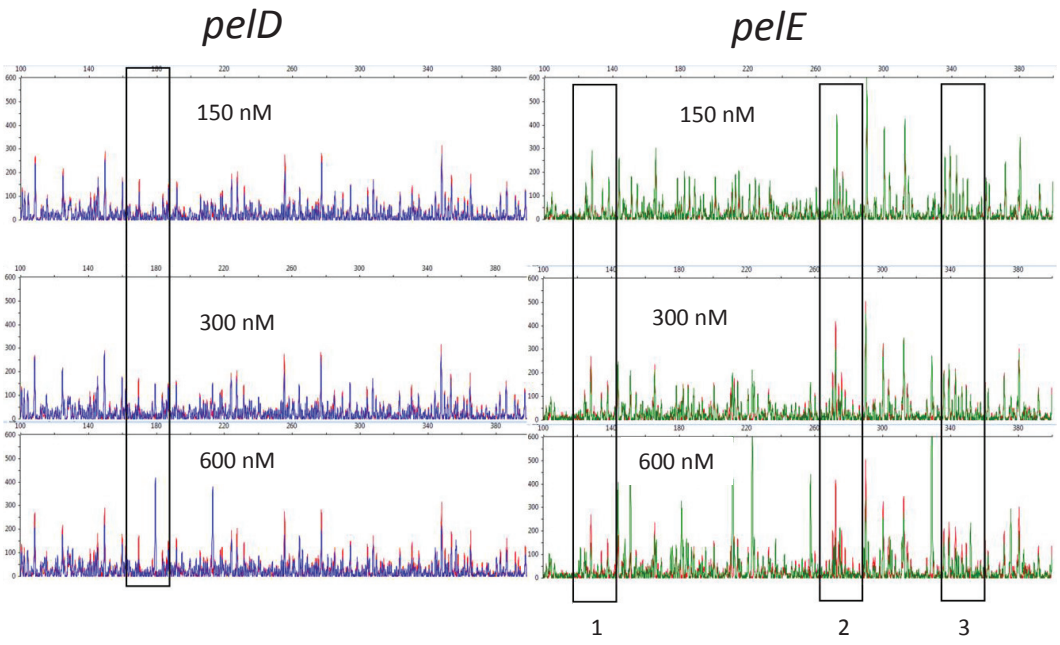


Figure 67 : empreinte à la DNase I causée par OmpR sur les régions régulatrices *pelD* et *pelE*.
Une empreinte à la DNase I a été réalisée sur les régions régulatrices *pelD* et *pelE* marquées par fluorescence. Les fragments ont alors été séparés selon leur taille par électrophorèse capillaire. Le profil obtenu en l'absence de protéine (rouge) est superposé au profil obtenu en présence d'une concentration donnée d'OmpR (bleu et vert). Les zones protégées par OmpR sont encadrées, et numérotées dans le cas de *pelE*.

Les positions de ces empreintes peuvent servir de guide pour détecter des sites de fixation OmpR. Bien qu'aucune donnée ne soit disponible concernant le consensus OmpR, les protéines OmpR de *D. dadantii* et de *S. enterica* sont quasi-identiques (100% similarité, 99% identité). En termes de structure (Figure 68), les deux acides aminés variables entre *D. dadantii* et *S. enterica* causent des modifications mineures dans les feuillets β proches du linker entre les domaines de phosphorylation et de fixation à l'ADN d'OmpR. En revanche, il n'y a pas de changement dans le motif HTH nécessaire à la reconnaissance de l'ADN. Ces observations permettent de supposer une bonne conservation du motif de fixation reconnu par les protéines entre les deux espèces.

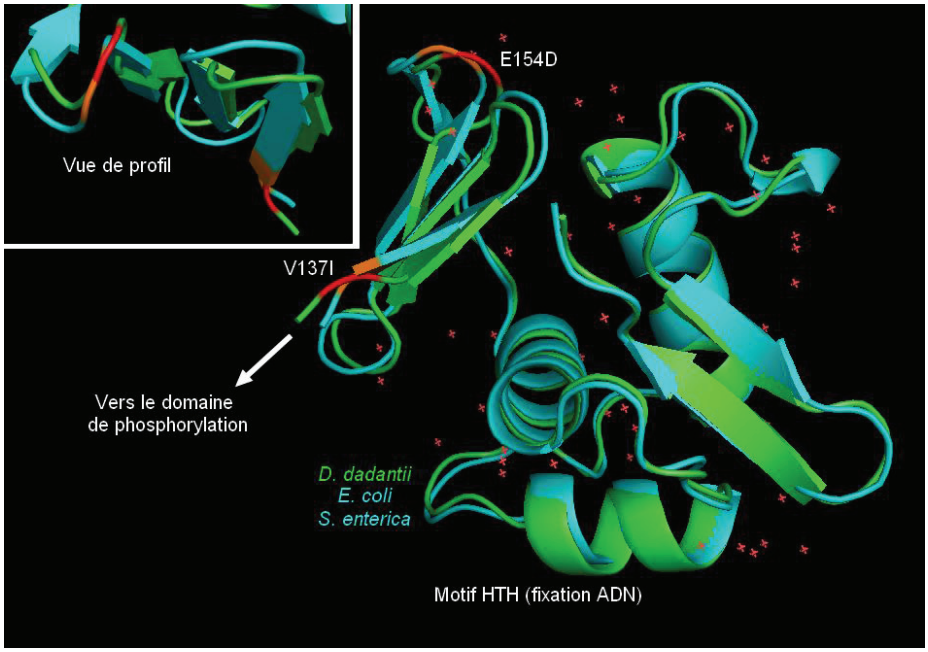


Figure 68 : Comparaison des structures des domaines C-terminaux d'OmpR chez *E. coli*, *S. enterica* et *D. dadantii*.

La structure aux rayons X du domaine C-terminal de OmpR de *E. coli* (pdb 10DD, en bleu) a été superposée à la structure prédite chez *D. dadantii* (voir section II-D, en vert). Les protéines de *S. enterica* et *E. coli* sont 100% identiques. Les deux résidus différant entre *E. coli* et *D. dadantii* sont indiqués en orange et rouge sur chaque structure. Les croix rouges représentent les molécules d'eau. HTH : Hélice-Tour-Hélice.

Les sites potentiels OmpR ont ainsi pu être identifiés à partir des empreintes et par similarité avec un consensus déterminé à partir des données de ChIP pour *S. enterica* (voir section II-E), comme montré en Figure 69. L'unique site OmpR putatif de *pelD* chevaucherait le site KdgR, ce qui pourrait mener à une compétition entre les deux régulateurs. Cette situation pourrait être à l'origine de l'expression plus précoce de *pelD* dans le mutant *ompR* comparé à la souche sauvage (Figure 65). Ces résultats restent préliminaires et la compétition entre OmpR et KdgR reste encore à être démontrée. Pour *pelE*, le site de forte affinité (OmpR2) est situé en aval du +1. Son rôle répresseur passe donc probablement par une inhibition de l'élongation du transcrit par barrage, similaire au rôle que remplit KdgR sur *pelD*. Les sites OmpR1 et OmpR3 sont situés à des positions inhabituelles (loin en amont et en aval, respectivement). La présence de pics d'hypersensibilité non associés à des sites OmpR laisse cependant penser à une courbure de la région régulatrice induite par OmpR à laquelle ces sites pourraient participer. Une régulation passant par la topologie de l'ADN similaire à ceux décrits sur les gènes *ssrA* et *hilD* de *S. enterica* ne peut pas être exclue [176].

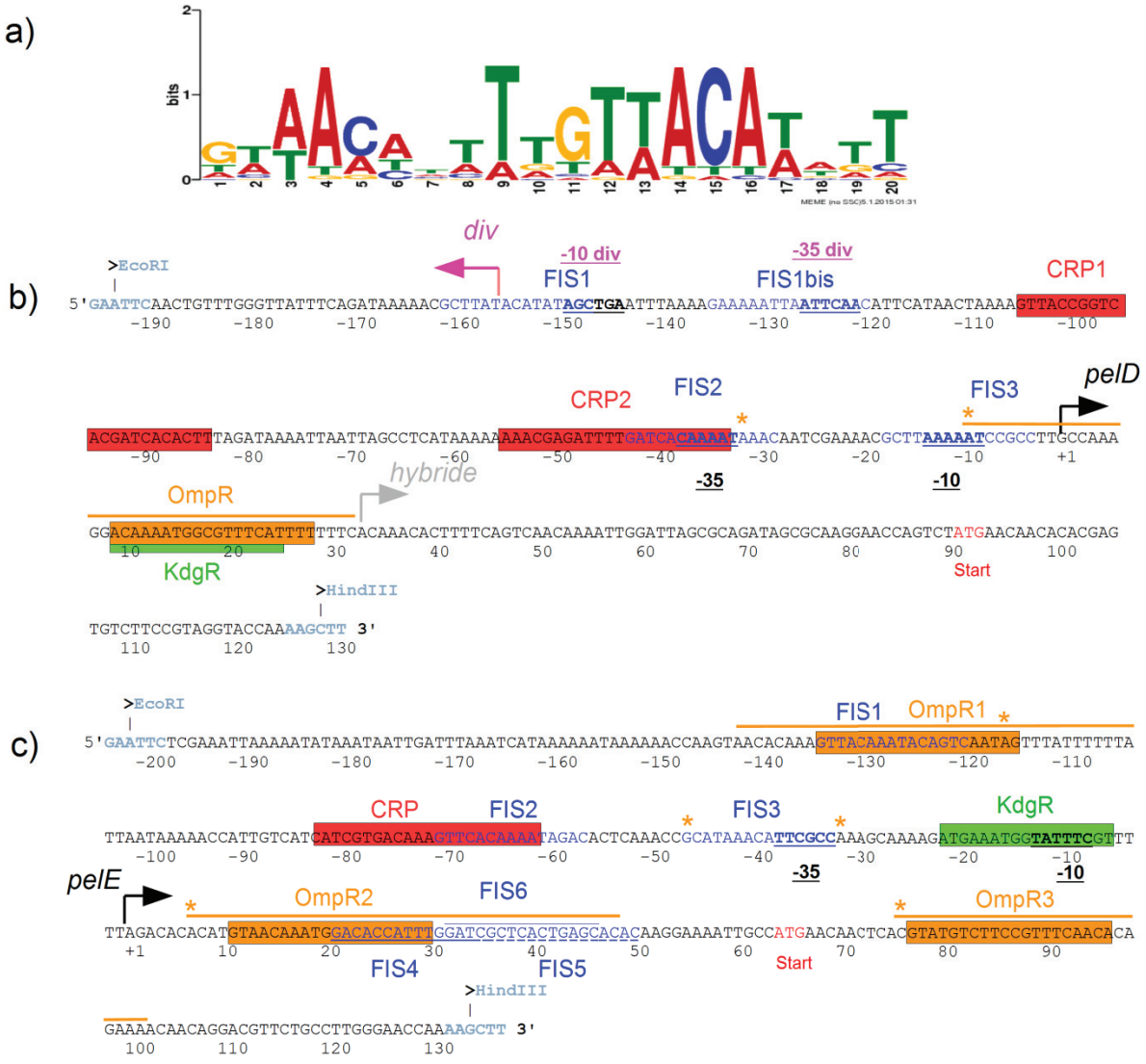


Figure 69 : Localisation des sites de fixation OmpR détectés par DNase I.

Le consensus de fixation OmpR a été déterminé à partir des données de ChIP obtenues chez *S. enterica* (a) puis positionné par rapport aux empreintes à la DNase I obtenues pour *pelD* (b) et *pelE* (c). Le code couleur est identique à la Figure 37. La position du promoteur hybride *pelD*/*pelE* est indiquée en gris sur la région régulatrice *pelD*. Les zones protégées par OmpR sont indiquées par des traits orange, les pics d'hypersensibilité majeurs par des étoiles orange et les consensus OmpR détectés par des rectangles orange. La numérotation des sites OmpR est identique à la Figure 67)

Cependant, et à l'inverse des autres sites de fixation des régulateurs présentés jusqu'à présent, les sites de fixation OmpR ne sont pas homologues entre *pelD* et *pelE*. Par exemple, le site OmpR2 de *pelE* qui possède un très bon début de consensus (GTAACA, position 11) n'est pas conservé chez *pelD* (CTTTTC, position 40). Cette observation implique que la régulation par OmpR s'est mise en place indépendamment sur *pelD* et *pelE*. Dans ces conditions, la réponse de l'hybride à OmpR est difficile à interpréter et pourrait simplement être due à la transition d'un mécanisme de répression par barrage du site OmpR de *pelD* vers une compétition directe pour l'hybride. Le site OmpR de *pelD* et OmpR2 de *pelE* ont en commun une position d'environ 10 pb après le TSS, ce qui laisse penser à une évolution convergente.

Chez *S. enterica*, OmpR est un régulateur global de la virulence en réponse à l'acidité par des fixations réparties à différents endroits du génome [176,226], mais aucune donnée n'était disponible pour un organisme phytopathogène. Nous avons donc exploré le rôle global de OmpR dans la virulence de *D. dadantii*. En l'absence de stress, les phénotypes du mutant *ompR* sont peu

marqués, avec uniquement une légère diminution de la mobilité et une augmentation de l'activité pectate lyase (Figure 70). Cela se traduit par une augmentation de la virulence sur des feuilles d'endive. En revanche, en présence de stress acide, la croissance du mutant *ompR* est fortement diminuée, de manière comparable au mutant *mfbR* (régulateur considéré comme principal dans la réponse au stress acide).

Phénotype	WT	*ompR*
Mobilité	510 UA	320 UA
Pectinase (boîte)	+	+
Cellulase	+	+
Protéase	+	+
Pectate lyase (liquide)	1,6 ± 0,4 U	4,0 ± 1,8 U
Tolérance pH 5	+	-
Virulence (masse endive macérée)	1,0 ± 0,4 g	1,6 ± 0,4 g

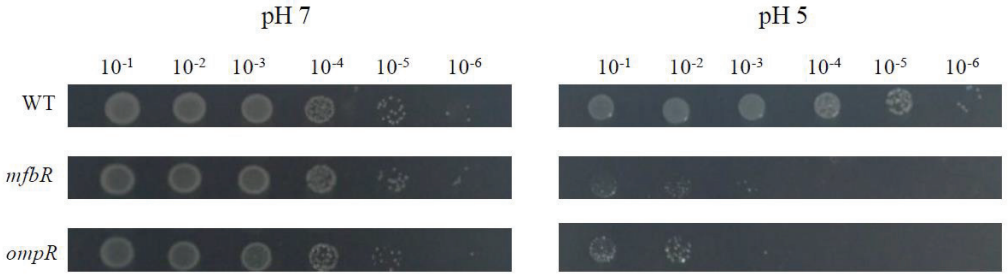


Figure 70 : Caractérisation phénotypique du mutant *ompR*

Les principaux facteurs de virulence ont été testés qualitativement sur milieu gélosé (voir section II-F), ainsi que l'activité pectate lyase en milieu liquide après 18h de culture. Les données de stress acide sont issues du travail de X. Jiang (communication personnelle).

Pour expliquer ces phénotypes, la distribution des sites OmpR sur le génome de *D. dadantii* a été explorée. Par manque de temps mais aussi à défaut de financement pour initier des analyses de transcriptome et d'interactome, cette étude n'a pu être réalisée que *in silico*. Le consensus de fixation OmpR déterminé à partir des données de ChIP de *S. enterica* (Figure 69a) a été utilisé pour chercher des sites de fixation chez *D. dadantii*. Les gènes cibles potentiels associés ont alors été classés par catégorie fonctionnelle et une analyse statistique a permis de déterminer les groupes fonctionnels différentiellement représentés parmi les cibles OmpR (Table 7). Les deux catégories les plus surreprésentées correspondent aux phénotypes observés : la répression de l'activité Pel correspond à la classe « catabolisme des cellules végétales » et l'activation de la mobilité à la classe « mobilité de la bactérie ». OmpR semble donc réguler la virulence directement. Par contre, aucun gène impliqué dans la réponse au stress ne possède de site OmpR prédit, malgré la sensibilité du mutant OmpR au stress acide. Cette régulation pourrait donc être indirecte ou s'exercer via des gènes de fonction inconnue.

Table 7 : Analyse fonctionnelle des cibles potentielles de OmpR sur le génome de *D. dadantii*.

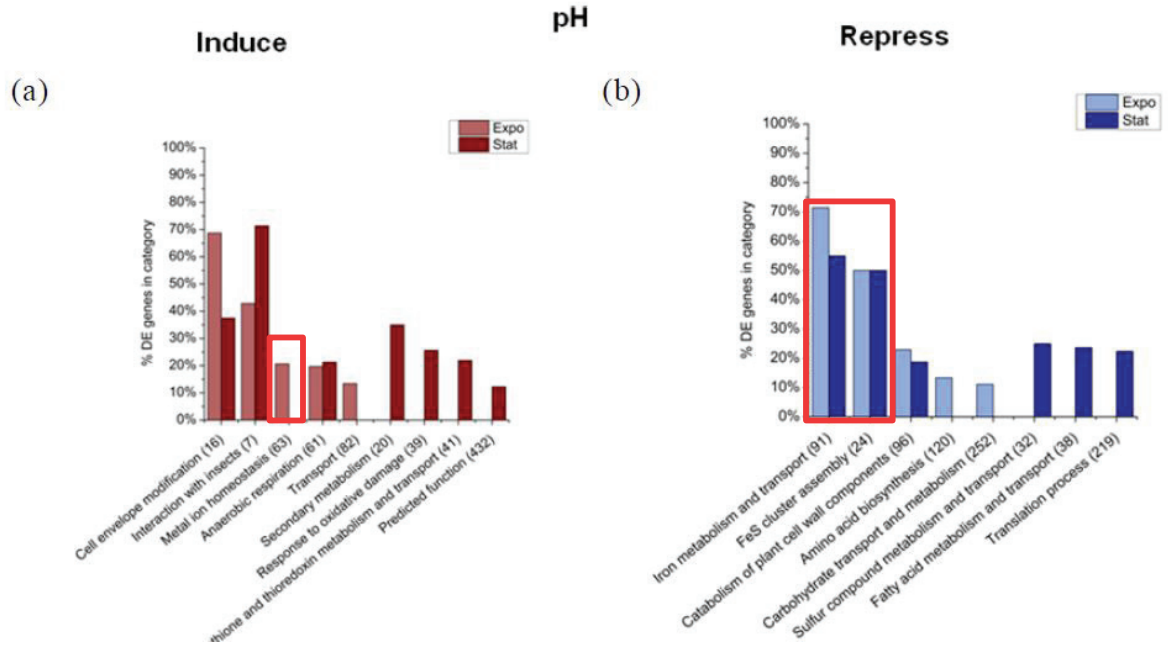
Classes de gènes	Nombre de gènes dans *D. dadantii*	Nombre de gènes régulés par OmpR prédits	Proportion %	p-value*
Catabolisme des cellules végétales	96	34	35,4	0,000
Mobilité de la bactérie	98	26	26,5	0,004
Fonction inconnue	579	113	19,5	0,007
Traduction et ribosomes	219	19	8,7	0,008

Métabolisme et transport de coenzymes	149	12	8,1	0,018
Cycle de Krebs	24	8	33,3	0,021
Détection et réponse au stress	23	0	0,0	0,039
Réplication, recombinaison et réparation de l'ADN	127	11	8,7	0,043
Protéines de membrane externe	14	5	35,7	0,048
Ensemble des gènes de *D.dadantii*	4555	687	15,0	

Les catégories sur-représentées sont indiquées en noir et les sous-représentées en gris

* : test exact de Fischer

Afin de prendre en compte également les effets indirects d'OmpR, les gènes codant pour des régulateurs et possédant les meilleurs sites de fixation OmpR ont été examinés individuellement. Le seul régulateur majeur trouvé est le régulateur général du métabolisme du fer Fur (62ème meilleur site du génome ; à titre de comparaison, *pelE* est en position 8 et *pelD* en 225). Ce lien potentiel avec Fur fait écho aux données de micropuces de X. Jiang (manuscrit en préparation) qui avait montré que les gènes liés au métabolisme du fer sont fortement réprimés par le stress acide (Figure 71). OmpR pourrait donc servir d'intermédiaire entre le stress acide et le métabolisme du fer via une régulation transcriptionnelle de Fur.



(c)

Name	pvalue	Description
pelE	4.54e-07	pectate lyase PelE
fur	3.83e-06	DNA-binding transcriptional dual regulator of siderophore biosynthesis and transport
pelD	2.26e-05	Pectate lyase PelD

Figure 71 : Le rôle potentiel de OmpR et Fur dans la réponse des gènes de *D. dadantii* à l'acidité.
Les catégories liées aux métaux dans les gènes induits (a) ou réprimés (b) par le stress acide sont encadrées en rouge (X. Jiang, manuscrit en préparation). (c) Position de fur dans les cibles OmpR prédites.

Afin de tester cette hypothèse, la réponse transcriptionnelle de *fur* et de plusieurs gènes du métabolisme du fer au stress acide a été analysée par qRT-PCR dans une souche sauvage et dans un mutant *ompR*. Une connexion entre OmpR et Fur devrait se traduire par une perte de la réponse transcriptionnelle de ces gènes à l'acidité dans le mutant OmpR. Plusieurs cibles ont été sélectionnées, à la fois des gènes induits et réprimés par le stress acide et possédant ou non un site OmpR prédit (Table 8). Aucune tendance claire permettant de confirmer ou d'infirmer l'hypothèse ne se dégage et les résultats obtenus sont incohérents avec les données de micropuces (par exemple, *ftnA* est réprimé par le stress acide au lieu d'être activé dans les données de micropuces), ce qui ne permet pas de conclure sur la connexion OmpR-Fur. L'espacement temporel entre les deux expériences, la sensibilité des gènes du métabolisme du fer au milieu et un changement de fournisseur des milieux entre les deux expériences ont probablement contribué à l'absence de reproductibilité des expériences.

Table 8 : Réponse des gènes du métabolisme du fer au stress acide

Gène	Fonction	Site OmpR prédit ?	Fold change WT (X. jiang)	Fold change WT	Fold change *ompR*
fur	Régulateur métabolisme fer	+	<2	+1.7	+1.1
zitB	Transporteur zinc (témoin)	-	+5	-1.2	-2
ftnA	Stockage fer	+	+30	-2	-2
efeU	Perméase fer	-	-10	-1.6	-9
yfeA	Transporteur fer	-	-12	-3	-6
tonB	Transporteur fer membranaire	+	-12	NA	NA

Les fold change sont entre les conditions sans stress et avec stress acide, en phase exponentielle de croissance et en absence de PGA. Une valeur positive indique une activation par le stress acide, une valeur négative une répression. NA : aucune expression n'a été détectée en condition de stress acide, empêchant un calcul des fold change.

De manière intéressante, des travaux récents sur *Yersinia* semblent également indiquer une connexion entre OmpR et le métabolisme du fer qui dépendrait de Fur [295]. En effet, chez cette bactérie, l'expression de gènes codant pour plusieurs transporteurs de fer est altérée dans un mutant *ompR* et aucune régulation directe forte n'a été identifiée. Les auteurs proposent une connexion indirecte passant par Fur, sans qu'il y ait pour l'instant de preuve expérimentale. Ce lien entre acidité, OmpR, Fur et métabolisme du fer pourrait donc être commun aux entérobactéries.

E. *paeY* et *pemA* : une association récente avec *pelA/pelE/pelD* (résultats non publiés)

Le troisième article s'est concentré sur l'évolution des gènes *pelA/pelE/pelD*. Cependant, ces gènes sont régulièrement colocalisés avec deux gènes codant pour des enzymes remplissant une fonction complémentaire aux Pel dans la dégradation de la pectine : *paeY* et *pemA* (Figure 56). PaeY et PemA éliminent en effet les groupements acetyl- et methyl- de la pectine, ce qui facilite l'action des pectate lyases (Figure 29). La colocalisation de gènes remplissant des fonctions proches ou complémentaires est fréquente et est souvent le signe d'un îlot génomique acquis par transfert horizontal [19].

Cette hypothèse a donc été testée sur ces gènes en reconstruisant la phylogénie des protéines homologues à PaeY ou PemA de la même manière que pour les Pel (Figure 72 et Figure 73). Tout d'abord et contrairement aux protéines Pel, les protéines PaeY et PemA de *Dickeya* et de *Pectobacterium* sont groupées de manière robuste (BV = 74% et PP = 1.00 pour PaeY, BV = 96% et PP = 1.00 pour PemA), comme attendu d'après la proximité des deux espèces. Cependant, la phylogénie de PaeY et PemA est aussi incohérente avec la phylogénie des espèces, avec par exemple la présence de planctomycètes (*R. baltica* et *R. brasiliensis*) au milieu des protéobactéries ou la dispersion des clostridiales pour PaeY. De même, la présence de firmicutes (*Paenibacillus*, *Clostridium*) et de protéobactéries (*Alliglaciecola*) avec les bacteroides pour PemA indique également des transferts horizontaux.

De plus, ces transferts horizontaux semblent être indépendants, puisque les associations proches d'espèces varient entre les arbres. Par exemple, les *Marinomonas* sont groupés avec les firmicutes et les actinobactéries dans l'arbre Pel, alors qu'ils sont proches des entérobactéries dans les arbres PaeY et PemA. Les gènes *pel*, *paeY* et *pemA* semblent donc avoir suivi un chemin

évolutif distinct des gènes *pel* et ne pas faire partie d'un îlot génomique.

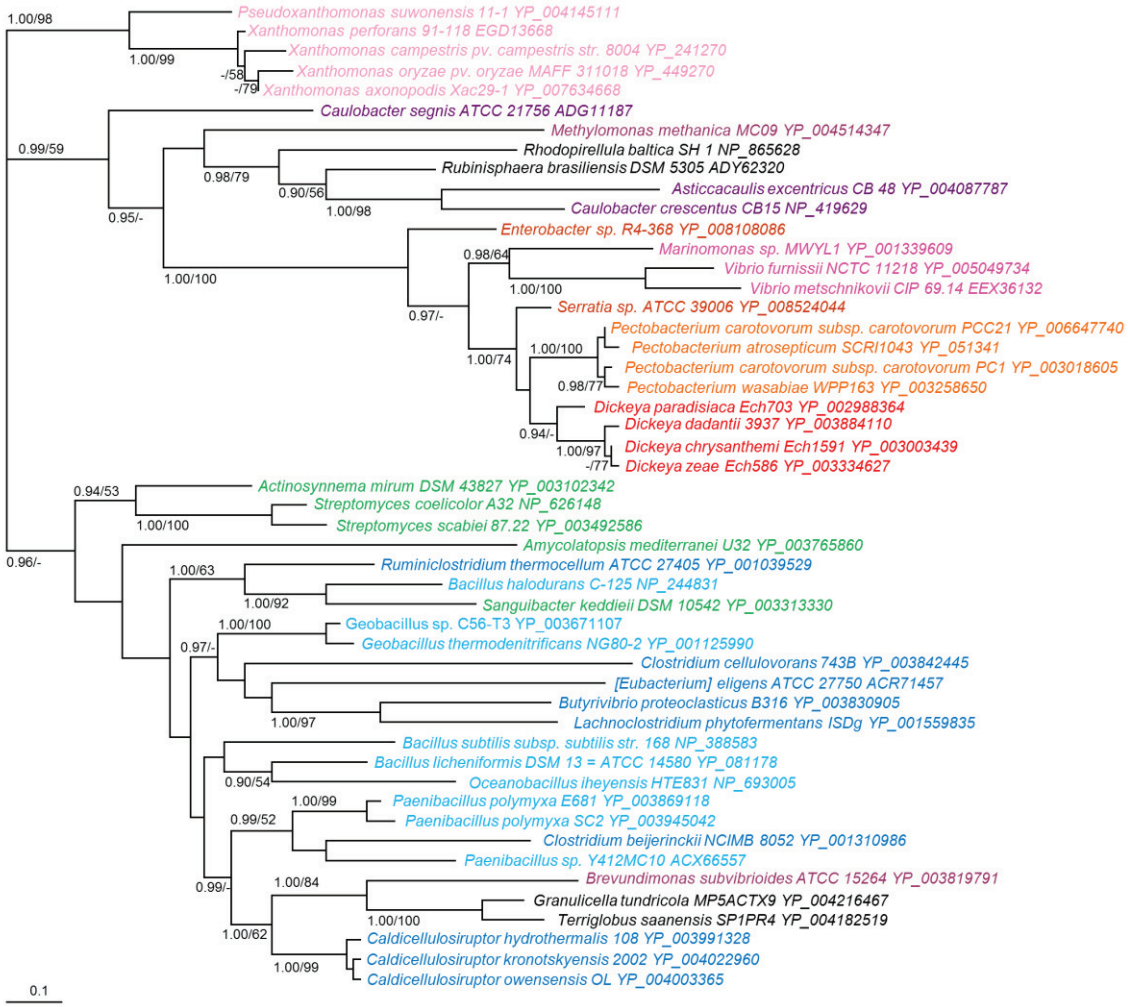


Figure 72 : Phylogénie bayésienne de 50 protéines PaeY (170 positions d'acides aminés)

L'arbre a été inféré avec MrBayes en utilisant un modèle d'évolution mixte et une distribution gamma avec 4 catégories (Γ4). La barre d'échelle représente le nombre moyen de substitutions par site. Le premier nombre (entre 0 et 1) correspond aux probabilités postérieures bayésiennes et le deuxième nombre (entre 0 et 100) aux valeurs de bootstrap d'une phylogénie par maximum de vraisemblance (ML) avec le modèle LG+Γ4 sur le même jeu de données (100 répliques). Un – est indiqué si la probabilité est inférieure à 0.9 ou si la valeur de bootstrap est inférieure à 50. Par clarté, rien n'est indiqué si la probabilité est inférieure à 0.9 et le bootstrap inférieur à 50. Le code couleur est identique à la Figure 55.

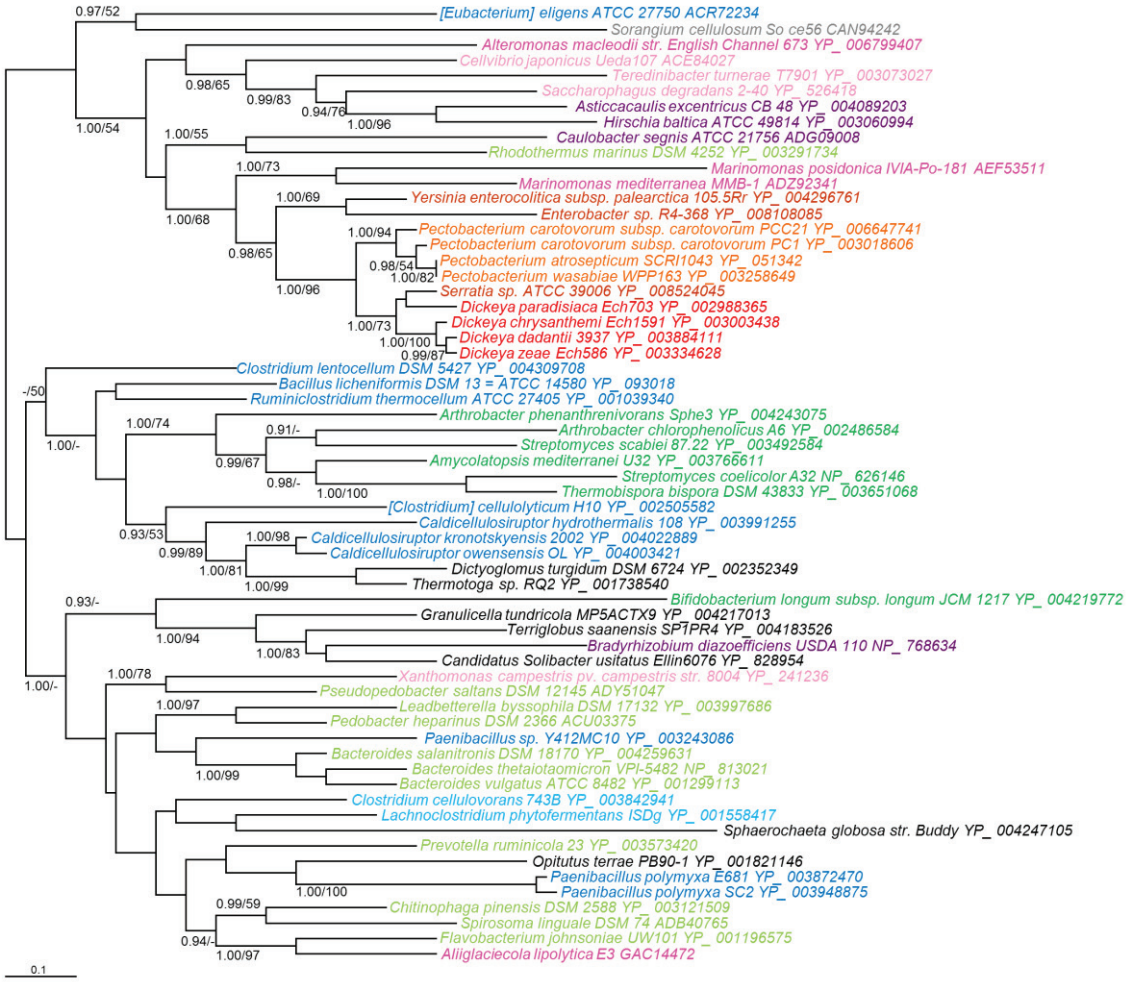


Figure 73 : Phylogénie bayésienne de 63 protéines PemA (214 positions d'acides aminés)
La légende est identique à la Figure 72.

Afin de clarifier ces observations, un arbre phylogénétique de toutes les espèces possédant au moins un homologue identifié de *pelA, pelE, pelD, paeY* ou *pemA* a été construit à partir des données de la littérature et la présence ainsi que la colocalisation des 5 gènes a été examinée pour chaque espèce (Figure 74). La duplication des gènes *pel* étant spécifique des *Dickeya*, elle n'a pas été considérée dans cette analyse à grande échelle.

Tout d'abord, on constate la présence fréquente d'un gène dans un embranchement donné. *pemA* est présent dans presque tous les génomes de bacteroides étudiés, alors que *pel* est fréquent chez les actinobactéries. Cette présence semble indiquer une acquisition ancienne de ces gènes dans ces embranchements. En revanche, chez les protéobactéries, la présence des trois gènes est sporadique, ce qui laisse supposer des transferts horizontaux fréquents et plus récents. De plus, deux genres (*Dickeya* et *Streptomyces*) possèdent une colocalisation des gènes *pel-paeY-pemA*. D'autres colocalisations sont observées à des endroits divers de l'arbre au sein de différents genres (*Pedobacter*, *Dicyoglomus*, *Actinosynnema*, *Marinomonas*, *Vibrio*…). La diversité taxonomique des espèces possédant de telles colocalisations de gènes impliqués dans une même voir métabolique suggère ainsi des évolutions convergentes répétées, ce qui implique que la colocalisation de ces gènes est évolutivement avantageuse.

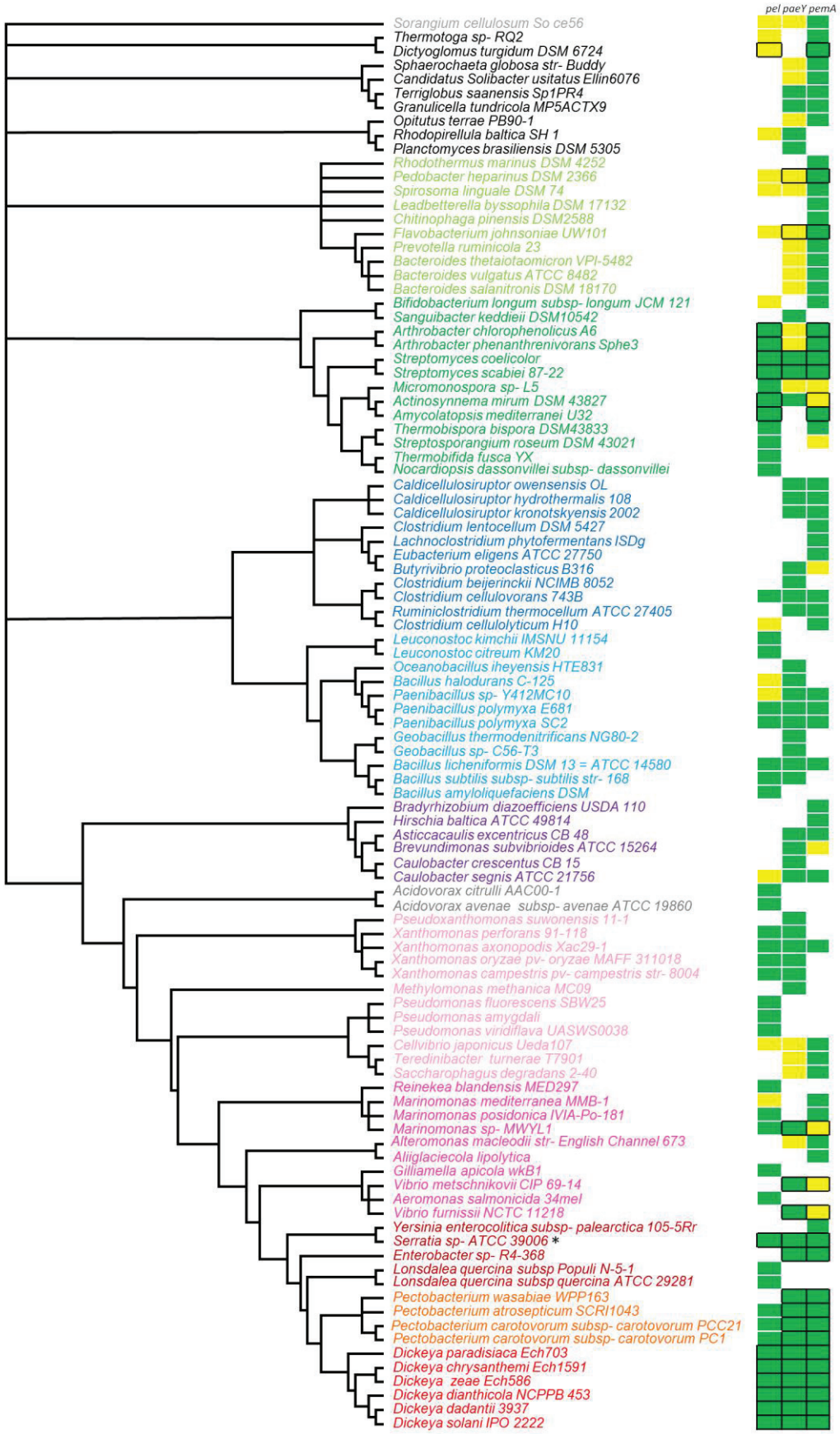

pel paeY pemA
Sorangium cellulosum So ce56
Thermotoga sp- RQ2
Dictyoglomus turgidum DSM 6724
Sphaerochaeta globosa str- Buddy
Candidatus Solibacter usitatus Ellin6076
Terriglobus saanensis Sp1PR4
Granulicella tundricola MP5ACTX9
Opitutus terrae PB90-1
Rhodopirellula baltica SH 1
Planctomyces brasiliensis DSM 5305
Rhodothermus marinus DSM 4252
Pedobacter heparinus DSM 2366
Spirosoma linguale DSM 74
Leadbetterella byssophila DSM 17132
Chitinophaga pinensis DSM2588
Flavobacterium johnsoniae UW101
Prevotella ruminicola 23
Bacteroides thetaiotaomicron VPI-5482
Bacteroides vulgatus ATCC 8482
Bacteroides salanitronis DSM 18170
Bifidobacterium longum subsp- longum JCM 121
Sanguibacter keddieii DSM10542
Arthrobacter chlorophenolicus A6
Arthrobacter phenanthrenivorans Sphe3
Streptomyces coelicolor
Streptomyces scabiei 87-22
Micromonospora sp- L5
Actinosynnema mirum DSM 43827
Amycolatopsis mediterranei U32
Thermobispora bispora DSM43833
Streptosporangium roseum DSM 43021
Thermobifida fusca YX
Nocardiopsis dassonvillei subsp- dassonvillei
Caldicellulosiruptor owensensis OL
Caldicellulosiruptor hydrothermalis 108
Caldicellulosiruptor kronotskyensis 2002
Clostridium lentocellum DSM 5427
Lachnoclostridium phytofermentans ISDg
Eubacterium eligens ATCC 27750
Butyrivibrio proteoclasticus B316
Clostridium beijerinckii NCIMB 8052
Clostridium cellulovorans 743B
Ruminiclostridium thermocellum ATCC 27405
Clostridium cellulolyticum H10
Leuconostoc kimchii IMSNU 11154
Leuconostoc citreum KM20
Oceanobacillus iheyensis HTE831
Bacillus halodurans C-125
Paenibacillus sp- Y412MC10
Paenibacillus polymyxa E681
Paenibacillus polymyxa SC2
Geobacillus thermodenitrificans NG80-2
Geobacillus sp- C56-T3
Bacillus licheniformis DSM 13 = ATCC 14580
Bacillus subtilis subsp- subtilis str- 168
Bacillus amyloliquefaciens DSM
Bradyrhizobium diazoefficiens USDA 110
Hirschia baltica ATCC 49814
Asticcacaulis excentricus CB 48
Brevundimonas subvibrioides ATCC 15264
Caulobacter crescentus CB 15
Caulobacter segnis ATCC 21756
Acidovorax citrulli AAC00-1
Acidovorax avenae subsp- avenae ATCC 19860
Pseudoxanthomonas suwonensis 11-1
Xanthomonas perforans 91-118
Xanthomonas axonopodis Xac29-1
Xanthomonas oryzae pv- oryzae MAFF 311018
Xanthomonas campestris pv- campestris str- 8004
Methylomonas methanica MC09
Pseudomonas fluorescens SBW25
Pseudomonas amygdali
Pseudomonas viridiflava UASWS0038
Cellvibrio japonicus Ueda107
Teredinibacter turnerae T7901
Saccharophagus degradans 2-40
Reinekea blandensis MED297
Marinomonas mediterranea MMB-1
Marinomonas posidonica IVIA-Po-181
Marinomonas sp- MWYL1
Alteromonas macleodii str- English Channel 673
Aliiglaciecola lipolytica
Gilliamella apicola wkB1
Vibrio metschnikovii CIP 69-14
Aeromonas salmonicida 34mel
Vibrio furnissii NCTC 11218
Yersinia enterocolitica subsp- palearctica 105-5Rr
Serratia sp- ATCC 39006 *
Enterobacter sp- R4-368
Lonsdalea quercina subsp Populi N-5-1
Lonsdalea quercina subsp quercina ATCC 29281
Pectobacterium wasabiae WPP163
Pectobacterium atrosepticum SCRI1043
Pectobacterium carotovorum subsp- carotovorum PCC21
Pectobacterium carotovorum subsp- carotovorum PC1
Dickeya paradisiaca Ech703
Dickeya chrysanthemi Ech1591
Dickeya zeae Ech586
Dickeya dianthicola NCPPB 453
Dickeya dadantii 3937
Dickeya solani IPO 2222

Figure 74 : Cladogramme des 95 bactéries possédant au moins un gène homologue à *pelA*, *pelE*, *pelD*, *paeY* ou *pemA* présenté sur les arbres précédents.

L'arbre a été construit à partir des données de la littérature [89,296–302]. Les carrés blancs ou colorés sur la droite indiquent la présence et l'absence d'un gène, respectivement. Une couleur jaune indique la présence d'un homologue distant non représenté sur les arbres précédents alors qu'une couleur verte correspond à un homologue présent sur ces arbres. Des bordures noires correspondent à des gènes colocalisés dans le génome correspondant. Les taxa sont colorés comme précédemment. * : La souche ATCC 39006 est ici placée selon sa taxonomie officielle qui la désigne comme un *Serratia*. Cependant, sa position est discutable et plusieurs observations suggèrent qu'il pourrait s'agit d'un *Dickeya* (voir section F).

F. Deux nouveaux *Dickeya* pour explorer l'acquisition des gènes *pel* : *D. aquatica* et ATCC 39006 (résultats non publiés)

Durant la thèse, une nouvelle espèce de *Dickeya* a été décrite : *D. aquatica* [225]. Cette espèce a été isolée à partir de l'eau de rivière et n'avait jamais été retrouvée associée à une plante. De plus, sa position phylogénétique indiquerait que cette espèce de *Dickeya* a divergé tôt, voire même avant *D. paradisiaca*. *D. aquatica* pouvait donc être un *Dickeya* n'ayant pas encore acquis des capacités d'infection de plante.

Afin de vérifier cette hypothèse, la virulence de la souche type de *D. aquatica* a été évaluée sur trois modèles végétaux courants : endive, pomme de terre et tomate (Table 9 et Figure 75). Une faible virulence sur pomme de terre et endive a été constatée, avec une infection se limitant au point d'entrée. En revanche, sur tomate, *D. aquatica* s'est montrée beaucoup plus efficace que *D. dadantii*, causant une infection généralisée et un éclatement du fruit après 42 h. Les défauts de virulence chez l'endive et la pomme de terre ne peuvent pas s'expliquer par un défaut de production de Pel, puisque l'activité Pel de *D. aquatica* est comparable et même légèrement plus précoce que celle de *D. dadantii* (Figure 76). *D. aquatica*, au moins en conditions de laboratoire, est donc pathogène mais semble avoir un spectre d'hôtes différent de celui de *D. dadantii*.

Table 9 : Masses de pourriture induite par *D. dadantii* ou *D. aquatica* lors de l'infection de différents végétaux

	D. dadantii	*D. aquatica*
Endive (feuille)	0.9 ± 0.5	0.02 ± 0.04
Pomme de terre (tubercule)	5.7 ± 2.4	1.7 ± 0.2
Tomate (fruit)	8.0 ± 4.6	36 ± 16

Les valeurs sont données en moyenne ± écart-type de 6 répliques par conditions, en g.

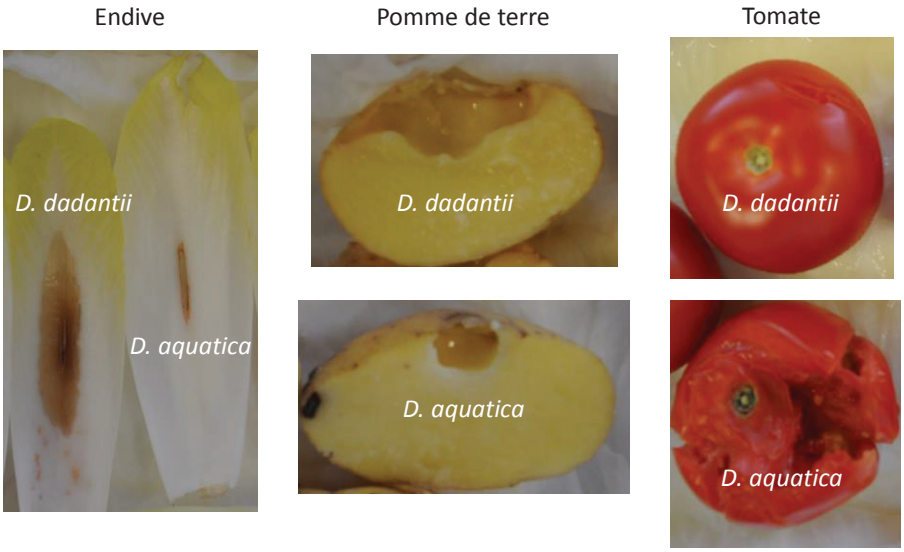


Figure 75 : Photographies des plantes infectées par *D. dadantii* et *D. aquatica*.
Les végétaux ont été infectés par *D. dadantii* ou *D. aquatica* comme décrit dans la section II-F-3 et les images prises après 18 h (endive) ou 42 h (pomme de terre et tomate) d'infection. Pour la pomme de terre, la zone pourrie a été prélevée avant de prendre la photographie.

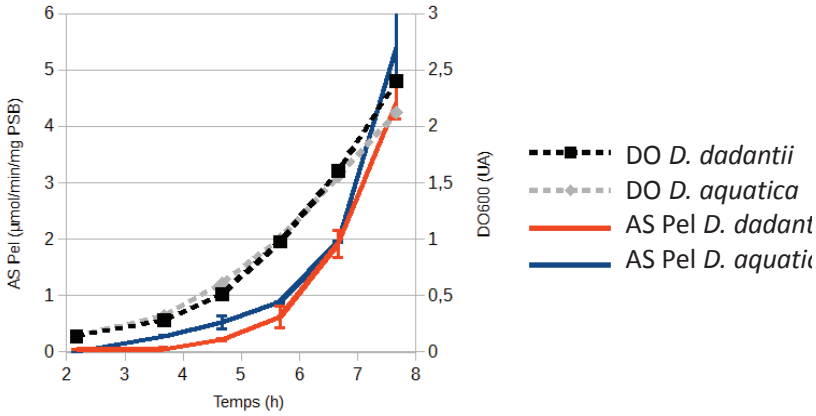


Figure 76 : Comparaison de l'activité pectate lyase précoce de *D. dadantii* et *D. aquatica*.
D. dadantii et *D. aquatica* ont été cultivés à 30°C en M63S + PGA et l'activité Pel a été suivie au cours du temps. Les graphiques représentent la moyenne de 2 répliques biologiques et les barres d'erreur représentent l'écart-type.

Comme discuté dans la section I-B, les Dickeya sont des pathogènes en général peu spécifiques qui utilisent des facteurs d'agression efficaces contre un large spectre d'hôtes. Les spécificités d'hôtes peuvent en revanche découler de différentes capacités de résistance aux stress des espèces de Dickeya [303]. La croissance de *D. aquatica* et *D. dadantii* a donc été comparée en présence des principaux stress rencontrés dans une plante (acide, osmotique et oxydatif), ainsi qu'en fonction de la température (Figure 77). Aucune différence majeure n'a été observée pour la réponse à la température, uniquement un léger défaut de croissance à 20°C pour *D. aquatica*. En revanche, la réponse aux stress est très différente. *D. aquatica* est plus sensible que *D. dadantii* aux stress osmotiques et oxydants, avec notamment une incapacité totale à pousser à 0.3 M NaCl. La croissance de *D. dadantii* et *D. aquatica* est comparable avec un stress acide à pH 5 mais le stress induit des modifications morphorlogiques de colonies chez *D.*

aquatica. Elles deviennent plus étendues (augmentation de diamètre d'un facteur 2 à 3) avec des formes rappelant une mobilité par saccades (twitching) ou glissement (sliding) [304]. *D. aquatica* est en revanche incapable de former des biofilms en conditions de culture usuelles et souffre d'un retard de croissance majeur par rapport à *D. dadantii* dans ces conditions de faible agitation (Figure 78).

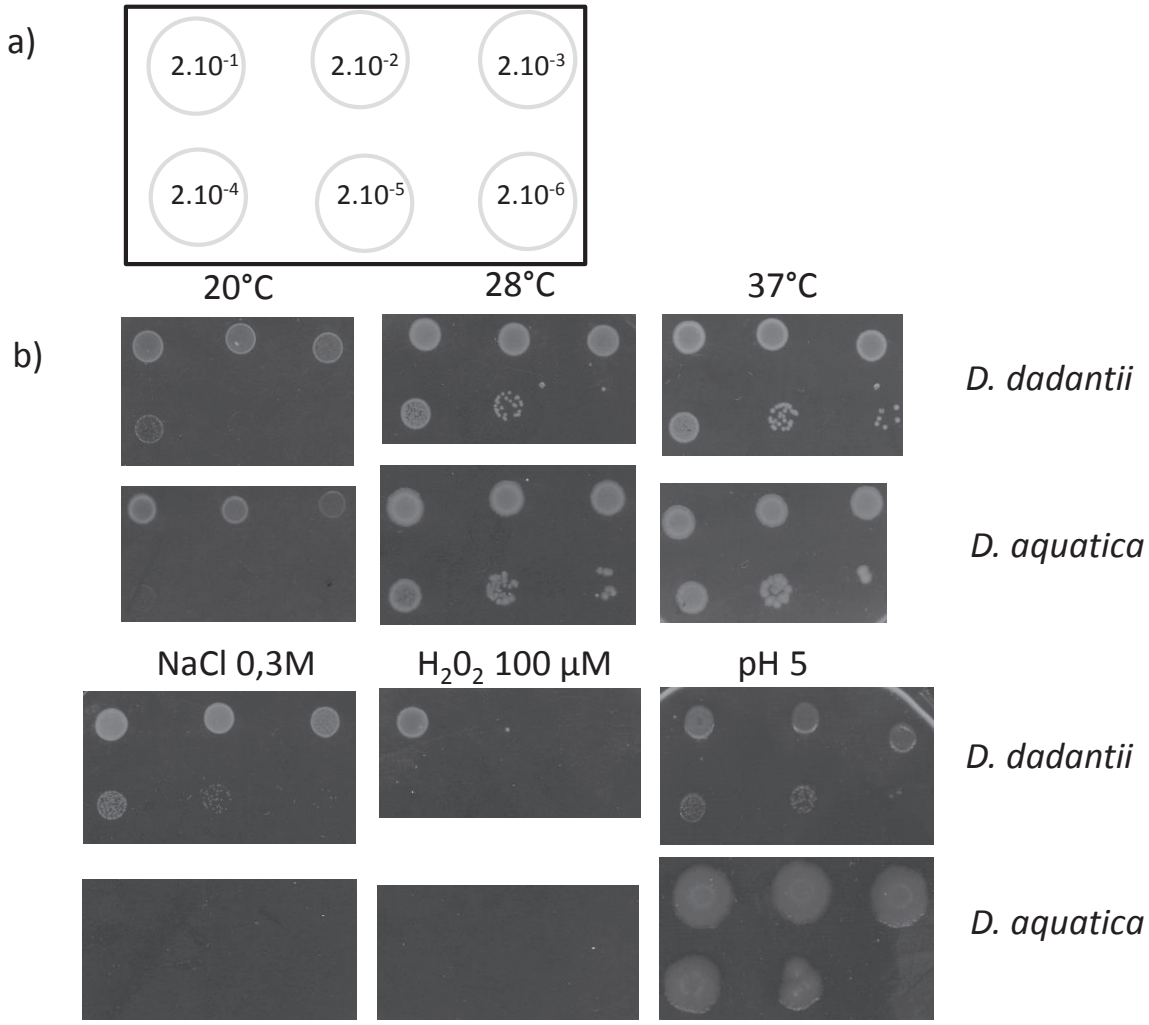


Figure 77 : Comparaison de la résistance aux stress de *D. dadantii* et *D. aquatica*

5 µL d'une suspension bactérienne à DO_{600} donnée (a) ont été déposés sur divers milieux gélosés M63G supplémentés éventuellement de NaCl (stress osmotique), H_2O_2 (stress oxydant) ou acide malique (stress acide) (b). Les boîtes avec stress ont été incubées à 28°C, les boîtes sans stress ont été incubées à la température indiquée. Le temps d'incubation est de 24h.

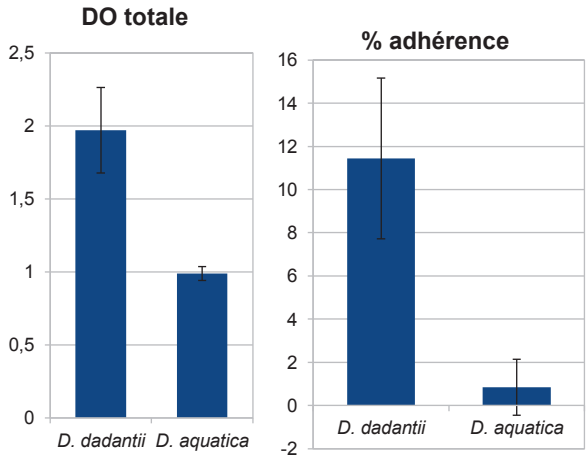


Figure 78 : Quantification de la formation de biofilms chez *D. dadantii* et *D. aquatica*
Les bactéries ont été cultivées en plaque 24 puits sans agitation en M63S. Le biofilm a été séparé du surnageant, rincé et la quantité de bactéries présentes dans chaque fraction a été quantifiée par mesure de la turbidité à 600 nm.

D. aquatica possède donc plusieurs caractéristiques intrigantes, notamment le twitching, l'hypersensibilité au stress osmotique et l'absence de biofilms, qui pourraient expliquer son spectre d'hôtes différent. D'autres facteurs peuvent également intervenir, comme par exemple l'absence d'un ou plusieurs gènes *pel* essentiels à la virulence dans certains hôtes seulement [237]. Aucun génome n'était disponible pour cette espèce, ce qui limitait fortement la compréhension des phénotypes observés. Nous avons donc entrepris le séquençage de la souche type de *D. aquatica* (174/2).

Le génome complet de *D. aquatica* consiste en un seul chromosome circulaire de 4 501 343 pb à un taux de G/C de 53.6 %. Une annotation automatique via RAST indique la présence de 4412 gènes et 98 ARN. Toutes ces valeurs sont habituelles pour les génomes de *Dickeya*.

La comparaison des gènes par famille a donné quelques points distinctifs. Tout d'abord, *D. aquatica* ne possède pas de gène *pelE* (uniquement *pelA* et *pelD*). En revanche, il possède le gène *pel10*, commun avec *D. chrysanthemi* (qui ne possède pas de gène *pelE* non plus). Ensuite, à l'inverse de tous les *Dickeya* sauf *D. paradisiaca*, il ne possède pas de système de sécrétion de type VI, mais comporte un système de sécrétion de type VII unique chez les *Dickeya*. Le twitching observé à pH 5 est une forme de mobilité qui est causée par des élongations et contractions successives d'un pilus de type IV. Or, le système de sécrétion de type VII unique chez *D. aquatica* est un homologue du système de pilus de type I Fim de *E. coli*, connu pour être un antagoniste du twitching [305]. Le pilus de type IV possédé par *D. aquatica* (tout comme *D. dadantii*) est homologue de PppD, un système fonctionnel mais non exprimé chez *E. coli* [306]. On peut alors supposer que, chez *D. aquatica*, *pppD* est exprimé mais que l'acquisition de *fim* permet de réprimer le twitching en conditions de pH neutre. Une délétion de *pppD* et/ou *fimA* chez *D. aquatica* serait nécessaire pour tester cette hypothèse. Alternativement, si la mobilité observée est due à un glissement, il pourrait alors s'agir d'un mécanisme passif ne nécessitant pas de pilus [307].

Dans le 3ème article, une souche de *Serratia* (ATCC 39006) se trouvait placée avec les *Dickeya* dans les phylogénies Pel, et à nouveau pour les phylogénies PaeY et PemA. Cette position est inattendue considérant l'éloignement phylogénétique relatif de *Dickeya* et *Serratia* comparé à

Dickeya et *Pectobacterium* par exemple. De plus, le génome de cette souche est séquencé et présente des anomalies montrées dans la Table 10. Cette souche est en effet le seul *Serratia* séquencé à posséder les gènes du locus *pelA-pelE-pelD-paeY-pemA* (avec un seul exemplaire du gène *pel*), mais également des gènes situés à d'autres loci comme *pelB-pelC* (qui sont spécifiques de *Dickeya* et *Pectobacterium*) ou *vfm* (qui est strictement spécifique de *Dickeya*). Ces observations laissent penser à une erreur de classification. En effet, la taxonomie de cette souche a été établie dans l'ère pré-génomique sur trois critères phénotypiques : production de prodigiosine, résistance au stress osmotique (4% m/v NaCl maximum, soit 0.7 M, sachant que *D. dadantii* peut pousser jusqu'à 0.5 M NaCl par exemple) et mésophilie (température optimale 30°C) [308]. La mésophilie et l'halotolérance sont compatibles avec une souche de *Dickeya* et l'acquisition des gènes de biosynthèse de prodigiosine pourrait résulter d'un transfert horizontal.

Table 10 : Présence ou absence de gènes caractéristiques de *Dickeya* et *Serratia* dans la souche ATCC 39006

Gène	*Dickeya*	ATCC 39006 = SC 11482	*Serratia*
pelA/E/D	+	+	-
pelB/C	+	+	-
paeY-pemA	+	+	-
vfm	+ (sauf *D. paradisiaca*)	+	-
Prodigiosine	-	+	+

+ indique la présence du gène, - son absence.

Afin de confirmer ces observations, une phylogénie des espèces appartenant aux genres *Dickeya*, *Pectobacterium* et *Serratia* a été construite sur la base des séquences des protéines ribosomales concaténées (Figure 79). Comme attendu d'après les données publiées, *D. aquatica* se place près de la racine des *Dickeya*. En revanche, sa divergence est postérieure à celle de *D. paradisiaca* (BV = 100%). ATCC 39006, quant à elle, ne fait clairement pas partie des *Serratia* (BV = 100%) et a divergé entre la divergence *Pectobacterium-Dickeya* et la divergence de *D. paradisiaca*. On peut donc s'attendre à retrouver des caractères ancestraux des *Dickeya* chez cette souche. On remarque également un groupe de *Dickeya* sans nom d'espèce (MK7 et NCPPB 3274) distincts des espèces connues et qui pourraient représenter une autre espèce de *Dickeya*.

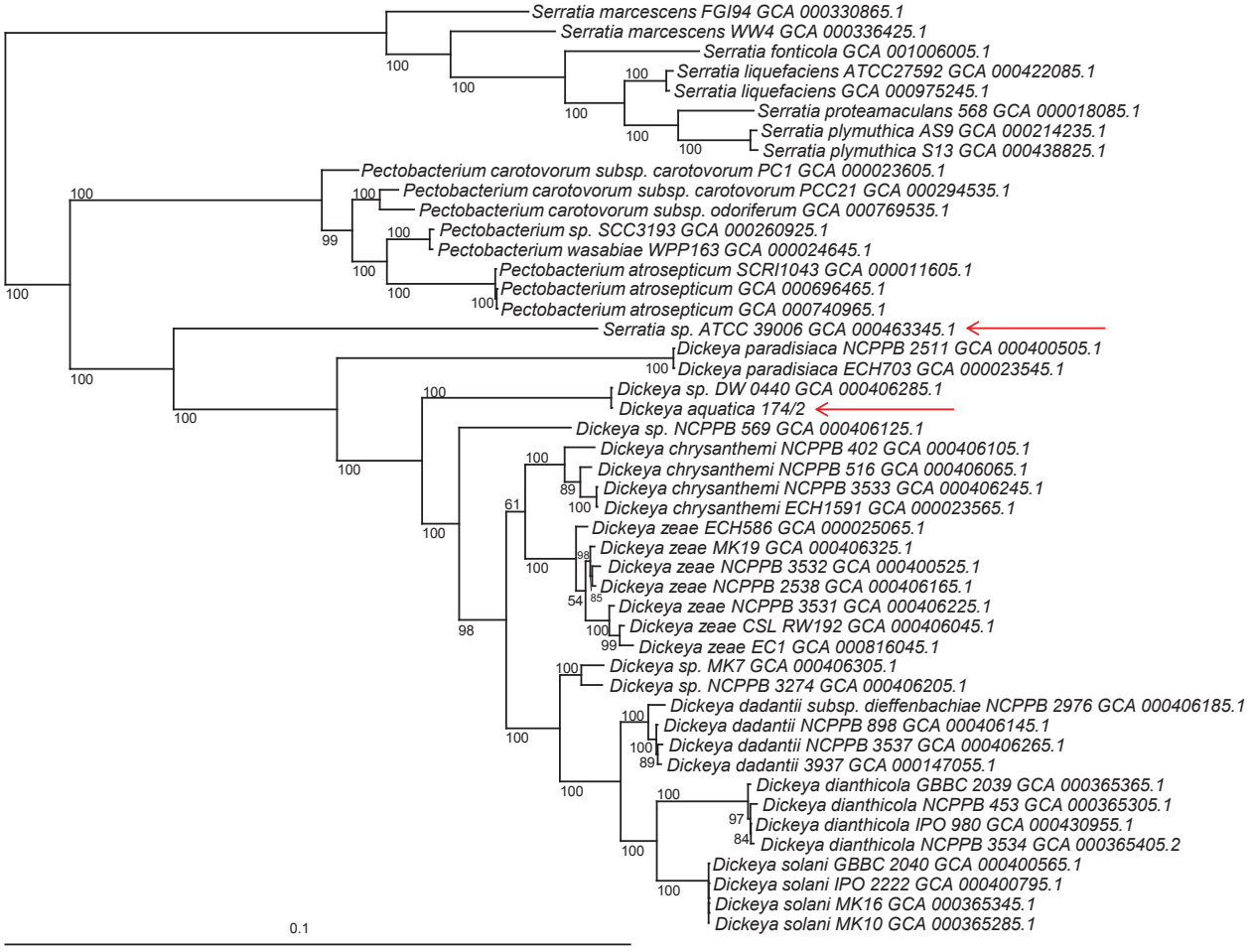


Figure 79 : Phylogénie des Dickeya, Serratia et pectobacterium établie sur la base des séquences des gènes codant pour les protéines ribosomales concaténées (57 protéines ribosomales pour 47 espèces)

L'arbre a été construit par IqTree avec le modèle GTR. La barre d'échelle représente 0.1 substitutions/site. Les chiffres correspondent aux valeurs de bootstrap (100 répliques du jeu de données original). Les flèches rouges indiquent la position des deux souches discutées dans le texte.

La structure des régions régulatrices *pelD/pelE* de *D. aquatica* et ATCC 39006 a alors été examinée. *D. aquatica* possède un seul gène *pelD* différencié (ainsi qu'un exemplaire de *pelA*), alors que ATCC 39006 possède un seul gène *pelAED* non différencié (Figure 56). La région régulatrice *pelD* de *D. aquatica* possède les caractéristiques typiques d'un gène *pelD* (*div*, CRP1, promoteur à une position de type *pelD*) à quelques exceptions près. Le site FIS1bis est absent, ce qui n'a probablement pas de conséquences régulatrices majeures (voir 2ème article), et un site FIS chevauche CRP1, ce qui pourrait introduire une compétition FIS-CRP au niveau de ce site, contrairement à ce qui est observé chez *D. dadantii* (2ème article). La région régulatrice *pel* de ATCC 39006 est très différent des régions régulatrices *pelD* et *pelE*. En effet, seuls CRP1 et KdgR sont conservés, ce qui confirme leur origine ancienne comme discuté dans le 3ème article. Il y a notamment un indel au niveau de la position habituelle des promoteurs *pelD*, ce qui cause un déplacement du promoteur prédit au niveau du site CRP1. Dans cette position, CRP est probablement répresseur de la transcription. Il est impossible de savoir si cette situation reflète l'état ancestral du gène ou si elle résulte d'adaptations spécifiques chez ATCC 39006.

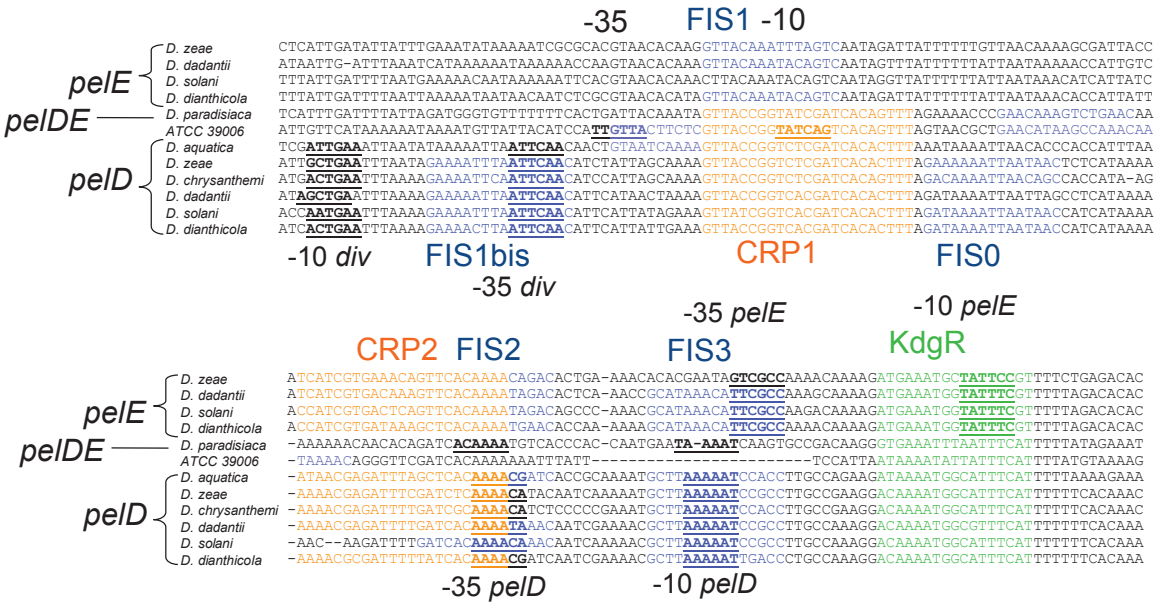


Figure 80 : Structures des régions régulatrices des gènes *pelE* et *pelD* incluant ATCC 39006 et *D. aquatica*

Les régions régulatrices des gènes *pelD-pelE* de *D. aquatica* et ATCC 39006 ont été ajoutées à la Figure 60. Le code couleur est identique.

IV. Discussion générale

La régulation transcriptionnelle de l'expression des gènes était connue pour être indispensable à l'acclimatation des bactéries à un environnement variable. Elle est flexible, hiérarchisée et multifactorielle, ce qui permet des ajustements fins de l'expression des gènes en fonction d'un grand nombre de paramètres externes et internes à la bactérie. Cependant, une bactérie possède des gènes d'origines diverses (ancestraux, dupliqués, acquis par transfert horizontal…) dont les expressions doivent être coordonnées.

Nous avons montré que la régulation transcriptionnelle participe à l'acquisition de la fonction d'un gène et peut même avoir plus d'importance que la fonction de la protéine pour déterminer la conservation ou l'élimination du gène. *pelD*, d'un côté, est régulé par divers mécanismes complémentaires (effet direct de FIS, compétition FIS-CRP, dépendance simultanée des deux sites CRP et activation par *div*) pour assurer à la fois une expression quasi-nulle durant les phases précoces d'infection et une expression maximale durant les phases tardives. Le modèle de régulation de *pelD* est cependant pour l'instant construit à partir de mesures biochimiques et génétiques. Ces résultats pourraient être vérifiés par des méthodes biophysiques telle que la microscopie à force atomique pour visualiser entre autres la formation du coude induit par CRP (Figure 54) ou la fixation simultanée de l'ARNP sur *pelD* et *div* (Figure 53). De plus, une validation in vivo de la fixation des différents régulateurs sur *pelD* au cours de la croissance est envisageable afin d'avoir une idée plus précise de la dynamique de la formation du complexe nucléoprotéique au niveau de la région régulatrice de ce gène.

A l'inverse, *pelE* s'affranchit de cette régulation en gardant un nombre minimum de sites FIS et CRP communs avec *pelD* (2 sites FIS, 1 site CRP) et en repositionnant son promoteur pour

diminuer l'efficacité de KdgR, le régulateur responsable de l'induction par les composés pectiques. Ces modifications, malgré la proximité biochimique des deux enzymes, permet à PelE et PelD de remplir des rôles distincts : initier la dégradation de la pectine et la poursuivre, respectivement. Une étude mécanistique de la régulation de *pelE* similaire à celle de *pelD* pourrait permettre de clarifier le rôle des sites de fixation non conservés avec *pelD*, notamment l'amas FIS4/5/6 situé en aval du promoteur et les sites OmpR1 et OmpR3, afin d'avoir une vision intégrative des innovations régulatrices spécifiques à la région régulatrice *pelE*.

Etrangement, deux espèces de *Dickeya* (*D. chrysanthemi* et *D. aquatica*) ont perdu *pelE* après son acquisition, en plus de *D. paradisiaca* et ATCC 39006 qui utilisent un gène *pel* non différencié. Malgré l'importance de *pelE* pour l'initiation de la dégradation de la pectine, *D. chrysanthemi*, *D. aquatica* et *D. paradisiaca* sont virulents. Utilisent-ils un autre gène *pel* spécialisé dans l'initiation, par exemple le gène *pel10* ? Une somme d'activité basale faible de l'ensemble des gènes *pel* ? A l'inverse, *D. solani* semble avoir compensé une diminution d'expression de *pelD* par une augmentation de celle de *pelE*. Est-ce que cela reflète une infection plus rapide ? Des études complémentaires similaires à celles réalisées dans l'article 3 associés à des tests d'infection en cinétique pourraient fournir des informations pertinentes pour répondre à ces questions.

Bien que ces observations de divergence de gènes induits par le décalage de TSS semblent être applicables à l'ensemble du règne bactérien, les règles générales gouvernant ces divergences sont encore inconnues. Des études chez d'autres modèles pourraient permettre de dégager des règles générales sur ce mécanisme qui jusqu'à présent ne semblait être présent que chez les eucaryotes. De la même manière, malgré la présence fréquente de promoteurs multiples sur les régions régulatrices bactériennes (pour la majorité des espèces, environ 90% des régions régulatrices portent plus d'un promoteur [63]), leur fonction est mal comprise. L'étude de ces systèmes à promoteurs multiples pourrait permettre de dégager des règles nouvelles sur leur implication dans la régulation transcriptionnelle.

De manière générale, l'évolution des gènes via leur régulation transcriptionnelle pourrait être soumise à une logique très différente de l'évolution des protéines. Beaucoup de mutations dans les protéines sont létales, notamment les fréquentes disruptions de cadre de lecture ou apparitions d'un codon stop qui produisent une protéine tronquée. Une mutation régulatrice a rarement un effet aussi marqué puisque le « bruit » transcriptionnel peut permettre une transcription faible du gène considéré. De plus, un gène devenu silencieux suite à une telle mutation n'est pas forcément éliminé (*pppD* chez *E. coli* par exemple). La bactérie peut alors compenser en modulant l'expression d'autres gènes qui peuvent restaurer la fonction perdue. A l'inverse, la redondance du code génétique permet des mutations parfaitement neutres. La régulation transcriptionnelle repose principalement sur une zone d'environ 100 bp autour du promoteur où chaque nucléotide a son importance, que ce soit pour définir l'affinité d'un régulateur pour son site de fixation, les propriétés physiques qui gouvernent le surenroulement de l'ADN, l'existence d'un promoteur secondaire… Les règles gouvernant l'évolution de la régulation transcriptionnelle pourraient ainsi être très différentes de ce à quoi les mutations protéiques nous ont habitué.

L'interaction entre mutations protéiques et régulatrices se pose également. A quel point chaque type de mutation (protéine ou régulatrice) contribue-t-il à l'émergence de nouvelles fonctions ? Comment sont-elles hiérarchisées ? Sont-elles mutuellement exclusives et utilisées dans des situations différentes (variation rapide ou graduelle de l'environnement, différents types de stress…), ou un type de mutation arrive-t-il avant l'autre ? Avec l'augmentation des

connaissances sur la régulation des gènes, la cartographie de plus en plus précise et exhaustive des sites d'initiation de la transcription associés ou non à un gène et le développement de nouveaux modèles d'étude bactériens, il est probable que ces questions deviennent centrales pour les travaux futurs sur l'évolution des fonctions des bactéries.

V. Références

1. Gil R, Silva FJ, Peretó J, Moya A. Determination of the Core of a Minimal Bacterial Gene Set. Microbiol Mol Biol Rev. 2004;68: 518–537. doi:10.1128/MMBR.68.3.518-537.2004

2. Glick BR. Modulation of plant ethylene levels by the bacterial enzyme ACC deaminase. FEMS Microbiol Lett. 2005;251: 1–7. doi:10.1016/j.femsle.2005.07.030

3. Waters LS, Storz G. Regulatory RNAs in Bacteria. Cell. 2009;136: 615–628. doi:10.1016/j.cell.2009.01.043

4. Walsh CT, Garneau-Tsodikova S, Gatto GJ. Protein Posttranslational Modifications: The Chemistry of Proteome Diversifications. Angew Chem Int Ed. 2005;44: 7342–7372. doi:10.1002/anie.200501023

5. Macek B, Mijakovic I, Olsen JV, Gnad F, Kumar C, Jensen PR, et al. The Serine/Threonine/Tyrosine Phosphoproteome of the Model Bacterium Bacillus subtilis. Mol Cell Proteomics. 2007;6: 697–707. doi:10.1074/mcp.M600464-MCP200

6. Haugen SP, Ross W, Gourse RL. Advances in bacterial promoter recognition and its control by factors that do not bind DNA. Nat Rev Microbiol. 2008;6: 507–519. doi:10.1038/nrmicro1912

7. Ross W, Aiyar SE, Salomon J, Gourse RL. Escherichia coli Promoters with UP Elements of Different Strengths: Modular Structure of Bacterial Promoters. J Bacteriol. 1998;180: 5375–5383.

8. Russell DR, Bennett GN. Construction and analysis of in vivo activity of E. coll promoter hybrids and promoter mutants that alter the −35 to −10 spacing. Gene. 1982;20: 231–243. doi:10.1016/0378-1119(82)90042-7

9. Murakami KS, Darst SA. Bacterial RNA polymerases: the wholo story. Curr Opin Struct Biol. 2003;13: 31–39. doi:10.1016/S0959-440X(02)00005-2

10. Vassylyev DG, Sekine S, Laptenko O, Lee J, Vassylyeva MN, Borukhov S, et al. Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 \AA resolution. Nature. 2002;417: 712–719.

11. Mekler V, Kortkhonjia E, Mukhopadhyay J, Knight J, Revyakin A, Kapanidis AN, et al. Structural Organization of Bacterial RNA Polymerase Holoenzyme and the RNA Polymerase-Promoter Open Complex. Cell. 2002;108: 599–614. doi:10.1016/S0092-8674(02)00667-0

12. Zhang G, Campbell EA, Minakhin L, Richter C, Severinov K, Darst SA. Crystal Structure of Thermus aquaticus Core RNA Polymerase at 3.3 Å Resolution. Cell. 1999;98: 811–824. doi:10.1016/S0092-8674(00)81515-9

13. Gaal T, Ross W, Blatter EE, Tang H, Jia X, Krishnan VV, et al. DNA-binding determinants of the alpha subunit of RNA polymerase: novel DNA-binding domain architecture. Genes Dev. 1996;10: 16–26.

14. Gruber TM, Gross CA. Multiple Sigma Subunits and the Partitioning of Bacterial Transcription

Space. Annu Rev Microbiol. 2003;57: 441–466. doi:10.1146/annurev.micro.57.030502.090913

15. Dong T, Yu R, Schellhorn H. Antagonistic regulation of motility and transcriptome expression by RpoN and RpoS in Escherichia coli. Mol Microbiol. 2011;79: 375–386. doi:10.1111/j.1365-2958.2010.07449.x

16. Barrios H, Valderrama B, Morett E. Compilation and analysis of sigma(54)-dependent promoter sequences. Nucleic Acids Res. 1999;27: 4305–4313.

17. Caspi R, Altman T, Billington R, Dreher K, Foerster H, Fulcher CA, et al. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases. Nucleic Acids Res. 2014;42: D459–D471. doi:10.1093/nar/gkt1103

18. Yang S, Peng Q, Zhang Q, Zou L, Li Y, Robert C, et al. Genome-Wide Identification of HrpL-Regulated Genes in the Necrotrophic Phytopathogen Dickeya dadantii 3937. Butler G, editor. PLoS ONE. 2010;5: e13472. doi:10.1371/journal.pone.0013472

19. Hacker J, Kaper JB. Pathogenicity Islands and the Evolution of Microbes. Annu Rev Microbiol. 2000;54: 641–679. doi:10.1146/annurev.micro.54.1.641

20. Helmann JD. Anti-sigma factors. Curr Opin Microbiol. 1999;2: 135–141. doi:10.1016/S1369-5274(99)80024-1

21. Wendrich TM, Blaha G, Wilson DN, Marahiel MA, Nierhaus KH. Dissection of the Mechanism for the Stringent Factor RelA. Mol Cell. 2002;10: 779–788. doi:10.1016/S1097-2765(02)00656-1

22. Srivatsan A, Wang JD. Control of bacterial transcription, translation and replication by (p)ppGpp. Curr Opin Microbiol. 2008;11: 100–105. doi:10.1016/j.mib.2008.02.001

23. Jishage M, Kvint K, Shingler V, Nyström T. Regulation of ς factor competition by the alarmone ppGpp. Genes Dev. 2002;16: 1260–1270. doi:10.1101/gad.227902

24. Pérez-Rueda E, Collado-Vides J. The repertoire of DNA-binding transcriptional regulators in Escherichia coli K-12. Nucleic Acids Res. 2000;28: 1838–1847. doi:10.1093/nar/28.8.1838

25. Salgado H, Peralta-Gil M, Gama-Castro S, Santos-Zavaleta A, Muñiz-Rascado L, García-Sotelo JS, et al. RegulonDB v8.0: omics data sets, evolutionary conservation, regulatory phrases, cross-validated gold standards and more. Nucleic Acids Res. 2013;41: D203-213. doi:10.1093/nar/gks1201

26. Martínez-Antonio A, Collado-Vides J. Identifying global regulators in transcriptional regulatory networks in bacteria. Curr Opin Microbiol. 2003;6: 482–489. doi:10.1016/j.mib.2003.09.002

27. Browning DF, Busby SJW. The regulation of bacterial transcription initiation. Nat Rev Microbiol. 2004;2: 57–65. doi:10.1038/nrmicro787

28. Busby S, Ebright RH. Transcription activation by catabolite activator protein (CAP). J Mol Biol. 1999;293: 199–213. doi:10.1006/jmbi.1999.3161

29. Rojo F. Mechanisms of transcriptional repression. Curr Opin Microbiol. 2001;4: 145–151. doi:10.1016/S1369-5274(00)00180-6

30. Escolar L, Pérez-Martín J, Lorenzo V de. Evidence of an Unusually Long Operator for the Fur Repressor in the Aerobactin Promoter of Escherichia coli. J Biol Chem. 2000;275: 24709–24714. doi:10.1074/jbc.M002839200

31. Bouffartigues E, Buckle M, Badaut C, Travers A, Rimsky S. H-NS cooperative binding to high-affinity sites in a regulatory element results in transcriptional silencing. Nat Struct Mol Biol. 2007;14: 441–448. doi:10.1038/nsmb1233

32. Wilkinson SP, Grove A. Ligand-responsive transcriptional regulation by members of the MarR family of winged helix proteins. Curr Issues Mol Biol. 2006;8: 51.

33. Camilli A, Bassler BL. Bacterial Small-Molecule Signaling Pathways. Science. 2006;311: 1113–1116. doi:10.1126/science.1121357

34. Fleischhacker AS, Kiley PJ. Iron-containing transcription factors and their roles as sensors. Curr Opin Chem Biol. 2011;15: 335–341. doi:10.1016/j.cbpa.2011.01.006

35. Wilson CJ, Zhan H, Swint-Kruse L, Matthews KS. The lactose repressor system: paradigms for regulation, allosteric behavior and protein folding. Cell Mol Life Sci. 2006;64: 3–16. doi:10.1007/s00018-006-6296-z

36. Randall GL, Zechiedrich L, Pettitt BM. In the absence of writhe, DNA relieves torsional stress with localized, sequence-dependent structural failure to preserve B-form. Nucleic Acids Res. 2009; gkp556. doi:10.1093/nar/gkp556

37. Gilbert N, Allan J. Supercoiling in DNA and chromatin. Curr Opin Genet Dev. 2014;25: 15–21. doi:10.1016/j.gde.2013.10.013

38. Travers A, Muskhelishvili G. DNA supercoiling — a global transcriptional regulator for enterobacterial growth? Nat Rev Microbiol. 2005;3: 157–169. doi:10.1038/nrmicro1088

39. Witz G, Stasiak A. DNA supercoiling and its role in DNA decatenation and unknotting. Nucleic Acids Res. 2010;38: 2119–2133. doi:10.1093/nar/gkp1161

40. Forterre P, Gribaldo S, Gadelle D, Serre M-C. Origin and evolution of DNA topoisomerases. Biochimie. 2007;89: 427–446. doi:10.1016/j.biochi.2006.12.009

41. Pommier Y, Leo E, Zhang H, Marchand C. DNA Topoisomerases and Their Poisoning by Anticancer and Antibacterial Drugs. Chem Biol. 2010;17: 421–433. doi:10.1016/j.chembiol.2010.04.012

42. Viard T, de la Tour CB. Type IA topoisomerases: A simple puzzle? Biochimie. 2007;89: 456–467. doi:10.1016/j.biochi.2006.10.013

43. Meyer S, Beslon G. Torsion-Mediated Interaction between Adjacent Genes. PLoS Comput Biol. 2014;10. doi:10.1371/journal.pcbi.1003785

44. Tse-Dinh Y-C, Beran RK. Multiple promoters for transcription of the Escherichia coli DNA topoisomerase I gene and their regulation by DNA supercoiling. J Mol Biol. 1988;202: 735–742. doi:10.1016/0022-2836(88)90554-2

45. Menzel R, Gellert M. Modulation of transcription by DNA supercoiling: a deletion analysis of the Escherichia coli gyrA and gyrB promoters. Proc Natl Acad Sci. 1987;84: 4185–4189.

46. Straney R, Krah R, Menzel R. Mutations in the -10 TATAAT sequence of the gyrA promoter affect both promoter strength and sensitivity to DNA supercoiling. J Bacteriol. 1994;176: 5999–6006.

47. Unniraman S, Nagaraja V. Axial distortion as a sensor of supercoil changes: a molecular model for the homeostatic regulation of DNA gyrase. J Genet. 2001;80: 119–124.

48. Dillon SC, Dorman CJ. Bacterial nucleoid-associated proteins, nucleoid structure and gene

expression. Nat Rev Microbiol. 2010;8: 185–195. doi:10.1038/nrmicro2261

49. Luijsterburg MS, Noom MC, Wuite GJL, Dame RT. The architectural role of nucleoid-associated proteins in the organization of bacterial chromatin: A molecular perspective. J Struct Biol. 2006;156: 262–272. doi:10.1016/j.jsb.2006.05.006

50. Ali Azam T, Iwata A, Nishimura A, Ueda S, Ishihama A. Growth Phase-Dependent Variation in Protein Composition of the Escherichia coli Nucleoid. J Bacteriol. 1999;181: 6361–6370.

51. Dame RT. The role of nucleoid-associated proteins in the organization and compaction of bacterial chromatin. Mol Microbiol. 2005;56: 858–870. doi:10.1111/j.1365-2958.2005.04598.x

52. de los Rios S, Perona JJ. Structure of the Escherichia coli Leucine-responsive Regulatory Protein Lrp Reveals a Novel Octameric Assembly. J Mol Biol. 2007;366: 1589–1602. doi:10.1016/j.jmb.2006.12.032

53. Wang H, Liu B, Wang Q, Wang L. Genome-Wide Analysis of the Salmonella Fis Regulon and Its Regulatory Mechanism on Pathogenicity Islands. Chakravortty D, editor. PLoS ONE. 2013;8: e64688. doi:10.1371/journal.pone.0064688

54. Lin S-H, Lee JC. Determinants of DNA Bending in the DNA−Cyclic AMP Receptor Protein Complexes in Escherichia coli. Biochemistry (Mosc). 2003;42: 4809–4818. doi:10.1021/bi027259+

55. Browning DF, Cole JA, Busby SJW. Suppression of FNR-dependent transcription activation at the Escherichia coli nir promoter by Fis, IHF and H-NS: modulation of transcription initiation by a complex nucleo-protein assembly. Mol Microbiol. 2000;37: 1258–1269. doi:10.1046/j.1365-2958.2000.02087.x

56. Hanamura A, Aiba H. Molecular mechanism of negative autoregulation of Escherichia coli crp gene. Nucleic Acids Res. 1991;19: 4413–4419.

57. El-Robh MS, Busby SJW. The Escherichia coli cAMP receptor protein bound at a single target can activate transcription initiation at divergent promoters: a systematic study that exploits new promoter probe plasmids. Biochem J. 2002;368: 835–843. doi:10.1042/BJ20021003

58. Strainic MG, Sullivan JJ, Collado-Vides J, deHaseth PL. Promoter Interference in a Bacteriophage Lambda Control Region: Effects of a Range of Interpromoter Distances. J Bacteriol. 2000;182: 216–220.

59. Sneppen K, Dodd IB, Shearwin KE, Palmer AC, Schubert RA, Callen BP, et al. A Mathematical Model for Transcriptional Interference by RNA Polymerase Traffic in Escherichia coli. J Mol Biol. 2005;346: 399–409. doi:10.1016/j.jmb.2004.11.075

60. Palmer AC, Ahlgren-Berg A, Egan JB, Dodd IB, Shearwin KE. Potent Transcriptional Interference by Pausing of RNA Polymerases over a Downstream Promoter. Mol Cell. 2009;34: 545–555. doi:10.1016/j.molcel.2009.04.018

61. Opel ML, Hatfield G. DNA supercoiling-dependent transcriptional coupling between the divergently transcribed promoters of the ilvYC operon of Escherichia coli is proportional to promoter strengths and transcript lengths. Mol Microbiol. 2001;39: 191–198.

62. Gerganova V, Maurer S, Stoliar L, Japaridze A, Dietler G, Nasser W, et al. Upstream Binding of Idling RNA Polymerase Modulates Transcription Initiation from a Nearby Promoter. J Biol Chem. 2015; jbc.M114.628131. doi:10.1074/jbc.M114.628131

63. Huerta AM, Francino MP, Morett E, Collado-Vides J. Selection for Unequal Densities of σ70

Promoter-Like Signals in Different Regions of Large Bacterial Genomes. PLoS Genet. 2006;2. doi:10.1371/journal.pgen.0020185

64. Perez JC, Groisman EA. Evolution of Transcriptional Regulatory Circuits in Bacteria. Cell. 2009;138: 233–244. doi:10.1016/j.cell.2009.07.002

65. Mushegian AR, Koonin EV. A minimal gene set for cellular life derived by comparison of complete bacterial genomes. Proc Natl Acad Sci. 1996;93: 10268–10273.

66. Santos CL, Tavares F, Thioulouse J, Normand P. A phylogenomic analysis of bacterial helix–turn–helix transcription factors. FEMS Microbiol Rev. 2009;33: 411–429. doi:10.1111/j.1574-6976.2008.00154.x

67. Hershberg R, Margalit H. Co-evolution of transcription factors and their targets depends on mode of regulation. Genome Biol. 2006;7: R62. doi:10.1186/gb-2006-7-7-r62

68. Galán-Vásquez E, Sánchez-Osorio I, Martínez-Antonio A. Transcription Factors Exhibit Differential Conservation in Bacteria with Reduced Genomes. Robinson-Rechavi M, editor. PLOS ONE. 2016;11: e0146901. doi:10.1371/journal.pone.0146901

69. Groisman EA. The Pleiotropic Two-Component Regulatory System PhoP-PhoQ. J Bacteriol. 2001;183: 1835–1842. doi:10.1128/JB.183.6.1835-1842.2001

70. Perez JC, Shin D, Zwir I, Latifi T, Hadley TJ, Groisman EA. Evolution of a Bacterial Regulon Controlling Virulence and Mg2+ Homeostasis. PLoS Genet. 2009;5. doi:10.1371/journal.pgen.1000428

71. Duprey A, Reverchon S, Nasser W. Bacterial virulence and Fis: adapting regulatory networks to the host environment. Trends Microbiol. 2014;22: 92–99. doi:10.1016/j.tim.2013.11.008

72. Rhodius VA, Suh WC, Nonaka G, West J, Gross CA. Conserved and Variable Functions of the σ E Stress Response in Related Genomes. PLOS Biol. 2005;4: e2. doi:10.1371/journal.pbio.0040002

73. Isalan M, Lemerle C, Michalodimitrakis K, Horn C, Beltrao P, Raineri E, et al. Evolvability and hierarchy in rewired bacterial gene networks. Nature. 2008;452: 840–845. doi:10.1038/nature06847

74. Kato A, Mitrophanov AY, Groisman EA. A connector of two-component regulatory systems promotes signal amplification and persistence of expression. Proc Natl Acad Sci U S A. 2007;104: 12063–12068. doi:10.1073/pnas.0704462104

75. Treangen TJ, Rocha EPC. Horizontal Transfer, Not Duplication, Drives the Expansion of Protein Families in Prokaryotes. Moran NA, editor. PLoS Genet. 2011;7: e1001284. doi:10.1371/journal.pgen.1001284

76. Navarre WW, Porwollik S, Wang Y, McClelland M, Rosen H, Libby SJ, et al. Selective Silencing of Foreign DNA with Low GC Content by the H-NS Protein in Salmonella. Science. 2006;313: 236–238. doi:10.1126/science.1128794

77. Will WR, Bale DH, Reid PJ, Libby SJ, Fang FC. Evolutionary expansion of a regulatory network by counter-silencing. Nat Commun. 2014;5: 5270. doi:10.1038/ncomms6270

78. Will WR, Navarre WW, Fang FC. Integrated circuits: how transcriptional silencing and counter-silencing facilitate bacterial evolution. Curr Opin Microbiol. 2015;23: 8–13. doi:10.1016/j.mib.2014.10.005

79. Smits WK, Grossman AD. The Transcriptional Regulator Rok Binds A+T-Rich DNA and Is Involved in Repression of a Mobile Genetic Element in Bacillus subtilis. PLoS Genet. 2010;6. doi:10.1371/journal.pgen.1001207

80. Fass E, Groisman EA. Control of Salmonella pathogenicity island-2 gene expression. Curr Opin Microbiol. 2009;12: 199–204. doi:10.1016/j.mib.2009.01.004

81. Sandegren L, Andersson DI. Bacterial gene amplification: implications for the evolution of antibiotic resistance. Nat Rev Microbiol. 2009;7: 578–588. doi:10.1038/nrmicro2174

82. Conant GC, Wolfe KH. Turning a hobby into a job: How duplicated genes find new functions. Nat Rev Genet. 2008;9: 938–950. doi:10.1038/nrg2482

83. Gu X, Zhang Z, Huang W. Rapid evolution of expression and regulatory divergences after yeast gene duplication. Proc Natl Acad Sci U S A. 2005;102: 707–712. doi:10.1073/pnas.0409186102

84. Martínez-Núñez MA, Pérez-Rueda E, Gutiérrez-Ríos RM, Merino E. New insights into the regulatory networks of paralogous genes in bacteria. Microbiology. 2010;156: 14–22. doi:10.1099/mic.0.033266-0

85. Cooper TF, Rozen DE, Lenski RE. Parallel changes in gene expression after 20,000 generations of evolution in Escherichiacoli. Proc Natl Acad Sci U S A. 2003;100: 1072–1077. doi:10.1073/pnas.0334340100

86. Philippe N, Crozat E, Lenski RE, Schneider D. Evolution of global regulatory networks during a long-term experiment with Escherichia coli. BioEssays. 2007;29: 846–860. doi:10.1002/bies.20629

87. Cooper TF, Remold SK, Lenski RE, Schneider D. Expression Profiles Reveal Parallel Evolution of Epistatic Interactions Involving the CRP Regulon in Escherichia coli. PLoS Genet. 2008;4. doi:10.1371/journal.pgen.0040035

88. Hauben L, Moore ERB, Vauterin L, Steenackers M, Mergaert J, Verdonck L, et al. Phylogenetic Position of Phytopathogens within the Enterobacteriaceae. Syst Appl Microbiol. 1998;21: 384–397. doi:10.1016/S0723-2020(98)80048-9

89. Brady CL, Cleenwerck I, Denman S, Venter SN, Rodriguez-Palenzuela P, Coutinho TA, et al. Proposal to reclassify Brenneria quercina (Hildebrand and Schroth 1967) Hauben et al. 1999 into a new genus, Lonsdalea gen. nov., as Lonsdalea quercina comb. nov., descriptions of Lonsdalea quercina subsp. quercina comb. nov., Lonsdalea quercina subsp. iberica subsp. nov. and Lonsdalea quercina subsp. britannica subsp. nov., emendation of the description of the genus Brenneria, reclassification of Dickeya dieffenbachiae as Dickeya dadantii subsp. dieffenbachiae comb. nov., and emendation of the description of Dickeya dadantii. Int J Syst Evol Microbiol. 2012;62: 1592–1602. doi:10.1099/ijs.0.035055-0

90. Winslow C-EA, Broadhurst J, Buchanan RE, Krumwiede C, Rogers LA, Smith GH. The Families and Genera of the Bacteria. J Bacteriol. 1917;2: 505–566.

91. Jiang X. Modulation de l'expression génique chez la bactérie phytopathogène Dickeya dadantii en réponse aux conditions de stress rencontrées au cours l'infection et rôle de la structuration du chromosome bactérien. INSA Lyon. 2015.

92. Hormaeche E, Edwards PR. A proposed genus Enterobacter. Int J Syst Evol Microbiol. 1960;10: 71–74.

93. Brenner DJ, McWhorter AC, Kai A, Steigerwalt AG, Farmer JJ. Enterobacter asburiae sp. nov., a

new species found in clinical specimens, and reassignment of Erwinia dissolvens and Erwinia nimipressuralis to the genus Enterobacter as Enterobacter dissolvens comb. nov. and Enterobacter nimipressuralis comb. nov. J Clin Microbiol. 1986;23: 1114–1120.

94. Hoffmann H, Stindl S, Ludwig W, Stumpf A, Mehlen A, Heesemann J, et al. Reassignment of Enterobacter dissolvens to Enterobacter cloacae as E. cloacae subspecies dissolvens comb. nov. and emended description of Enterobacter asburiae and Enterobacter kobei. Syst Appl Microbiol. 2005;28: 196–205. doi:10.1016/j.syapm.2004.12.010

95. Sanders WE, Sanders CC. Enterobacter spp.: pathogens poised to flourish at the turn of the century. Clin Microbiol Rev. 1997;10: 220–241.

96. Bishop AL, Davis RM. Internal decay of onions caused by Enterobacter cloacae. Plant Dis. 1990;74: 692–694.

97. Coutinho TA, Venter SN. Pantoea ananatis: an unconventional plant pathogen. Mol Plant Pathol. 2009;10: 325–335. doi:10.1111/j.1364-3703.2009.00542.x

98. Oh C-S, Beer SV. Molecular genetics of Erwinia amylovora involved in the development of fire blight. FEMS Microbiol Lett. 2005;253: 185–192. doi:10.1016/j.femsle.2005.09.051

99. Brady C, Cleenwerck I, Venter S, Vancanneyt M, Swings J, Coutinho T. Phylogeny and identification of Pantoea species associated with plants, humans and the natural environment based on multilocus sequence analysis (MLSA). Syst Appl Microbiol. 2008;31: 447–460. doi:10.1016/j.syapm.2008.09.004

100. EPPO. EPPO New PQR [Internet]. 2016 [cited 3 Mar 2016]. Available: http://newpqr.eppo.int/download2.php

101. Dutkiewicz J, Mackiewicz B, Kinga Lemieszek M, Golec M, Milanowski J. Pantoea agglomerans: a mysterious bacterium of evil and good. Part III. Deleterious effects: infections of humans, animals and plants. Ann Agric Environ Med AAEM. 2016;23: 197–205. doi:10.5604/12321966.1203878

102. Smits TH, Jaenicke S, Rezzonico F, Kamber T, Goesmann A, Frey JE, et al. Complete genome sequence of the fire blight pathogen Erwinia pyrifoliae DSM 12163T and comparative genomic insights into plant pathogenicity. BMC Genomics. 2010;11: 2. doi:10.1186/1471-2164-11-2

103. Gardan L, Christen R, Achouak W, Prior P. Erwinia papayae sp. nov., a pathogen of papaya (Carica papaya). Int J Syst Evol Microbiol. 2004;54: 107–113. doi:10.1099/ijs.0.02718-0

104. Sutra L, Christen R, Bollet C, Simoneau P, Gardan L. Samsonia erythrinae gen. nov., sp. nov., isolated from bark necrotic lesions of Erythrina sp., and discrimination of plant-pathogenic Enterobacteriaceae by phenotypic features. Int J Syst Evol Microbiol. 2001;51: 1291–1304.

105. Yang L, Xu B, He W, Zhang L. The HrpW protein of Lonsdalea quercina N-5-1 has pectate lyase activity and is required for full bacterial virulence. J Basic Microbiol. 2014;54: 1126–1135. doi:10.1002/jobm.201300342

106. Huvenne H, Messens E, Maes M. Willow wood sap promotes the density-dependent pathogenesis of Brenneria salicis. Environ Microbiol. 2009;11: 1463–1472. doi:10.1111/j.1462-2920.2009.01874.x

107. Charkowski A, Blanco C, Condemine G, Expert D, Franza T, Hayes C, et al. The Role of Secretion Systems and Small Molecules in Soft-Rot *Enterobacteriaceae* Pathogenicity. Annu Rev Phytopathol. 2012;50: 425–449. doi:10.1146/annurev-phyto-081211-173013

108. Ma B, Hibbing ME, Kim H-S, Reedy RM, Yedidia I, Breuer J, et al. Host Range and Molecular Phylogenies of the Soft Rot Enterobacterial Genera Pectobacterium and Dickeya. Phytopathology. 2007;97: 1150–1163. doi:10.1094/PHYTO-97-9-1150

109. Potrykus M, Golanowska M, Hugouvieux-Cotte-Pattat N, Lojkowska E. Regulators Involved in Dickeya solani Virulence, Genetic Conservation, and Functional Variability. Mol Plant Microbe Interact. 2014;27: 700–711. doi:10.1094/MPMI-09-13-0270-R

110. Pérombelon MCM. Potato diseases caused by soft rot erwinias: an overview of pathogenesis. Plant Pathol. 2002;51: 1–12.

111. Hugouvieux-Cotte-Pattat N, Condemine G, Nasser W, Reverchon S. Regulation of Pectinolysis in Erwinia Chrysanthemi. Annu Rev Microbiol. 1996;50: 213–257. doi:10.1146/annurev.micro.50.1.213

112. Hugouvieux-Cotte-Pattat N, Condemine G, Shevchik VE. Bacterial pectate lyases, structural and functional diversity. Environ Microbiol Rep. 2014;6: 427–440.

113. Hugouvieux-Cotte-Pattat N, Robert-Baudouy J. Analysis of the regulation of the pelBC genes in Erwinia chrysanthemi 3937. Mol Microbiol. 1992;6: 2363–2376. doi:10.1111/j.1365-2958.1992.tb01411.x

114. Toth IK, Bell KS, Holeva MC, Birch PRJ. Soft rot erwiniae: from genes to genomes. Mol Plant Pathol. 2002;4: 17–30.

115. Toth IK, Birch PR. Rotting softly and stealthily. Curr Opin Plant Biol. 2005;8: 424–429. doi:10.1016/j.pbi.2005.04.001

116. Reverchon S, Nasser W. *Dickeya* ecology, environment sensing and regulation of virulence programme: *Dickeya dadantii* pathogenicity. Environ Microbiol Rep. 2013; n/a-n/a. doi:10.1111/1758-2229.12073

117. Prigent-Combaret C, Zghidi-Abouzid O, Effantin G, Lejeune P, Reverchon S, Nasser W. The nucleoid-associated protein Fis directly modulates the synthesis of cellulose, an essential component of pellicle-biofilms in the phytopathogenic bacterium Dickeya dadantii. Mol Microbiol. 2012;86: 172–186. doi:10.1111/j.1365-2958.2012.08182.x

118. Jahn CE, Willis DK, Charkowski AO. The Flagellar Sigma Factor FliA Is Required for Dickeya dadantii Virulence. Mol Plant Microbe Interact. 2008;21: 1431–1442. doi:10.1094/MPMI-21-11-1431

119. Antúnez-Lamas M, Cabrera-Ordóñez E, López-Solanilla E, Raposo R, Trelles-Salazar O, Rodríguez-Moreno A, et al. Role of motility and chemotaxis in the pathogenesis of Dickeya dadantii 3937 (ex Erwinia chrysanthemi 3937). Microbiol Read Engl. 2009;155: 434–442. doi:10.1099/mic.0.022244-0

120. Koskiniemi S, Lamoureux JG, Nikolakakis KC, t'Kint de Roodenbeke C, Kaplan MD, Low DA, et al. Rhs proteins from diverse bacteria mediate intercellular competition. Proc Natl Acad Sci U S A. 2013;110: 7032–7037. doi:10.1073/pnas.1300627110

121. Reverchon S, Rouanet C, Expert D, Nasser W. Characterization of Indigoidine Biosynthetic Genes in Erwinia chrysanthemi and Role of This Blue Pigment in Pathogenicity. J Bacteriol. 2002;184: 654–665. doi:10.1128/JB.184.3.654-665.2002

122. Nasser W, Dorel C, Wawrzyniak J, Van Gijsegem F, Groleau M-C, Déziel E, et al. Vfm a new quorum sensing system controls the virulence of *Dickeya dadantii*. Environ Microbiol. 2013;15:

865–880. doi:10.1111/1462-2920.12049

123. Bauer DW, Bogdanove AJ, Beer SV, Collmer A. Erwinia chrysanthemi hrp genes and their involvement in soft rot pathogenesis and elicitation of the hypersensitive response. Mol Plant-Microbe Interact MPMI. 1994;7: 573–581.

124. Yang C-H, Gavilanes-Ruiz M, Okinaka Y, Vedel R, Berthuy I, Boccara M, et al. hrp Genes of Erwinia chrysanthemi 3937 Are Important Virulence Factors. Mol Plant Microbe Interact. 2002;15: 472–480. doi:10.1094/MPMI.2002.15.5.472

125. Lee JS, Ryu HR, Cha JY, Baik HS. The hrp pathogenicity island of Pseudomonas syringae pv. tomato DC3000 is induced by plant phenolic acids. J Microbiol. 2015;53: 725–731. doi:10.1007/s12275-015-5256-4

126. Schulte R, Bonas U. Expression of the Xanthomonas campestris pv. vesicatoria hrp gene cluster, which determines pathogenicity and hypersensitivity on pepper and tomato, is plant inducible. J Bacteriol. 1992;174: 815–823.

127. Van Gijsegem F, Vasse J, De Rycke R, Castello P, Boucher C. Genetic dissection of Ralstonia solanacearum hrp gene cluster reveals that the HrpV and HrpX proteins are required for Hrp pilus assembly. Mol Microbiol. 2002;44: 935–946.

128. Andersson RA, Eriksson ARB, Heikinheimo R, Mäe A, Pirhonen M, Kõiv V, et al. Quorum Sensing in the Plant Pathogen Erwinia carotovora subsp. carotovora: The Role of expREcc. Mol Plant Microbe Interact. 2000;13: 384–393. doi:10.1094/MPMI.2000.13.4.384

129. Rodionov DA, Gelfand MS, Hugouvieux-Cotte-Pattat N. Comparative genomics of the KdgR regulon in Erwinia chrysanthemi 3937 and other gamma-proteobacteria. Microbiology. 2004;150: 3571–3590. doi:10.1099/mic.0.27041-0

130. Nasser W, Reverchon S, Condemine G, Robert-Baudouy J. Specific Interactions of Erwinia chrysanthemi KdgR Repressor with Different Operators of Genes Involved in Pectinolysis. J Mol Biol. 1994;236: 427–440. doi:10.1006/jmbi.1994.1155

131. Matsumoto H, Jitareerat P, Baba Y, Tsuyumu S. Comparative Study of Regulatory Mechanisms for Pectinase Production by Erwinia carotovora subsp. carotovora and Erwinia chrysanthemi. Mol Plant Microbe Interact. 2003;16: 226–237. doi:10.1094/MPMI.2003.16.3.226

132. Lu Y, Rashidul IM, Hirata H, Tsuyumu S. KdgR, an IClR Family Transcriptional Regulator, Inhibits Virulence Mainly by Repression of hrp Genes in Xanthomonas oryzae pv. oryzae. J Bacteriol. 2011;193: 6674–6682. doi:10.1128/JB.05714-11

133. Goudeau DM, Parker CT, Zhou Y, Sela S, Kroupitski Y, Brandl MT. The Salmonella Transcriptome in Lettuce and Cilantro Soft Rot Reveals a Niche Overlap with the Animal Host Intestine. Appl Environ Microbiol. 2012;79: 250–262. doi:10.1128/AEM.02290-12

134. Kwan G, Charkowski AO, Barak JD. Salmonella enterica Suppresses Pectobacterium carotovorum subsp. carotovorum Population and Soft Rot Progression by Acidifying the Microaerophilic Environment. mBio. 2013;4: e00557-12. doi:10.1128/mBio.00557-12

135. George AS, González IS, Lorca GL, Teplitski M. Contribution of the Salmonella enterica KdgR Regulon to Persistence of the Pathogen in Vegetable Soft Rots. Appl Environ Microbiol. 2016;82: 1353–1360. doi:10.1128/AEM.03355-15

136. Reverchon S, Expert D, Robert-Baudouy J, Nasser W. The cyclic AMP receptor protein is the main activator of pectinolysis genes in Erwinia chrysanthemi. J Bacteriol. 1997;179: 3500–3508.

137. Botsford JL, Harman JG. Cyclic AMP in prokaryotes. Microbiol Rev. 1992;56: 100–122.

138. Görke B, Stülke J. Carbon catabolite repression in bacteria: many ways to make the most out of nutrients. Nat Rev Microbiol. 2008;6: 613–624. doi:10.1038/nrmicro1932

139. Imamura R, Yamanaka K, Ogura T, Hiraga S, Fujita N, Ishihama A, et al. Identification of the cpdA Gene Encoding Cyclic 3',5'-Adenosine Monophosphate Phosphodiesterase in Escherichia coli. J Biol Chem. 1996;271: 25423–25429. doi:10.1074/jbc.271.41.25423

140. Gosset G, Zhang Z, Nayyar S, Cuevas WA, Saier MH. Transcriptome Analysis of Crp-Dependent Catabolite Control of Gene Expression in Escherichia coli. J Bacteriol. 2004;186: 3516–3524. doi:10.1128/JB.186.11.3516-3524.2004

141. Grainger DC, Hurd D, Harrison M, Holdstock J, Busby SJW. Studies of the distribution of Escherichia coli cAMP-receptor protein and RNA polymerase along the E. coli chromosome. Proc Natl Acad Sci U S A. 2005;102: 17693–17698. doi:10.1073/pnas.0506687102

142. Zhou Y, Kolb A, Busby SJW, Wang Y-P. Spacing requirements for Class I transcription activation in bacteria are set by promoter elements. Nucleic Acids Res. 2014;42: 9209–9216. doi:10.1093/nar/gku625

143. Niu W, Kim Y, Tau G, Heyduk T, Ebright RH. Transcription Activation at Class II CAP-Dependent Promoters: Two Interactions between CAP and RNA Polymerase. Cell. 1996;87: 1123–1134.

144. Richet E, Søgaard-Andersen L. CRP induces the repositioning of MalT at the Escherichia coli malKp promoter primarily through DNA bending. EMBO J. 1994;13: 4558–4567.

145. Nasser W, Robert-Baudouy J, Reverchon S. Antagonistic effect of CRP and KdgR in the transcription control of the Erwinia chrysanthemi pectinolysis genes. Mol Microbiol. 1997;26: 1071–1082. doi:10.1046/j.1365-2958.1997.6472020.x

146. Rouanet C, Nomura K, Tsuyumu S, Nasser W. Regulation of pelD and pelE, Encoding Major Alkaline Pectate Lyases in Erwinia chrysanthemi: Involvement of the Main Transcriptional Factors. J Bacteriol. 1999;181: 5948–5957.

147. de Crecy-Lagard V, Glaser P, Lejeune P, Sismeiro O, Barber CE, Daniels MJ, et al. A Xanthomonas campestris pv. campestris protein similar to catabolite activation factor is involved in regulation of phytopathogenicity. J Bacteriol. 1990;172: 5877–5883.

148. Moreno R, Ruiz-Manzano A, Yuste L, Rojo F. The Pseudomonas putida Crc global regulator is an RNA binding protein that inhibits translation of the AlkS transcriptional regulator. Mol Microbiol. 2007;64: 665–675. doi:10.1111/j.1365-2958.2007.05685.x

149. Colton DM, Stabb EV. Rethinking the roles of CRP, cAMP, and sugar-mediated global regulation in the Vibrionaceae. Curr Genet. 2015;62: 39–45. doi:10.1007/s00294-015-0508-8

150. Fuchs EL, Brutinel ED, Jones AK, Fulcher NB, Urbanowski ML, Yahr TL, et al. The Pseudomonas aeruginosa Vfr Regulator Controls Global Virulence Factor Expression through Cyclic AMP-Dependent and -Independent Mechanisms. J Bacteriol. 2010;192: 3553–3564. doi:10.1128/JB.00363-10

151. Smith RS, Wolfgang MC, Lory S. An Adenylate Cyclase-Controlled Signaling Network Regulates Pseudomonas aeruginosa Virulence in a Mouse Model of Acute Pneumonia. Infect Immun. 2004;72: 1677–1684. doi:10.1128/IAI.72.3.1677-1684.2004

152. Milanesio P, Arce-Rodríguez A, Muñoz A, Calles B, de Lorenzo V. Regulatory exaptation of the catabolite repression protein (Crp)–cAMP system in Pseudomonas putida. Environ Microbiol. 2011;13: 324–339. doi:10.1111/j.1462-2920.2010.02331.x

153. Arce-Rodríguez A, Durante-Rodríguez G, Platero R, Krell T, Calles B, de Lorenzo V. The Crp regulator of Pseudomonas putida: evidence of an unusually high affinity for its physiological effector, cAMP. Environ Microbiol. 2012;14: 702–713. doi:10.1111/j.1462-2920.2011.02622.x

154. Donovan GT, Norton JP, Bower JM, Mulvey MA. Adenylate Cyclase and the Cyclic AMP Receptor Protein Modulate Stress Resistance and Virulence Capacity of Uropathogenic Escherichia coli. Infect Immun. 2013;81: 249–258. doi:10.1128/IAI.00796-12

155. Lathem WW, Schroeder JA, Bellows LE, Ritzert JT, Koo JT, Price PA, et al. Posttranscriptional Regulation of the Yersinia pestis Cyclic AMP Receptor Protein Crp and Impact on Virulence. mBio. 2014;5. doi:10.1128/mBio.01038-13

156. Rohmer L, Hocquet D, Miller SI. Are pathogenic bacteria just looking for food? Metabolism and microbial pathogenesis. Trends Microbiol. 2011;19: 341–348. doi:10.1016/j.tim.2011.04.003

157. Winfield MD, Groisman EA. Role of Nonhost Environments in the Lifestyles of Salmonella and Escherichia coli. Appl Environ Microbiol. 2003;69: 3687–3694. doi:10.1128/AEM.69.7.3687-3694.2003

158. Hornef MW, Wick MJ, Rhen M, Normark S. Bacterial strategies for overcoming host innate and adaptive immune responses. Nat Immunol. 2002;3: 1033–1040. doi:10.1038/ni1102-1033

159. Latunde-Dada GO. Iron metabolism: microbes, mouse, and man. BioEssays. 2009;31: 1309–1317. doi:10.1002/bies.200900101

160. Shen S, Fang FC. Integrated stress responses in Salmonella. Int J Food Microbiol. 2012;152: 75–81. doi:10.1016/j.ijfoodmicro.2011.04.017

161. Brown SP, Cornforth DM, Mideo N. Evolution of virulence in opportunistic pathogens: generalism, plasticity, and control. Trends Microbiol. 2012;20: 336–342. doi:10.1016/j.tim.2012.04.005

162. Goldberg MD, Johnson M, Hinton JC, Williams PH. Role of the nucleoid-associated protein Fis in the regulation of virulence properties of enteropathogenic Escherichia coli. Mol Microbiol. 2001;41: 549–559.

163. Falconi M, Prosseda G, Giangrossi M, Beghetto E, Colonna B. Involvement of FIS in the H-NS-mediated regulation of virF gene of Shigella and enteroinvasive Escherichia coli. Mol Microbiol. 2001;42: 439–452.

164. Kelly A, Goldberg MD, Carroll RK, Danino V, Hinton JCD, Dorman CJ. A global role for Fis in the transcriptional control of metabolism and type III secretion in Salmonella enterica serovar Typhimurium. Microbiology. 2004;150: 2037–2053. doi:10.1099/mic.0.27209-0

165. Schechter LM, Jain S, Akbar S, Lee CA. The small nucleoid-binding proteins H-NS, HU, and Fis affect hilA expression in Salmonella enterica serovar Typhimurium. Infect Immun. 2003;71: 5432–5435.

166. Lenz DH, Bassler BL. The small nucleoid protein Fis is involved in Vibrio cholerae quorum sensing. Mol Microbiol. 2007;63: 859–871. doi:10.1111/j.1365-2958.2006.05545.x

167. Lautier T, Nasser W. The DNA nucleoid-associated protein Fis co-ordinates the expression of

the main virulence genes in the phytopathogenic bacterium Erwinia chrysanthemi. Mol Microbiol. 2007;66: 1474–1490. doi:10.1111/j.1365-2958.2007.06012.x

168. Finkel SE, Johnson RC. The Fis protein: it's not just for DNA inversion anymore. Mol Microbiol. 1992;6: 3257–3265. doi:10.1111/j.1365-2958.1992.tb02193.x

169. Ge J, Lou Z, Cui H, Shang L, Harshey RM. Analysis of phage Mu DNA transposition by whole-genome Escherichia coli tiling arrays reveals a complex relationship to distribution of target selection protein B, transcription and chromosome architectural elements. J Biosci. 2011;36: 587–601.

170. Choi H-S, Kim K-S, Park JW, Jung YH, Lee Y. Effects of FIS protein on rnpB transcription in Escherichia coli. Mol Cells. 2005;19: 239–245.

171. Schneider R, Travers A, Kutateladze T, Muskhelishvili G. A DNA architectural protein couples cellular physiology and DNA topology in Escherichia coli. Mol Microbiol. 1999;34: 953–964. doi:10.1046/j.1365-2958.1999.01656.x

172. Bradley MD, Beach MB, de Koning APJ, Pratt TS, Osuna R. Effects of Fis on Escherichia coli gene expression during different growth stages. Microbiol Read Engl. 2007;153: 2922–2940. doi:10.1099/mic.0.2007/008565-0

173. Barnhart MM, Chapman MR. Curli Biogenesis and Function. Annu Rev Microbiol. 2006;60: 131–147. doi:10.1146/annurev.micro.60.080805.142106

174. Rossiter AE, Browning DF, Leyton DL, Johnson MD, Godfrey RE, Wardius CA, et al. Transcription of the plasmid-encoded toxin gene from Enteroaggregative Escherichia coli is regulated by a novel co-activation mechanism involving CRP and Fis. Mol Microbiol. 2011;81: 179–191. doi:10.1111/j.1365-2958.2011.07685.x

175. Weinstein-Fischer D, Altuvia S. Differential regulation of Escherichia coli topoisomerase I by Fis. Mol Microbiol. 2007;63: 1131–1144. doi:10.1111/j.1365-2958.2006.05569.x

176. Cameron ADS, Dorman CJ. A Fundamental Regulatory Mechanism Operating through OmpR and DNA Topology Controls Expression of Salmonella Pathogenicity Islands SPI-1 and SPI-2. PLoS Genet. 2012;8: e1002615. doi:10.1371/journal.pgen.1002615

177. Saldana Z, Xicohtencatl-Cortes J, Avelino F, Phillips AD, Kaper JB, Puente JL, et al. Synergistic role of curli and cellulose in cell adherence and biofilm formation of attaching and effacing Escherichia coli and identification of Fis as a negative regulator of curli. Environ Microbiol. 2009;11: 992–1006. doi:10.1111/j.1462-2920.2008.01824.x

178. Steen JA, Steen JA, Harrison P, Seemann T, Wilkie I, Harper M, et al. Fis is essential for capsule production in Pasteurella multocida and regulates expression of other important virulence factors. PLoS Pathog. 2010;6: e1000750. doi:10.1371/journal.ppat.1000750

179. Ó Cróinín T, Carroll RK, Kelly A, Dorman CJ. Roles for DNA supercoiling and the Fis protein in modulating expression of virulence genes during intracellular growth of Salmonella enterica serovar Typhimurium. Mol Microbiol. 2006;62: 869–882. doi:10.1111/j.1365-2958.2006.05416.x

180. Steinmann R, Dersch P. Thermosensing to adjust bacterial virulence in a fluctuating environment. Future Microbiol. 2013;8: 85–105. doi:10.2217/fmb.12.129

181. Czajkowski R, Perombelon MC, van Veen JA, van der Wolf JM. Control of blackleg and tuber soft rot of potato caused by Pectobacterium and Dickeya species: a review. Plant Pathol.

2011;60: 999–1013.

182. Lebeau A, Reverchon S, Gaubert S, Kraepiel Y, Simond-Côte E, Nasser W, et al. The GacA global regulator is required for the appropriate expression of Erwinia chrysanthemi 3937 pathogenicity genes during plant infection. Environ Microbiol. 2008;10: 545–559. doi:10.1111/j.1462-2920.2007.01473.x

183. Mansfield J, Genin S, Magori S, Citovsky V, Sriariyanum M, Ronald P, et al. Top 10 plant pathogenic bacteria in molecular plant pathology. Mol Plant Pathol. 2012;13: 614–629. doi:10.1111/j.1364-3703.2012.00804.x

184. Toth IK, van der Wolf JM, Saddler G, Lojkowska E, Hélias V, Pirhonen M, et al. Dickeya species: an emerging problem for potato production in Europe. Plant Pathol. 2011;60: 385–399. doi:10.1111/j.1365-3059.2011.02427.x

185. Sepulchre J-A, Reverchon S, Nasser W. Modeling the onset of virulence in a pectinolytic bacterium. J Theor Biol. 2007;244: 239–257. doi:10.1016/j.jtbi.2006.08.010

186. Kepseu WD, Sepulchre J-A, Reverchon S, Nasser W. Toward a quantitative modeling of the synthesis of the pectate lyases, essential virulence factors in Dickeya dadantii. J Biol Chem. 2010;285: 28565–28576. doi:10.1074/jbc.M110.114710

187. Hommais F, Oger-Desfeux C, Van Gijsegem F, Castang S, Ligori S, Expert D, et al. PecS Is a Global Regulator of the Symptomatic Phase in the Phytopathogenic Bacterium Erwinia chrysanthemi 3937. J Bacteriol. 2008;190: 7508–7522. doi:10.1128/JB.00553-08

188. Reverchon S, Van Gijsegem F, Effantin G, Zghidi-Abouzid O, Nasser W. Systematic targeted mutagenesis of the MarR/SlyA family members of Dickeya dadantii 3937 reveals a role for MfbR in the modulation of virulence gene expression in response to acidic pH. Mol Microbiol. 2010;78: 1018–1037. doi:10.1111/j.1365-2958.2010.07388.x

189. Franza T, Expert D. Role of iron homeostasis in the virulence of phytopathogenic bacteria: an “à la carte” menu. Mol Plant Pathol. 2013;14: 429–438. doi:10.1111/mpp.12007

190. Nachin L, Barras F. External pH: an environmental signal that helps to rationalize pel gene duplication in Erwinia chrysanthemi. Mol Plant Microbe Interact. 2000;13: 882–886.

191. Lopez-Solanilla E, Garcia-Olmedo F, Rodriguez-Palenzuela P. Inactivation of the sapA to sapF locus of Erwinia chrysanthemi reveals common features in plant and animal bacterial pathogenesis. Plant Cell. 1998;10: 917–924.

192. Lautier T, Blot N, Muskhelishvili G, Nasser W. Integration of two essential virulence modulating signals at the Erwinia chrysanthemi pel gene promoters: a role for Fis in the growth-phase regulation. Mol Microbiol. 2007;66: 1491–1505. doi:10.1111/j.1365-2958.2007.06010.x

193. Ouafa Z-A, Reverchon S, Lautier T, Muskhelishvili G, Nasser W. The nucleoid-associated proteins H-NS and FIS modulate the DNA supercoiling response of the pel genes, the major virulence factors in the plant pathogen bacterium Dickeya dadantii. Nucleic Acids Res. 2012;40: 4306–4319. doi:10.1093/nar/gks014

194. Garai P, Gnanadhas DP, Chakravortty D. Salmonella enterica serovars Typhimurium and Typhi as model organisms. Virulence. 2012;3: 377–388. doi:10.4161/viru.21087

195. Figueira R, Holden DW. Functions of the Salmonella pathogenicity island 2 (SPI-2) type III secretion system effectors. Microbiology. 2012;158: 1147–1161. doi:10.1099/mic.0.058115-0

196. Chen LM, Kaniga K, Galán JE. Salmonella spp. are cytotoxic for cultured macrophages. Mol Microbiol. 1996;21: 1101–1115. doi:10.1046/j.1365-2958.1996.471410.x

197. Brennan MA, Cookson BT. Salmonella induces macrophage death by caspase-1-dependent necrosis. Mol Microbiol. 2000;38: 31–40. doi:10.1046/j.1365-2958.2000.02103.x

198. Que F, Wu S, Huang R. Salmonella Pathogenicity Island 1(SPI-1) at Work. Curr Microbiol. 2013;66: 582–587. doi:10.1007/s00284-013-0307-8

199. Ellermeier CD, Ellermeier JR, Slauch JM. HilD, HilC and RtsA constitute a feed forward loop that controls expression of the SPI1 type three secretion system regulator hilA in Salmonella enterica serovar Typhimurium. Mol Microbiol. 2005;57: 691–705. doi:10.1111/j.1365-2958.2005.04737.x

200. Dorman CJ, Corcoran CP. Bacterial DNA topology and infectious disease. Nucleic Acids Res. 2009;37: 672–678. doi:10.1093/nar/gkn996

201. Osborne SE, Coombes BK. Transcriptional Priming of Salmonella Pathogenicity Island-2 Precedes Cellular Invasion. PLoS ONE. 2011;6. doi:10.1371/journal.pone.0021648

202. Dorman CJ. Co-operative roles for DNA supercoiling and nucleoid-associated proteins in the regulation of bacterial transcription. Biochem Soc Trans. 2013;41: 542–547. doi:10.1042/BST20120222

203. Mitchell A, Pilpel Y. A mathematical model for adaptive prediction of environmental changes by microorganisms. Proc Natl Acad Sci U S A. 2011;108: 7271–7276. doi:10.1073/pnas.1019754108

204. Jakovleva J, Teppo A, Velts A, Saumaa S, Moor H, Kivisaar M, et al. Fis regulates the competitiveness of Pseudomonas putida on barley roots by inducing biofilm formation. Microbiol Read Engl. 2012;158: 708–720. doi:10.1099/mic.0.053355-0

205. Cameron ADS, Kröger C, Quinn HJ, Scally IK, Daly AJ, Kary SC, et al. Transmission of an Oxygen Availability Signal at the Salmonella enterica Serovar Typhimurium fis Promoter. Hayes F, editor. PLoS ONE. 2013;8: e84382. doi:10.1371/journal.pone.0084382

206. Mhedbi-Hajri N, Malfatti P, Pédron J, Gaubert S, Reverchon S, Van Gijsegem F. PecS is an important player in the regulatory network governing the coordinated expression of virulence genes during the interaction between Dickeya dadantii 3937 and plants. Environ Microbiol. 2011;13: 2901–2914. doi:10.1111/j.1462-2920.2011.02566.x

207. Perera IC, Grove A. Urate Is a Ligand for the Transcriptional Regulator PecS. J Mol Biol. 2010;402: 539–551. doi:10.1016/j.jmb.2010.07.053

208. Becker BF. Towards the physiological function of uric acid. Free Radic Biol Med. 1993;14: 615–631.

209. Río LA del, Corpas FJ, Sandalio LM, Palma JM, Gómez M, Barroso JB. Reactive oxygen species, antioxidant systems and nitric oxide in peroxisomes. J Exp Bot. 2002;53: 1255–1272. doi:10.1093/jexbot/53.372.1255

210. Wang Z-C, Liu C-J, Huang Y-J, Wang Y-S, Peng H-L. PecS regulates the urate-responsive expression of type 1 fimbriae in Klebsiella pneumoniae CG43. Microbiology. 2015;161: 2395–2409. doi:10.1099/mic.0.000185

211. Perera IC, Grove A. MarR homologs with urate-binding signature. Protein Sci Publ Protein

Soc. 2011;20: 621–629. doi:10.1002/pro.588

212. Huang H, Mackel BJ, Grove A. Streptomyces coelicolor encodes a urate-responsive transcriptional regulator with homology to PecS from plant pathogens. J Bacteriol. 2013;195: 4954–4965. doi:10.1128/JB.00854-13

213. Rouanet C, Nasser W. The PecM protein of the phytopathogenic bacterium Erwinia chrysanthemi, membrane topology and possible involvement in the efflux of the blue pigment indigoidine. J Mol Microbiol Biotechnol. 2001;3: 309–318.

214. Hérault E, Reverchon S, Nasser W. Role of the LysR-type transcriptional regulator PecT and DNA supercoiling in the thermoregulation of *pel* genes, the major virulence factors in *Dickeya dadantii*: *Dickeya dadantii* PecT protein and virulence thermoregulation. Environ Microbiol. 2014;16: 734–745. doi:10.1111/1462-2920.12198

215. Hérault E. Régulation de la synthèse des facteurs de virulence par la température chez la bactérie phytopathogène Dickeya dadantii [Internet]. Lyon 1. 2013. Available: http://www.theses.fr/2013LYO10259

216. Nasser W, Reverchon S. H-NS-dependent activation of pectate lyases synthesis in the phytopathogenic bacterium Erwinia chrysanthemi is mediated by the PecT repressor. Mol Microbiol. 2002;43: 733–748.

217. Surgey N, Robert-Baudouy J, Condemine G. The Erwinia chrysanthemi pecT gene regulates pectinase gene expression. J Bacteriol. 1996;178: 1593–1599.

218. Mukherjee A, Cui Y, Ma W, Liu Y, Chatterjee AK. hexA of Erwinia carotovora ssp. carotovora strain Ecc71 negatively regulates production of RpoS and rsmB RNA, a global regulator of extracellular proteins, plant virulence and the quorum-sensing signal, N-(3-oxohexanoyl)- l-homoserine lactone. Environ Microbiol. 2000;2: 203–215. doi:10.1046/j.1462-2920.2000.00093.x

219. Kernell Burke A, Duong DA, Jensen RV, Stevens AM. Analyzing the Transcriptomes of Two Quorum-Sensing Controlled Transcription Factors, RcsA and LrhA, Important for Pantoea stewartii Virulence. PLoS ONE. 2015;10. doi:10.1371/journal.pone.0145358

220. Honda N, Iyoda S, Yamamoto S, Terajima J, Watanabe H. LrhA positively controls the expression of the locus of enterocyte effacement genes in enterohemorrhagic Escherichia coli by differential regulation of their master regulators PchA and PchB. Mol Microbiol. 2009;74: 1393–1341. doi:10.1111/j.1365-2958.2009.06937.x

221. Richards GR, Herbert EE, Park Y, Goodrich-Blair H. Xenorhabdus nematophila lrhA Is Necessary for Motility, Lipase Activity, Toxin Expression, and Virulence in Manduca sexta Insects. J Bacteriol. 2008;190: 4870–4879. doi:10.1128/JB.00358-08

222. Lehnen D, Blumer C, Polen T, Wackwitz B, Wendisch VF, Unden G. LrhA as a new transcriptional key regulator of flagella, motility and chemotaxis genes in Escherichia coli. Mol Microbiol. 2002;45: 521–532. doi:10.1046/j.1365-2958.2002.03032.x

223. Mouslim C, Hughes KT. The Effect of Cell Growth Phase on the Regulatory Cross-Talk between Flagellar and Spi1 Virulence Gene Expression. PLoS Pathog. 2014;10. doi:10.1371/journal.ppat.1003987

224. Maeda K, Imae Y, Shioi JI, Oosawa F. Effect of temperature on motility and chemotaxis of Escherichia coli. J Bacteriol. 1976;127: 1039–1046.

225. Parkinson N, DeVos P, Pirhonen M, Elphinstone J. Dickeya aquatica sp. nov., isolated from

waterways. Int J Syst Evol Microbiol. 2014;64: 2264–2266. doi:10.1099/ijs.0.058693-0

226. Perkins TT, Davies MR, Klemm EJ, Rowley G, Wileman T, James K, et al. ChIP-seq and transcriptome analysis of the OmpR regulon of *Salmonella enterica* serovars Typhi and Typhimurium reveals accessory genes implicated in host colonization: ChIP-seq analysis of OmpR regulon. Mol Microbiol. 2013;87: 526–538. doi:10.1111/mmi.12111

227. Tardy F, Nasser W, Robert-Baudouy J, Hugouvieux-Cotte-Pattat N. Comparative analysis of the five major Erwinia chrysanthemi pectate lyases: enzyme characteristics and potential inhibitors. J Bacteriol. 1997;179: 2503–2511.

228. Duprey A, Muskhelishvili G, Reverchon S, Nasser W. Temporal control of Dickeya dadantii main virulence gene expression by growth phase-dependent alteration of regulatory nucleoprotein complexes. Biochim Biophys Acta. 2016; doi:10.1016/j.bbagrm.2016.08.001

229. Bhat A, Shin M, Jeong J-H, Kim H-J, Lim H-J, Rhee JH, et al. DNA looping-dependent autorepression of LEE1 P1 promoters by Ler in enteropathogenic Escherichia coli (EPEC). Proc Natl Acad Sci U S A. 2014;111: E2586–E2595. doi:10.1073/pnas.1322033111

230. Skoko D, Yoo D, Bai H, Schnurr B, Yan J, McLeod SM, et al. Mechanism of chromosome compaction and looping by the E. coli nucleoid protein Fis. J Mol Biol. 2006;364: 777–798. doi:10.1016/j.jmb.2006.09.043

231. Jin DJ, Cagliero C, Zhou YN. Role of RNA Polymerase and Transcription in the Organization of the Bacterial Nucleoid. Chem Rev. 2013; 130813134324001. doi:10.1021/cr4001429

232. Travers A, Muskhelishvili G. Bacterial chromatin. Curr Opin Genet Dev. 2005;15: 507–514. doi:10.1016/j.gde.2005.08.006

233. Schneider R, Travers A, Muskhelishvili G. FIS modulates growth phase-dependent topological transitions of DNA in Escherichia coli. Mol Microbiol. 1997;26: 519–530.

234. Browning DF, Grainger DC, Busby SJ. Effects of nucleoid-associated proteins on bacterial chromosome structure and gene expression. Curr Opin Microbiol. 2010;13: 773–780. doi:10.1016/j.mib.2010.09.013

235. Hassan S, Shevchik VE, Robert X, Hugouvieux-Cotte-Pattat N. PelN Is a New Pectate Lyase of Dickeya dadantii with Unusual Characteristics. J Bacteriol. 2013;195: 2197–2206. doi:10.1128/JB.02118-12

236. Reverchon S, Nasser W, Robert-Baudouy J. Characterization of kdgR, a gene of Erwinia chrysanthemi that regulates pectin degradation. Mol Microbiol. 1991;5: 2203–2216.

237. Beaulieu C, Boccara M, Van Gijsegem F. Pathogenic Behavior of Pectinase-Defective *Erwinia chrysanthemi* Mutants on Different Plants. Mol Plant Microbe Interact. 1993;6: 197. doi:10.1094/MPMI-6-197

238. Boccara M, Diolez A, Rouve M, Kotoujansky A. The role of individual pectate lyases of Erwinia chrysanthemi strain 3937 in pathogenicity on saintpaulia plants. Physiol Mol Plant Pathol. 1988;33: 95–104. doi:10.1016/0885-5765(88)90046-X

239. Miller JH. Experiment in Molecular Genetics. Cold Spring Harbor Laboratory Press; 1972.

240. Duprey A, Chansavang V, Fremion F, Gonthier C, Louis Y, Lejeune P, et al. “NiCo Buster”: engineering E. coli for fast and efficient capture of cobalt and nickel. J Biol Eng. 2014;8: 19. doi:10.1186/1754-1611-8-19

241. Marschall C, Labrousse V, Kreimer M, Weichart D, Kolb A, Hengge-Aronis R. Molecular analysis of the regulation of csiD, a carbon starvation-inducible gene in Escherichia coli that is exclusively dependent on σS and requires activation by cAMP-CRP. J Mol Biol. 1998;276: 339–353. doi:10.1006/jmbi.1997.1533

242. Sambrook J, Russell DW. Molecular cloning : a laboratory manual. 3rd ed. Cold Spring Harbor Laboratory Press; 2001.

243. Kotoujansky A, Lemattre M, Boistard P. Utilization of a thermosensitive episome bearing transposon TN10 to isolate Hfr donor strains of Erwinia carotovora subsp. chrysanthemi. J Bacteriol. 1982;150: 122–131.

244. González-Gil G, Bringmann P, Kahmann R. FIS is a regulator of metabolism in Escherichia coli. Mol Microbiol. 1996;22: 21–29. doi:10.1111/j.1365-2958.1996.tb02652.x

245. Bardonnet N, Trautwetter A, Couchoux-Luthaud G, Blanco C. Plasmids with the uidA reporter gene for the detection of promoters and transcription signals. Mol Gen Genet MGG. 1988;212: 390–392.

246. Miller WG, Leveau JH, Lindow SE. Improved gfp and inaZ broad-host-range promoter-probe vectors. Mol Plant-Microbe Interact MPMI. 2000;13: 1243–1250. doi:10.1094/MPMI.2000.13.11.1243

247. Jiang X, Sobetzko P, Nasser W, Reverchon S, Muskhelishvili G. Chromosomal "Stress-Response" Domains Govern the Spatiotemporal Expression of the Bacterial Virulence Program. mBio. 2015;6: e00353-15. doi:10.1128/mBio.00353-15

248. De Jong H, Ranquet C, Ropers D, Pinel C, Geiselmann J. Experimental and computational validation of models of fluorescent and luminescent reporter genes in bacteria. BMC Syst Biol. 2010;4: 55.

249. Verkhusha VV, Kuznetsova IM, Stepanenko OV, Zaraisky AG, Shavlovsky MM, Turoverov KK, et al. High Stability of Discosoma DsRed As Compared to Aequorea EGFP†. Biochemistry (Mosc). 2003;42: 7879–7884. doi:10.1021/bi034555t

250. Hatoum A, Roberts J. Prevalence of RNA polymerase stalling at Escherichia coli promoters after open complex formation. Mol Microbiol. 2008;68: 17–28. doi:10.1111/j.1365-2958.2008.06138.x

251. Münch R, Hiller K, Grote A, Scheer M, Klein J, Schobert M, et al. Virtual Footprint and PRODORIC: an integrative framework for regulon prediction in prokaryotes. Bioinformatics. 2005;21: 4187–4189. doi:10.1093/bioinformatics/bti635

252. Karambelkar S, Swapna G, Nagaraja V. Silencing of toxic gene expression by Fis. Nucleic Acids Res. 2012;40: 4358–4367. doi:10.1093/nar/gks037

253. James V, Hugouvieux-Cotte-Pattat N. Regulatory systems modulating the transcription of the pectinase genes of Erwinia chrysanthemi are conserved in Escherichia coli. Microbiol Read Engl. 1996;142 (Pt 9): 2613–2619.

254. Eriksson J, Nordström T, Nyrén P. Method enabling firefly luciferase-based bioluminometric assays at elevated temperatures. Anal Biochem. 2003;314: 158–161. doi:10.1016/S0003-2697(02)00647-4

255. Maxwell A. DNA gyrase as a drug target. Trends Microbiol. 1997;5: 102–109. doi:10.1016/S0966-842X(96)10085-8

256. Balke VL, Gralla JD. Changes in the linking number of supercoiled DNA accompany growth transitions in Escherichia coli. J Bacteriol. 1987;169: 4499–4506.

257. McLeod SM, Aiyar SE, Gourse RL, Johnson RC. The C-terminal domains of the RNA polymerase α subunits: contact site with fis and localization during co-activation with CRP at the Escherichia coli proP P2 promoter. J Mol Biol. 2002;316: 517–529. doi:10.1006/jmbi.2001.5391

258. Davidsson PR, Kariola T, Niemi O, Palva ET. Pathogenicity of and plant immunity to soft rot pectobacteria. Front Plant Sci. 2013;4. doi:10.3389/fpls.2013.00191

259. Nasser W, Schneider R, Travers A, Muskhelishvili G. CRP Modulates fis Transcription by Alternate Formation of Activating and Repressing Nucleoprotein Complexes. J Biol Chem. 2001;276: 17878–17886. doi:10.1074/jbc.M100632200

260. Ball CA, Osuna R, Ferguson KC, Johnson RC. Dramatic changes in Fis levels upon nutrient upshift in Escherichia coli. J Bacteriol. 1992;174: 8043–8056.

261. Keane OM, Dorman CJ. The gyr genes of Salmonella enterica serovar Typhimurium are repressed by the factor for inversion stimulation, Fis. Mol Genet Genomics. 2003;270: 56–65. doi:10.1007/s00438-003-0896-1

262. Chastanet A, Losick R. Just-in-Time Control of Spo0A Synthesis in Bacillus subtilis by Multiple Regulatory Mechanisms ▽. J Bacteriol. 2011;193: 6366–6374. doi:10.1128/JB.06057-11

263. Squire DJP, Xu M, Cole JA, Busby SJW, Browning DF. Competition between NarL-dependent activation and Fis-dependent repression controls expression from the *Escherichia coli yeaR* and *ogt* promoters. Biochem J. 2009;420: 249–257. doi:10.1042/BJ20090183

264. González-Gil G, Kahmann R, Muskhelishvili G. Regulation of crp transcription by oscillation between distinct nucleoprotein complexes. EMBO J. 1998;17: 2877–2885.

265. Rivetti C, Guthold M, Bustamante C. Wrapping of DNA around the E.coli RNA polymerase open promoter complex. EMBO J. 1999;18: 4464–4475. doi:10.1093/emboj/18.16.4464

266. Maurer S, Fritz J, Muskhelishvili G, Travers A. RNA polymerase and an activator form discrete subcomplexes in a transcription initiation complex. EMBO J. 2006;25: 3784–3790. doi:10.1038/sj.emboj.7601261

267. Kroll JS. Bacterial virulence: an environmental response. Arch Dis Child. 1991;66: 361–363.

268. Yahr TL, Wolfgang MC. Transcriptional regulation of the Pseudomonas aeruginosa type III secretion system. Mol Microbiol. 2006;62: 631–640. doi:10.1111/j.1365-2958.2006.05412.x

269. Altier C. Genetic and environmental control of Salmonella invasion. J Microbiol. 2005;43: 85–92.

270. Elena SF, Lenski RE. Microbial genetics: Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. Nat Rev Genet. 2003;4: 457–469. doi:10.1038/nrg1088

271. Wray GA. The evolutionary significance of cis-regulatory mutations. Nat Rev Genet. 2007;8: 206–216. doi:10.1038/nrg2063

272. Barrick JE, Yu DS, Yoon SH, Jeong H, Oh TK, Schneider D, et al. Genome evolution and adaptation in a long-term experiment with Escherichia coli. Nature. 2009;461: 1243–1247. doi:10.1038/nature08480

273. Park C, Makova KD. Coding region structural heterogeneity and turnover of transcription start sites contribute to divergence in expression between duplicate genes. Genome Biol. 2009;10: R10.

274. Main BJ, Smith AD, Jang H, Nuzhdin SV. Transcription Start Site Evolution in Drosophila. Mol Biol Evol. 2013;30: 1966–1974. doi:10.1093/molbev/mst085

275. Hugouvieux-Cotte-Pattat N, Dominguez H, Robert-Baudouy J. Environmental conditions affect transcription of the pectinase genes of Erwinia chrysanthemi 3937. J Bacteriol. 1992;174: 7807–7818.

276. Pritchard L, Humphris S, Baeyen S, Maes M, Van Vaerenbergh J, Elphinstone J, et al. Draft Genome Sequences of Four Dickeya dianthicola and Four Dickeya solani Strains. Genome Announc. 2013;1. doi:10.1128/genomeA.00087-12

277. Blount ZD, Barrick JE, Davidson CJ, Lenski RE. Genomic analysis of a key innovation in an experimental Escherichia coli population. Nature. 2012;489: 513–518. doi:10.1038/nature11514

278. Medina-Rivera A, Defrance M, Sand O, Herrmann C, Castro-Mondragon JA, Delerce J, et al. RSAT 2015: Regulatory Sequence Analysis Tools. Nucleic Acids Res. 2015; gkv362. doi:10.1093/nar/gkv362

279. Gouy M, Guindon S, Gascuel O. SeaView Version 4: A Multiplatform Graphical User Interface for Sequence Alignment and Phylogenetic Tree Building. Mol Biol Evol. 2010;27: 221–224. doi:10.1093/molbev/msp259

280. Talavera G, Castresana J. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Syst Biol. 2007;56: 564–577. doi:10.1080/10635150701472164

281. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. Syst Biol. 2010;59: 307–321. doi:10.1093/sysbio/syq010

282. Le SQ, Gascuel O. An Improved General Amino Acid Replacement Matrix. Mol Biol Evol. 2008;25: 1307–1320. doi:10.1093/molbev/msn067

283. Boussau B, Gouy M. Efficient Likelihood Computations with Nonreversible Models of Evolution. Syst Biol. 2006;55: 756–768. doi:10.1080/10635150600975218

284. Galtier N, Gouy M. Inferring pattern and process: maximum-likelihood implementation of a nonhomogeneous model of DNA sequence evolution for phylogenetic analysis. Mol Biol Evol. 1998;15: 871–879.

285. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, et al. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol. 2012;61: 539–542. doi:10.1093/sysbio/sys029

286. Tebbutt J, Rhodius VA, Webster CL, Busby SJW. Architectural requirements for optimal activation by tandem CRP molecules at a class I CRP-dependent promoter. FEMS Microbiol Lett. 2002;210: 55–60. doi:10.1111/j.1574-6968.2002.tb11159.x

287. Ranwez V, Harispe S, Delsuc F, Douzery EJP. MACSE: Multiple Alignment of Coding SEquences Accounting for Frameshifts and Stop Codons. PLOS ONE. 2011;6: e22594. doi:10.1371/journal.pone.0022594

288. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. BMC Bioinformatics. 2009;10: 421. doi:10.1186/1471-2105-10-421

289. Mäder U, Nicolas P, Depke M, Pané-Farré J, Debarbouille M, van der Kooi-Pol MM, et al. Staphylococcus aureus Transcriptome Architecture: From Laboratory to Infection-Mimicking Conditions. PLoS Genet. 2016;12: e1005962.

290. Alkhateeb RS, Vorhölter F-J, Rückert C, Mentz A, Wibberg D, Hublik G, et al. Genome wide transcription start sites analysis of Xanthomonas campestris pv. campestris B100 with insights into the gum gene cluster directing the biosynthesis of the exopolysaccharide xanthan. J Biotechnol. 2016;225: 18–28. doi:10.1016/j.jbiotec.2016.03.020

291. Quinn HJ, Cameron ADS, Dorman CJ. Bacterial Regulon Evolution: Distinct Responses and Roles for the Identical OmpR Proteins of Salmonella Typhimurium and Escherichia coli in the Acid Stress Response. Hughes D, editor. PLoS Genet. 2014;10: e1004215. doi:10.1371/journal.pgen.1004215

292. Gao H, Zhang Y, Han Y, Yang L, Liu X, Guo Z, et al. Phenotypic and transcriptional analysis of the osmotic regulator OmpR in Yersinia pestis. BMC Microbiol. 2011;11: 39. doi:10.1186/1471-2180-11-39

293. Condemine G, Ghazi A. Differential Regulation of Two Oligogalacturonate Outer Membrane Channels, KdgN and KdgM, of Dickeya dadantii (Erwinia chrysanthemi). J Bacteriol. 2007;189: 5955–5962. doi:10.1128/JB.00218-07

294. Rentschler AE, Lovrich SD, Fitton R, Enos-Berlage J, Schwan WR. OmpR regulation of the uropathogenic Escherichia coli fimB gene in an acidic/high osmolality environment. Microbiology. 2013;159: 316–327. doi:10.1099/mic.0.059386-0

295. Nieckarz M, Raczkowska A, Dębski J, Kistowski M, Dadlez M, Heesemann J, et al. Impact of OmpR on the membrane proteome of Yersinia enterocolitica in different environments: repression of major adhesin YadA and heme receptor HemR. Environ Microbiol. 2015; n/a-n/a. doi:10.1111/1462-2920.13165

296. Paster BJ, Dewhirst FE, Olsen I, Fraser GJ. Phylogeny of Bacteroides, Prevotella, and Porphyromonas spp. and related bacteria. J Bacteriol. 1994;176: 725–732.

297. Williams KP, Sobral BW, Dickerman AW. A Robust Species Tree for the Alphaproteobacteria. J Bacteriol. 2007;189: 4578–4586. doi:10.1128/JB.00269-07

298. Munoz R, Yarza P, Ludwig W, Euzéby J, Amann R, Schleifer K-H, et al. Release LTPs104 of the All-Species Living Tree. Syst Appl Microbiol. 2011;34: 169–170. doi:10.1016/j.syapm.2011.03.001

299. Gao B, Gupta RS. Phylogenetic Framework and Molecular Signatures for the Main Clades of the Phylum Actinobacteria. Microbiol Mol Biol Rev. 2012;76: 66–112. doi:10.1128/MMBR.05011-11

300. Yutin N, Galperin MY. A genomic update on clostridial phylogeny: Gram-negative spore formers and other misplaced clostridia. Environ Microbiol. 2013;15: 2631–2641. doi:10.1111/1462-2920.12173

301. Wattam AR, Abraham D, Dalay O, Disz TL, Driscoll T, Gabbard JL, et al. PATRIC, the bacterial bioinformatics database and analysis resource. Nucleic Acids Res. 2014;42: D581-591. doi:10.1093/nar/gkt1099

302. Ramulu HG, Groussin M, Talla E, Planel R, Daubin V, Brochier-Armanet C. Ribosomal proteins:

Toward a next generation standard for prokaryotic systematics? Mol Phylogenet Evol. 2014;75: 103–117. doi:10.1016/j.ympev.2014.02.013

303. Jiang X, Zghidi-Abouzid O, Oger-Desfeux C, Hommais F, Greliche N, Muskhelishvili G, et al. Global transcriptional response of Dickeya dadantii to environmental stimuli relevant to the plant infection. Environ Microbiol. 2016; doi:10.1111/1462-2920.13267

304. Henrichsen J. Bacterial surface translocation: a survey and a classification. Bacteriol Rev. 1972;36: 478.

305. Lane MC, Simms AN, Mobley HLT. Complex Interplay between Type 1 Fimbrial Expression and Flagellum-Mediated Motility of Uropathogenic Escherichia coli. J Bacteriol. 2007;189: 5523–5533. doi:10.1128/JB.00434-07

306. Sauvonnet N, Gounon P, Pugsley AP. PpdD type IV pilin of Escherichia coli K-12 can Be assembled into pili in Pseudomonas aeruginosa. J Bacteriol. 2000;182: 848–854.

307. Harshey RM. Bacterial Motility on a Surface: Many Ways to a Common Goal. Annu Rev Microbiol. 2003;57: 249–273. doi:10.1146/annurev.micro.57.030502.091014

308. Parker WL, RATHNUM ML, WELLS JJS, TREJO WH, PRINCIPE PA, SYKES RB. SQ 27,860, a simple carbapenem produced by species of Serratia and Erwinia. J Antibiot (Tokyo). 1982;35: 653–660.