



THÈSE

en vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

délivré par :

l'Université Toulouse 3 Paul Sabatier

présentée et soutenue par :

Christophe GARCIE

le mercredi 14 décembre 2016

Titre :

**Modulation atypique de la biosynthèse de la colibactine,
une génotoxine de *Escherichia coli*,
ou comment un îlot génomique est en symbiose avec le chromosome bactérien**

École doctorale et discipline ou spécialité :

ED Biologie-Santé-Biotechnologies : Microbiologie

Unité de recherche :

Institut de Recherche en Santé Digestive (IRSD)
INSERM UMR1220 – INRA UMR1416 – ENVT – Université Toulouse 3

Directeur/trice(s) de Thèse :

Dr. Patricia MARTIN
Pr. Éric OSWALD

Jury :

Dr. Catherine SCHOULER, DR, Tours
Pr. Stéphane BONACORSI, PU-PH, Paris
Dr. Patricia MARTIN, Chercheur, Toulouse
Pr. Éric OSWALD, PU-PH, Toulouse

Président du jury – Rapporteur
Rapporteur
Co-directrice de Thèse
Co-directeur de Thèse



THÈSE

en vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

délivré par :

l'Université Toulouse 3 Paul Sabatier

présentée et soutenue par :

Christophe GARCIE

le mercredi 14 décembre 2016

Titre :

**Modulation atypique de la biosynthèse de la colibactine,
une génotoxine de *Escherichia coli*,
ou comment un îlot génomique est en symbiose avec le chromosome bactérien**

École doctorale et discipline ou spécialité :

ED Biologie-Santé-Biotechnologies : Microbiologie

Unité de recherche :

Institut de Recherche en Santé Digestive (IRSD)
INSERM UMR1220 – INRA UMR1416 – ENVT – Université Toulouse 3

Directeur/trice(s) de Thèse :

Dr. Patricia MARTIN
Pr. Éric OSWALD

Jury :

Dr. Catherine SCHOULER, DR, Tours
Pr. Stéphane BONACORSI, PU-PH, Paris
Dr. Patricia MARTIN, Chercheur, Toulouse
Pr. Éric OSWALD, PU-PH, Toulouse

Président du jury – Rapporteur
Rapporteur
Co-directrice de Thèse
Co-directeur de Thèse

REMERCIEMENTS

À Madame le Docteur Catherine Schouler, président du jury et rapporteur

Je vous remercie d'avoir aimablement accepté de juger ce travail de thèse en tant que rapporteur, et de me faire l'honneur de présider ce jury. Veuillez trouver ici la marque de mon plus grand respect.

À Monsieur le Professeur Stéphane Bonacorsi, rapporteur

Je vous remercie d'avoir aimablement accepté d'être rapporteur de cette thèse, et de me faire l'honneur de participer à ce jury. Soyez assuré de ma profonde reconnaissance.

À Madame le Docteur Patricia Martin, co-directrice de thèse

Je te remercie d'avoir, un jour, accepté d'être ma co-directrice de thèse. Pour ta disponibilité sans faille, tes précieux conseils, ta rigueur, ton encadrement au quotidien, et pour ta relecture attentive de cette thèse, du premier au dernier mot, un immense merci ! Trouve ici toute ma gratitude et mon respect sincère.

À Monsieur le Professeur Éric Oswald, co-directeur de thèse

Je vous remercie de m'avoir accueilli dans votre équipe de recherche pour réaliser mon Master 2 Recherche puis ma thèse d'Université, et pour avoir accepté de co-diriger cette thèse. Merci pour votre encadrement, votre disponibilité et votre confiance.

Je tiens également à remercier :

Tous les membres de l'équipe CPTP11, devenu l'équipe IRSD2, pour m'avoir accueilli dans le monde de la recherche en 2011, et pour m'avoir accompagné, soutenu, aidé, conseillé, ... Vive *Escherichia coli* ! Et vive la colibactine (que l'on isolera... un jour) !

Les personnes avec qui j'ai eu la chance de collaborer pendant ma thèse : Pierre Genevaux, pour son expertise sur les *Heat shock proteins*, Fabien Létisse pour nos échanges sur le métabolisme des polyamines, Alexander Brachmann pour ses connaissances sur les PKS-NRPS (mais aussi pour son accueil lors de mon séjour à l'ETH à Zurich) et l'équipe de Jörn Piel, et enfin Amélie Garénaux et Sébastien Houle de l'équipe de Charles Dozois au Canada pour leur expertise sur la quantification des sidérophores.

L'ensemble des personnes ayant contribué au cours de mon cursus à ma formation, loin d'être achevée, de pharmacien, de biologiste médical et de chercheur.

Tout le personnel du Plateau Technique Infectiologie du CHU de Toulouse, en particulier les biologistes et les techniciens du laboratoire de Bactériologie-Hygiène.

Tous mes amis rencontrés au cours de mes études (ma petite Châtaigne en particulier, je te souhaite le meilleur), de mon internat (Agnès, Alexandre, Aude, Camille, Céline, David, Guillaume, Jérôme, Julie, Marion, Michaël et Sarah), de mon M2R (Alix, Elise et Stevie) ou ailleurs (François, Meryem, Nina, Pierre et Sandrinette).

Caroline, Johann et Louis, mes amis si précieux. Merci de faire partie de ma vie. Que nos routes se suivent toujours, malgré la distance. Je vous souhaite le meilleur.

Tous les membres de ma famille, en particulier mes grands-parents, pour leur amour. Et les membres de ma belle-famille, pour leur gentillesse, leur générosité et leur soutien.

Mon petit frère adoré, Florent. Je suis fier du jeune homme que tu deviens jour après jour. Crois en tes projets, ne laisse personne te détourner de tes objectifs, et tu verras que rien n'est impossible. Je serai à tes côtés pour t'aider de mon mieux à accomplir tes projets. Ze t'M.

Mes parents, vous sans qui rien ne serait possible, merci pour votre amour et votre soutien. Merci pour tout ce que vous faites pour moi, merci de m'avoir permis d'aller toujours plus loin. Et parce qu'on ne le dit jamais assez, je vous aime !

Dani ma chérie, pour ton amour et ton soutien sans faille depuis 11 ans... Merci de m'avoir soutenu pour : un premier concours, un deuxième concours, une première thèse, une deuxième thèse, et j'en passe... Promis, je m'arrête là !!! Merci de me combler de bonheur tous les jours ! Merci de me supporter au quotidien ! Merci d'être l'élément essentiel à mon équilibre ! Et merci pour ce cadeau de la vie ! Je suis fier, comblé et heureux d'être ton chéri doudou ! Je t'aime plindé fort mon grand chaton !

Et enfin, mon petit chaton, Mathilde. Tu es ce que j'ai de plus précieux au monde. A l'heure où j'écris ces lignes, tu ne sais pas encore lire mais tu commences déjà à être bien bavarde ! Ta venue m'a permis de beaucoup relativiser et de me recentrer sur l'essentiel. Tu n'imagines pas à quel point mon rôle de papa me rend heureux et épanouit. Et je t'aime plus que le ciel !

TABLE DES MATIÈRES

TABLE DES ILLUSTRATIONS.....	10
TABLE DES SIGLES ET ABRÉVIATIONS	12
PRÉAMBULE.....	15
SYNTHÈSE BIBLIOGRAPHIQUE.....	17
CHAPITRE I : <i>Escherichia coli</i>, UNE BACTÉRIE AUX MULTIPLES FACETTES	19
1. <i>Escherichia coli</i> : une espèce versatile	19
2. Diversité génétique	21
3. Plasticité génomique.....	22
3.1. Transferts génétiques horizontaux.....	23
3.2. Éléments génétiques mobiles	24
3.3. Focus sur les îlots génomiques	26
CHAPITRE II : MÉTABOLITES SECONDAIRES DE TYPE POLYCÉTIDES ET PEPTIDES NON-RIBOSOMAUx SYNTHÉTISÉS PAR <i>Escherichia coli</i>	27
1. Introduction aux polycétides et peptides non-ribosomaux.....	27
1.1. Métabolites secondaires	27
1.2. Synthèse des composés PK et NRP	29
1.2.1. Principe général.....	29
1.2.2. Structure des machineries enzymatiques.....	31
1.2.3. Logique réactionnelle.....	32
1.2.4. Taille des machineries enzymatiques.....	34
2. Les sidérophores, des systèmes de captation du fer	35
2.1. Généralités sur l'acquisition du fer par les bactéries	35
2.2. Étapes d'acquisition du fer par les sidérophores.....	36
2.3. Rôle des différents sidérophores dans la virulence de <i>E. coli</i>	39
3. La colibactine, une génotoxine bactérienne	41
3.1. Données moléculaires.....	41
3.2. Éléments de biosynthèse.....	42
3.3. Relations structure-activité.....	46
3.4. Phénotypes associés.....	46
3.5. Impacts sur la santé humaine	48

CHAPITRE III : CONTRÔLE QUALITÉ DES PROTÉINES CHEZ *Escherichia coli* 49

1. Principes généraux.....	49
2. Contrôle qualité des protéines néo-synthétisées	51
2.1. Machinerie de traduction	51
2.2. Trigger Factor	51
2.3. Système Hsp70 : complexe [DnaK–DnaJ–GrpE]	53
2.4. Système Hsp60 : complexe [GroEL–GroES]	55
2.5. Coopération entre les systèmes TF, DnaKJE et GroESL.....	57
3. Contrôle qualité des protéines en situations de stress.....	59
3.1. Mesures préventives	60
3.1.1. Prévention de la dénaturation.....	60
3.1.2. Prévention de l'oxydation	60
3.1.3. Modulation de l'agrégation	61
3.2. Mesures curatives	61
3.2.1. Désagrégation protéique	61
3.2.2. Repliement conformationnel.....	62
3.2.3. Élimination des protéines mal conformées.....	62
3.2.4. Élimination des agrégats protéiques.....	63
4. Focus sur la machinerie Hsp90	65
4.1. Généralités	65
4.2. Structure et cycle réactionnel.....	66
4.3. Fonctions cellulaires et protéines clientes	67
4.4. Collaboration avec d'autres systèmes chaperons.....	70
4.5. Pathologies associées et cible thérapeutique.....	71

CHAPITRE IV : MÉTABOLISME ET FONCTIONS DES POLYAMINES CHEZ *Escherichia coli* 73

1. Notions générales.....	73
1.1. Structure des polyamines.....	73
1.2. Distribution des polyamines	74
2. Métabolisme des polyamines.....	74
2.1. Biosynthèse.....	74
2.2. Transport.....	76
2.3. Régulation de l'homéostasie intracellulaire des polyamines.....	77
3. Fonctions des polyamines.....	78
3.1. Régulation de la traduction	78
3.2. Incorporation dans la structure de métabolites secondaires	79

RÉSULTATS EXPÉRIMENTAUX 81

RATIONNEL 83

CHAPITRE I : LA PROTÉINE CHAPERON HSP90_{Ec} EST NÉCESSAIRE À LA BIOSYNTHÈSE DE COLIBACTINE ET DE YERSINIABACTINE 87

Article 1. The bacterial stress-responsive Hsp90 chaperone (HtpG) is required for the production of the genotoxin colibactin and the siderophore yersiniabactin in *Escherichia coli* 89

CHAPITRE II : LA SPERMIDINE MODULE L'ACTIVITÉ GÉNOTOXIQUE DE LA COLIBACTINE 101

Article 2. Connection between primary and secondary metabolism in *Escherichia coli*: the polyamine spermidine modulates the production of the genotoxin colibactin 103

DISCUSSION & PERSPECTIVES 127

À LA DÉCOUVERTE DE LA RÉGULATION DE L'ÎLOT *pks* 129

UN ÎLOT *pks*, MAIS DES COLIBACTINES 133

ÉVOLUTION DE DEUX ÎLOTS INTERCONNECTÉS 135

PERSPECTIVES 137

RÉFÉRENCES BIBLIOGRAPHIQUES 141

ANNEXES 159

Article 3. Interplay between siderophores and colibactin genotoxin biosynthetic pathways in *Escherichia coli* 161

Article 4. Colibactin biosynthesis and biological activity depend on the rare aminomalonyl polyketide precursor 175

Article 5. Iron homeostasis regulates the genotoxicity of *Escherichia coli* that produces colibactin 179

ABSTRACT 191

TABLE DES ILLUSTRATIONS

Figures

<u>Figure 1.</u> Tropisme des différents pathovars de <i>Escherichia coli</i>	20
<u>Figure 2.</u> Diversité phylogénétique des souches de <i>E. coli</i>	21
<u>Figure 3.</u> Plasticité génomique de <i>E. coli</i>	23
<u>Figure 4.</u> Mécanismes de transferts génétiques horizontaux.....	24
<u>Figure 5.</u> Contribution des éléments génétiques mobiles à l'évolution de l'espèce <i>E. coli</i>	25
<u>Figure 6.</u> Structure générale et caractéristiques communes des îlots génomiques	26
<u>Figure 7.</u> Exemples de composés de type PK, NRP et PK-NRP utilisés en thérapeutique.....	28
<u>Figure 8.</u> Organisation générale des mégasynthèses de type PKS ou NRPS	29
<u>Figure 9.</u> Cycle réactionnel des mégasynthèses PKS et NRPS	33
<u>Figure 10.</u> Structure des sidérophores synthétisés par <i>E. coli</i>	36
<u>Figure 11.</u> Étapes de l'acquisition du fer par les sidérophores	37
<u>Figure 12.</u> Organisation génique des <i>clusters</i> codant les sidérophores de <i>E. coli</i>	37
<u>Figure 13.</u> Organisation de l'îlot génomique <i>pks</i>	41
<u>Figure 14.</u> Biosynthèse de la pré-colibactine	43
<u>Figure 15.</u> Modèle de synthèse et d'export de la colibactine	45
<u>Figure 16.</u> Structure des pré-colibactines.....	46
<u>Figure 17.</u> Dommages cellulaires induits par des souches produisant la colibactine	47
<u>Figure 18.</u> Principes généraux du contrôle qualité protéique chez <i>E. coli</i>	50
<u>Figure 19.</u> Structure et cycle réactionnel de la protéine chaperon Trigger Factor.....	52
<u>Figure 20.</u> Structure et cycle réactionnel du système chaperon Hsp70 (DnaKJE).....	54
<u>Figure 21.</u> Structure et cycle réactionnel du système chaperon Hsp60 (GroESL).....	56
<u>Figure 22.</u> Vue d'ensemble du contrôle qualité des protéines néo-synthétisées	58

<u>Figure 23.</u> Vue d'ensemble du contrôle qualité protéique en situations de stress.....	59
<u>Figure 24.</u> Mécanisme de désagrégation protéique par ClpB et DnaKJE	62
<u>Figure 25.</u> Mécanisme de dégradation protéique par les protéases.....	63
<u>Figure 26.</u> Élimination des agrégats protéiques par réorganisation asymétrique	64
<u>Figure 27.</u> Structure et cycle réactionnel de la protéine chaperon Hsp90 _{Ec}	66
<u>Figure 28.</u> Modèle du contrôle qualité des protéines dénaturées intégrant Hsp90 _{Ec}	70
<u>Figure 29.</u> Exemples de structure de polyamines.....	73
<u>Figure 30.</u> Métabolisme des polyamines chez <i>E. coli</i>	75
<u>Figure 31.</u> Systèmes de transport des polyamines chez <i>E. coli</i> [210]	76
<u>Figure 32.</u> Mécanismes de stimulation de la traduction par les polyamines chez <i>E. coli</i>	79
<u>Figure 33.</u> Exemples de métabolites secondaires comportant un motif spermidine	80
<u>Figure 34.</u> Protocole de construction, criblage et analyse de la banque de mutants.....	84
<u>Figure 35.</u> Principes généraux du maintien de l'homéostasie protéique	87
<u>Figure 36.</u> Schéma-bilan : régulateurs et métabolites associés à l'îlot génomique <i>pks</i>	134
<u>Figure 37.</u> Interconnexions entre les îlots génomiques <i>pks</i> et HPI.....	135

Tableaux

<u>Tableau I.</u> Domaines et monomères des mégasynthèses de type PKS et NRPS	30
<u>Tableau II.</u> Composants spécifiques des quatre systèmes de sidérophores de <i>E. coli</i>	37
<u>Tableau III.</u> Distribution des polyamines chez <i>E. coli</i> [205].....	74
<u>Tableau IV.</u> Synthèse de l'analyse des mutants d'intérêt situés hors de l'îlot <i>pks</i>	86

TABLE DES SIGLES ET ABRÉVIATIONS

- σ : facteur sigma
- A : Adénylation
- ABC : *ATP-Binding Cassette*
- ACP : *Acyl Carrier Protein*
- ADN : Acide DésoxyriboNucléique
- ADP : Adénosine DiPhosphate
- AMP : Adénosine MonoPhosphate
- ANOVA : *ANalysis Of VAriance*
- APEC : *Avian Pathogenic E. coli*
- ARN : Acide RiboNucléique
- ARNm : ARN messenger
- ARNt : ARN de transfert
- Asn : Asparagine
- AT : Acyltransférase
- ATP : Adénosine TriPhosphate
- ATPase : Adénosine TriPhosphatase
- C : Condensation
- Cas : *CRISPR-associated protein*
- CFU : *Colony Forming Unit*
- CoA : Coenzyme A
- CRC : *ColoRectal Cancer*
- CRISPR : *Clustered Regularly Interspaced Short Palindromic Repeat*
- CTD : *C-Terminal Domain*
- Cy : Cyclodéshydratase
- DAEC : *Diffusely Adherent E. coli*
- DH : DésHydratase
- DMEM : *Dulbecco's Modified Eagle's Medium*
- DNA : *DesoxyriboNucleic Acid*
- DnaKJE : complexe [DnaK–DnaJ–GrpE]
- DR : *Direct Repeat*
- E : Épimérase
- *E. coli* : *Escherichia coli*
- EAEC : *EnteroAggregative E. coli*
- EcN : *E. coli* Nissle 1917
- ECOR : *E. coli* *COllection Reference*
- EHEC : *EnteroHaemorrhagic E. coli*
- EIEC : *Enteroinvasive E. coli*
- EPEC : *EnteroPathogenic E. coli*
- ER : ÉnoylRéductase
- ETEC : *EnteroToxicogenic E. coli*
- ExPEC : *Extraintestinal Pathogenic E. coli*
- FCS : *Fetal Calf Serum*
- Fur : *Ferric uptake regulatory protein*
- GroESL : complexe [GroES–GroEL]
- GTP : Guanosine TriPhosphate
- H-NS : *Histone-like Nucleoid-Structuring Protein*
- HPI : *High Pathogenicity Island*
- Hsp : *Heat shock protein*
- HtpG : *High-temperature protein G*
- IPEC : *Intestinal Pathogenic E. coli*
- IS : *Insertion Sequence*
- kb : kilobases
- kDa : kiloDaltons
- KR : KétoRéductase
- KS : KétoSynthase
- LB : *Lysogeny Broth*

-
- LC/MS : *Liquid Chromatography/Mass Spectrometry*
 - LEE : *Locus d'Effacement des Entérocytes*
 - MD : *Middle Domain*
 - MDa : *MégaDaltons*
 - MFS : *Major Facilitation Superfamily*
 - MLEE : *MultiLocus Enzyme Electrophoresis*
 - MLST : *MultiLocus Sequence Typing*
 - MOI : *Multiplicity Of Infection*
 - NBD : *Nucleotide Binding Domain*
 - NGAL : *Neutrophil Gelatinase-Associated Lipocalin*
 - NMEC : *Neonatal Meningitis E. coli*
 - NRP : *NonRibosomal Peptide*
 - NRPS : *NonRibosomal Peptide Synthetase*
 - NS : *NonSignificant*
 - NTD : *N-Terminal Domain*
 - OD : *Optical Density*
 - Ox : *Oxydase*
 - PBS : *Phosphate-Buffered Saline*
 - PCP : *Peptidyl Carrier Protein*
 - PCR : *Polymerase Chain Reaction*
 - PK : *PolyKetide*
 - PKS : *PolyKetide Synthase*
 - ppGpp : *Guanosine tétraphosphate*
 - PPlase : *Peptidyl-Propyl Isomérase*
 - PPTase : *PhosphoPantéthéinyl Transférase*
 - qRT-PCR : *quantitative Reverse Transcription-PCR*
 - RLU : *Relative Luminescence Unit*
 - RNA : *RiboNucleic Acid*
 - RND : *Resistance Nodulation and cell Division*
 - ROS : *Reactive Oxygen Species*
 - RT : *Reverse-Transcriptase*
 - SASP : *Senescent-Associated Secretory Pattern*
 - SBD : *Substrate-Binding Domain*
 - SD : *Shine-Dalgarno*
 - SEM : *Standard Errors of the Mean*
 - sHsp : *small Heat shock protein*
 - SHU : *Syndrome Hémolytique et Urémique*
 - sRNA : *small RNA*
 - STEC : *Shiga Toxin-producing E. coli*
 - TCS : *Two-Component System*
 - TE : *ThioEstérase*
 - TF : *Trigger Factor*
 - UPEC : *UroPathogenic E. coli*
 - WT : *Wild-Type*

PRÉAMBULE

Escherichia coli (*E. coli*) est une espèce bactérienne à la fois commensale du tube digestif de l'Homme et pathogène responsable de diverses infections [1]. La diversité phénotypique de cette souche s'explique par l'acquisition de nombreux facteurs d'adaptation et de virulence portés par des éléments génétiques mobiles, dont les îlots génomiques [1].

Notre équipe de recherche a identifié en 2006 un îlot génomique, l'îlot *pks*, qui code une machinerie enzymatique complexe synthétisant la colibactine, une génotoxine produite par certaines souches de *E. coli* [2]. Cette génotoxine induit des cassures double-brin de l'ADN sur les cellules eucaryotes *in vitro* et *in vivo* [2,3]. La colibactine n'est pas une protéine, mais un métabolite secondaire de type polycétide/peptide non-ribosomal.

L'objectif de cette thèse était d'élucider le mode de régulation de cet îlot génomique, dans le but d'identifier d'éventuelles nouvelles cibles thérapeutiques antibactériennes. Des résultats préliminaires de l'équipe semblaient indiquer que certains gènes du *core genome* de *E. coli* moduleraient la production de la colibactine. Afin d'identifier les gènes non essentiels de *E. coli*, situés hors de l'îlot génomique *pks*, impliqués dans la synthèse de cette génotoxine, nous avons construit une banque de mutants par insertion de transposons. Ce criblage a permis d'identifier 29 gènes candidats, mais deux groupes de gènes ont été particulièrement étudiés dans le cadre de cette thèse : trois gènes codants des protéines chaperons, et trois gènes codant des enzymes du métabolisme des polyamines.

Ce travail se propose donc d'étudier, d'une part le rôle des protéines chaperons, et d'autre part l'implication du métabolisme des polyamines, dans la production de la colibactine.

Ce manuscrit de thèse s'articule donc autour : d'une synthèse bibliographique portant sur la versatilité de l'espèce *E. coli*, les métabolites secondaires de type polycétides/peptides non-ribosomiaux qu'elle synthétise, le contrôle qualité des protéines assuré par les chaperons moléculaires et les protéases pour maintenir l'homéostasie protéique, et le métabolisme et les fonctions des polyamines ; d'une présentation des résultats expérimentaux, sous forme d'article ou d'ébauche de manuscrit ; et d'une discussion générale des résultats venant compléter les points qui auront déjà été évoqués dans la partie précédente. Enfin, les publications auxquelles j'ai participé en marge de ce travail sont annexées en fin de manuscrit.

SYNTHÈSE
BIBLIOGRAPHIQUE

CHAPITRE I : *Escherichia coli*, UNE BACTÉRIE AUX MULTIPLES FACETTES

1. *Escherichia coli* : une espèce versatile

Escherichia coli a été découvert en 1885 par le pédiatre Theodor Escherich dans des selles de nourrissons. D'abord appelée *Bacterium coli commune*, elle a été rebaptisée *Escherichia coli* en 1919 [4]. *E. coli* est une bactérie mobile à Gram négatif qui appartient à la famille des *Enterobacteriaceae*, au sein du phylum des γ -Protéobactéries.

Hôte commun de la microflore intestinale de l'Homme et des animaux à sang chaud, *E. coli* s'établit dans le tractus digestif dès les premières heures après la naissance [1]. Cette espèce a adapté son métabolisme à cette niche écologique, parvenant ainsi à s'implanter durablement face à plus de 500 autres espèces bactériennes [5]. Elle perdure ensuite tout au long de la vie d'un individu en tant que bactérie anaérobie facultative sous-dominante du microbiote intestinal.

E. coli n'est pas seulement une bactérie commensale du tube digestif, puisqu'elle est capable de survivre dans de nombreux habitats écologiques différents. Grâce à leurs capacités d'adaptation à l'environnement, certaines souches de *E. coli* peuvent être retrouvées dans le sol, l'eau, les sédiments ou les aliments [6]. D'autres souches sont devenues pathogènes par l'acquisition de facteurs de virulence, et sont responsables de diverses pathologies infectieuses de gravité variable [1].

Les souches de *E. coli* pathogènes peuvent être divisées en deux catégories, selon que les manifestations infectieuses sont intra- ou extra-intestinales. Chaque catégorie peut être subdivisée en pathovars, définis comme des groupements de souches pathogènes d'une même espèce catégorisées en fonction des symptômes occasionnés et de caractéristiques pathogéniques propres (Figure 1) [1,7] :

- les souches pathogènes intestinales (IPEC) sont subdivisées en six pathovars : entéropathogènes (EPEC), entérohémorragiques (EHEC), entéroinvasives (EIEC), entérotoxinogènes (ETEC), entégroagrégatives (EAEC) et à adhésion diffuse (DAEC) ;
- les souches pathogènes extra-intestinales (ExPEC) sont classées en trois pathovars : uropathogènes (UPEC) et méningitiques néonatales (NMEC) chez l'Homme, et pathogènes aviaires (APEC) chez la volaille.

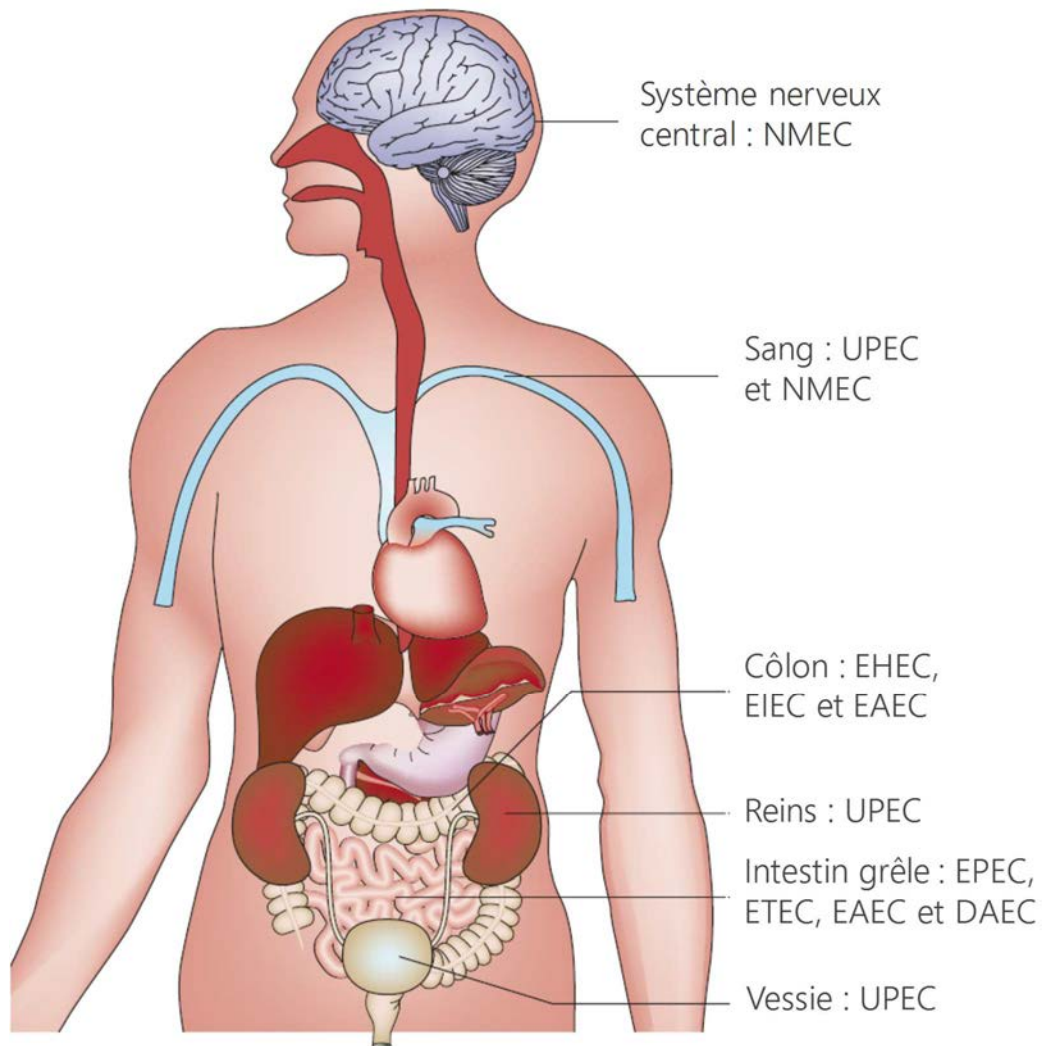


Figure 1. Tropisme des différents pathovars de *Escherichia coli*

Les souches de E. coli pathogènes peuvent être responsables d'infections intestinales (IPEC) ou extra-intestinales (ExPEC). Selon les symptômes occasionnés, elles sont classées en pathovars : EPEC = E. coli entéro-pathogènes ; EHEC = E. coli entéro-hémorragiques ; EIEC = E. coli entéro-invasives ; ETEC = E. coli entéro-toxinogènes ; EAEC = E. coli entéro-agrégatives ; DAEC = E. coli à adhésion diffuse ; UPEC : E. coli uropathogènes ; NMEC = E. coli méningitiques néonatales. D'après [8].

La versatilité phénotypique de cette espèce bactérienne résulte d'une importante diversité génétique, qui repose principalement sur la plasticité de son génome et sur l'acquisition de facteurs d'adaptation et de virulence portés par des éléments génétiques mobiles, acquis notamment grâce aux transferts génétiques horizontaux. Ces transferts génétiques, mais aussi les mutations spontanées, les réarrangements chromosomiques, ou encore la perte d'information génétique, sont autant d'éléments qui contribuent aux variations observées dans le contenu et la structure du génome de *E. coli*.

2. Diversité génétique

La première classification proposée pour l'espèce *E. coli* est basée sur la distribution non aléatoire des antigènes de surface, ou sérotypage, et permet de classer les souches en sérotypes selon la combinaison des antigènes somatiques O, flagellaires H et de capsule K [9].

Dans les années 1990, une collection de référence représentative de la diversité des souches de *E. coli*, appelée ECOR (pour *Escherichia coli collection reference*) est établie, puis analysée par méthode MLEE (pour *multilocus enzyme electrophoresis*) [10]. Les avancées technologiques en biologie moléculaire ont permis de faire évoluer cette technique vers la méthode MLST (pour *multilocus sequence typing*). L'analyse de la collection ECOR en MLST, grâce au séquençage de huit gènes de ménage, a permis de définir sept groupes phylogénétiques distincts : A, B1, B2, C, D, E et F (Figure 2) [11,12].

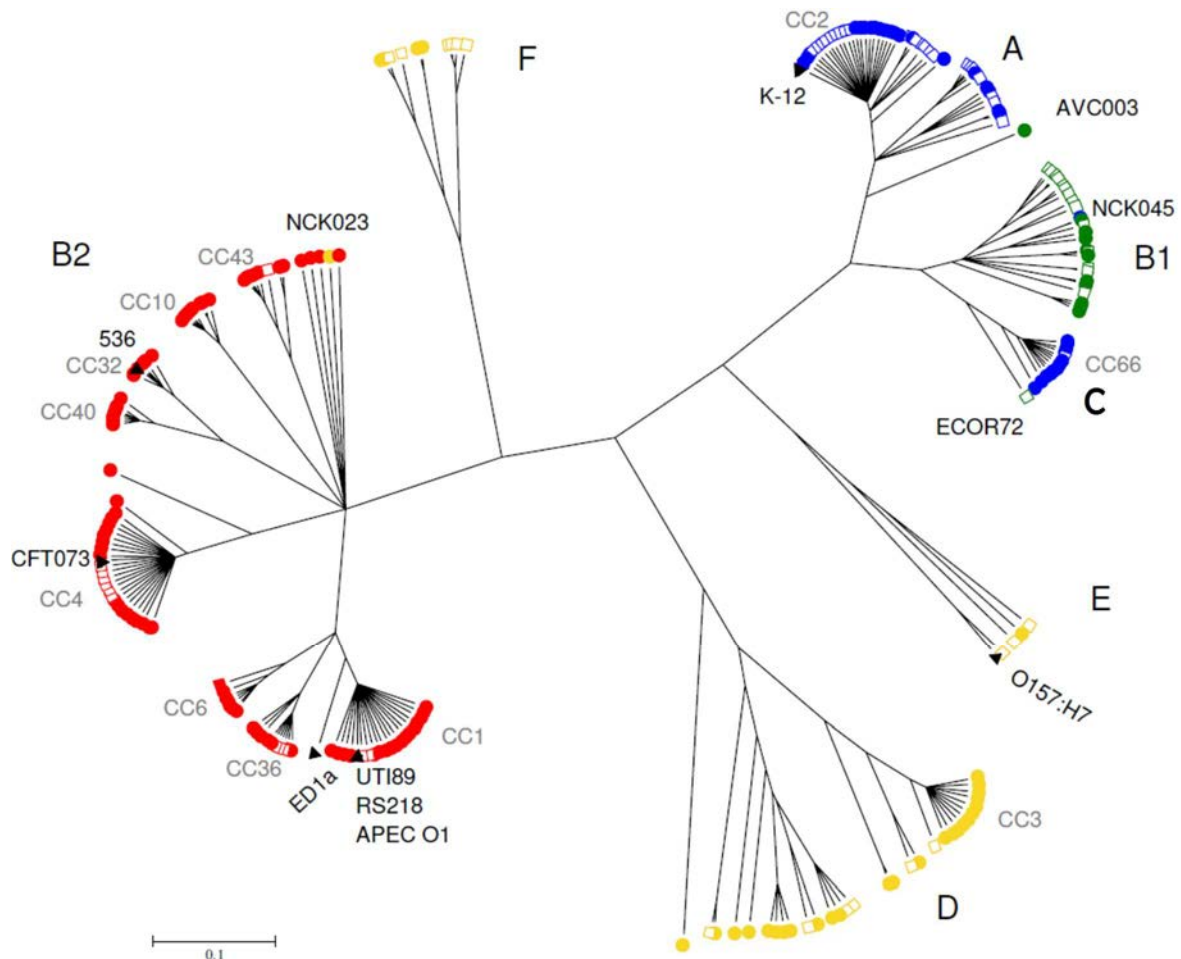


Figure 2. Diversité phylogénétique des souches de *E. coli*

Analyse phylogénétique par MLST de 161 souches isolées de bactériémies (cercles), 67 souches de la collection ECOR (carrés) et 7 souches de référence (triangles), définissant 7 phylogroupes [12].

De manière générale, les souches commensales et les souches pathogènes intestinales se répartissent au sein de l'ensemble de l'arbre phylogénétique [12]. Les souches responsables d'infections extra-intestinales appartiennent majoritairement au groupe B2 et minoritairement au groupe D. La prévalence des souches du groupe B2 ne cesse d'augmenter dans les flores intestinales des populations humaines des pays industrialisés et/ou à climat tempéré [13,14].

3. Plasticité génomique

De manière générale, le génome d'une espèce peut être scindé en deux catégories [15] :

- d'une part, le *core genome* est défini comme l'ensemble des gènes présents chez toutes les souches d'une même espèce bactérienne [16]. Il inclut notamment les gènes essentiels à la réplication, la transcription, la traduction... Le *core genome* est ainsi le pilier génomique de l'espèce, définissant ses fonctions métaboliques de base ;
- d'autre part, le *dispensable genome* comprend les gènes non-essentiels, ou accessoires, qui sont présents seulement chez certaines souches. Ces gènes sont responsables de la diversité phénotypique des souches et de leur capacité d'adaptation aux conditions environnementales. Ils sont pour la plupart portés par des éléments génétiques mobiles, parfois appelé *mobilome*, tels que des plasmides, des bactériophages, des transposons, ou encore des îlots génomiques [1].

La somme des gènes du *core genome* et du *flexible genome* définit le *pangenome*, c'est-à-dire le répertoire total des gènes d'une espèce, dont le nombre dépasse largement le contenu moyen du génome d'une souche donnée [15,17]. Selon plusieurs études récentes de comparaison de génomes de diverses souches de l'espèce *E. coli*, le *pangenome* comprend entre 16 000 et 18 000 gènes, alors que le *core genome* n'est constitué que de 1 000 à 2 000 gènes, soit seulement 5 à 10% du nombre total de gènes retrouvés au sein de l'ensemble des souches séquencées (Figure 3) [14,18,19]. Possédant en moyenne 4 700 gènes, une souche de *E. coli* ne peut donc représenter à elle seule l'ensemble des caractéristiques de l'espèce [20].

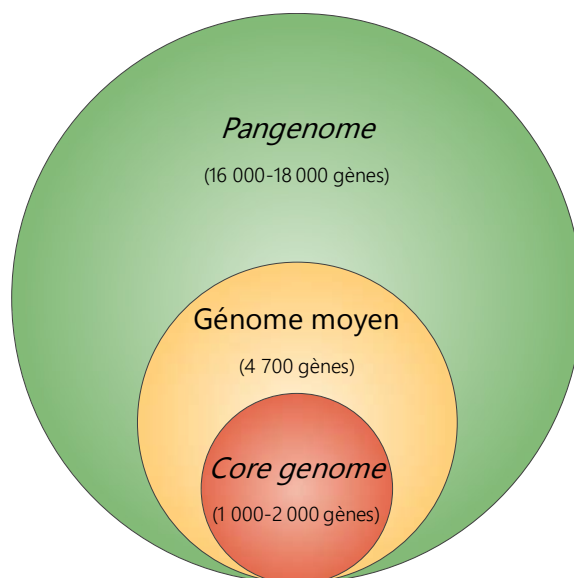


Figure 3. Plasticité génomique de *E. coli*

Une souche de *E. coli* contient en moyenne 4 700 gènes, combinaison des gènes du core genome et de gènes du dispensable genome, l'ensemble constituant le pangenome. D'après [20].

3.1. Transferts génétiques horizontaux

Les transferts génétiques horizontaux contribuent fortement à l'évolution des espèces bactériennes. L'acquisition de certains gènes lors de ces transferts peut radicalement modifier les capacités d'une souche à coloniser une niche écologique ou à provoquer une infection [21]. Ce matériel génétique étranger, véhiculé par un large éventail d'éléments génétiques mobiles, peut être acquis grâce à trois principaux mécanismes (Figure 4) [22] :

- la **transformation** correspond à la captation d'un fragment d'ADN libre présent dans le milieu extérieur à travers la membrane d'une bactérie dite compétente ;
- la **transduction** se traduit par l'inoculation par un bactériophage de matériel génétique exogène d'origine bactérienne dans une autre bactérie dite sensible ;
- la **conjugaison** est un transfert de matériel génétique d'une bactérie donneuse, possédant des pili sexuels, vers une bactérie réceptrice.

Le matériel génétique transféré doit alors franchir plusieurs étapes avant de s'intégrer durablement au sein du génome de la bactérie réceptrice [23] :

- être reconnu et toléré par les systèmes régulant l'acquisition d'ADN étranger ;
- s'établir dans le génome bactérien, soit par intégration dans le chromosome après recombinaison homologue, soit en tant qu'élément mobile à répllication autonome ;
- être non seulement exprimé mais surtout régulé en coordination avec les autres gènes impliqués dans les mêmes fonctions d'adaptation et/ou de virulence ;

- s'adapter au cours de l'évolution en fonction de la pression de sélection environnementale exercée sur la population bactérienne.

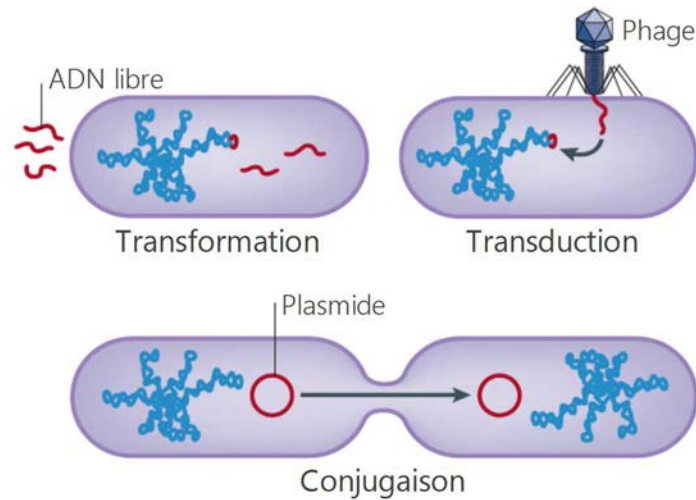


Figure 4. Mécanismes de transferts génétiques horizontaux

*Transformation : captation d'ADN libre dans le milieu par une bactérie compétente.
 Transduction : injection d'ADN bactérien étranger par l'intermédiaire d'un bactériophage.
 Conjugaison : transfert d'ADN lors d'un contact direct entre une bactérie donneuse, dotée de pili sexuels, et une bactérie réceptrice. D'après [24].*

3.2. Éléments génétiques mobiles

Le matériel génétique étranger acquis grâce aux transferts horizontaux est véhiculé par un large éventail d'éléments génétiques mobiles [22] :

- les **plasmides** sont définis comme des structures d'ADN double-brin circulaire à répllication autonome vis-à-vis du chromosome bactérien. Ils sont transmis à chaque cellule fille après division bactérienne. Leur taille varie de quelques kilobases (kb) à quelques centaines de kb. Ils sont transférés horizontalement par conjugaison entre bactéries d'une même espèce ou non. Ils codent une variété de gènes non essentiels, tels que des facteurs d'adaptation, de virulence ou de résistance aux antibiotiques ;
- les **transposons** sont des séquences d'ADN qui peuvent changer fréquemment leur localisation chromosomique ou être transférées vers des plasmides. Ils sont associés à une transposase qui catalyse les événements de transposition au sein du génome ;
- les **bactériophages** sont des virus infectant les bactéries et dont le génome peut s'incorporer au chromosome bactérien sous forme de prophage ;
- les **îlots génomiques** sont de longs segments d'ADN exogène capables de s'intégrer au chromosome bactérien et qui regroupent un bloc indissociable de gènes. Ils présentent certains caractères spécifiques qui seront détaillés plus loin (cf. § 3.3.).

Grâce à la plasticité de son génome, une souche de *E. coli* commensale peut acquérir de nombreux facteurs d'adaptation et de virulence portés par des éléments génétiques mobiles, qui lui permettent de coloniser de nouvelles niches et lui confèrent un pouvoir pathogène. La [Figure 5](#) présente le rôle des éléments génétiques mobiles dans l'émergence des différents pathovars de *E. coli* en fonction de l'acquisition de facteurs de virulence spécifiques [1].

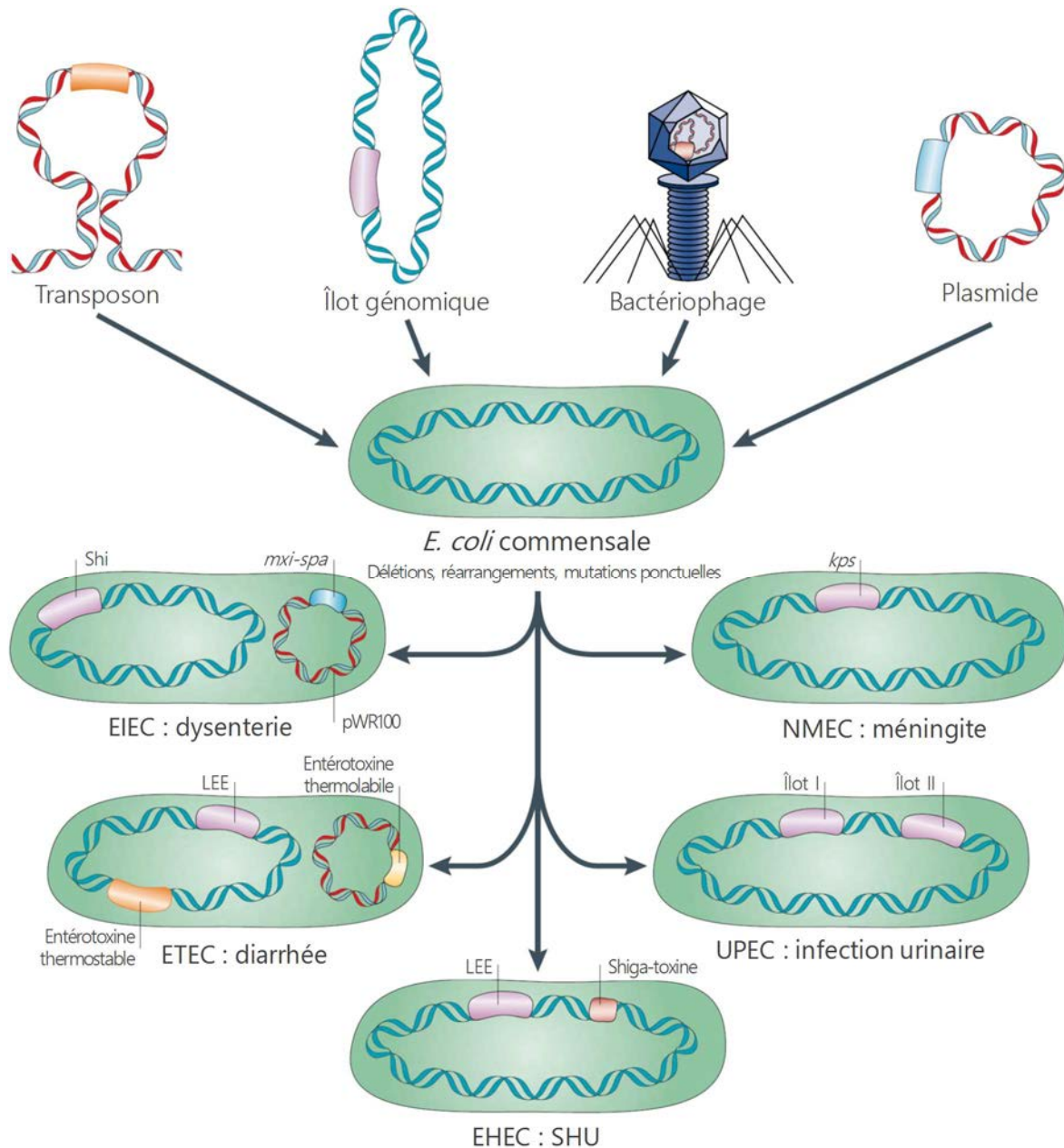


Figure 5. Contribution des éléments génétiques mobiles à l'évolution de l'espèce *E. coli*

De nombreux facteurs de virulence de *E. coli* sont codés par des éléments génétiques mobiles : transposons (ex. : entérotoxine thermostable des ETEC) ; plasmides (ex. : entérotoxine thermolabile des ETEC) ; bactériophages (ex. : Shiga toxine des EHEC) ; îlots génomiques (ex. : locus d'effacement des entérocytes [LEE] des EPEC/EHEC). Les souches commensales subissent également des délétions, des mutations ponctuelles et d'autres réarrangements génétiques pouvant contribuer à l'acquisition d'un pouvoir pathogène. L'ensemble de ces mécanismes sont à l'origine de l'émergence des différents pathovars de *E. coli*, responsables de diarrhée, dysenterie, syndrome hémolytique et urémique (SHU), infection urinaire ou encore méningite néonale [1].

3.3. Focus sur les îlots génomiques

Les îlots génomiques se définissent comme un bloc indissociable de gènes portés par le chromosome bactérien. Ces gènes codent divers facteurs d'adaptation ou de virulence, tels que des toxines, des adhésines ou des invasines. Les îlots génomiques possèdent les caractéristiques communes suivantes (Figure 6) [25–27] :

- leur taille excède 10 kb et peut aller jusqu'à 200 kb ;
- leur contenu en bases G+C, ainsi que les codons utilisés, diffèrent de ceux de la bactérie hôte, témoignant d'une origine étrangère ;
- ils présentent aux deux extrémités des séquences répétées (ou DR, pour *direct repeat*), générées lors d'évènements de recombinaison participant à leur intégration dans le génome de la bactérie hôte ;
- leur insertion chromosomique se situe généralement à proximité d'un gène d'ARN de transfert (ARNt), qui sont considérés comme des sites préférentiels d'insertion et d'échange d'éléments génétiques mobiles ;
- outre leurs gènes d'intérêt, ils comportent généralement des séquences d'insertion (ou IS, pour *insertion sequence*), complètes ou partielles, ainsi que des gènes de mobilité, cryptiques ou exprimés, codant des fonctions d'intégrase ou de transposase nécessaires aux mécanismes d'insertion et/ou d'excision chromosomique ;
- ils peuvent s'exciser du chromosome, à une fréquence variable selon les îlots.

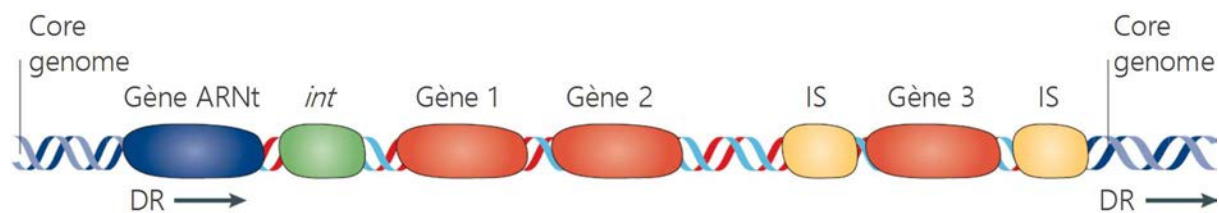


Figure 6. Structure générale et caractéristiques communes des îlots génomiques

Les îlots génomiques sont des séquences d'ADN acquises par transferts horizontaux et intégrées au chromosome bactérien à proximité d'un gène codant un ARNt. Ils sont classiquement flanqués de séquences répétées (DR) à leurs extrémités. Ils codent des gènes conférant des capacités d'adaptation ou de virulence à la bactérie hôte, des gènes de mobilité codant intégrase (*int*) et/ou transposase permettant leur insertion ou excision chromosomique, et de multiples séquences d'insertion (IS) [27].

Certains métabolites secondaires de *E. coli*, les sidérophores et la colibactine, sont des facteurs de virulence codés par des îlots génomiques, et font l'objet du prochain chapitre.

CHAPITRE II : MÉTABOLITES SECONDAIRES DE TYPE POLYCÉTIDES ET PEPTIDES NON-RIBOSOMAux SYNTHÉTISÉS PAR *Escherichia coli*

1. Introduction aux polycétides et peptides non-ribosomaux

1.1. Métabolites secondaires

Le métabolisme d'un organisme comprend le métabolisme primaire, essentiel à la croissance et au développement, et le métabolisme secondaire, qui est souvent nécessaire à l'interaction de l'organisme avec son environnement [28]. Les voies de biosynthèse du métabolisme secondaire aboutissent, à partir de composés issus du métabolisme primaire et par des réactions complexes et souvent hautement spécifiques, à des produits finis aux structures chimiques très élaborées et aux fonctions variées.

Les six voies principales de métabolisme secondaire synthétisent les terpénoïdes, les alcaloïdes, les peptides, les oligosaccharides, les **polycétides** (ou **PK** pour *polyketides*) et les **peptides non-ribosomaux** (ou **NRP** pour *nonribosomal peptides*) [29]. Des combinaisons entre ces six voies majeures peuvent se produire et permettent la synthèse de nouveaux composés sans créer de nouvelles voies de biosynthèse, comme par exemples les molécules hybrides de type PK-NRP [30].

Les structures de plus de 200 000 métabolites secondaires ont été élucidées à ce jour, fournissant une complexité et une diversité colossale de squelettes carbonés uniques. Les métabolites secondaires des micro-organismes, des plantes et des organismes marins forment ainsi un immense arsenal de produits naturels aux propriétés spécifiques, et sont à l'origine de nombreux médicaments [29]. En 1990, environ 80% des médicaments étaient en effet soit des produits naturels, soit des analogues hémi-synthétiques de ces derniers [31]. Ils appartiennent pour l'essentiel aux composés de type PK, NRP ou PK-NRP, tels que de nombreux antibiotiques, des antifongiques, des antiparasitaires, des immunosuppresseurs, des hypocholestérolémiantes et des médicaments anticancéreux (Figure 7). Aussi, malgré l'expansion de la chimie médicinale synthétique dans les années 1990, 50% des nouveaux médicaments restent aujourd'hui encore basés sur des produits naturels [32].

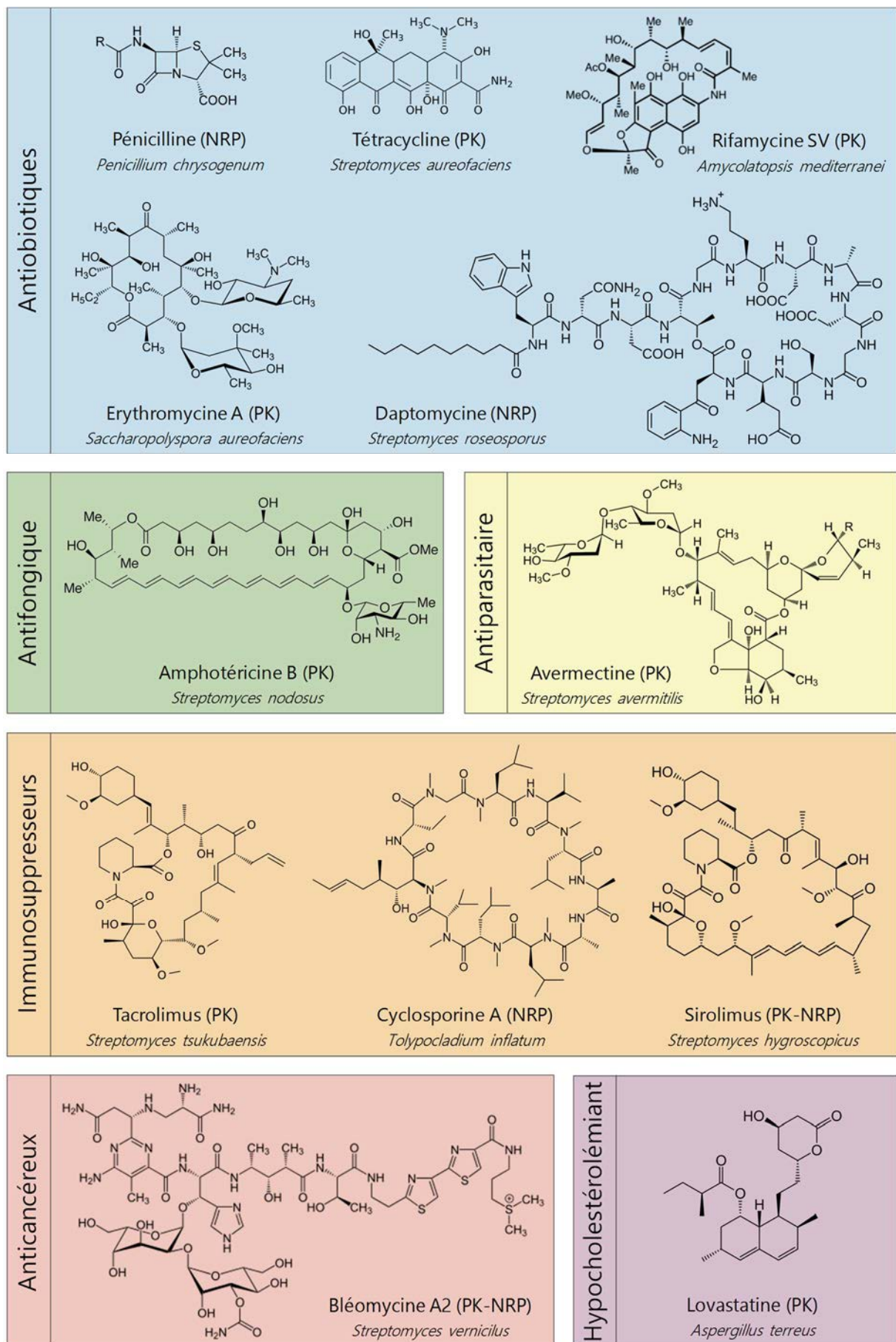


Figure 7. Exemples de composés de type PK, NRP et PK-NRP utilisés en thérapeutique
Le type de composé, le micro-organisme producteur et la classe thérapeutique sont indiqués.

1.2. Synthèse des composés PK et NRP

1.2.1. Principe général

La biosynthèse des composés PK et NRP a été principalement observée chez les micro-organismes bactériens et fongiques [33]. Chez *E. coli*, seuls quatre composés de cette famille ont été décrits à ce jour : des sidérophores, l'entérobactine (NRP), les salmochélines (NRP glycosylés) et la yersiniabactine (PK-NRP) [34], ainsi que la génotoxine colibactine (PK-NRP) [2].

Les composés de type PK, NRP et PK-NRP sont des macromolécules dont la synthèse est indépendante des ribosomes, mais est assurée par des complexes multi-enzymatiques, appelés **mégasynthèses**, de type **PKS** (pour *polyketide synthase*), **NRPS** (pour *nonribosomal peptide synthetase*) ou hybrides PKS-NRPS [35]. Les gènes de ces mégasynthèses, généralement organisés en *clusters* pour permettre une régulation coordonnée de la biosynthèse, codent un ou plusieurs **modules**, eux-mêmes subdivisés en **domaines**, chacun catalysant une réaction enzymatique spécifique (Figure 8) [36,37].

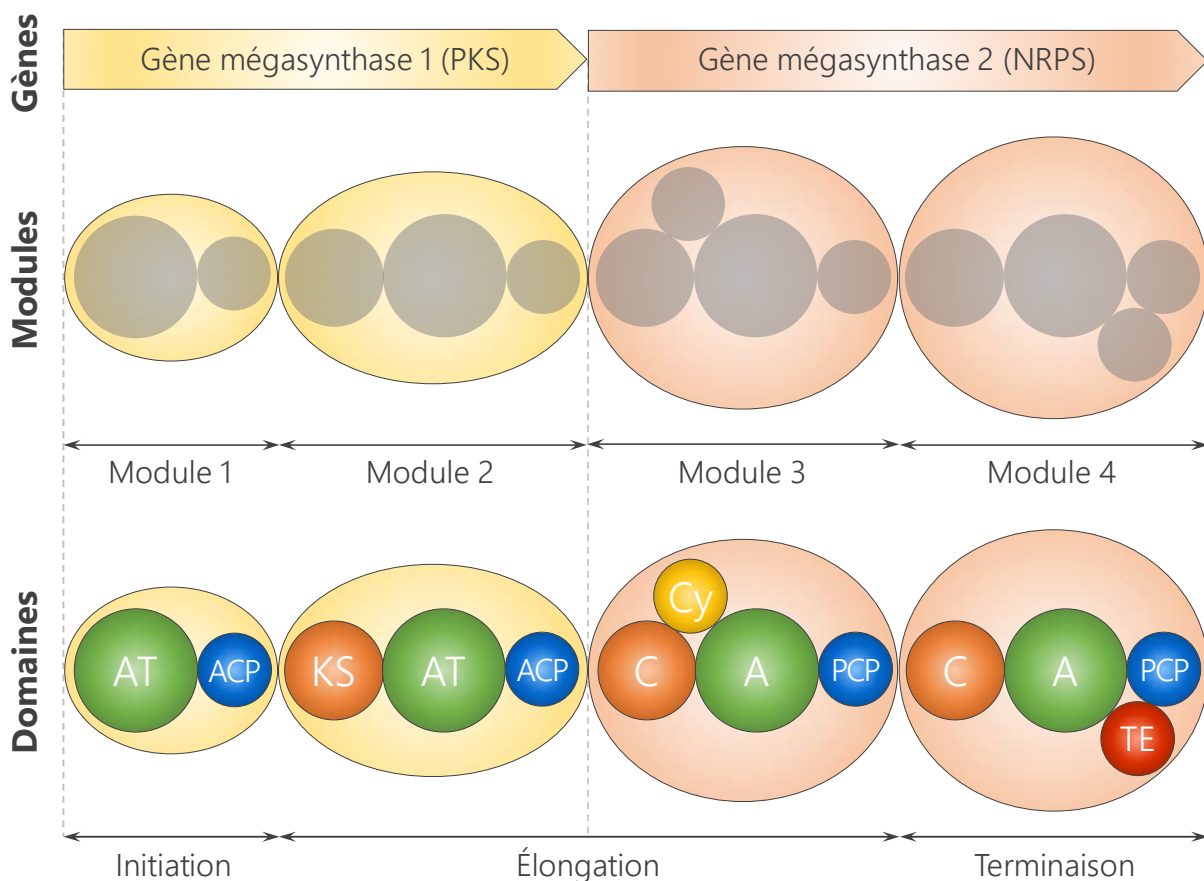


Figure 8. Organisation générale des mégasynthèses de type PKS ou NRPS

Les gènes des mégasynthèses codent des modules enzymatiques. Chaque module est organisé en plusieurs domaines. Chaque domaine possède une activité catalytique spécifique (voir aussi le Tableau I page suivante). D'après [37].

Les mégasynthèses recrutent et activent des **monomères** issus du métabolisme primaire, qui peuvent être des acides carboxyliques (pour les PK) ou des acides aminés (pour les NRP) (Tableau I). Elles assurent ensuite l'assemblage successif de ces monomères par condensation chimique, selon un ordre généralement colinéaire à l'organisation génique. Elles peuvent également modifier l'oligomère en cours de synthèse grâce à des domaines optionnels (Tableau I) [35]. En outre, des **enzymes accessoires**, monofonctionnelles et souvent codées dans le *cluster* de la machinerie, catalysent des modifications complémentaires des monomères et/ou des oligomères pendant ou après l'assemblage, favorisant la diversité des composés [35].

Tableau I. Domaines et monomères des mégasynthèses de type PKS et NRPS

	PKS		NRPS
Domaines de base	Acyl carrier protein (ACP)	Thiolation	Peptidyl carrier protein (PCP)
	Acyltransférase (AT)	Activation	Adénylation (A)
	Kétosynthase (KS)	Condensation	Condensation (C)
Domaines facultatifs	KR : ketoréductase DH : déshydratase ER : énoylréductase Thioestérase (TE)	Optionnels	E : épimérase Ox : oxydase Cy : cyclodéshydratase Thioestérase (TE)
	Monomères		
	Acétyl-CoA <chem>CC(=O)SCoA</chem>		Acides aminés - protéinogènes ou non - série -L ou -D
	Malonyl-CoA <chem>OC(=O)CC(=O)SCoA</chem>		Acides aryles <chem>Rc1ccc(cc1)C(=O)O</chem>

L'organisation de ces machineries enzymatiques évoque une chaîne d'assemblage d'automobiles : les monomères correspondent aux pièces détachées ; ils sont assemblés au niveau de différents postes de travail, les modules ; chaque module effectue plusieurs tâches, grâce aux différents domaines enzymatiques ; les intermédiaires de synthèse sont transférés d'un module au suivant, passant ainsi successivement par tous les postes de travail selon un ordre déterminé, jusqu'à obtenir le composé final. Celui-ci peut éventuellement être modifié en post-production par les enzymes accessoires.

Ainsi, la structure unique des composés PK et NRP repose sur le nombre de modules enzymatiques mis en jeu, l'ordre d'action des enzymes et les réactions chimiques catalysées par chaque enzyme dans et après la chaîne d'assemblage [35].

1.2.2. Structure des machineries enzymatiques

L'activité des mégasynthèses PKS, NRPS ou hybrides PKS-NRPS repose sur trois domaines de base, dont l'association triptyque définit un module enzymatique, et auxquels peuvent s'ajouter des domaines optionnels (Tableau I et Figure 8) [35,37] :

- le domaine de **thiolation** porte l'oligomère en cours d'élongation ;
- le domaine d'**activation** reconnaît et active spécifiquement le monomère à condenser ;
- le domaine de **condensation** assure le transfert du monomère activé sur l'oligomère ;
- les domaines **optionnels** modifient l'oligomère en cours d'élongation ;
- le domaine de **terminaison** libère l'oligomère synthétisé de la machinerie.

▪ Organisation structurale des PKS

Pour les PKS, le domaine de thiolation est appelé ACP (pour *acyl carrier protein*), le domaine d'activation est nommé AT (pour acyltransférase) et le domaine de condensation est annoté KS (pour kétosynthase) (Tableau I). L'ordre typique de ces trois domaines sur un module est **KS-AT-ACP**, mais le premier module (ou module d'initiation) d'une chaîne d'assemblage PKS ne comporte généralement pas de domaine KS (Figure 8). Les domaines optionnels catalysent des réactions de réduction (KR ou ER), de déshydratation (DH)... Le dernier module (ou module de terminaison) comporte un domaine thioestérase (TE), qui assure la libération de l'oligomère synthétisé par hydrolyse ou cyclisation (Tableau I) [35].

▪ Organisation structurale des NRPS

Pour les NRPS, le domaine de thiolation est appelé PCP (pour *peptidyl carrier protein*), le domaine d'activation est nommé A (pour adénylation) et le domaine de condensation est annoté C (pour condensation) (Tableau I). L'ordre typique des trois domaines sur un module est **C-A-PCP** (Figure 8). Le module d'initiation d'une chaîne d'assemblage NRPS ne comporte généralement pas de domaine C. Les domaines optionnels catalysent des réactions d'épimérisation (E), de cyclodéshydratation (Cy), d'oxydation (O)... Comme pour les PKS, le module de terminaison comporte un domaine TE à activité hydrolytique ou cyclisante, chargé de libérer l'oligomère néosynthétisé de la machinerie enzymatique (Tableau I et Figure 8) [37].

▪ Organisation structurale des NRPS

Les chaînes d'assemblage hybrides contiennent des modules enzymatiques PKS et NRPS, qui peuvent aussi coexister sur une même protéine, alors appelée PKS-NRPS [35].

1.2.3. Logique réactionnelle

▪ **Activation des mégasynthèses**

Avant toute étape de biosynthèse, les mégasynthèses doivent être préalablement activées post-traductionnellement, par fixation d'un résidu pantéthéinyl terminé par un groupe thiol, au niveau de leur domaine de thiolation ([Figure 9](#)). L'amorçage des domaines *apo*-ACP (PKS) ou *apo*-PCP (NRPS) en domaines *holo*-ACP ou *holo*-PCP fonctionnels est assuré par des phosphopantéthéinyl transférases (PPTases), souvent codées dans les *clusters* de gènes codant l'ensemble de la machinerie. Chaque domaine de thiolation ainsi activé est en mesure d'accueillir un monomère activé. Les mégasynthèses démarrent alors un cycle de synthèse, qui comprend l'activation des monomères puis élongation des oligomères ([Figure 9](#)) [38,39].

▪ **Activation des monomères**

Les monomères qui composent les PK et les NRP sont issus du métabolisme primaire, ou bien synthétisés à partir des métabolites primaires par des voies métaboliques secondaires, fréquemment codées dans le *cluster* comprenant les gènes des mégasynthèses. L'utilisation de ces monomères repose sur leur activation préalable [35,40].

Les composés PK sont constitués de monomères d'acides carboxyliques, activés grâce à une liaison thioester avec le coenzyme A (CoA), tels que l'acétyl-CoA et le malonyl-CoA ([Tableau I](#)). Ils sont chargés sur les PKS grâce aux domaines AT [35,41].

Les NRP sont synthétisés à partir des 20 acides aminés protéinogènes, d'acides aminés non protéinogènes et d'acides aryles ([Tableau I](#)). Plus de 500 monomères différents sont dénombrés et à l'origine de la diversité des composés NRP. Leur activation repose sur une réaction d'adénylation avec de l'ATP, catalysée par le domaine A des NRPS [35,37].

Les monomères activés des PK et des NRP sont ensuite chargés sur les domaines de thiolation, ACP pour les PKS et PCP pour les NRPS, pour former respectivement des acyl- et des aminoacyl-thioesters conservant leur énergie d'activation ([Figure 9](#)) [35].

▪ **Élongation des oligomères**

L'élongation repose sur le transfert de l'oligomère de la mégasynthase $n - 1$ sur l'acyl- ou l'aminoacyl-thioester de la mégasynthase n , qui catalyse la formation d'une liaison covalente entre les deux grâce à son domaine de condensation ([Figure 9](#)). Ce processus est répété tout au long de la chaîne d'assemblage selon l'ordre défini par l'agencement des mégasynthèses [35,40].

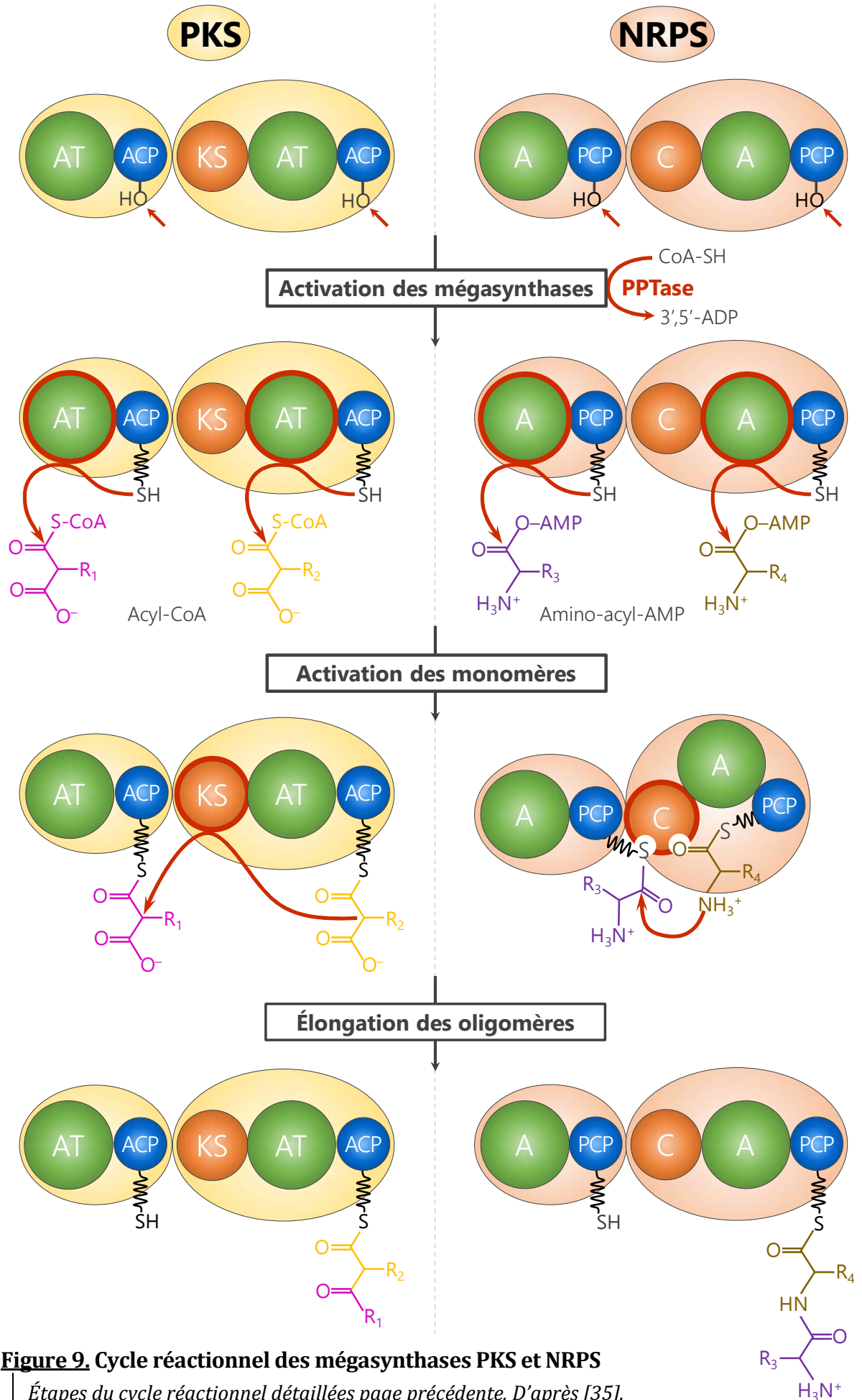


Figure 9. Cycle réactionnel des mégasynthases PKS et NRPS

Étapes du cycle réactionnel détaillées page précédente. D'après [35].

▪ Modification des oligomères

L'oligomère peut être modifié sur la chaîne d'assemblage multienzymatique au cours des différentes phases d'élongation ou en fin de biosynthèse. Ces modifications sont assurées au cours de l'élongation par les domaines enzymatiques optionnels des mégasynthèses. Des enzymes accessoires, dites de « finition », et fréquemment codées dans le *cluster* comprenant les gènes des mégasynthèses, peuvent également modifier l'oligomère au cours de sa synthèse ou après son assemblage. Ces modifications, nombreuses et complexes, sont à l'origine de la diversité des composés et essentielles à leur activité biologique [35,40].

▪ Terminaison de la biosynthèse

La libération des oligomères en fin de chaîne de synthèse s'effectue grâce à un domaine TE généralement présent sur le module de terminaison. Ce domaine à activité thioestérase, dite de type I, peut être hydrolytique ou cyclisant. La cyclisation présente souvent un caractère fondamental, car elle rigidifie la conformation du composé, permet des interactions spécifiques avec les cibles biologiques, et confère aux NRP une résistance aux protéases [35,40].

▪ Déblocage des machineries

Les chaînes d'assemblage PKS et NRPS peuvent parfois faire des erreurs et incorporer des monomères inadaptés sur les domaines de thiolation, bloquant ainsi la chaîne d'assemblage. Les *clusters* de gènes de biosynthèse codent fréquemment des enzymes de réparation, appelées thioestérases de type II, qui éliminent les monomères inadaptés par hydrolyse [35,42].

1.2.4. Taille des machineries enzymatiques

Les chaînes d'assemblage PKS/NRPS se composent parfois de complexes protéiques pouvant atteindre plusieurs mégaDaltons (MDa). Par exemple, la mégasynthétase de la cyclosporine possède 41 domaines répartis sur 11 modules, regroupés en une seule protéine de 15 000 résidus correspondant à une masse de 1,5 MDa, alors que le peptide cyclique qu'elle synthétise n'est long que de 11 acides aminés [43]. De si grandes protéines sont vraisemblablement sensibles aux problèmes de repliement, vulnérables à la protéolyse, et enclines à l'inactivation par mutation [35]. L'organisation des modules sur plusieurs sous-unités protéiques peut aider à éviter ces problèmes. Cependant, le fractionnement d'une chaîne d'assemblage en de multiples protéines crée d'autres écueils, comme la reconnaissance inter-protéines lors du transfert des oligomères entre les modules enzymatiques [35,40].

Pour illustrer ces voies de biosynthèse PKS et NRPS, les quatre PK/NRP synthétisés par *E. coli* (entérobactine, salmochélines, yersiniabactine et colibactine) sont maintenant présentés.

2. Les sidérophores, des systèmes de captation du fer

2.1. Généralités sur l'acquisition du fer par les bactéries

Le fer est un ion métallique essentiel pour presque tous les organismes vivants [44]. Au niveau biologique, le fer est un constituant majeur de nombreuses protéines, sous forme de groupement héminique ou fer-soufre [45]. La capacité du fer à facilement échanger des électrons en fait également un cofacteur de choix dans divers processus biologiques vitaux, tels que le cycle de Krebs, le transport d'oxygène ou encore la réplication de l'ADN [46].

Le fer est retrouvé sous deux états d'oxydo-réduction interchangeables : la forme réduite Fe^{2+} , appelée fer ferreux, et la forme oxydée Fe^{3+} , appelée fer ferrique. L'état du fer dépend principalement de la teneur en oxygène du milieu environnant. En conditions d'anaérobiose, le fer se retrouve principalement sous sa forme réduite Fe^{2+} et est alors relativement soluble. À l'inverse, en présence d'oxygène, il est majoritairement présent sous forme oxydée Fe^{3+} et présente une solubilité très faible [47].

Le fer est indispensable à la survie et à la virulence des micro-organismes bactériens, mais son acquisition est une réelle problématique. Chez l'hôte, le fer est en effet principalement retrouvé sous forme ferrique très peu soluble (10^{-9} M à pH 7). Sa biodisponibilité est d'autant plus faible que la majorité du fer est complexé aux protéines de l'hôte, abaissant ainsi la concentration sérique en fer libre à environ 10^{-24} M [48]. Chez les vertébrés, les deux tiers du pool ferrique sont complexés à l'hémoglobine, près d'un tiers est stocké dans les cellules par la ferritine, et une infime proportion est séquestrée par des protéines circulantes liant le fer, principalement la transferrine et la lactoferrine [46].

Afin de maintenir une concentration intracellulaire d'environ 10^{-6} M nécessaire à leur survie chez l'hôte, les bactéries ont développé différents mécanismes d'acquisition du fer. Elles expriment ainsi de multiples transporteurs, capable de capter le fer circulant ou complexé à l'hémoglobine [49], ou récepteurs, favorisant l'extraction du fer associé aux protéines circulantes de l'hôte (transferrine et lactoferrine) [50]. Elles synthétisent également des sidérophores, petites molécules sécrétées présentant une très forte affinité pour le fer ferrique, et capables de récupérer le fer séquestré par les protéines de l'hôte [46,51].

Cette partie se focalisera sur l'acquisition du fer par les sidérophores synthétisés par *E. coli*, que sont l'entérobactine, les salmochélines et la yersiniabactine, composés PK et/ou NRP, ainsi que l'aérobactine, de nature chimique différente, et sur leur rôle dans la virulence bactérienne.

2.2. Étapes d'acquisition du fer par les sidérophores

Les sidérophores sont des chélateurs spécifiques du fer, aptes à capturer le fer ferrique complexé à la transferrine ou la lactoferrine de l'hôte grâce à leur affinité extrêmement élevée [52]. Ils sont synthétisés par les bactéries, les champignons et certaines plantes [53]. Presque toutes les souches de *E. coli*, ainsi que de nombreuses entérobactéries, synthétisent l'entérobactine, un composé NRP de la famille des catéchols (Figure 10) [48]. Les souches d'ExPEC peuvent produire trois sidérophores additionnels : les salmochélines, dérivés glycosylés de l'entérobactine ; la yersiniabactine, composé PK-NRP de la famille des phénols ; et l'aérobactine, composé non-PK et non-NRP de type citrate-hydroxamate (Figure 10) [48].

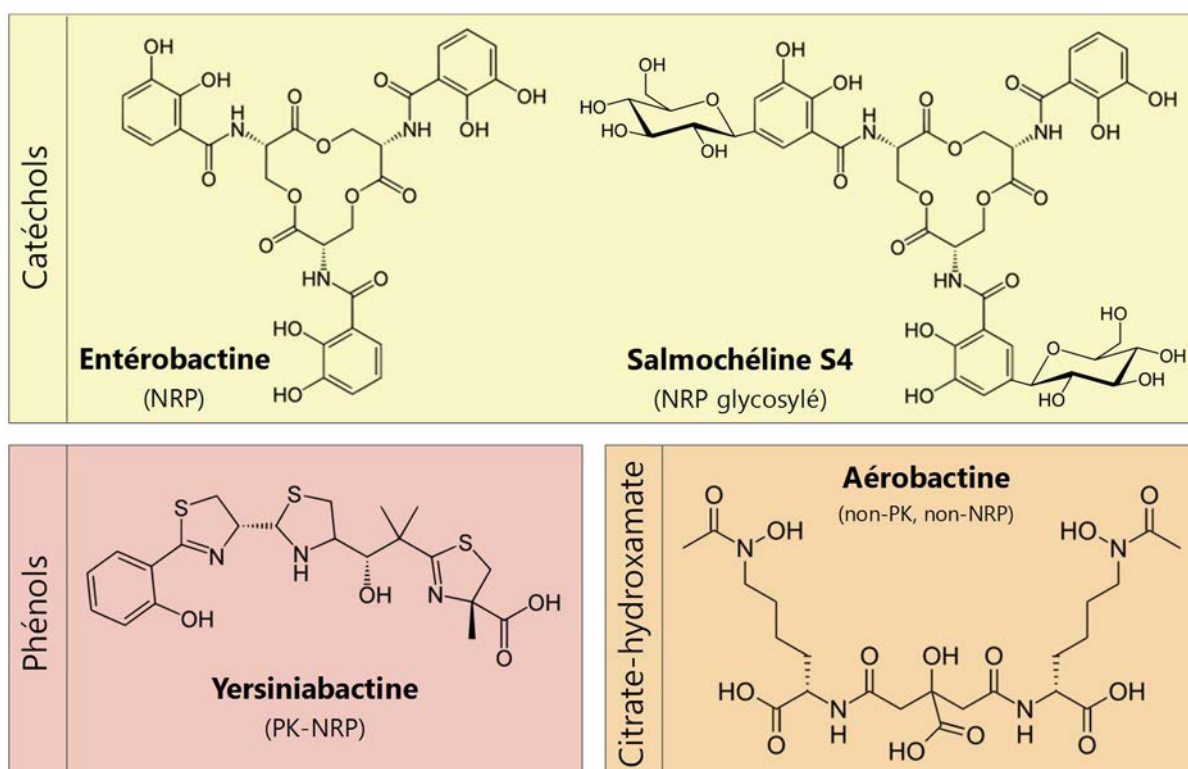


Figure 10. Structure des sidérophores synthétisés par *E. coli*

Si ces sidérophores diffèrent par leur structure chimique, les étapes spécifiques de l'acquisition du fer sont identiques et comprennent (Figure 11) [51] : ① la biosynthèse et ② l'export du sidérophore, puis ③ la recapture du complexe $[\text{Fe}^{3+}\text{-sidérophore}]$ (ou ferri-sidérophore) à la surface bactérienne et ④ son internalisation, et enfin ⑤ la libération du fer dans le cytoplasme. Les composants spécifiques nécessaires à la biosynthèse, l'export, la recapture, l'internalisation et la libération pour chacun des quatre sidérophores de *E. coli* sont détaillés dans le Tableau II. Pour chaque système, les gènes codant les différents composants nécessaires à l'acquisition du fer sont organisés en *clusters*, représentés dans la Figure 12 [51].

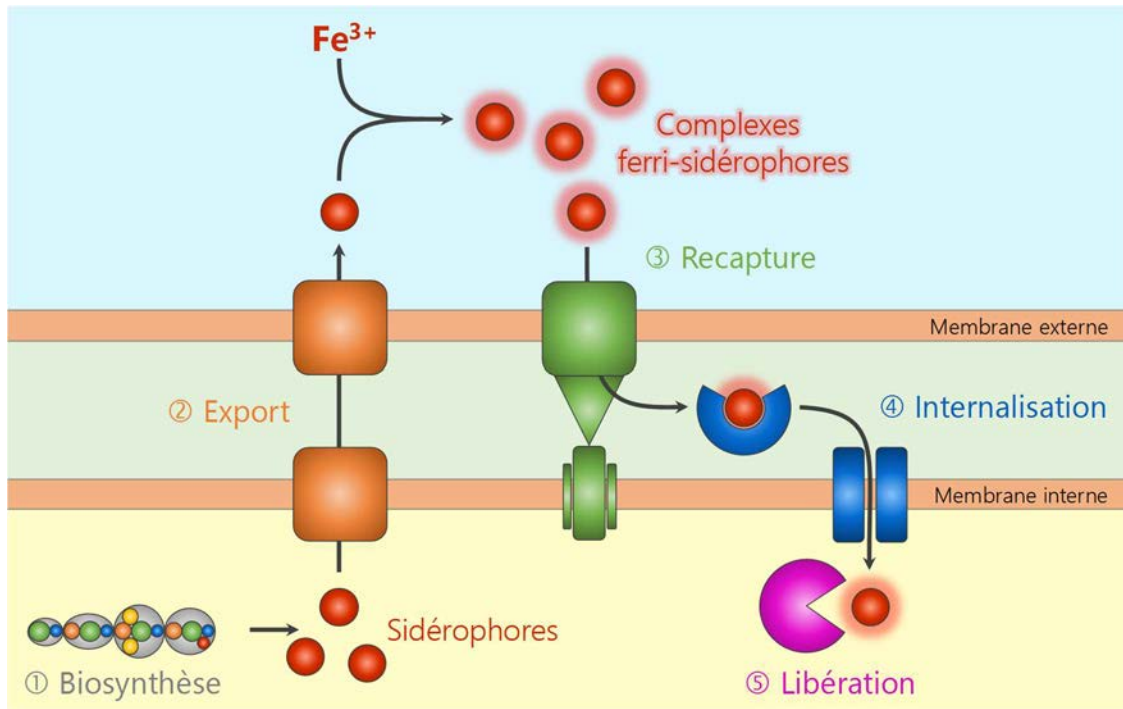


Figure 11. Étapes de l'acquisition du fer par les sidérophores
 Étapes de l'acquisition du fer détaillées page précédente. D'après [51].

Tableau II. Composants spécifiques des quatre systèmes de sidérophores de *E. coli*

	Entérobactine (NRP)	Salmochélines (NRP glycosylés)	Yersiniabactine (PK-NRP)	Aérobactine (autre)
1. Biosynthèse	EntABCDE	EntABCDE + IroC	YbtSEUT + Irp1/Irp2	IucABCD
2. Export	EntS + TolC	IroC + ?	?	?
3. Recapture	FepA	IroN	FyuA	IutA
4. Internalisation	FepB + FepCDG	FepB + FepCDG	YbtPQ	FhuD + FhuBC
5. Libération	Fes	IroD/IroE	?	Réductases

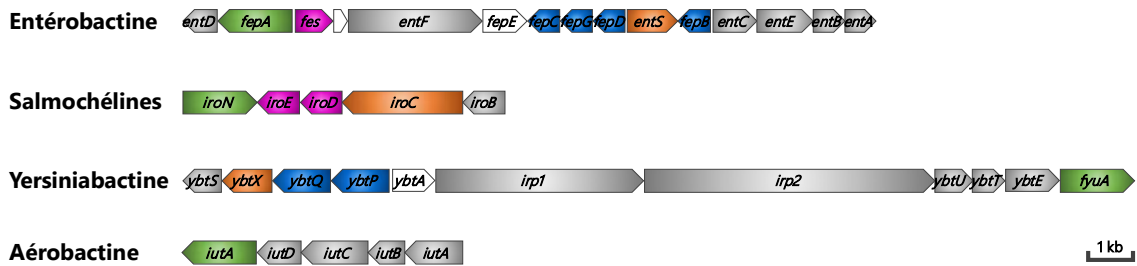


Figure 12. Organisation génique des clusters codant les sidérophores de *E. coli*
 Biosynthèse ■ ; export ■ ; recapture ■ ; internalisation ■ ; libération ■. D'après [51].

▪ **Biosynthèse**

La biosynthèse des sidérophores par les bactéries est induite en conditions de carence en fer intracellulaire. Les types de voies de biosynthèse sont différents selon la famille chimique. De manière générale, la biosynthèse des catéchols (entérobactine) et phénols (yersiniabactine) requiert des machineries enzymatiques de type NRPS et/ou PKS, alors que la synthèse des hydroxamates (aérobactine) est PKS/NRPS-indépendante [51].

▪ **Export**

À l'issue de leur synthèse, les sidérophores sont exportés dans le milieu extracellulaire par des transporteurs dédiés afin de remplir leur fonction d'acquisition du fer. Trois familles de transporteurs sont impliquées : les MFS (pour *major facilitation superfamily*), les RND (pour *resistance nodulation and cell division*) et les ABC (pour *ATP-binding cassette*) (Tableau II) [51].

▪ **Recapture**

La recapture des complexes ferri-sidérophores fait intervenir des récepteurs spécifiques localisés sur la membrane externe [51]. Les différents récepteurs sont dépendants du système de transfert d'énergie TonB–ExbB₇–ExbD₂. Le complexe [ExbB₇–ExbD₂] utilise la force proton-motrice de la membrane plasmique interne pour transférer l'énergie à la protéine TonB. La liaison d'un complexe ferri-sidérophore à son récepteur spécifique déclenche le recrutement de TonB, induisant alors par transfert d'énergie un changement conformationnel du récepteur et l'import périplasmique du complexe ferri-sidérophore [54]. Certaines bactéries sont en outre dotées de récepteurs pour des sidérophores synthétisés par d'autres micro-organismes (ou xénosidérophores) et présents dans l'environnement [51].

▪ **Internalisation**

Une fois dans le périplasme, les complexes ferri-sidérophores vont traverser la membrane interne grâce à des systèmes de transport ATP-dépendants. Ils peuvent éventuellement être pris en charge par des protéines périplasmiques de transfert, qui facilite leur adressage vers les transporteurs et accélère leur internalisation vers le cytoplasme [51].

▪ **Libération**

Pour des sidérophores comme l'aérobactine, la libération du fer du complexe se produit par réduction du sidérophore ferrique en sidérophore ferreux sous l'action de réductases, suivie d'une libération spontanée du fer [48]. Mais pour d'autres sidérophores, tels que l'entérobactine et les salmochélines, l'action des réductases ne suffit pas et la libération du fer nécessite l'action d'estérases capables d'hydrolyser le sidérophore [51].

2.3. Rôle des différents sidérophores dans la virulence de *E. coli*

Les sidérophores, en plus d'être d'excellents chélateurs du fer, participent à la virulence des bactéries pathogènes. En effet, lors d'une infection, une réponse rapide de l'hôte visant à protéger le fer de l'organisme est très rapidement mise en place pour empêcher son utilisation par la bactérie. La sécrétion de sidérophores par les souches pathogènes dans ces conditions est donc un élément essentiel à la survie [51].

▪ Entérobactine et salmochélines

L'entérobactine est un sidérophore commun aux souches commensales et pathogènes de *E. coli*, et son implication dans la virulence semble être négligeable. Une souche d'ExPEC produisant l'entérobactine, mais dont la production de salmochélines et d'aérobactine a été abolie, est complètement atténuée dans un modèle d'infection systémique chez le poulet [55]. L'inefficacité de l'entérobactine au cours d'une infection a été attribuée au fait qu'elle peut être séquestrée par des protéines de défense immunitaire de l'hôte, comme par exemple la lipocaline-2 (encore appelée NGAL pour *neutrophil gelatinase-associated lipocalin*) [56].

Cependant, les dérivés glycosylés de l'entérobactine, que sont les salmochélines, peuvent échapper à cette séquestration. Le processus de glycosylation assuré par IroB est donc important dans la virulence [56]. Mais les autres gènes codant la synthèse et le transport des salmochélines, regroupés en *cluster* sur les plasmides de virulence ColV/M ou sur un îlot génomique, sont également impliqués dans la virulence. Le récepteur IroN des salmochélines est requis pour la virulence de certaines souches d'UPEC [57,58] et de NMEC [59,60]. Enfin, la délétion complète du locus *iroA* résulte en une diminution de la virulence de souches de *E. coli* pathogènes aviaires dans un modèle d'infection systémique chez le poulet [55].

▪ Yersiniabactine

Les gènes codant la biosynthèse et le transport de la yersiniabactine sont regroupés au sein de l'îlot génomique HPI (pour *high pathogenicity island*), fréquemment présent dans le génome des souches des différents pathovars de *E. coli* [61,62]. Le rôle de la yersiniabactine dans la virulence a été démontré dans la physiopathologie des infections urinaires, où elle joue un rôle crucial dans la captation du fer *in vivo* [63] et dans la formation du biofilm à la surface des cellules urothéliales [64]. La yersiniabactine modulerait également la réponse immunitaire innée de l'hôte. Par son rôle chélateur, la yersiniabactine diminue en effet la disponibilité du fer pour les cellules immunitaires, les privant d'un élément indispensable à la production d'espèces réactives de l'oxygène (ou ROS, pour *reactive oxygen species*) par la réaction de Fenton [65].

▪ Aérobactine

Les gènes codant la biosynthèse et le transport de l'aérobactine, regroupés comme pour les salmochélines en *cluster* sur les plasmides de virulence ColV/ColM ou sur un îlot génomique, sont plus fréquemment retrouvés dans les souches de *E. coli* pathogènes que dans les commensales [66]. La virulence de souches de *E. coli* pathogènes aviaires déficientes pour la production d'aérobactine est atténuée dans un modèle d'infection systémique chez le poulet [55]. De plus, l'aérobactine participe également à la recapture du fer *in vivo* dans un modèle murin d'infection urinaire [63].

Enfin, il a été montré que les conditions environnementales modulent la production de sidérophores par *E. coli* [67]. Chaque sidérophore procure un avantage adaptatif aux souches qui les produisent, et la proportion relative du rôle de chacun diffère selon les situations. Certains sidérophores auront un rôle exclusif dans la captation du fer, tandis que d'autres agiront en modulant la réponse de l'hôte vis-à-vis de l'infection [51]. Par exemple, Garcia *et al.* ont montré dans un modèle d'infection urinaire que le rôle des sidérophores de type catéchols (entérobactine et salmochélines) était redondant, alors que les sidérophores non-catéchols (yersiniabactine et aérobactine), ainsi que les récepteur hémiques, contribuaient chacun individuellement à la virulence des souches [63]. Le rôle respectif de chaque sidérophore peut être complètement différent dans d'autres modèles d'infections, puisqu'il a été démontré que les salmochélines et l'aérobactine ont chacun un rôle déterminant dans la virulence de souches de *E. coli* pathogènes aviaires dans un modèle d'infection systémique chez le poulet [68].

3. La colibactine, une génotoxine bactérienne

La colibactine est un composé hybride de type polycétide/peptide non-ribosomal (PK-NRP) produit par plusieurs espèces bactériennes de la famille des *Enterobacteriaceae*, dont *E. coli*. Ce métabolite secondaire, identifié en 2006 mais dont la structure complète est encore inconnue aujourd'hui, induit des cassures double-brin de l'ADN sur les cellules eucaryotes [2].

3.1. Données moléculaires

▪ Support génétique

Le support génétique permettant la synthèse de la colibactine est un îlot génomique de 54 kb, l'îlot *pks*, comprenant 19 gènes (*clbA* à *clbS*) qui codent une machinerie de biosynthèse de type PKS/NRPS, ainsi que des enzymes et protéines accessoires (Figure 12) [2,69] :

- le gène *clbA* code une phosphopantéthéinyl transférase (PPTase) ;
- les gènes *clbC*, *clbI* et *clbO* codent des enzymes PKS ;
- les gènes *clbH*, *clbJ* et *clbN* codent des enzymes NRPS ;
- les gènes *clbB* et *clbK* codent des enzymes hybrides PKS-NRPS ;
- les gènes *clbD*, *clbE*, *clbF*, *clbG*, *clbL* et *clbQ* codent des enzymes accessoires ;
- le gène *clbM* code une pompe à efflux [70] ;
- le gène *clbP* code une peptidase périplasmique [71] ;
- le gène *clbR* code une protéine comportant un domaine de liaison à l'ADN ;
- le gène *clbS* code une protéine de résistance à l'activité génotoxique [72].

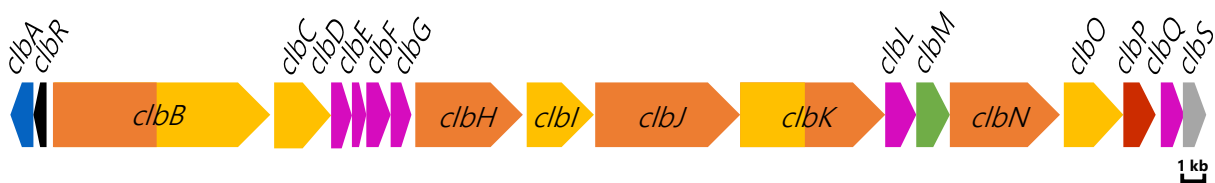


Figure 13. Organisation de l'îlot génomique *pks*

L'îlot génomique *pks* comprend 19 gènes codant des PKS (jaune), des NRPS (orange), ou des hybrides PKS-NRPS, une PPTase (bleu), des enzymes accessoires (violet), une pompe à efflux (vert), une peptidase périplasmique (rouge), une protéine avec un domaine de liaison à l'ADN (noir), et une protéine de résistance (gris). D'après [69].

L'îlot *pks* possède les caractéristiques d'un îlot génomique, avec une taille comprise en 10 et 200 kb, un contenu moyen en G+C différent du reste du génome, des séquences répétées (DR) aux extrémités, deux séquences d'insertion (IS), une transposase et une intégrase [2].

▪ **Distribution de l'îlot *pks***

L'îlot *pks* est présent chez certaines souches commensales et pathogènes extra-intestinales de *E. coli* appartenant quasi-exclusivement au groupe phylogénétique B2. Il est absent des souches de *E. coli* K-12 de laboratoire ou pathogènes intestinales [2,73,74]. L'îlot est également retrouvé chez d'autres espèces de la famille des *Enterobacteriaceae*, telles que *Klebsiella pneumoniae*, *Citrobacter koseri* et *Enterobacter aerogenes* [73,74]. La séquence de l'îlot est très conservée entre ces espèces, avec 98% d'identité de séquence nucléotidique. L'îlot a également été retrouvé dans le génome d'autres protéobactéries, notamment des espèces bactériennes vivant en association avec des éponges marines [75] ou des abeilles [76].

▪ **Épidémiologie des souches *pks*⁺**

Concernant les souches pathogènes, les études épidémiologiques montrent qu'approximativement 40% des souches d'ExPEC possèdent l'îlot *pks* [74], atteignant même 58% pour les souches responsables de bactériémie [73], 59% pour les isolats d'urosepsis [77], et 72% des ExPEC retrouvées dans les prostatites aiguës [78].

Concernant les souches commensales, la prévalence de l'îlot *pks* a été rapportée dans 38 à 58% des souches du phylogroupe B2. Dans la population des pays industrialisés, le portage de souches possédant l'îlot *pks* a été évalué à environ 25% chez les adultes sains [73,74,77], et 15 à 18% des nouveaux-nés sont colonisés dès la naissance par des souches *pks*⁺ [79,80].

▪ **Environnement génétique**

Dans les souches de *E. coli* du phylogroupe B2, l'insertion chromosomique de l'îlot *pks* est située au niveau du locus de l'ARNt *asnW*, un site privilégié d'intégration et d'échanges d'éléments génétiques mobiles [81]. Lorsqu'il est présent, l'îlot *pks* est systématiquement associé et inséré à proximité de l'îlot génomique HPI codant les gènes nécessaires à la biosynthèse et au transport de la yersiniabactine [74].

3.2. Éléments de biosynthèse

La synthèse de la colibactine est assurée par une machinerie enzymatique composée de mégasynthèses, dont trois PKS (ClbC, ClbI et ClbO), trois NRPS (ClbH, ClbJ et ClbN) et deux hybrides PKS-NRPS (ClbB et ClbK), ainsi que d'enzymes accessoires (ClbA, ClbD-G, ClbL, ClbP-Q). À ce jour, la structure complète de la colibactine est inconnue, mais plusieurs études publiées ces trois dernières années ont identifié de nombreux intermédiaires de synthèse [82–91].

▪ Activation de la machinerie

CibA est une PPTase, enzyme essentielle pour l'activation des mégasynthèses, qui catalyse la fixation d'un bras pantéthéinyl terminé par un groupe thiol sur les domaines de thiolation [2,39]. CibA participe également à la biosynthèse des autres composés PK/NRP synthétisés par *E. coli*, l'entérobactine, les salmochélines et la yersiniabactine [92]. Il faut donc garder à l'esprit qu'une souche mutée pour *clbA* ne sera certes plus capable de produire la colibactine, mais pourra également être altérée dans sa capacité à produire les sidérophores.

▪ Étapes de biosynthèse

Une fois activées, les mégasynthèses recrutent leurs monomères spécifiques. La première enzyme intervenant dans la synthèse est la NRPS **CibN**. Cette mégasynthase, dont le substrat est l'asparagine (Asn), génère du ***N*-myristoyl-D-asparagine**, ou produit de clivage (*cf.* § Clivage page suivante) [82,83]. L'assemblage se poursuit ensuite avec l'intervention successive des enzymes **CibB-C-H-I-J-K** qui utilisent divers substrats, conventionnels (malonyl-CoA, alanine, sérine, glycine, cystéine) ou non (acide aminocyclopropane-carboxylique) (Figure 14) [84–91].

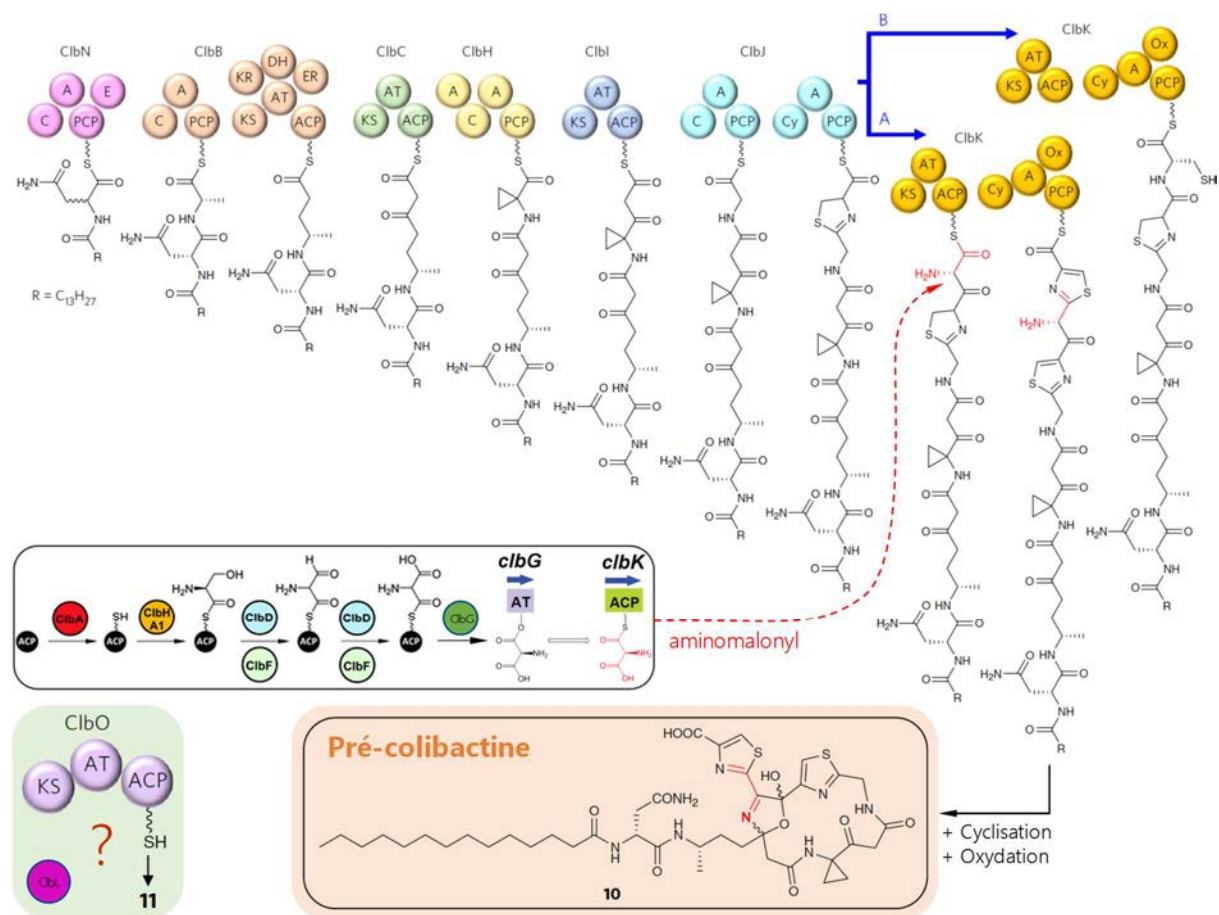


Figure 14. Biosynthèse de la pré-colibactine

Étapes de biosynthèse détaillées dans le texte. D'après [91].

Par ailleurs, il a été récemment démontré que les enzymes accessoires **ClbD-E-F**, en collaboration avec l'un des domaines A de ClbH, génèrent à partir de L-sérine un monomère inhabituel utilisable par les PKS de l'îlot, l'aminomalonyl-ACP [93]. Ce monomère est ensuite transféré aux différents modules PKS par **ClbG**, enzyme monofonctionnelle à activité acyltransférase [89]. La dernière étude publiée sur la biosynthèse de la colibactine par Li *et al.* a pour la première fois identifié une pré-colibactine comportant un résidu aminomalonyl [91]. L'intégration de ce monomère atypique modifie drastiquement la conformation de la molécule. La pré-colibactine caractérisée dans cette étude diffère d'ailleurs des autres intermédiaires de synthèse proposés jusqu'ici. Les auteurs proposent ainsi deux voies alternatives de synthèse, intégrant (voie **A**) ou non (voie **B**) une unité aminomalonyl par ClbK (Figure 14) [91].

Enfin, trois enzymes ont encore une fonction inconnue dans la biosynthèse de colibactine :

- **ClbL** serait une amidase d'après une analyse *in silico*, mais son rôle est énigmatique [2] ;
- **ClbO** est une PKS qui pourrait, comme ClbK, intégrer un résidu aminomalonyl. Li *et al.* évoquent dans leur dernière étude qu'ils auraient détecté un intermédiaire issu de l'action de ClbO, en cours de caractérisation structurale (Figure 14) [91] ;
- **ClbQ** est une thioestérase de type II, dont le rôle habituel est de débloquer les chaînes d'assemblage en hydrolysant les intermédiaires de synthèse aberrants. ClbQ pourrait cependant avoir un fonctionnement atypique et libérerait les intermédiaires précoces conformes de la machinerie de biosynthèse [91].

▪ **Export**

Le métabolite final produit par la chaîne d'assemblage, dont la structure complète reste inconnue, est une pré-colibactine inactive. Cette pré-colibactine est ensuite exportée dans le périplasme de la bactérie grâce à **ClbM** (Figure 15) [70]. Cette pompe à efflux est insérée dans la membrane interne par douze domaines transmembranaires, dont l'inactivation par mutation du gène *clbM* conduit à une réduction de l'activité génotoxique. L'hypothèse actuelle pour ce phénotype atténué suppose que d'autres pompes à efflux complètent partiellement l'absence de ClbM [2,70].

▪ **Clivage**

Une fois dans le périplasme, la pré-colibactine subit l'action de **ClbP**. Cette protéine périplasmique ancrée à la membrane interne hydrolyse, par son activité D-amino-peptidase, le motif pro-drogue de la pré-colibactine, libérant la forme active de la colibactine ainsi que le produit de clivage, ou **N-myristoyl-D-asparagine** (Figure 15) [71,82,94].

L'ensemble des intermédiaires de synthèse proposés jusqu'ici ont été identifiés dans des souches mutées pour *clbP*. Leur relevance biologique est donc inconnue. De plus, la perte du produit de clivage grâce à l'intervention de ClbP modifie très probablement la structure chimique et la réactivité de la colibactine active.

▪ Systèmes de protection

La stratégie pro-drogue, précédemment décrite pour deux antibiotiques PK/NRP, la zwittermicine [95] et la xénocoumacine [96], serait un mécanisme de protection de la bactérie productrice de l'activité de ces composés toxiques.

L'îlot *pks* code un système de protection additionnel, la protéine **ClbS**, qui protégerait la bactérie productrice de l'activité génotoxique de la colibactine en modifiant ou en séquestrant les molécules de colibactine active qui pourraient se trouver dans le cytoplasme, par exemple lors d'une ré-internalisation après l'étape de clivage périplasmique (Figure 15) [72]. De plus, l'expression hétérologue de ce système de protection dans des cellules eucaryotes leur confère une résistance partielle à la génotoxicité induite par la colibactine, suggérant une pénétration de la toxine dans la cellule hôte [72].

▪ Système de régulation ?

La protéine **ClbR** codée par l'îlot *pks* comporte un domaine de liaison à l'ADN comportant un motif hélice-tour-hélice [2]. Bien que la fonction de cette protéine de la famille LuxR soit inconnue, elle pourrait intervenir dans la régulation de l'expression des gènes de l'îlot.

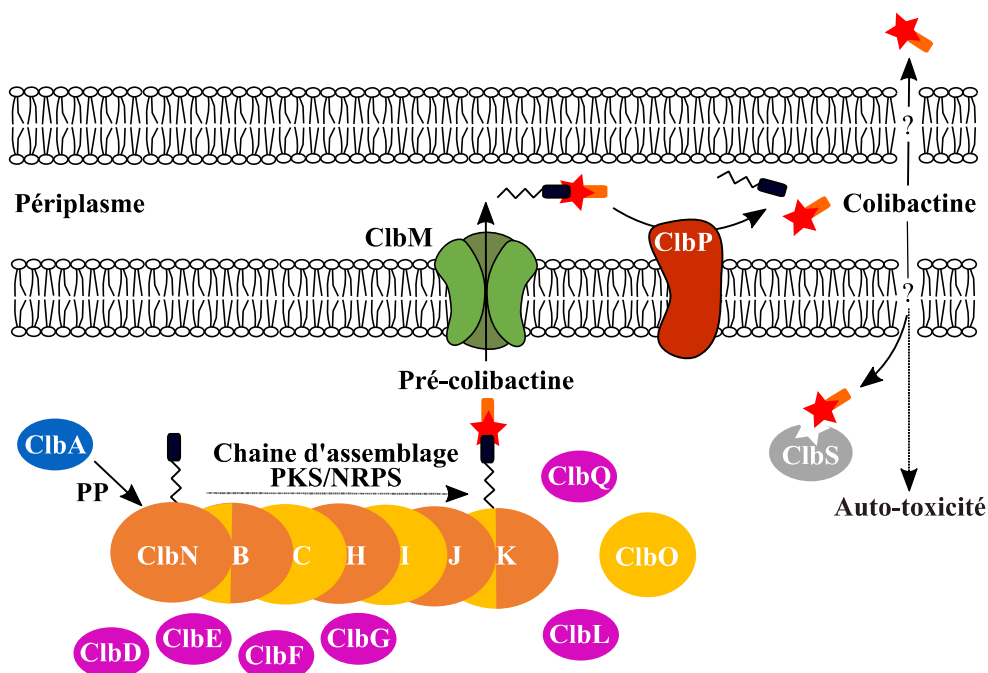


Figure 15. Modèle de synthèse et d'export de la colibactine

PPtase ■ ; PKS ■ ; NRPS ■ ; enzyme accessoire ■ ; pompe à efflux ■ ; peptidase ■ ; protéine de résistance ■ [69].

3.3. Relations structure-activité

Les intermédiaires de synthèse identifiés en 2015 dans les souches mutées pour *clbP* ont révélé la présence un motif particulier dans la structure de la colibactine, baptisé « ogive » (ou "warhead") [85,87,88]. Après clivage, la colibactine active présente en effet un cycle spiro-cyclopropane électrophile extrêmement réactif, dont la structure évoque celle d'agents alkylants de l'ADN, tels que les illudines, les duocarmycines ou la yatakémicine [97]. D'après la structure d'un intermédiaire isolé dans un double mutant $\Delta clbP \Delta clbG$, ce motif caractéristique semble être connecté à un groupement bithiazole (Figure 16 – Pré-colibactine 795) [89]. Il a été montré que ce double noyau aromatique, retrouvé dans la structure des bléomycines et des phléomycines, confère des propriétés d'intercalation entre les brins d'ADN.

Cette hypothèse a été récemment remise en question par Li *et al.* [91]. La structure du nouvel intermédiaire de synthèse identifié dans un double mutant $\Delta clbP \Delta clbQ$ inclue un résidu aminomalonyl supplémentaire, qui modifie drastiquement le réarrangement moléculaire proposé jusqu'ici. Si le cycle spiro-cyclopropane est toujours présent, le motif bithiazole a laissé sa place à un macrocycle incluant un hétérocycle, structure sans équivalent dans les composés naturels selon les auteurs (Figure 16 – Pré-colibactine 886) [91].

En attendant la caractérisation de la structure complète de la colibactine active, il convient de rester critique sur les données apportées par la structure des intermédiaires et rester prudent sur les hypothèses de relation structure-activité.

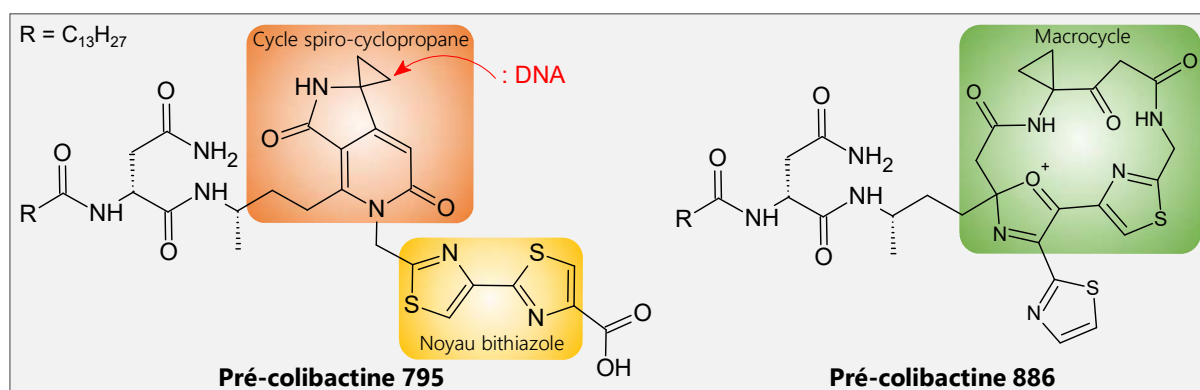


Figure 16. Structure des pré-colibactines

Structure des pré-colibactines 795 (voie de synthèse A) et 886 (voie de synthèse B). D'après [91].

3.4. Phénotypes associés

L'infection transitoire de cellules eucaryotes par des souches de *E. coli* possédant l'îlot *pks* conduit à l'apparition de cassures double-brin de l'ADN des cellules infectées (Figure 17A) [2].

Ces dommages déclenchent l'activation de la réponse aux dommages à l'ADN, qui implique la voie de signalisation ATM/Chk2/Cdc25, entraînent secondairement une **phosphorylation de l'histone H2AX** et un arrêt du cycle cellulaire en phase G2/M stoppant la prolifération (Figure 17A). L'activité génotoxique de la colibactine est dépendante d'une interaction entre la bactérie productrice et la cellule eucaryote cible, ainsi que de l'inoculum bactérien, mais n'est pas induite par un surnageant de culture, un lysat bactérien ou des bactéries tuées [2].

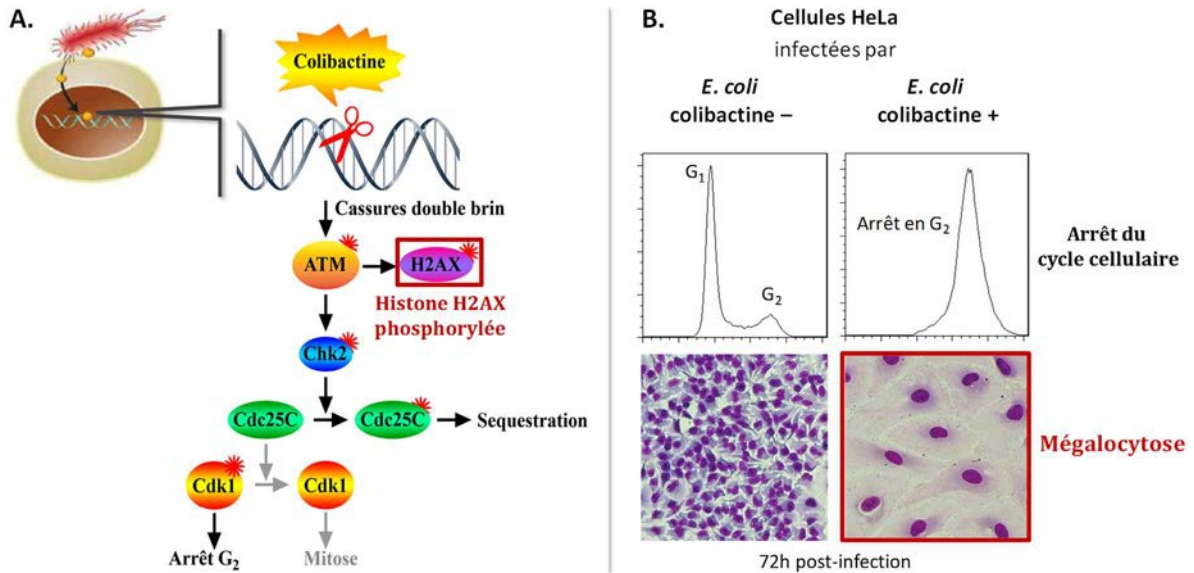


Figure 17. Dommages cellulaires induits par des souches produisant la colibactine

- A. Mécanismes cellulaires activés en réponse aux cassures double-brin de l'ADN induites par des souches de *E. coli* produisant la colibactine au contact de cellules eucaryotes [2].
- B. Des cellules eucaryotes de type HeLa infectées par une souche de *E. coli* produisant la colibactine présentent un arrêt du cycle cellulaire en phase G2/M. Lorsqu'il est prolongé, il se traduit par un phénotype de mégaloctyose après 72 heures d'incubation post-infection [2].

Le nombre de cassures double-brin va déterminer le devenir des cellules infectées. Une quantité massive bloquera le cycle cellulaire et induira rapidement un phénomène d'apoptose [2]. Une dose modérée résultera en une sénescence cellulaire prématurée, caractérisée par un élargissement cellulaire appelé **mégaloctyose**, et un arrêt du cycle cellulaire (Figure 17B) [98]. Ces cellules en sénescence sécrètent dans le milieu des cytokines pro-inflammatoires, des chimiokines, des facteurs de croissance, des ROS et des protéases. Ce phénotype de sécrétion associé à la sénescence (ou SASP, pour *senescent-associated secretory pattern*) induit par la colibactine peut favoriser la croissance de cellules tumorales par sécrétion de facteur de croissance hépatocytaire [99]. Enfin, les cellules qui survivent à un faible nombre de cassures double-brin peuvent réparer leur ADN, puis poursuivre leur cycle cellulaire. Cependant, une sous-population peut proliférer avec des dommages à l'ADN résiduels et persistants, qui conduisent à l'apparition d'anomalies chromosomiques et d'une instabilité génétique, pouvant conduire à la transformation oncogénique des cellules eucaryotes infectées [3].

3.5. Impacts sur la santé humaine

▪ **Colibactine et cancer**

Les dommages à l'ADN induits par les souches *pks*⁺ sur les cellules intestinales peuvent aboutir à des mutations géniques et à une initiation de la transformation cellulaire [3]. Il a été démontré que ces souches peuvent contribuer à la survenue d'un cancer colorectal (ou CRC, pour *colorectal cancer*) dans un contexte inflammatoire [100]. De plus, les souches produisant la colibactine sont capables de favoriser la croissance de tumeurs dans un modèle murin de CRC par l'induction d'une sénescence cellulaire et la production de médiateurs du SASP [99]. S'il a été mis en évidence une plus forte prévalence des souches *pks*⁺ dans des biopsies digestives de patients atteints de CRC [101], le rôle exact de la colibactine dans le CRC reste à démontrer.

▪ **Colibactine et virulence**

Le rôle de l'îlot *pks* dans la virulence a été démontré dans un modèle murin de sepsis avec une souche d'ExPEC isolée chez un patient atteint d'une méningite néonatale. La production de colibactine est associée à une aggravation de la lymphopénie et une mortalité supérieure, en comparaison à une souche ne produisant plus la génotoxine [102]. Ces résultats ont été confirmés dans un modèle d'infection systémique néonatale chez le rat, où la mutation de la peptidase ClbP a été associée à une altération de la virulence, mais dans une moindre mesure qu'une souche mutée pour la PPTase ClbA [103]. Ces résultats soulignent le rôle de ClbA dans la biosynthèse de colibactine et des sidérophores, donc dans la virulence de souches d'ExPEC [92].

▪ **Colibactine et homéostasie de la barrière intestinale**

Dans un modèle de colonisation néonatale du tube digestif, il a été montré que des souches commensales *pks*⁺, acquises par transmission verticale, induisent des dommages à l'ADN transitoires sur les entérocytes des rongeurs, qui se traduisent à l'âge adulte par une instabilité chromosomique persistante [80]. Les animaux adultes colonisés par des souches *pks*⁺ depuis la naissance présentent également une altération de la barrière intestinale, plus perméable, qui est responsable d'une augmentation de la réponse immunitaire aux antigènes de la lumière digestive, pouvant favoriser l'apparition de manifestations allergiques ou auto-immunes [80,104].

▪ **Colibactine et activité probiotique**

L'îlot *pks* est également présent dans la souche probiotique *E. coli* Nissle 1917, utilisée de manière empirique et démontrée cliniquement efficace dans le traitement des maladies inflammatoires chroniques de l'intestin [105]. L'interruption du gène *clbA* dans cette souche, et donc l'inactivation de la machinerie de synthèse codée par l'îlot *pks*, abolit ses propriétés anti-inflammatoires dans un modèle de colite chez le rongeur [105].

Les éléments nécessaires à la production de la colibactine identifiés dans ce travail de thèse, la protéine chaperon Hsp90_{Ec} et la spermidine, font l'objet des deux prochains chapitres.

CHAPITRE III : CONTRÔLE QUALITÉ DES PROTÉINES CHEZ *Escherichia coli*

1. Principes généraux

Les protéines sont des macromolécules biologiques aux fonctions extrêmement variées et aux structures particulièrement complexes. Elles sont impliquées dans la quasi-totalité des processus biologiques chez tous les organismes vivants.

Synthétisées par le ribosome sous forme de longues chaînes linéaires d'acides aminés, les protéines sont libérées par la machinerie de synthèse dans un environnement cytosolique surchargé en molécules de nature diverse. Elles sont ainsi exposées à un nombre important d'interactions moléculaires non spécifiques, parfois responsables d'altérations structurales et/ou fonctionnelles [106]. Afin de s'en protéger, elles doivent subir différents repliements conformationnels, sous la forme d'interactions intra-moléculaires spécifiques, afin d'acquérir une structure tridimensionnelle unique, thermodynamiquement stable, fonctionnelle, et devenir ainsi biologiquement actives [106]. Cette conformation définitive est déterminée par la séquence primaire en acides aminés. Cependant, le repliement de la majorité des protéines est relativement lent, et nécessite de nombreuses étapes intermédiaires. Ce processus devient alors excessivement complexe *in vivo* en raison de l'encombrement moléculaire. Au sein de la cellule, la concentration en protéines cytosoliques peut en effet atteindre 400 g/L, freinant ainsi le repliement adéquat des protéines, et favorisant la formation d'agrégats [107].

Pour favoriser le repliement conformationnel et éviter l'agrégation des protéines, tous les organismes vivants, dont les bactéries, ont développé des systèmes de contrôle qualité des protéines particulièrement sophistiqués. Ils sont constitués de molécules chaperons et de protéases, appartenant pour la plupart à la famille des protéines de choc thermique (ou Hsp, pour *heat shock protein*), qui assurent le maintien de l'homéostasie protéique grâce à de nombreuses fonctions, énumérées ci-dessous et résumées dans la [Figure 18](#) [108] :

- elles favorisent le repliement des protéines nouvellement synthétisées ① ;
- elles orientent les protéines vers le compartiment cellulaire approprié ② ;
- elles favorisent le repliement des protéines ③ dénaturées lors d'un stress ④ ;

- elles dégradent les protéines mal conformées ayant subi une dénaturation trop importante et non réversible ⑤ ;
- elles préviennent la dénaturation des protéines conformées, mais également l'agrégation des protéines non ou mal conformées ⑥ ;
- elles résolvent les agrégats protéiques ⑦ et permettent leur solubilisation ⑧ ;

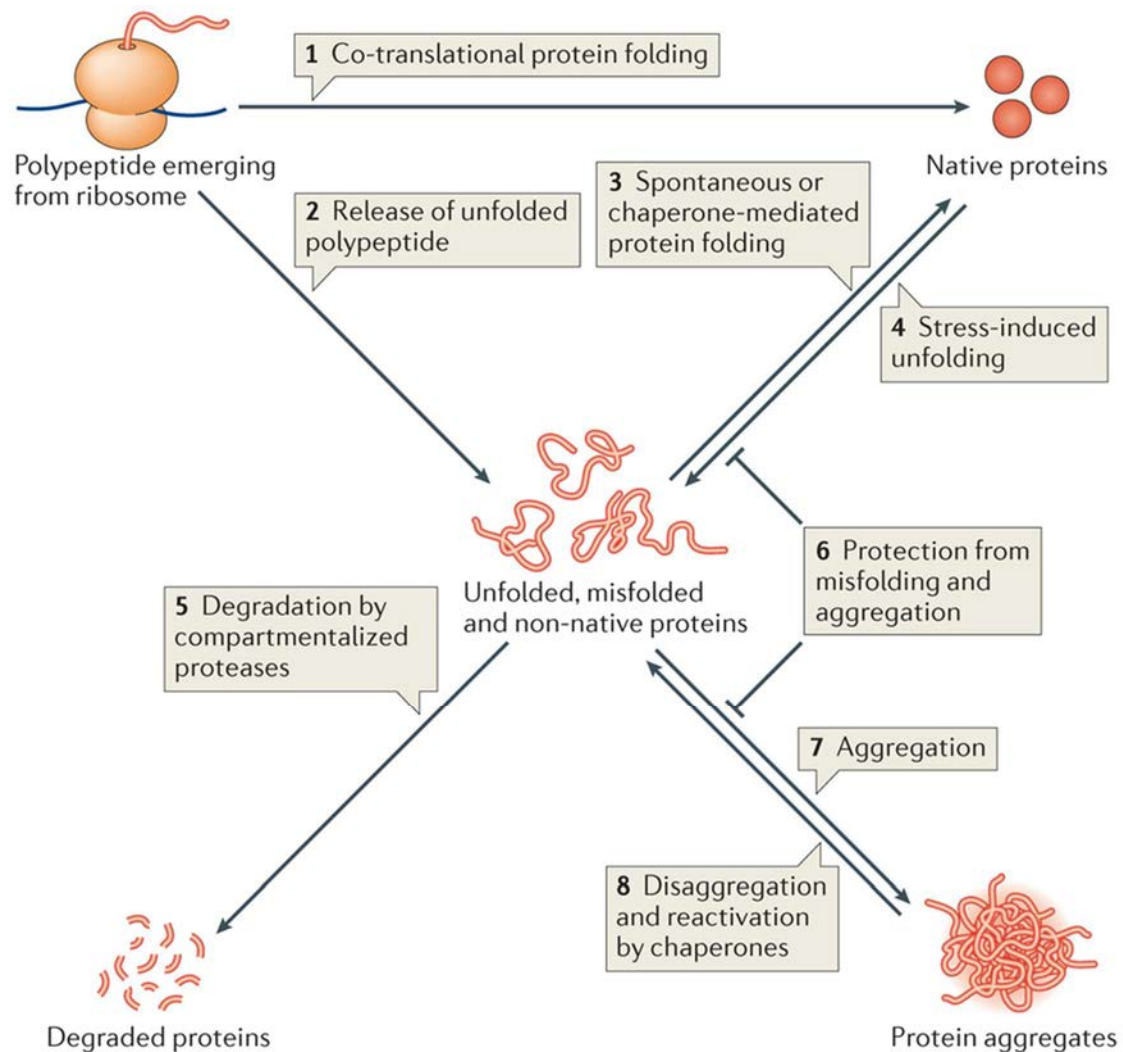


Figure 18. Principes généraux du contrôle qualité protéique chez *E. coli*

Les mécanismes mis en jeu sont détaillés dans le paragraphe précédent [109].

Ce chapitre va explorer différentes facettes du contrôle qualité des protéines chez *E. coli*. Nous présenterons d'abord les systèmes chaperons impliqués dans le repliement conformationnel des protéines néo-synthétisées. Nous aborderons ensuite les différentes stratégies préventives et curatives mises en place pour assurer le maintien de l'homéostasie protéique en situations de stress. Et nous terminerons cette partie par un focus sur le système chaperon Hsp90.

2. Contrôle qualité des protéines néo-synthétisées

Les protéines subissent un contrôle qualité dès l'étape de la traduction par l'action de la machinerie de synthèse protéique. Différents systèmes chaperons prennent ensuite en charge ces protéines immatures dès la sortie du ribosome pour poursuivre leur maturation.

2.1. Machinerie de traduction

Le premier facteur influençant le repliement des protéines est la séquence de leur ARN messager (ARNm). Ils comportent dans certaines régions des codons rares, qui correspondent à des complexes [acide aminé-ARNt] présents en faible concentration intracellulaire, ce qui va ralentir le rythme de la traduction, et favoriser les premières étapes de repliement protéique au cours de l'élongation de la chaîne peptidique naissante [110].

Le ribosome bactérien participe également aux premières étapes du repliement. Il présente dans sa sous-unité 50S un tunnel moléculaire, par lequel sort la chaîne peptidique en cours d'élongation. Elle subit dans ce tunnel de nombreuses interactions, qui peuvent moduler la cinétique de la traduction et favoriser la formation de structures secondaires spécifiques, telles que des hélices α ou des feuillets β [111].

Le rôle de la machinerie de traduction (ribosome, ARNm et ARNt), s'il est essentiel pour des protéines de petite taille, inférieure à 100 acides aminés, est cependant limité pour des protéines de taille plus conséquente, qui vont requérir pour leur repliement l'intervention de différentes protéines chaperons.

2.2. Trigger Factor

▪ Généralités

La première protéine chaperon à interagir avec une protéine néo-synthétisée au cours de son élongation est la protéine *Trigger Factor* (TF) associée au ribosome [112]. Son activité de chaperon, ATP-indépendante, ne se résume pas à maintenir les protéines dans leur état natif, mais permet très efficacement le repliement conformationnel des protéines néo-synthétisées.

▪ Structure

TF est une protéine de 48 kiloDaltons (kDa) comportant trois domaines ([Figure 19A](#)) [113] :

- la partie N-terminale correspond au domaine de liaison au ribosome (motif GFRxGxxP) ;

- la partie intermédiaire porte l'activité peptidyl-prolyl isomérase (PPIase) ;
- la partie C-terminale assure l'activité chaperon de TF.

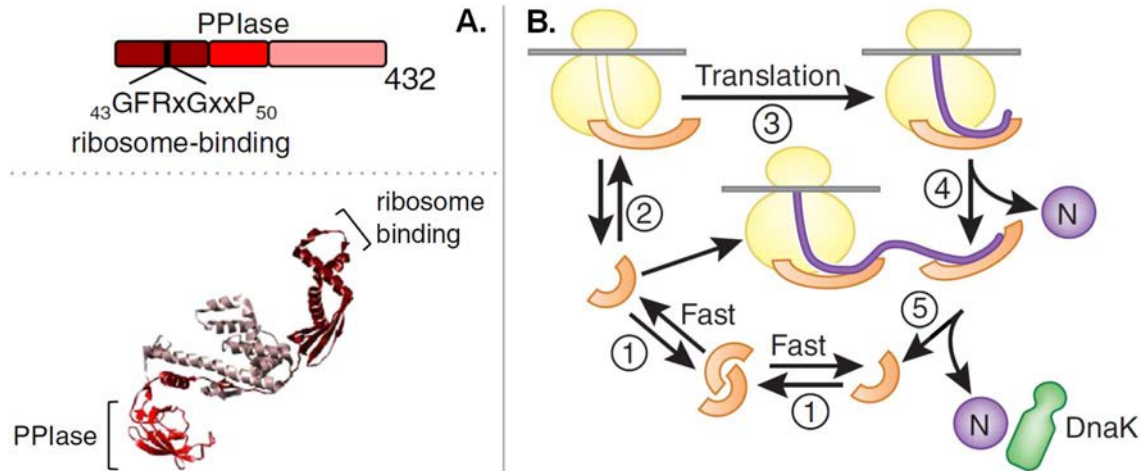


Figure 19. Structure et cycle réactionnel de la protéine chaperon Trigger Factor

A. Organisation des domaines protéiques et structure tridimensionnelle de TF [111].

B. Cycle réactionnel de TF [114]. ① La forme libre de TF est en équilibre entre état monomérique et dimérique. ② TF monomérique libre se lie au ribosome en l'absence de peptide avec une faible affinité. ③ La présence d'un peptide en cours d'élongation augmente l'affinité de TF pour le ribosome, qui favorise les premières étapes du repliement conformationnel ou empêche les interactions non spécifiques dans l'attente d'informations suffisantes sur la conformation définitive. ④ Après sa dissociation du ribosome, TF peut soit se détacher de la chaîne peptidique si elle a atteint sa conformation native, soit y rester lié pour favoriser son transfert à d'autres systèmes chaperons ou l'assemblage de protéines multimériques. ⑤ Après sa dissociation de la protéine cliente, la forme libre de TF se trouve à nouveau en équilibre entre état monomérique et dimérique.

▪ Cycle réactionnel

TF est cycliquement lié ou non au ribosome, de manière ATP-indépendante, et forme une arche protectrice au niveau du tunnel de sortie protéique du ribosome. Son cycle réactionnel est décrit la [Figure 19B](#). La liaison de TF au ribosome est indispensable à son activité chaperon. Cette fixation peut être relativement tardive au cours de l'élongation protéique. En effet, la présence d'une chaîne peptidique en cours de traduction augmente significativement l'affinité de TF pour le ribosome [115]. Suite à son détachement du ribosome, TF peut néanmoins rester lié à la chaîne peptidique en cours d'élongation, de façon à faciliter le transfert de son substrat aux systèmes chaperons intervenant en aval ([Figure 19B](#)) [115].

▪ Substrats

TF interagit de manière co-translationnelle avec la majorité des protéines synthétisées, et 60 à 70% d'entre elles acquièrent leur conformation définitive sans la participation d'autres systèmes chaperons [116].

▪ **Fonction cellulaire**

De récents travaux confirment que TF agit majoritairement selon 2 mécanismes, non mutuellement exclusifs, pour favoriser le repliement des protéines néo-synthétisées [117] :

- la cavité formée par les domaines N- et C-terminaux de TF fait office d'espace de repliement protéique privilégié permettant l'assemblage de petits domaines ;
- la fixation de TF sur des longues chaînes peptidiques en cours d'élongation retarde leur repliement jusqu'à l'obtention d'informations suffisantes sur la conformation définitive, évitant ainsi la formation d'intermédiaires protéiques incorrects.

▪ **Mutation et phénotypes associés**

La mutation du gène *tig* codant TF chez *E. coli* est tolérée et n'induit pas d'effet majeur sur la croissance bactérienne [116]. L'absence d'altération phénotypique majeure s'explique par le fait que l'activité de TF est efficacement substituée par le système DnaKJE [118].

2.3. Système Hsp70 : complexe [DnaK–DnaJ–GrpE]

▪ **Généralités**

Le système Hsp70 est une machinerie complexe comprenant la protéine chaperon DnaK (Hsp70), sa protéine co-chaperon DnaJ (analogue de Hsp40), et le facteur d'échange nucléotidique GrpE. Ce complexe [DnaK–DnaJ–GrpE], encore appelé DnaKJE, est exprimé à un niveau basal relativement important, et peut être induit en réponse à différents stress [119].

▪ **Structure**

DnaK est une protéine de 638 acides aminés composée de 3 domaines (Figure 20A) [120] :

- un domaine N-terminal de liaison au nucléotide (ou NBD, pour *nucleotide binding domain*) portant l'activité ATPasique ;
- une région *linker* hautement conservée relie les domaines N- et C-terminaux ;
- un domaine C-terminal de liaison au substrat (ou SBD, pour *substrate binding domain*).

DnaJ est une protéine de 376 acides aminés, fonctionnelle sous forme dimérique, qui possède un domaine J caractéristique, indispensable à son interaction avec DnaK (Figure 20A). Elle permet le recrutement des protéines clientes et stimule l'activité ATPase de DnaK [121].

GrpE est une protéine de 197 acides aminés, appartenant à la famille des facteurs d'échanges nucléotidiques, qui favorise la dissociation de l'ADP de DnaK, ainsi que la libération des protéines clientes conformées (Figure 20A) [121].

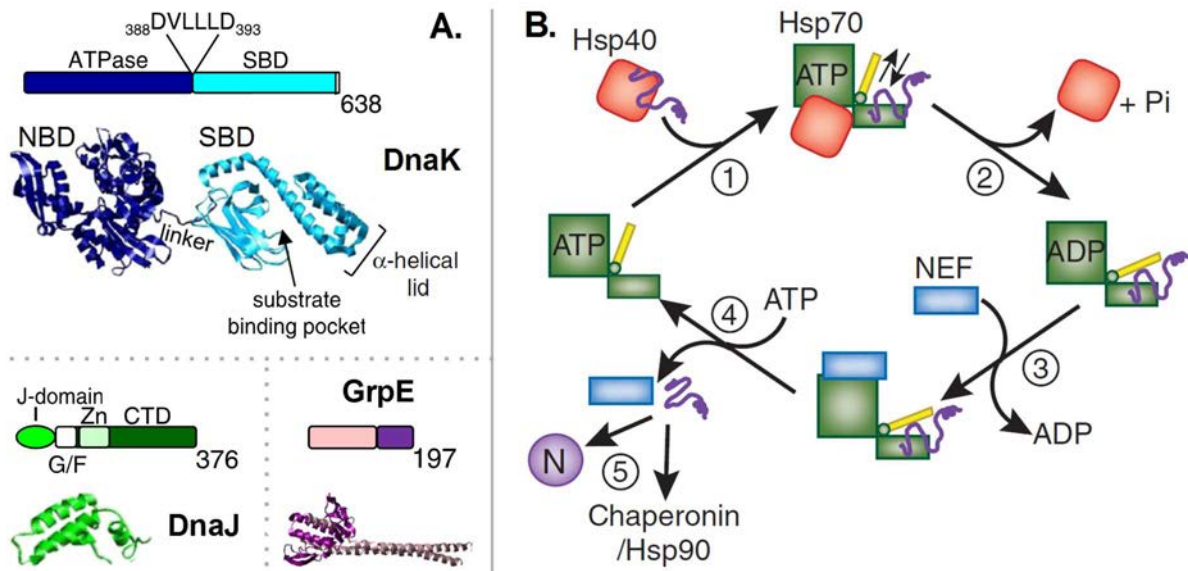


Figure 20. Structure et cycle réactionnel du système chaperon Hsp70 (DnaKJE)

A. Organisation des domaines protéiques et structure tridimensionnelle de DnaK, DnaJ et GrpE [111].

B. Cycle réactionnel du système DnaKJE [114]. ① DnaK, lié à l'ATP et en conformation ouverte, présente une faible affinité pour ses substrats. Les protéines clientes sont donc recrutées par la protéine co-chaperon DnaJ. ② La liaison de DnaJ sur DnaK catalyse l'hydrolyse de l'ATP en ADP, puis modifie l'état de DnaK en conformation fermée, et favorise le repliement conformationnel de la protéine cliente. DnaJ se dissocie alors de DnaK. ③ La fixation du facteur d'échange de nucléotide GrpE entraîne la dissociation de l'ADP. ④ La liaison d'une nouvelle molécule d'ATP sur DnaK permet sa transition en conformation ouverte, et la libération de la protéine cliente et de GrpE. ⑤ La protéine cliente libérée peut atteindre sa conformation définitive, être prise en charge par d'autres systèmes chaperons, ou se lier de nouveau à DnaJ pour subir un nouveau cycle.

▪ Cycle réactionnel

Durant le cycle de la protéine chaperon DnaK, détaillé dans la [Figure 20B](#), le domaine de liaison aux substrats alterne entre un état ouvert ou fermé selon la nature du nucléotide fixé sur le domaine ATPase. Lorsqu'une molécule d'ATP est liée à DnaK, le SBD en conformation ouverte présente une faible affinité pour ses substrats, avec une cinétique d'échange rapide. Réciproquement, lorsqu'une molécule d'ADP est liée au NBD, le SBD en conformation fermée possède une affinité élevée et une cinétique d'échange ralentie [122].

Si l'hydrolyse de l'ATP est primordiale à l'activité chaperon de DnaK, celle-ci possède pourtant une faible activité ATPasique. Son cycle réactionnel est ainsi contrôlé par son co-chaperon DnaJ, qui cible les substrats mais surtout stimule l'hydrolyse de l'ATP par le domaine ATPase de DnaK. Le cofacteur GrpE favorise pour sa part la dissociation de l'ADP et la libération des substrats du SBD de DnaK [122].

▪ Substrats

DnaK se lie préférentiellement à de courtes séquences polypeptidiques hydrophobes, accessibles sur des protéines néo-synthétisées en cours de repliement, sur des protéines devant être exportées, ou sur des protéines dénaturées [123]. Une récente analyse de l'interactome de DnaK chez *E. coli* a ainsi révélé plus de 600 substrats potentiels en conditions physiologiques, dont 80% sont retrouvés au niveau du cytosol, 11% de la membrane interne, 3% de la membrane externe et 3% du périplasme [124].

▪ Fonctions cellulaires

Il est admis que le système DnaKJE intervient dans le repliement co- et post-translationnel d'environ 15% des protéines néo-synthétisées [116]. DnaK participe également au transport des protéines sécrétées, à l'assemblage de complexes multimériques, ainsi qu'au repliement des protéines dénaturées, en collaboration avec d'autres systèmes chaperons [118].

▪ Mutation et phénotypes associés

Le rôle central du système DnaKJE dans le contrôle qualité protéique est confirmé par la délétion du gène *dnaK*. Un mutant $\Delta dnaK$ présente plusieurs caractéristiques phénotypiques notables : une cryosensibilité en dessous de 20°C, une croissance lente à 30°C avec filamentation et perte de mobilité, et une thermosensibilité au-dessus de 35°C [119].

2.4. Système Hsp60 : complexe [GroEL–GroES]

▪ Généralités

Le système Hsp60 est le seul système chaperon essentiel à la viabilité de *E. coli* [125]. Composé de GroEL, complexe protéique en forme de tonneau, et de GroES, co-chaperon obligatoire semblable à un couvercle, l'ensemble forme la machinerie GroESL, véritable cage hydrophobe permettant le repliement protéique à l'abri de toute interaction non spécifique.

▪ Structure et cycle réactionnel

GroEL est un complexe protéique en forme de tonneau composé de deux anneaux collés dos-à-dos, qui forment deux cavités pouvant accueillir des protéines non conformées [126]. Chaque sous-unité comprend trois domaines, de l'extérieur vers l'intérieur (Figure 21A) [111] :

- le domaine apical est impliqué dans le recrutement des substrats et la liaison à GroES ;
- le domaine intermédiaire permet la communication allostérique entre les domaines ;
- le domaine équatorial est nécessaire à la liaison de l'ATP.

GroES est un complexe protéique en forme de couvercle qui, en obturant les cavités de GroEL, fournit un environnement protégé aux substrats de ce système (Figure 21A) [126].

Le cycle réactionnel du système GroESL est présenté dans la Figure 21B [114].

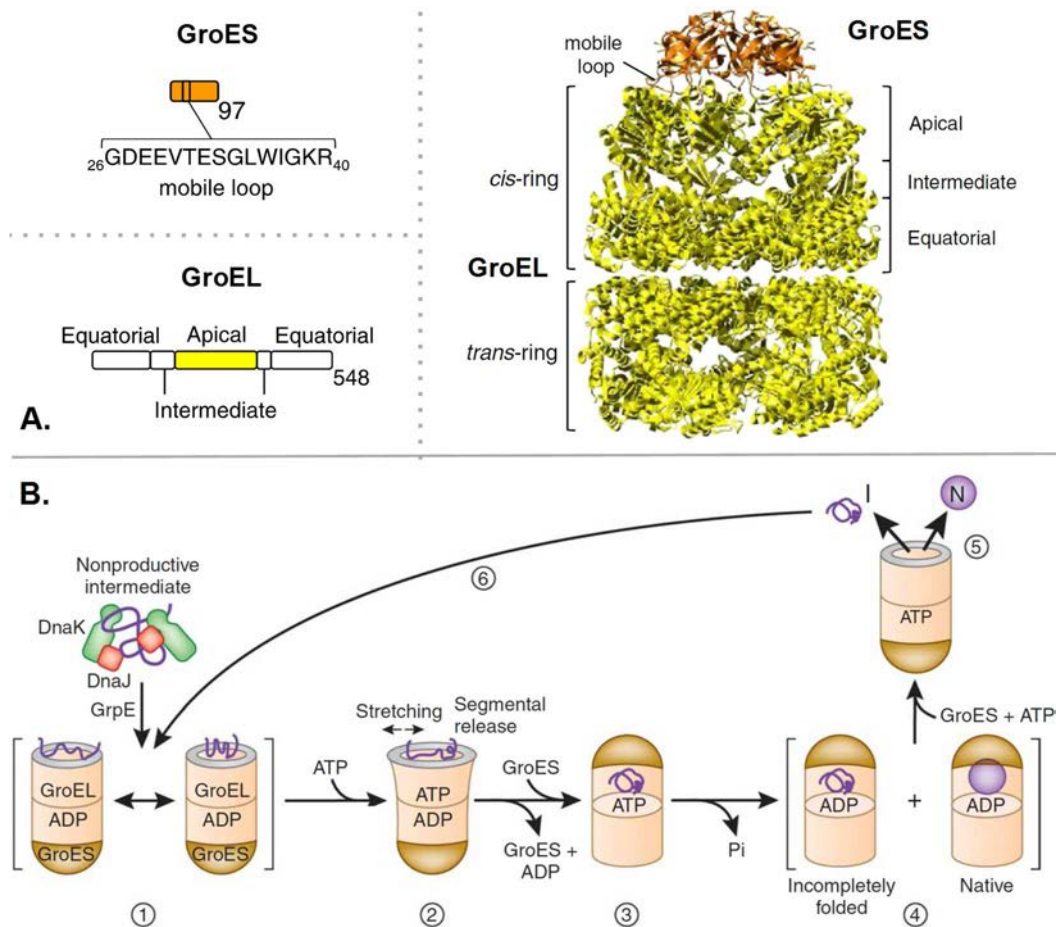


Figure 21. Structure et cycle réactionnel du système chaperon Hsp60 (GroESL)

A. Organisation des domaines protéiques et structure tridimensionnelle de GroEL et GroES [111].

B. Cycle réactionnel du système GroESL [114]. ① Les protéines clientes sont généralement transférées du système DnaKJE vers le système GroESL. Lors de sa liaison à un complexe GroEL, la protéine cliente oscille entre différents états conformationnels, lâches ou compacts. ② La fixation de l'ATP sur ce complexe entraîne un élargissement de la cavité GroEL, permettant à la protéine cliente de pénétrer complètement dans le tunnel. ③ La liaison du complexe GroES termine alors l'encapsulation du substrat. ④ La protéine cliente peut alors se replier dans l'espace isolé que constitue la cavité du complexe GroESL, jusqu'à hydrolyse de l'ATP. ⑤ La dissociation de GroES permet la libération du substrat, dans sa conformation définitive ou non. ⑥ Les protéines imparfaitement conformées peuvent nécessiter un nouveau cycle réactionnel.

▪ Substrats

Il a été montré que le système GroESL interagit avec 250 substrats différents, soit approximativement 10% de l'ensemble des protéines cytosoliques [127]. Ces 250 protéines ont été réparties en différentes classes selon leur dépendance vis-à-vis du système chaperon.

▪ Fonctions cellulaires

La machinerie GroESL est impliquée dans le repliement conformationnel des protéines néo-synthétisées en aval des systèmes TF et DnaKJE. Deux mécanismes ont été proposés pour le repliement protéique assuré par le système chaperon Hsp60 :

- la cavité centrale agit comme une cage inerte, au sein de laquelle le repliement des protéines s'exécute à l'abri d'interactions non spécifiques avec le contenu cytosolique pouvant aboutir à la formation d'intermédiaires mal conformés ou d'agrégats [128] ;
- la cavité centrale permet de dénaturer activement les intermédiaires protéiques mal engagés dans le repliement, et de favoriser grâce à plusieurs cycles réactionnels l'obtention de conformations protéiques adéquates [129].

▪ Mutation et phénotypes associés

GroESL possèdent 85 substrats obligatoires identifiés *in vitro*, et 57 d'entre eux ont été confirmés *in vivo*. Parmi ces substrats, six sont des protéines essentielles chez *E. coli*. Ceci peut expliquer le caractère essentiel du système GroESL pour la viabilité bactérienne quelles que soient les conditions [125]. Les mutations viables de l'opéron *groESL* induisent généralement une thermosensibilité, un ralentissement de la croissance et une filamentation [130].

2.5. Coopération entre les systèmes TF, DnaKJE et GroESL

Les protéines néo-synthétisées peuvent donc être successivement prises en charge par les trois systèmes chaperons TF, DnaKJE et GroESL, qui favorisent le repliement protéique *de novo* et forment ainsi un réseau interactif assurant un contrôle qualité co- et post-traductionnel efficace, où l'acteur central est le système DnaKJE [114].

De nombreuses études ont mis en évidence les redondances existantes entre ces trois systèmes. En l'absence de TF, le nombre de substrats de DnaK augmente considérablement, passant de 654 à 998 [124]. De plus, si la délétion de *tig* ou *dnaK* est tolérée par *E. coli*, une souche mutée pour ces deux gènes n'est pas viable à des températures excédant 30°C [118]. Il a par ailleurs été montré que le système GroESL est surexprimé en l'absence de TF et DnaK, et interagit avec 150 substrats supplémentaires [127]. Inversement, un mutant viable de GroESL est caractérisé par une surexpression de DnaKJE et une modification de l'interactome de DnaK [124]. Ces résultats confirment l'interconnexion des trois systèmes.

La Figure 22 synthétise le contrôle qualité des protéines néo-synthétisées assuré par le réseau interactif des systèmes chaperons TF, DnaKJE et GroESL.

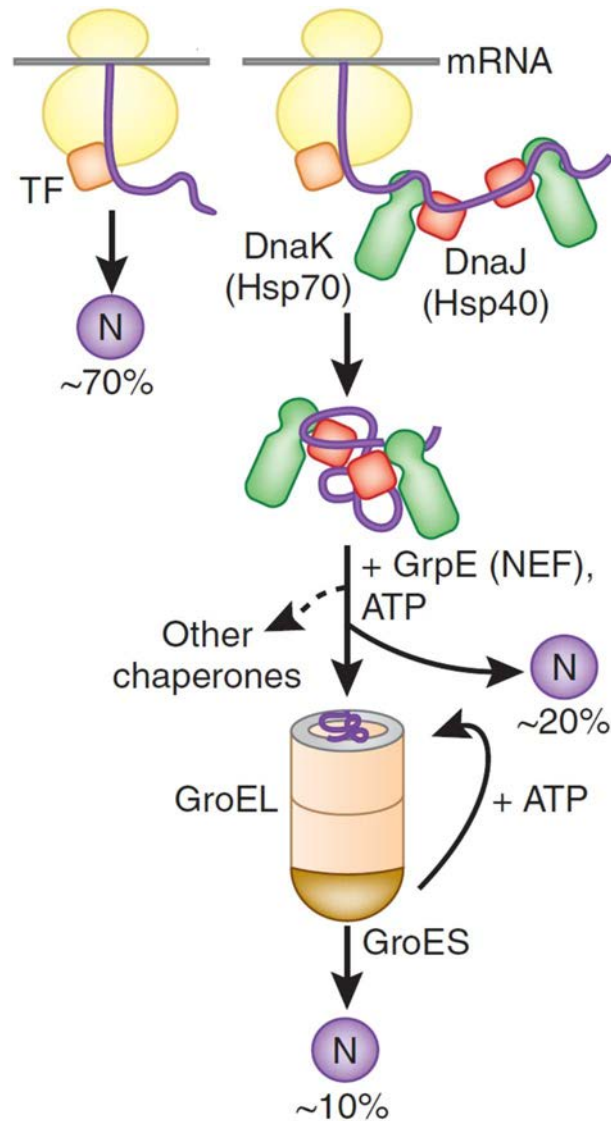


Figure 22. Vue d'ensemble du contrôle qualité des protéines néo-synthétisées

La majorité des protéines néo-synthétisées interagit de manière co-traductionnelle avec la protéine TF. 70% des protéines sont alors capables d'atteindre leur conformation définitive rapidement après leur synthèse, sans l'intervention d'autres systèmes chaperons. Les protéines de taille plus conséquentes (environ 30%) vont être prises en charge par la machinerie DnaKJE, qui va permettre le repliement de deux tiers d'entre elles grâce à un ou plusieurs cycles réactionnels ATP-dépendants. Les protéines n'ayant pas encore réussi à atteindre leur état natif vont alors être encapsulées par le système GroESL pour leur repliement [114].

3. Contrôle qualité des protéines en situations de stress

Les bactéries sont confrontées en permanence à diverses conditions environnementales défavorables. Elles doivent donc s'adapter rapidement pour faire face à un choc thermique, un stress oxydatif, une variation de pH... ; autant de situations pouvant entraîner la dénaturation voire l'agrégation des protéines, et perturber l'homéostasie protéique [131]. Les bactéries, dont *E. coli*, ont ainsi développé diverses stratégies préventives ou curatives pour répondre aux altérations que peuvent subir les protéines en situations de stress (Figure 23).

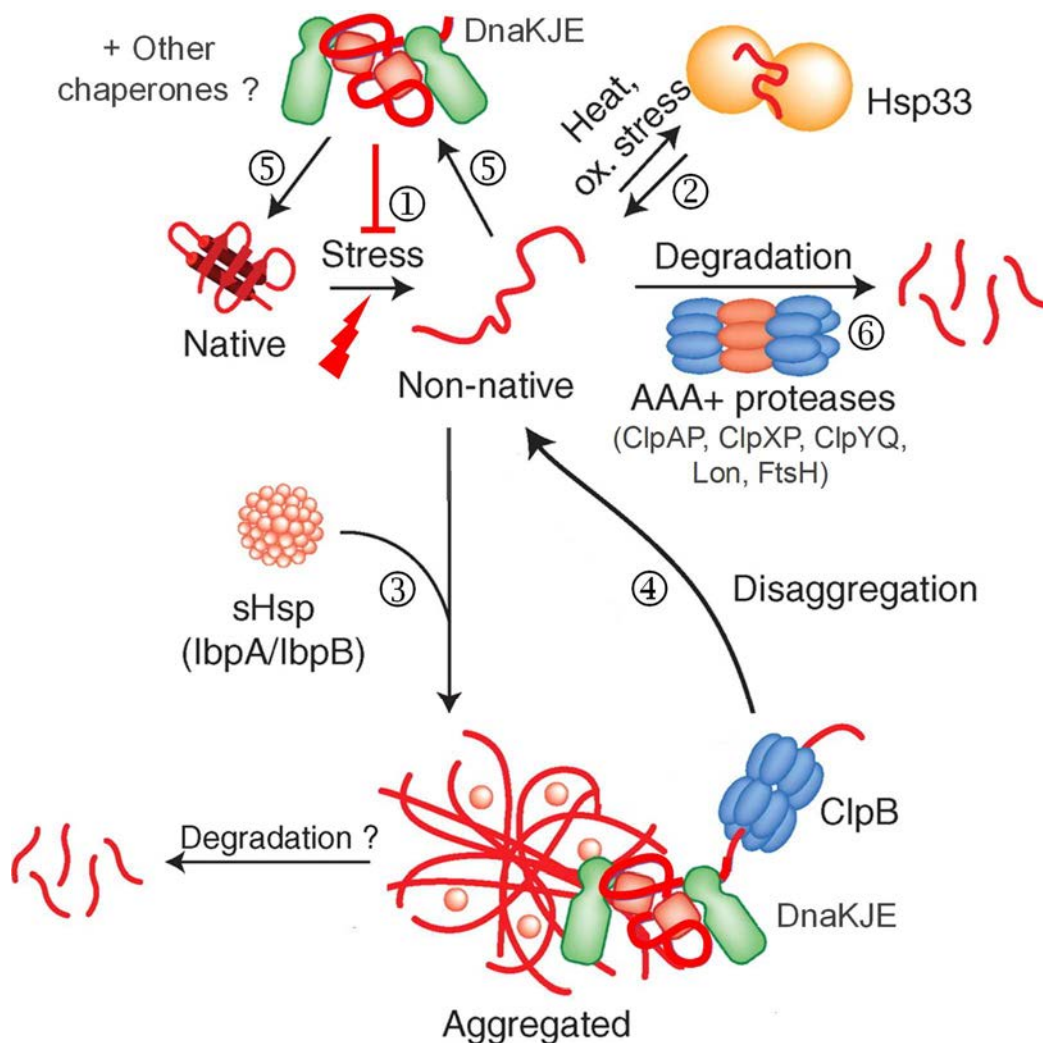


Figure 23. Vue d'ensemble du contrôle qualité protéique en situations de stress

Stratégies de maintien de l'homéostasie protéique en situations de stress détaillées pages suivantes.
D'après [108].

3.1. Mesures préventives

3.1.1. Prévention de la dénaturation

La première stratégie mise en place par le contrôle qualité protéique lors d'un choc thermique consiste à maintenir les protéines en conformation soluble ([Figure 23 – ①](#)) [108].

Chez *E. coli*, le système DnaKJE joue un rôle majeur dans la prévention de la dénaturation, probablement grâce à sa concentration cellulaire élevée et son importante capacité de liaison aux substrats [132]. De plus, le cofacteur GrpE, facteur d'échange de nucléotide de DnaK, est inactivé de manière réversible lors d'un choc thermique. Cette inactivation entraîne un important retard de libération des protéines clientes en ralentissant les échanges ATP/ADP, car DnaK reste figé dans sa conformation fermée à haute affinité pour son substrat [133]. GrpE agit ainsi comme un senseur qui favorise le maintien des protéines liées à DnaK durant le stress thermique.

3.1.2. Prévention de l'oxydation

In vivo, un choc thermique est souvent accompagné d'un stress oxydatif. Le maintien de l'homéostasie protéique dans cette situation devient particulièrement complexe, car les systèmes DnaKJE et GroESL sont inactivés en conditions oxydantes [134]. Cette inactivation s'explique par la diminution rapide et massive de la concentration intracellulaire en ATP.

En réponse au stress oxydatif, la protéine Hsp33 va prendre le relais du système DnaKJE ([Figure 23 – ②](#)) [134]. Cette protéine chaperon possède un mécanisme d'activation particulier. Elle présente dans son site actif quatre résidus cystéine qui stabilisent un atome de zinc en coordination tétraédrique en conditions physiologiques. Lors d'une exposition concomitante à une température élevée et un stress oxydatif, Hsp33 subit d'importantes modifications structurales secondaires à l'oxydation des cystéines en ponts disulfures intermoléculaires. La protéine dimérisée présente ainsi une activité chaperon ATP-indépendante par formation de complexes stables avec ses substrats [135].

En conditions redevenues normales, le site actif de Hsp33 va être réduit grâce à l'action de réductases cellulaires. Pour autant, elle reste complexée à ses substrats. Leur libération va alors nécessiter l'action du système DnaKJE, qui intervient dès que le taux d'ATP cellulaire atteint de nouveau des concentrations suffisantes pour permettre le repliement protéique [134].

3.1.3. Modulation de l'agrégation

Lorsque les capacités des systèmes DnaKJE et Hsp33 sont dépassées, les protéines vont être potentiellement dénaturées. En situation de stress prolongé, ces protéines altérées peuvent même former des agrégats protéiques, dont la dissociation est particulièrement complexe et énergivore pour la bactérie.

Afin de limiter ce processus, le contrôle qualité met en jeu les protéines IbpA et IbpB, deux chaperons moléculaires de 16 kDa qui constituent la famille des petites protéines de choc thermique (ou sHsp, pour *small heat shock protein*) ([Figure 23 – ③](#)) [136]. Ces deux protéines fonctionnent conjointement sous forme hétérodimérique, en fixant des protéines clientes mal conformées et en les séquestrant dans un état non conformé. Au fur et à mesure que les substrats sont recrutés, les complexes s'agrandissent jusqu'à former des agrégats protéiques organisés, dont la dissolution par le système ClpB/DnaK sera ensuite facilitée (*cf.* § 3.2.1.).

La production des protéines IbpA et IbpB est fortement induite lors d'un choc thermique. Leur activation est de plus facilitée par l'élévation de température, qui favorise leur oligomérisation et augmente leur affinité pour les protéines clientes altérées [137].

3.2. Mesures curatives

3.2.1. Désagrégation protéique

Des conditions de stress sévères ou prolongées vont conduire à l'accumulation d'agrégats protéiques après saturation des systèmes et des stratégies préventives évoqués ci-dessus. Cependant, l'action combinée de la protéine chaperon ClpB et du système DnaKJE assure le sauvetage des protéines contenues dans ces agrégats, d'abord par resolubilisation, puis grâce à une nouvelle étape de repliement conformationnel ([Figure 23 – ④](#)) [138].

Le processus de désagrégation requiert l'intervention séquentielle de DnaKJE puis ClpB ([Figure 24](#)). La fixation initiale de DnaK sur l'agrégat protéique en limite l'accès pour les protéases et les autres systèmes chaperons, et facilite l'aiguillage des peptides agglomérés vers ClpB (i) [139]. ClpB exerce alors une force de traction sur la protéine à solubiliser, l'obligeant à traverser son pore central et à s'extraire de l'agrégat protéique dans lequel elle était piégée (ii). Ce processus actif ATP-dépendant sera ainsi répété afin d'extraire les peptides un par un jusqu'à dissolution complète de l'agrégat (iii) [140]. Ces peptides subiront ensuite l'action d'autres systèmes chaperons pour retrouver leur conformation active (iv).

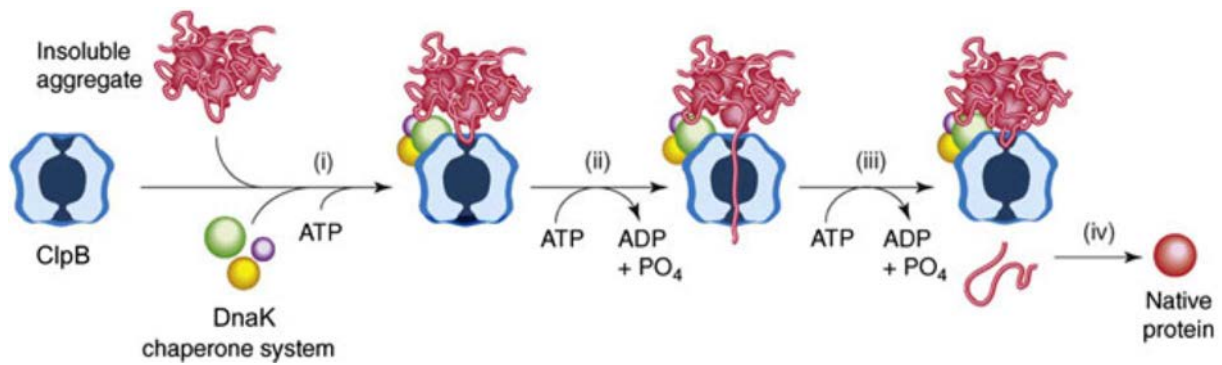


Figure 24. Mécanisme de désagrégation protéique par ClpB et DnaKJE

Étapes du processus de désagrégation protéique détaillées page précédente [140].

3.2.2. Repliement conformationnel

Si les protéines dénaturées lors d'un stress et celles resolubilisées par ClpB et DnaKJE n'ont pas subi d'altérations trop importantes, elles peuvent être de nouveau conformées par l'action d'un ou plusieurs systèmes chaperons (Figure 23 – ©).

Les systèmes DnaKJE et GroESL, impliqués dans le repliement des protéines *de novo*, sont également nécessaires au repliement des protéines dénaturées pour acquérir à nouveau une conformation adéquate [141]. Ils ont été largement décrits plus haut (*cf.* § 2.3. et § 2.4.).

La protéine chaperon Hsp90_{EC}, encore appelée HtpG, semble également impliquée dans le repliement des protéines mal conformées suite à un stress [142]. Les fonctions de cette protéine chaperon ATP-dépendante, analogue procaryote du système Hsp90 largement étudié chez les organismes eucaryotes, sont peu connues à ce jour. La plupart des articles faisant une revue de la littérature sur le contrôle qualité protéique chez les bactéries excluent Hsp90_{EC} dans la description du réseau de protéines chaperons [108,111,141,143]. Nous faisons le choix de lui consacrer une section entière (*cf.* § 4.).

3.2.3. Élimination des protéines mal conformées

Il est fréquent que les protéines aient subi des altérations majeures ou irréversibles, auquel cas les systèmes chaperons ne seront pas capables de restaurer leur conformation. Ces protéines endommagées sont alors prises en charge par l'une des cinq protéases de *E. coli* : ClpAP, ClpXP, ClpYQ, Lon ou FtsH (Figure 23 – ©) [144].

Ces 5 systèmes protéasiques sont tous composés de 2 unités à activités distinctes [144] :

- une unité à activité chaperon recrutant les substrats ;
- une unité à activité protéasique en forme de tunnel dégradant les substrats.

Ces deux activités peuvent être portées par une seule chaîne protéique, comme pour les protéases Lon et FtsH. Les autres systèmes, ClpAP, ClpXP et ClpYQ, comprennent deux chaînes : ClpA, ClpX et ClpY correspondent à l'activité chaperon, ClpP et ClpQ à l'activité protéasique [144].

Les protéases assurent la dégradation des protéines trop altérées pour être de nouveau conformées, mais aussi dans certains cas la dissolution des agrégats protéiques, ou encore la protéolyse de nombreux régulateurs de voies de signalisation [108]. Elles sont donc impliquées dans la régulation de divers processus cellulaires, et notamment dans la régulation du contrôle qualité protéique et la réponse au choc thermique.

Le processus protéolytique débute par la reconnaissance de la protéine cliente à dégrader par l'unité à activité chaperon (Figure 25). Elle procède au dépliement conformationnel du substrat, qui traverse alors un canal protéique formé par cette unité, jusqu'à pénétrer dans la chambre catalytique de l'unité protéasique [144]. La chaîne peptidique est ainsi fragmentée en petits peptides et acides aminés, qui seront recyclés pour la synthèse d'autres protéines.

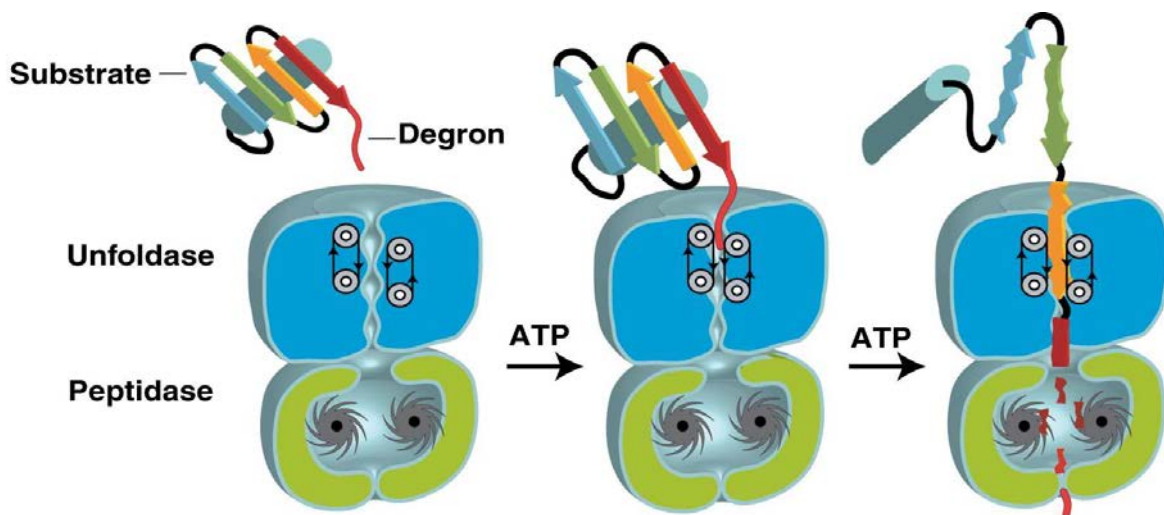


Figure 25. Mécanisme de dégradation protéique par les protéases

Étapes du processus de protéolyse détaillées dans le paragraphe précédent [137].

3.2.4. Élimination des agrégats protéiques

Il peut arriver qu'un agrégat protéique ne puisse pas être résolu par l'action des protéines chaperons ou des protéases [145]. Ces agrégats, composés de nombreuses protéines altérées, ont alors tendance à grossir en interagissant avec d'autres protéines mal conformées. Leur persistance est potentiellement dangereuse pour la cellule, car ils servent de support à l'agrégation de nouvelles protéines, séquestrent généralement divers acteurs du contrôle qualité protéique, et ralentissent la multiplication cellulaire [146].

Les organismes vivants ont ainsi développé une stratégie face à cette agrégation terminale : la réorganisation asymétrique (Figure 26) [147]. L'agrégat protéique insoluble va d'abord être délocalisé vers l'un des pôles de la cellule. Lors de la division cellulaire, une seule des deux cellules filles obtenues possédera cet agrégat. Cette cellule présentera une croissance lente, alors que l'autre ne sera pas impactée puisqu'aucun agrégat protéique ne ralentira sa division [146]. La délocalisation de l'agrégat à l'un des pôles cellulaires implique vraisemblablement un processus énergétique actif qui n'a pour le moment pas été élucidé.

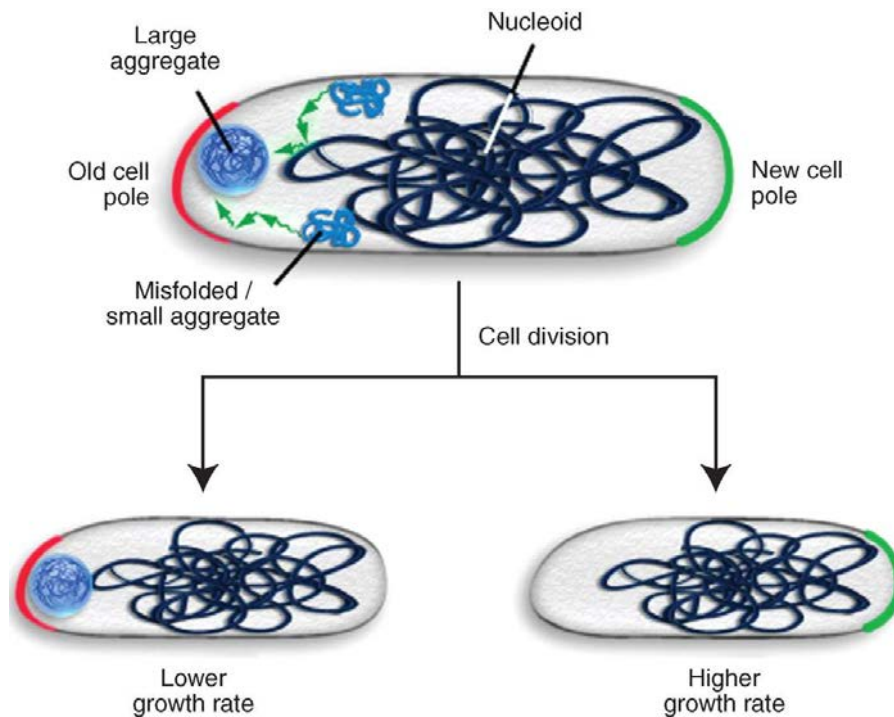


Figure 26. Élimination des agrégats protéiques par réorganisation asymétrique

Étapes du processus de réorganisation asymétrique détaillées dans le paragraphe précédent [108].

4. Focus sur la machinerie Hsp90

4.1. Généralités

Hsp90 est une protéine chaperon ATP-dépendante ubiquitaire, hautement conservée, et abondamment exprimée. Elle participe au contrôle qualité des protéines cytosoliques en favorisant leur repliement conformationnel. Hsp90 appartient aux protéines de choc thermique Hsp, induites en réponse à une élévation de température ou d'autres situations de stress pour assurer le maintien de l'homéostasie protéique intracellulaire.

La famille Hsp90 est retrouvée chez tous les organismes eucaryotes et procaryotes, et peut ainsi être divisée en 5 sous-familles selon la localisation : Hsp90 cytosolique eucaryote, Grp94 dans le réticulum endoplasmique, Trap1 mitochondriale, Hsp90c au niveau des chloroplastes, et HtpG présente chez les bactéries [148]. Cette famille est de plus hautement conservée. La comparaison des séquences entre les différents analogues montre par exemple 40% d'identité de séquence et 55% de similitudes entre Hsp90 humaine et HtpG [149].

Chez *E. coli*, l'analogue procaryote de Hsp90, codé par le gène *htpG*, est appelé HtpG (pour *high-temperature protein G*), ou encore Hsp90_{Ec}, notation qui sera utilisée ici. Cette protéine est abondamment exprimée en conditions normales de croissance, et est en outre induite au cours d'un choc thermique [150]. Son taux peut également être modifié selon la disponibilité des nutriments dans le milieu extracellulaire [151].

Si Hsp90 est essentielle chez les organismes eucaryotes, Hsp90_{Ec} n'est pas indispensable à la viabilité et la survie chez *E. coli*, même en conditions de stress. Un ralentissement de la croissance est cependant observé lorsque la température de culture dépasse 40°C [152]. L'induction d'un stress thermique augmente aussi considérablement le nombre de protéines agrégées chez une souche de *E. coli* mutée pour *htpG* [142].

Cette partie propose de faire une revue de la littérature sur la protéine chaperon Hsp90_{Ec}. Si les données structurales sont relativement abondantes, le mécanisme d'action, les substrats et les fonctions cellulaires de cet analogue procaryote sont moins bien connus.

4.2. Structure et cycle réactionnel

Hsp90_{Ec} est une protéine active sous forme dimérique. Chaque homodimère est constitué par 3 domaines protéiques (Figure 27A) [148] :

- le domaine N-terminal (ou NTD, pour *N-terminal domain*) comporte le site de liaison au nucléotide (NBD), impliqué dans la liaison et l'hydrolyse de l'ATP, ainsi qu'une région formant un couvercle pour la cavité protéique du NBD [153] ;
- le domaine intermédiaire (ou MD, pour *middle domain*) assure la jonction entre les domaines N et C-terminaux, et est responsable de la flexibilité de la molécule [153] ;
- le domaine C-terminal (ou CTD, pour *C-terminal domain*) est nécessaire à la dimérisation de la protéine chaperon [154].

La structure tridimensionnelle de Hsp90 est conservée chez les organismes eucaryotes et procaryotes [155]. Cependant, contrairement à son homologue eucaryote, Hsp90_{Ec} ne possède pas de domaine de liaison *linker* entre les domaines NTD et MD, ni de motif conservé MEEVD à l'extrémité N-terminale, deux éléments essentiels aux interactions entre la machinerie Hsp90 et ses co-chaperons [148]. Ceci explique que Hsp90_{Ec} fonctionne de manière indépendante, à la différence des analogues eucaryotes de Hsp90, qui nécessitent l'intervention de nombreuses protéines co-chaperons pour fonctionner correctement [156].

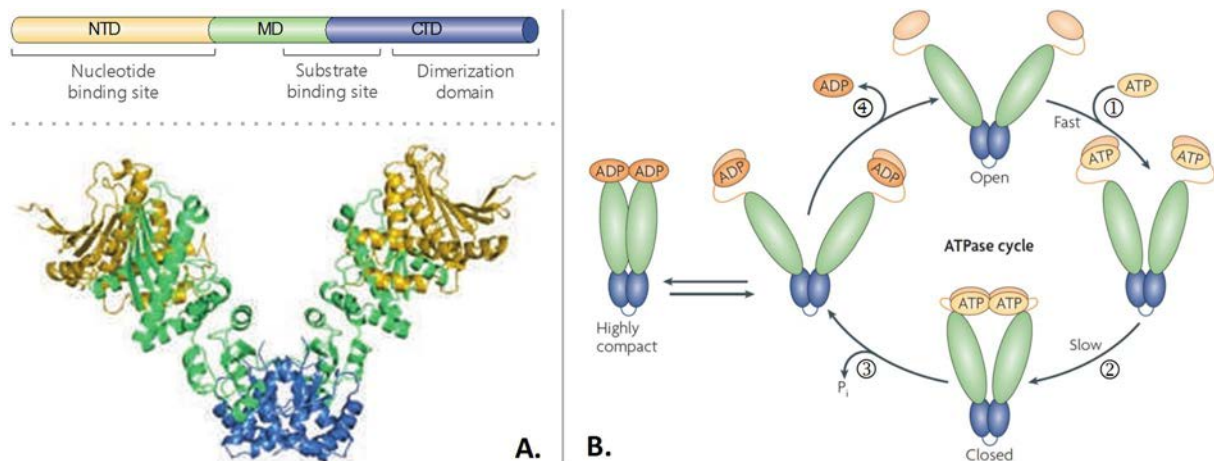


Figure 27. Structure et cycle réactionnel de la protéine chaperon Hsp90_{Ec}

A. Organisation des domaines protéiques et structure tridimensionnelle de Hsp90_{Ec}. D'après [148].

B. Cycle ATPasique de Hsp90_{Ec} [148]. ① & ② La fixation d'ATP sur Hsp90_{Ec} induit une transition en 2 étapes entre les conformations ouverte et fermée de la protéine chaperon. ③ L'hydrolyse de l'ATP en ADP entraîne une compaction transitoire suivie d'un relâchement conformationnel de Hsp90_{Ec}. ④ La dissociation de l'ADP permet le retour de Hsp90_{Ec} en conformation ouverte.

Une récente étude a permis de mettre en évidence une région très conservée de Hsp90_{Ec}, essentielle pour la liaison des protéines clientes, composée par plusieurs résidus d'acides aminés des domaines MD et CTD, et exposée à la surface de la protéine [157]. Ces résidus semblent importants à la fois pour l'interaction avec la protéine cliente, mais également pour son remodelage.

La protéine chaperon Hsp90_{Ec} est une molécule très flexible, pouvant subir de nombreux réarrangements, qui compliquent particulièrement la détermination de sa structure [158,159]. Elle présente une conformation en forme de « V » lorsqu'elle n'est pas liée à un nucléotide ou un substrat [155]. L'angle d'ouverture de la protéine libre est d'environ 75°, mais peut aller jusqu'à 120° dans certaines conditions [160].

Hsp90_{Ec} est une protéine chaperon ATP-dépendante, et présente un cycle réactionnel qui passe par plusieurs états structuraux relativement conservés entre les organismes vivants (Figure 27B) [161]. La liaison d'ATP sur Hsp90_{Ec} induit une réduction de l'angle d'ouverture, jusqu'à aboutir à une conformation fermée de la molécule, dans laquelle les NTD de chaque monomère interagissent [155]. Hsp90_{Ec} présente également une activité ATPasique, bien que relativement faible, au niveau du domaine N-terminal [162]. L'hydrolyse de l'ATP en ADP induit alors une compaction importante puis un relâchement de la molécule. Ceci permet la libération de l'ADP et la restauration de la conformation ouverte de Hsp90_{Ec} non liée [155,163]. La liaison d'une protéine cliente pourrait également induire des réarrangements moléculaires, non caractérisés pour le moment, mais aussi potentialiser l'activité ATPasique de Hsp90_{Ec} [164].

4.3. Fonctions cellulaires et protéines clientes

Hsp90 est une protéine chaperon essentielle chez les organismes eucaryotes, impliquée dans divers processus cellulaires, tels que le trafic protéique, la transduction de signaux ou la maturation de récepteurs [165]. Elle agit en facilitant le remodelage, l'activation et la régulation de plus de 200 protéines clientes, avec l'assistance de nombreux co-chaperons [56].

Les fonctions cellulaires et les protéines clientes de Hsp90_{Ec} sont beaucoup moins connues. Deux études de l'interactome de Hsp90 chez *E. coli*, réalisées par Butland *et al.* en 2005 [166], et par Arifuzzaman *et al.* en 2006 [167], ont pourtant permis d'identifier respectivement 15 et 54 candidats potentiels. Cependant, seuls deux substrats sont communs aux deux études, et aucune protéine cliente n'a pour le moment été confirmée.

▪ Fonctions cellulaires et substrats chez *E. coli*

Les deux seules protéines clientes interagissant avec Hsp90_{Ec} identifiées et confirmées chez *E. coli* sont la protéine ribosomale L2, un composant essentiel de la sous-unité ribosomale 50S [168], et la protéine DnaA, facteur d'initiation majeur de la réplication de l'ADN chromosomique [169]. Dans le cas de la protéine L2, Hsp90_{Ec} reconnaît spécifiquement sa protéine cliente grâce à une région comprise entre les domaines MD et CTD, et son activité ATPasique est augmentée d'un facteur 10 environ. Ces données sont en accord avec les études ayant respectivement caractérisé le site de liaison aux substrats [157] et montré que la liaison d'une protéine cliente potentialise l'activité ATPasique de la protéine chaperon Hsp90_{Ec} [164].

Par ailleurs, plusieurs études ont récemment démontré que Hsp90_{Ec} est nécessaire au bon fonctionnement de différents systèmes et processus cellulaires, sans pour autant clairement identifier le(s) partenaire(s) protéique(s) interagissant directement avec la protéine chaperon.

Hsp90_{Ec} est ainsi essentiel à l'activité du système CRISPR (pour *clustered regularly interspaced short palindromic repeats*), un important système de protection bactérien vis-à-vis des virus et des transferts horizontaux d'ADN et d'ARN [170]. Il confère aux bactéries une immunité adaptative qui module l'intégration d'acides nucléiques étrangers dans leur propre génome. Yosef *et al.* ont démontré dans cette étude que le système CRISPR était inactivé dans une souche mutée pour le gène *htpG*. Cette inactivation est liée à une diminution du niveau d'expression de la protéine Cas3 (protéine de la famille Cas, pour *CRISPR-associated protein*). Dans ce système, la protéine chaperon Hsp90_{Ec} semble donc jouer un rôle de régulateur positif.

Afin de déterminer d'autres fonctions cellulaires de Hsp90_{Ec} chez *E. coli*, Grudniak *et al.* ont récemment étudié les modifications phénotypiques observées chez une souche mutée pour *htpG* [171]. Ils ont démontré que cette souche présente une diminution de la synthèse et de la stabilité d'enzymes de types β -lactamases ou phosphatases alcalines, ainsi qu'une altération de la capacité à former des biofilms lorsque la température dépasse 42°C [171]. De précédentes études ont également mis en évidence le rôle de Hsp90_{Ec} dans divers processus cellulaires, tels que la protection contre le stress oxydatif ou la motilité, grâce à l'étude phénotypique de souches sur- ou sous-exprimant Hsp90_{Ec} [172–174].

Le rôle de Hsp90_{Ec} dans la motilité bactérienne est probablement lié à son implication dans l'assemblage du moteur flagellaire chez *E. coli* [175]. Li *et al.* ont en effet montré que Hsp90_{Ec} semble interagir avec FliI et FliN, deux protéines oligomériques intégrées au sein du complexe formant le moteur flagellaire. Ainsi, la protéine chaperon Hsp90 procaryote pourrait être impliquée dans l'assemblage de volumineux complexes multimériques.

▪ **Fonctions cellulaires et substrats caractérisés chez d'autres espèces bactériennes**

L'analogue procaryote de Hsp90 est également nécessaire au fonctionnement de systèmes et complexes multimériques chez d'autres espèces bactériennes. HtpG est en effet impliqué dans la formation du phycobilisome, un complexe protéique participant à la photosynthèse chez *Synechococcus elongatus* [176]. La protéine chaperon interagit directement avec la protéine de liaison L_R30, qui s'intercale entre deux protéines chargées de capter la lumière.

D'autres études ont également montré que HtpG était impliqué dans l'assemblage de molécules de type PK-NRP produites par des machineries enzymatiques complexes PKS/NRPS. HtpG est ainsi impliquée dans la synthèse de l'albicidine, un antibiotique synthétisé par *Xanthomonas albilineans* [177], et de l'arthrofactine, un lipopeptide de *Pseudomonas* sp. nécessaire à la formation du biofilm [178]. Le mécanisme d'action de la protéine chaperon dans la synthèse de ces deux composés n'a pas été identifié. Les auteurs proposent que HtpG pourrait participer à l'assemblage des machineries enzymatiques, favoriser le repliement de certaines enzymes de ces voies de biosynthèse, ou encore être impliqué dans la régulation de la production de ces molécules [177–179].

Le rôle de HtpG en tant que régulateur a précédemment été mis en évidence pour la synthèse des molécules à noyau tétrapyrrole [180]. Chez *Synechococcus elongatus*, HtpG interagit en effet *in vitro* et *in vivo* avec l'uroporphyrinogène décarboxylase, nommée HemE, et semble réguler l'activité de cette enzyme.

La stratégie visant à étudier les modifications phénotypiques induites par la mutation du gène *htpG* a été employée chez de nombreuses espèces bactériennes pour identifier d'autres fonctions cellulaires. HtpG est ainsi impliquée dans l'adaptation à diverses conditions de stress (choc thermique, acclimatation au froid, stress oxydatif, ...) chez les cyanobactéries [181,182]. Des résultats similaires ont été retrouvés pour d'autres espèces bactériennes, telles que *Bacillus subtilis* [183], *Clostridium acetobutylicum* [184], *Edwardsiella tarda* [185], *Vibrio vulnificus* [186], ou encore certaines γ -protéobactéries psychrophiles [187].

Enfin, quatre études récentes ont montré l'implication de HtpG dans la virulence de *Edwardsiella tarda* [185], de *Francisella tularensis* [188], de *Leptospira interrogans* [189] et de *Salmonella enterica* serovar Typhimurium [190] (cf. § 4.5.).

4.4. Collaboration avec d'autres systèmes chaperons

La présentation du contrôle qualité des protéines en début de chapitre a clairement démontré que les différents systèmes chaperons collaborent étroitement pour former un réseau interactif indispensable au maintien de l'homéostasie protéique chez *E. coli*.

La protéine chaperon Hsp90 n'échappe pas à cette règle. Chez les organismes eucaryotes, Hsp90 interagit avec de nombreux co-chaperons pour assurer le repliement de ses substrats. Chez *E. coli* pourtant, aucun co-chaperon de Hsp90_{Ec} n'a été identifiée à ce jour. Quatre études récentes ont cependant permis de mettre en évidence une collaboration entre les systèmes DnaKJE et Hsp90_{Ec} dans la réactivation de protéines dénaturées suite à un stress [162,191–193].

Genest *et al.* ont ainsi montré que la luciférase ayant subi une inactivation thermique peut être réactivée grâce à l'intervention séquentielle de DnaKJE puis de Hsp90_{Ec} [162]. Ils ont poursuivi leur travaux en démontrant l'interaction directe entre DnaK et Hsp90_{Ec}, dont la liaison est stabilisée par la formation du complexe [DnaK–Hsp90_{Ec}–protéine cliente] [191].

Ces données ont été renforcées par l'analyse du réseau de protéines chaperons de *E. coli* réalisée par Kumar *et al.*, qui ont confirmé l'interaction entre DnaK et Hsp90_{Ec} [192]. Ils ont ainsi proposé un modèle organisationnel du réseau de chaperons et de protéases dans la prise en charge des protéines dénaturées, qui intègre Hsp90_{Ec} (Figure 28). Des résultats similaires ont été obtenus chez *Synechococcus elongatus*, où HtpG interagit directement avec DnaK et DnaJ pour permettre le repliement de protéines dénaturées [193].

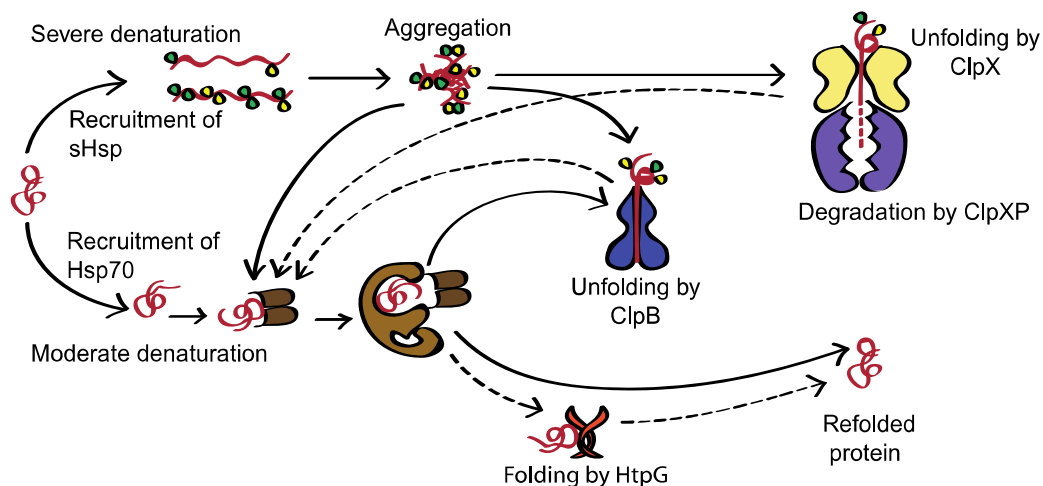


Figure 28. Modèle du contrôle qualité des protéines dénaturées intégrant Hsp90_{Ec}

En fonction de la sévérité des dommages induits par une dénaturation thermique, une protéine peut suivre deux voies majeures. Dans le cas d'une altération sévère, elle forme des agrégats organisés grâce à l'action des sHsps IbpA et IbpB. Ces agrégats sont ensuite soit dégradés par l'action de protéases (ClpXP par exemple), soit résolus par l'action de la désagrégase ClpB. Lorsque les dommages sont moins sévères, la protéine dénaturée est recrutée par le système DnaKJE, qui va favoriser son repliement, avec l'aide de ClpB ou encore de Hsp90_{Ec} [192].

4.5. Pathologies associées et cible thérapeutique

L'implication de Hsp90 dans diverses pathologies est aujourd'hui clairement établie. L'exemple le plus étudié actuellement est sans aucun doute le rôle de Hsp90 dans le cancer. Une multitude de protéines clientes de Hsp90 humaine est en effet impliquée dans de nombreuses voies de signalisation à potentiel oncogénique [194]. Hsp90 est désormais considérée comme une cible thérapeutique particulièrement séduisante, ce qui a justifié le développement de stratégies d'inhibition. Des inhibiteurs spécifiques de l'activité ATPasique de Hsp90 ont été identifiés, tels que la geldanamycine, le radicicol ou la novobiocine [195]. L'inhibition de l'activité du système chaperon favorise la dégradation des protéines clientes par le protéasome, neutralisant de fait les voies de signalisation où elles sont impliquées. De nombreux essais cliniques sont en cours afin d'évaluer le potentiel anti-tumoral et anti-prolifératif de ces inhibiteurs et de leurs dérivés.

Ces dernières années, le rôle de Hsp90 a également été rapporté dans les pathologies neuro-dégénératives, telles que la maladie d'Alzheimer ou la maladie de Parkinson, dont le développement est corrélé à l'accumulation de fibres amyloïdes [196]. La formation de ces composés amyloïdes semble associée à un dérèglement du contrôle qualité des protéines. De récents essais cliniques ont montré que l'inhibition de Hsp90 induit une activation secondaire d'autres systèmes chaperons, mais surtout favorise la dégradation des agrégats protéiques par le protéasome, diminuant la quantité de structures fibrillaires et donc la symptomatologie de ces pathologies neuro-dégénératives [194].

De nombreuses études évaluant le rôle des protéines chaperons dans la virulence de micro-organismes pathogènes ont également vu le jour [197]. L'induction de la réponse au stress lors du processus infectieux a en effet été démontrée chez divers agents pathogènes, bactériens, fongiques, parasitaires ou viraux. Le rôle de Hsp90 a été particulièrement étudié dans le processus infectieux de pathogènes protozoaires majeurs, dont *Leishmania donovani*, *Plasmodium falciparum*, *Toxoplasma gondii* ou *Trypanosoma cruzi* [198]. La protéine chaperon semble indispensable aux transitions entre les différents stades du cycle parasitaire de ces protozoaires. Les inhibiteurs décrits précédemment sont actuellement en phase de test pour confirmer le potentiel de Hsp90 comme cible thérapeutique dans les infections parasitaires.

Concernant les champignons pathogènes, un récent rapport présente l'inhibition de Hsp90 comme une stratégie thérapeutique intéressante pour potentialiser l'efficacité des molécules antifongiques existantes, mais aussi bloquer l'évolution des résistances aux traitements [199].

Quatre études récentes ont par ailleurs mis en évidence l'implication de l'analogue procaryote de Hsp90 dans la virulence de bactéries pathogènes [185,188–190]. HtpG est en effet primordiale à *Edwardsiella tarda*, bactérie pathogène des poissons, pour faire face aux conditions de stress rencontrées chez l'hôte lors du processus infectieux [185]. Dang *et al.* ont montré que la mutation du gène *htpG* entraîne une réduction majeure de la dissémination systémique de la souche et une moindre résistance à l'effet bactéricide des macrophages, dans un modèle d'infection *in vivo* chez le poisson plat *Paralichthys olivaceus*. HtpG est également impliquée dans la virulence de *Francisella tularensis*, bactérie capable de leurrer le système immunitaire inné de l'hôte et de se multiplier dans les macrophages [188]. Une souche mutée pour *htpG* présente une altération importante de la capacité à persister dans la rate et à induire une cytotoxicité macrophagique dans un modèle murin de sepsis. King *et al.* ont récemment décrit que la protéine chaperon HtpG est un facteur de virulence de *Leptospira interrogans* dans un modèle aiguë de leptospirose chez le hamster, bien que son rôle dans la pathogenèse reste inconnu [189]. Enfin, Verbrugghe *et al.* ont dernièrement démontré que la persistance tissulaire d'une souche de *Salmonella* Typhimurium mutée pour *htpG* est atténuée dans un modèle porcin, sans pour autant caractériser le mécanisme physiopathologique impliqué [190].

La littérature sur le rôle de Hsp90 procaryote dans la virulence bactérienne reste pour le moment anecdotique. Mais si son implication dans la pathogénicité de bactéries impliquées en santé humaine ou animale venait à s'enrichir dans un futur proche, à l'image de ce qu'il est maintenant décrit chez les micro-organismes pathogènes parasites ou fongiques, HtpG pourrait devenir une cible anti-infectieuse intéressante comme l'est aujourd'hui son analogue humain dans le cancer.

CHAPITRE IV : MÉTABOLISME ET FONCTIONS DES POLYAMINES CHEZ *Escherichia coli*

1. Notions générales

Les polyamines sont des molécules polycationiques linéaires à plusieurs fonctions amines retrouvées chez tous les organismes vivants [200]. Elles sont impliquées dans une large variété de processus biologiques, tels que la synthèse des acides nucléiques, la régulation de l'expression génique ou encore la traduction des protéines. Leur implication dans ces fonctions physiologiques majeures les rend indispensables à la croissance normale et à la multiplication des cellules procaryotes et eucaryotes [201].

1.1. Structure des polyamines

Les polyamines sont classées selon le nombre de fonctions amines ([Figure 29](#)) [202] :

- les diamines, comprenant : le 1,3-diaminopropane, la putrescine et la cadavérine ;
- les triamines, incluant : la norspermidine, la spermidine et l'homospermidine ;
- les tétra-amines, regroupant : la norspermine, la spermine et la thermospermine.

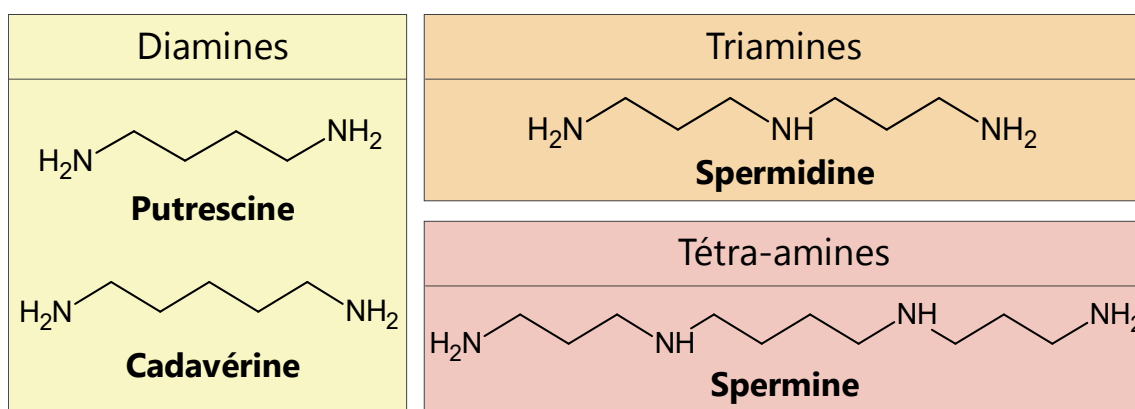


Figure 29. Exemples de structure de polyamines

À pH physiologique, les polyamines sont protonées au niveau de leurs différentes fonctions amines. Elles interagissent ainsi avec de nombreuses molécules chargées négativement, notamment les acides nucléiques, l'ATP, les phospholipides et certaines protéines [203].

1.2. Distribution des polyamines

La putrescine et la spermidine sont les polyamines les plus fréquemment retrouvées dans les différents règnes du vivant [204]. Chez *E. coli*, la concentration intracellulaire en putrescine (20-40 mM) est plus élevée que celle en spermidine (5-10 mM), contrairement à la plupart des bactéries [201]. Au sein de la cellule bactérienne, ces deux polyamines sont majoritairement associées aux acides nucléiques (Tableau III) [205]. La spermidine en particulier est liée à près de 90% aux ARN.

Tableau III. Distribution des polyamines chez *E. coli* [205]

	Putrescine		Spermidine	
	mM	(%)	mM	(%)
Libre	12,5	(38,8)	0,26	(3,8)
ADN	3,0	(9,3)	0,35	(5,1)
ARN	15,4	(47,9)	6,17	(89,7)
ATP	0,46	(1,4)	0,05	(0,7)
Phospholipides	0,84	(2,6)	0,05	(0,7)
Total	32,2		6,88	

2. Métabolisme des polyamines

Tous les organismes vivants ne sont pas en mesure de synthétiser l'ensemble du panel de polyamines. Les bactéries sont généralement capables de synthétiser la putrescine, la spermidine et la cadavérine, et certaines espèces bactériennes peuvent en outre produire de la spermine [200]. *E. coli* en particulier peut synthétiser la putrescine, la spermidine et la cadavérine, mais ne produit pas de spermine [206].

2.1. Biosynthèse

Chez les bactéries, la biosynthèse des polyamines résulte de la décarboxylation d'acides aminés précurseurs : l'ornithine, l'arginine, la méthionine ou la lysine [200,207].

Chez *E. coli*, la **putrescine** est synthétisée selon deux voies métaboliques (Figure 30) [200].

- la décarboxylation de l'arginine en agmatine grâce à l'arginine décarboxylase (**SpeA**), suivie de la conversion de l'agmatine en putrescine assurée par l'agmatinase (**SpeB**) ;
- la décarboxylation de l'ornithine en putrescine grâce à l'ornithine décarboxylase (**SpeC**).

La synthèse de **spermidine** nécessite ensuite deux étapes (Figure 30) [200] :

- la conversion de méthionine en S-adénosyl-méthionine par la méthionine adénosyl-transférase (**MetK**), puis sa décarboxylation en S-adénosyl-méthioninamine grâce à la S-adénosyl-méthionine décarboxylase (**SpeD**) ;
- la condensation de putrescine et de S-adénosyl-méthioninamine en spermidine, assurée par la spermidine synthase (**SpeE**).

Le catabolisme de la spermidine consiste en son acétylation en N^1 - ou N^8 -acétyl-spermidine grâce à la spermidine acétyl-transférase (**SpeG**) (Figure 30) [208].

Par ailleurs, la **cadavérine** est synthétisée à partir de lysine grâce à l'intervention de la lysine décarboxylase (**CadA**) (Figure 30) [200]. Cette diamine est très peu présente comparativement aux deux précédentes polyamines, et n'est produite chez *E. coli* qu'en cas d'excès de lysine dans le milieu extérieur ou lors d'un blocage de la synthèse de putrescine [209].

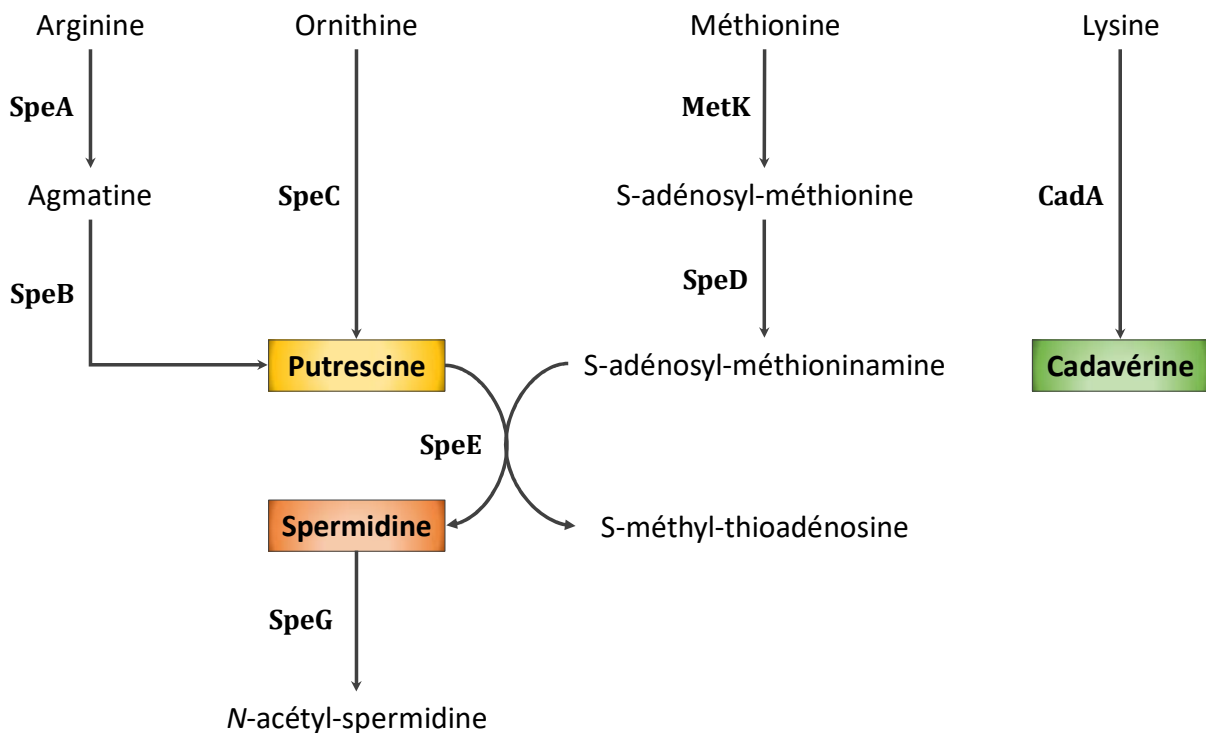


Figure 30. Métabolisme des polyamines chez *E. coli*

Voies de biosynthèse et de catabolisme des polyamines détaillées dans le paragraphe précédent. *SpeA* : arginine décarboxylase ; *SpeB* : agmatinase ; *SpeC* : ornithine décarboxylase ; *MetK* : méthionine adénosyl-transférase ; *SpeD* : S-adénosyl-méthionine décarboxylase ; *SpeE* : spermidine synthase ; *SpeG* : spermidine acétyl-transférase ; *CadA* : lysine décarboxylase. D'après [210].

2.2. Transport

En complément de la synthèse de polyamines *de novo*, les bactéries possèdent également des systèmes de transport afin de capter les polyamines présentes dans le milieu extérieur. Cinq systèmes de transport des polyamines ont été caractérisés chez *E. coli* (Figure 31) [211] :

- deux systèmes d'import : PotABCD captant préférentiellement la spermidine, et PotFGHI captant spécifiquement la putrescine ;
- deux systèmes antiports : PotE permettant l'échange d'ornithine et de putrescine, et CadB chargé de l'échange de lysine et cadavérine ;
- un système d'export : MdtIJ, chargé d'effluer la spermidine en cas d'accumulation.

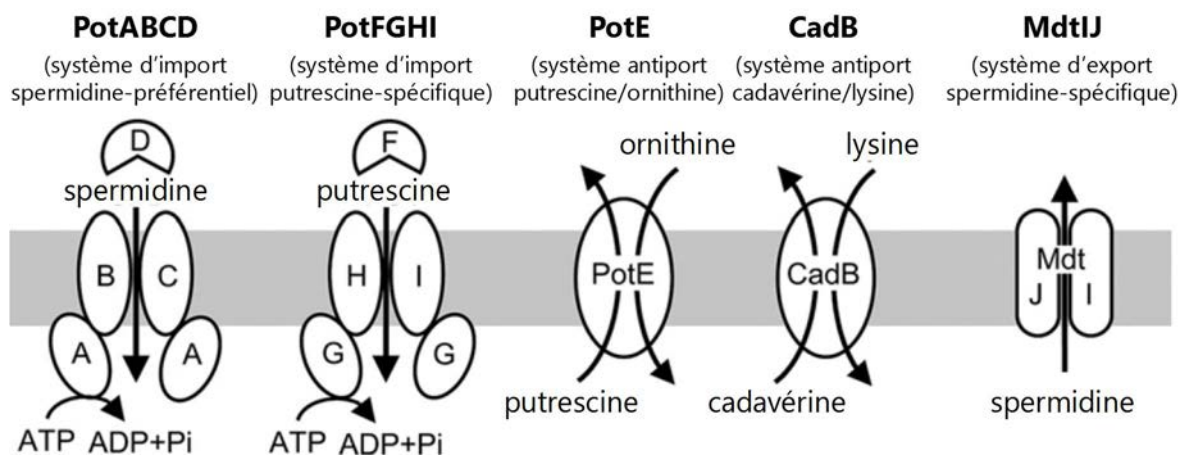


Figure 31. Systèmes de transport des polyamines chez *E. coli* [210]

Les systèmes PotABCD et PotFGHI appartiennent à la famille des transporteurs ABC, et leurs gènes respectifs sont organisés en opérons (*potABCD* et *potFGHI*, respectivement) [211]. L'organisation de ces deux systèmes est similaire, et consiste en un complexe protéique composé d'une sous-unité liant le substrat (PotD et PotF, respectivement), de deux sous-unités formant un canal membranaire (PotBC et PotHI) et d'une sous-unité ATPasique (PotA et PotG). Chaque sous-unité est indispensable au fonctionnement du système, car la mutation d'un seul gène de l'opéron abolit l'import de polyamines [212]. Dans ces systèmes, la liaison d'une polyamine extracellulaire sur la sous-unité liant le substrat entraîne un changement conformationnel du canal membranaire, qui à son tour modifie la structure de la sous-unité ATPasique [201]. Cette dernière peut alors hydrolyser une molécule d'ATP, ce qui permet l'entrée de la polyamine dans le cytoplasme. PotD capte préférentiellement la spermidine, mais peut aussi se lier à la putrescine. En revanche, PotF capte spécifiquement la putrescine, car son site de liaison est trop étroit pour permettre la liaison de la spermidine [213].

PotE et CadB sont des antiports à douze segments transmembranaires présentant leurs extrémités N- et C-terminales dans le cytoplasme. PotE catalyse à la fois l'import, à pH neutre, et l'export, à pH acide, de putrescine en échange d'ornithine. CadB, sur le même principe, permet en fonction du pH de capter ou d'excréter de la cadavérine en échange de lysine [211].

Enfin, le système MdtIJ permet l'excrétion de spermidine grâce à deux sous-unités protéiques à quatre segments transmembranaires, MdtI et MdtJ [214].

2.3. Régulation de l'homéostasie intracellulaire des polyamines

L'homéostasie des polyamines au sein de la cellule bactérienne est très finement régulée, au niveau de leur biosynthèse, leur dégradation, leur captation ou leur efflux [200,201,205].

Plusieurs enzymes de biosynthèse sont régulées par la concentration intracellulaire en polyamines. Ainsi, la putrescine et la spermidine inhibent réversiblement l'activité de SpeA et SpeC [200]. La spermidine exerce également un rétrocontrôle négatif sur l'expression et l'activité de SpeD [201]. SpeC est par ailleurs régulée par d'autres mécanismes impliquant des nucléotides. Son activité enzymatique est inhibée par le ppGpp, et à l'inverse stimulée par le GTP. De plus, son expression est réprimée par le niveau d'AMP cyclique [215].

La concentration intracellulaire en polyamines régule également les systèmes de transport. La spermidine et la putrescine se fixent respectivement sur les sous-unités PotA et PotG des systèmes de transport PotABCD et PotFGHI [201]. Elles inhibent ainsi leur activité ATPasique et diminuent l'import de polyamines extracellulaires. Par ailleurs, certaines sous-unités de ces systèmes de transport, comme PotD, inhibent la transcription de leur propre opéron [216].

Enfin, le métabolisme des polyamines peut être régulé par la concentration en précurseurs, et notamment en ornithine, ainsi que par l'acidité du milieu [200]. L'ornithine décarboxylase possède, outre son isoforme constitutive SpeC, une isoforme inductible **SpeF** [200,207]. Lorsque la concentration en ornithine est trop importante, ou en cas d'acidification du pH intracellulaire, la transcription de *speF* est induite. Ce gène est co-exprimé avec *potE*, codant l'antiport ornithine/putrescine PotE [201,211]. Dans ces conditions, l'ornithine est transformée par SpeF en putrescine et dioxyde de carbone, ce qui consomme des protons et permet une neutralisation du pH. La putrescine synthétisée peut alors être excrétée par l'antiport PotE afin de réguler l'homéostasie intracellulaire [211].

3. Fonctions des polyamines

Par leur structure polycationique, les polyamines interagissent avec les acides nucléiques, et particulièrement les ARN. Il n'est donc pas surprenant qu'elles soient principalement impliquées dans la régulation de la traduction, par formation de complexes [ARN–polyamine]. Elles peuvent par ailleurs être intégrées dans la structure de métabolites secondaires spécifiques.

3.1. Régulation de la traduction

Dans la cellule bactérienne, les polyamines sont liées à plus de 50% aux ARN, voire 90% dans le cas de la spermidine. Les protéines dont la synthèse est régulée par les polyamines ont été identifiées en utilisant une souche déficiente pour la production de polyamines [217]. Cette souche, incapable de synthétiser la putrescine, présente une altération de croissance en l'absence de putrescine exogène. L'ajout de putrescine dans le milieu de culture restaure un phénotype normal par captation de putrescine extracellulaire puis synthèse de spermidine [218].

En comparant les profils d'expression protéique de cette souche en présence ou non de polyamines exogènes, Igarashi et Kashiwagi ont mis en évidence 11 régulateurs protéiques, transcriptionnels ou traductionnels, dont l'expression est augmentée par les polyamines. Ce « modulon » des polyamines suggère donc que la traduction de nombreux ARNm est indirectement régulée par les polyamines [219].

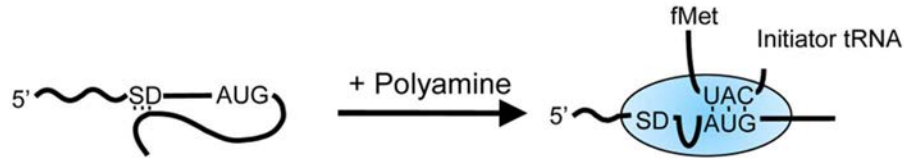
Les auteurs ont confirmé cette hypothèse en analysant le transcriptome de la souche déficiente pour la production de polyamines en phase exponentielle de croissance avec ou sans ajout de putrescine. Ils ont ainsi identifié 309 gènes induits, dont 240 appartiennent au régulon de l'un des 11 facteurs de transcription identifiés par l'analyse protéomique [219]. En régulant la traduction de plusieurs régulateurs transcriptionnels globaux, les polyamines modulent le niveau d'expression d'un grand nombre de protéines, permettant ainsi le maintien de conditions optimales à la croissance bactérienne [219].

Trois mécanismes de stimulation de la synthèse protéique ont été identifiés ([Figure 32](#)) [219] :

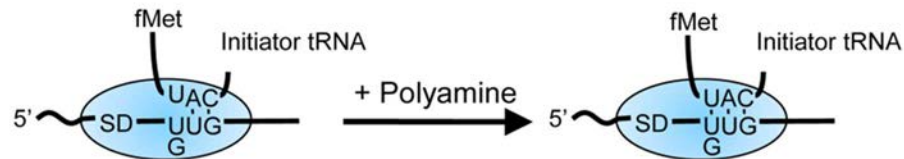
- les polyamines modifient la conformation d'un ARNm qui présente un éloignement trop important entre la séquence Shine-Dalgarno (SD) et le codon d'initiation AUG, facilitant ainsi la formation du complexe d'initiation de la traduction ;
- les polyamines facilitent l'initiation de la traduction lorsqu'un codon d'initiation est inhabituel (UUG ou GUG) ou inefficace ;

- les polyamines favorisent la lecture d'un codon stop par un ARNt suppresseur, ou permettent un décalage du cadre de lecture.

1. Long distance between SD sequence and initiation AUG



2. Initiation on inefficient initiation codon



3. Suppression and +1 frameshifting on nonsense codon

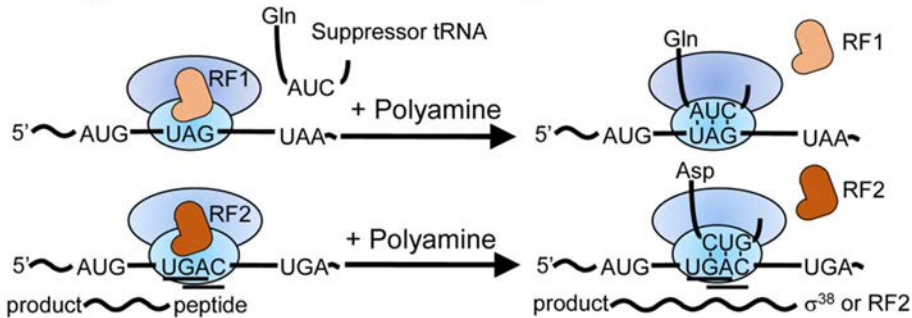


Figure 32. Mécanismes de stimulation de la traduction par les polyamines chez *E. coli*

Mécanismes de stimulation de la synthèse protéique détaillés dans le paragraphe précédent [219].

3.2. Incorporation dans la structure de métabolites secondaires

Trois sidérophores nécessitent pour leur synthèse l'incorporation de polyamines [202]. La vibriobactine, produite par *Vibrio cholerae*, et la vulnibactine, produite par *Vibrio vulnificus*, sont deux sidérophores NRP de la famille des catéchols, qui intègrent tous deux dans leur structure une molécule de norspermidine (Figure 33) [220,221]. La pétrobactine, synthétisée par *Bacillus anthracis*, est un sidérophore non-PK non-NRP de la famille des citrate-catéchols, dont la synthèse requiert deux molécules de spermidine (Figure 33) [222]. Par ailleurs, un isoforme spécifique de la bléomycine, la bléomycine A₅, contient dans sa structure une molécule de spermidine (Figure 33). Ce motif participe spécifiquement à l'entrée de l'isoforme A₅ dans les cellules eucaryotes *via* le transporteur SLC22A16, impliqué dans l'import de spermidine et de carnitine chez les mammifères [223]. Ce récepteur, particulièrement exprimé dans les cellules testiculaires, permet également l'entrée de doxorubicine [224].

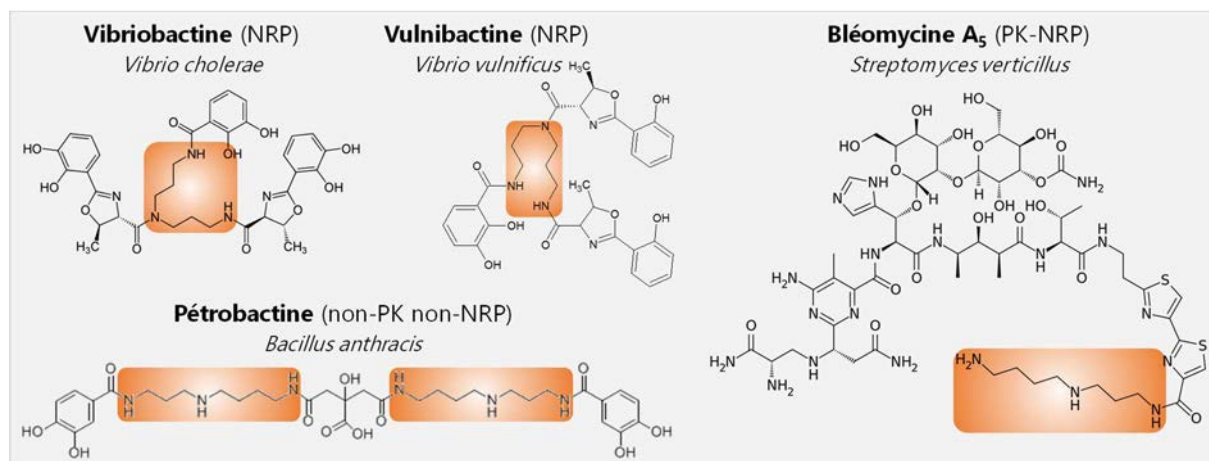


Figure 33. Exemples de métabolites secondaires comportant un motif spermidine

Le type de composé, le micro-organisme producteur et le(s) résidu(s) spermidine sont indiqués.

RÉSULTATS EXPÉRIMENTAUX

RATIONNEL

L'équipe de recherche dans laquelle j'ai réalisé ma thèse a identifié en 2006 une génotoxine, la colibactine, produite par certaines souches de *E. coli* et capable d'induire des cassures double-brin de l'ADN sur les cellules eucaryotes *in vitro* et *in vivo* [2,3]. En dépit des connaissances acquises depuis sa découverte il y a 10 ans, de nombreuses zones d'ombre persistent. La colibactine semble être un composé instable, et sa structure n'a pas encore été entièrement élucidée à ce jour, même si plusieurs études publiées ces trois dernières années ont permis de caractériser différents intermédiaires de synthèse [82–91], et de proposer un modèle de biosynthèse [69]. De plus, les systèmes de régulation de la production de la toxine n'ont pour le moment pas été identifiés. Nous ignorons également si la colibactine est exportée ou non dans la cellule eucaryote. Enfin, le mécanisme d'action moléculaire de la génotoxine sur les cellules eucaryotes, qui pourrait être direct ou indirect, reste inconnu.

La synthèse des PK-NRP codés par des îlots génomiques nécessite classiquement la participation de gènes situés hors de ces îlots, par exemple des gènes du métabolisme permettant la production de monomères [35]. Des résultats préliminaires de l'équipe semblaient de plus indiquer que certains gènes du *core genome* de *E. coli* seraient impliqués dans la production de la colibactine.

Dans le but de mieux appréhender les mécanismes de production de la colibactine, l'objectif de cette thèse était d'identifier les gènes non essentiels de *E. coli* situés hors de l'îlot génomique *pks* impliqués dans la biosynthèse de la génotoxine. Nous avons pour cela construit une banque de mutants par insertion aléatoire de transposons dans la souche de *E. coli* K-12 DH10B *pks*⁺ (portant l'îlot *pks* sur un chromosome bactérien artificiel), dont le protocole de criblage est présenté dans la [Figure 34](#).

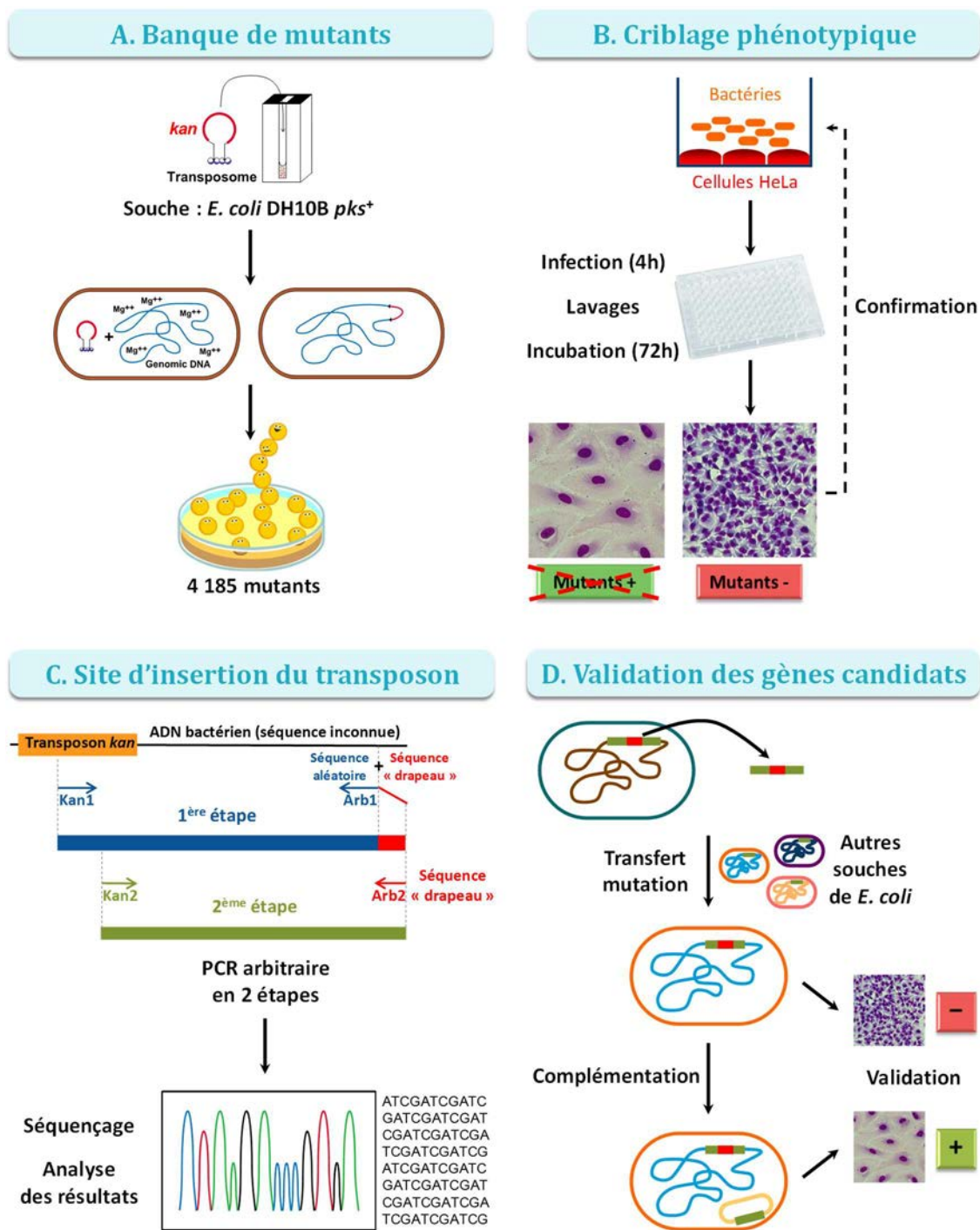


Figure 34. Protocole de construction, criblage et analyse de la banque de mutants

A. Afin d'identifier les gènes de *E. coli* non essentiels à la survie bactérienne mais impliqués dans la synthèse de la colibactine, une banque de mutants a été construite par insertion aléatoire de transposons dans la souche de laboratoire de *E. coli* K-12 DH10B *pks*⁺ (portant l'îlot *pks* sur un chromosome bactérien artificiel). 4 185 mutants ont été obtenus, isolés et conservés à -80°C. **B.** Tous les mutants ont été criblés pour la production de colibactine par un test d'interaction bactéries – cellules eucaryotes. Les mutants ayant perdu la capacité à induire un phénotype de mégaloctose ont été confirmés grâce au même test. **C.** Le site d'insertion du transposon a ensuite été recherché pour les mutants d'intérêt, par PCR arbitraire et séquençage de l'amplicon obtenu. L'analyse des résultats de séquence a permis de distinguer les mutants d'intérêt portant une insertion située hors de l'îlot *pks* de ceux présentant le transposon au sein de l'îlot. Les mutants hors îlot ont ainsi été regroupés selon leur fonction à l'aide d'une analyse globale *in silico*. **D.** L'implication des gènes les plus intéressants dans la synthèse de colibactine a été confirmée par test d'interaction bactéries-cellules, après mutation dans d'autres contextes génétiques et complémentation.

Sur 4 185 mutants criblés, 119 ont perdu la capacité à induire un phénotype de mégalocytose. 70 d'entre eux portaient une insertion du transposon au sein de l'îlot *pks* et ont été écartés. L'étude des 39 mutants d'intérêt restants a permis d'identifier 29 gènes candidats potentiellement impliqués dans la synthèse de colibactine. Le Tableau IV résume l'analyse des mutants d'intérêt en les regroupant par fonction.

Deux principaux groupes ont émergé de cette analyse globale (Tableau IV) :

- des gènes codant des protéines chaperons : *htpG* (Hsp90), *dnaJ* (Hsp70), *groL* (Hsp60) ;
- des gènes impliqués dans le métabolisme des polyamines : *speB*, *speC*, *puuR*.

Ces deux groupes de gènes ont été la base des axes de recherche de ma thèse, dont les résultats expérimentaux sont présentés dans les deux prochains chapitres.

Les autres mutants identifiés, plus difficiles à classer *a priori*, seront étudiés ultérieurement. Cette banque de mutants pour la production de colibactine constitue désormais une véritable mine d'informations à exploiter. La validation des mutants d'intérêt ouvre de nombreuses perspectives pour approfondir nos connaissances concernant la biosynthèse, la structure, l'export et les mécanismes de régulation de la production de cette génotoxine.

Tableau IV. Synthèse de l'analyse des mutants d'intérêt situés hors de l'îlot *pks*

Groupes	Mutant (n° plaque + position)	N° électroporatio n	Site d'insertion du Transposon		Position dans DH10B	Position dans MG1655	BLAST		Conservation du gène dans d'autres souches Colibactine +						Partenaires identifiés (STRING)	PCR <i>dhb/dhbN</i>	Criblage milieu CAS	Remarques
			Abréviation	Gène ou région intergénique interrompue			Homologie (DH10B)	Séquence obtenue	<i>E. coli</i> 536	<i>E. coli</i> AB08972	<i>E. coli</i> CF7073	<i>E. coli</i> IHE3034	<i>E. coli</i> UM146	<i>E. coli</i> UTB9				
Métabolisme	II-E8	1	<i>speB</i>	Agmatinase	3 176 716	3 081 647	74 nt	GCTGCGG	x	x	x	x	x	x	x			
	XVIII-H10	1	<i>speC</i>	Ornithine decarboxylase (constitutive)	3 209 500	3 106 819	673 nt	GCCTGGG	x	x	x	x	x	x	x			
	XVII-E10	1	<i>puuR</i>	DNA-binding transcriptional repressor	1 449 678	1 360 282	834 nt	GGCCACA	Faible homologie (15%)									
Protéines	IV-F11	1	<i>groL</i>	groEL, molecular chaperone HSP60 family, large subunit	4 471 039	4 370 677	746 nt	TGCCACC	x	x	x	x	x	x	x			
	XXXII-F6	3	<i>hspG</i>	HspG, molecular chaperone HSP90 family	434 517	495 186	603 nt	GTCCAGG	x	x	x	x	x	x	x			
	XXXIII-H9	3	<i>dnaJ</i>	DnaJ, molecular chaperone HSP40, co-chaperone with DnaK	14 371	14 371	368 + 286 nt	GATGTGTC	x	x	x	x	x	x	x			PCR <i>dhb/dhbN</i>
Efflux	XVIII-H11	1	<i>mdtE</i>	Multidrug resistance efflux transporter	3 757 524	3 657 979	814 nt	AGCGTAC	x	x	x	x	x	x	x			
	XXXVI-B4	4	<i>gadW</i>	DNA-binding transcriptional activator	3 760 117	3 662 372	122 nt	GGCATGG	x	x	x	x	x	x	x			
	XVII-G12	1	<i>hpaA</i>	Cytoplasmic membrane lipoprotein-28	3 934 825	3 837 241	975 nt	GCTTAGC	x	x	x	x	x	x	x			
	XXXIX-D1	4	<i>spr</i>	Predicted peptidase, outer membrane lipoprotein	2 359 426	2 268 438	212 nt	GACAGGC	x	x	x	x	x	x	x			
ATP synthase	XVII-H7	1	<i>atpE</i>	F0 sector of membrane-bound ATP synthase, subunit c	4 016 643	3 923 966	179 nt	CAGGTCCT	x	x	x	x	x	x	x			
	XXXII-H1	3	<i>atpE</i>	F0 sector of membrane-bound ATP synthase, subunit c	4 016 709	3 919 125	95 nt	TGCCGATA	x	x	x	x	x	x	x			
	V-H9	2	<i>atpD</i>	F1 sector of membrane-bound ATP synthase, subunit β	4 012 527	3 914 943	641 nt	GGGTGTAG	x	x	x	x	x	x	x			Faible pousse
	XIV-H1	2	<i>atpD</i>	F1 sector of membrane-bound ATP synthase, subunit β	4 012 526	3 914 942	633 nt	GGGTGTAG	x	x	x	x	x	x	x			Faible pousse
	XXI-C6	1	<i>glfA</i>	Glucose-inhibited cell-division protein	4 019 548	3 921 964	150 + 61 nt	AACAGAA	x	x	x	x	x	x	x			Mutant partiel
PEP	XXXVIII-F2	4	<i>ppc/argE</i>	PEP carboxylase / Acetylornithine decarboxylase	4 250 876	4 151 179	580 nt	GGCGCGA	x	x	x	x	x	x	x			Mutant partiel
	XXXVIII-B9	4	<i>ppc/argE</i>	PEP carboxylase / Acetylornithine decarboxylase	4 250 910	4 151 213	607 nt	CTTTTATA	x	x	x	x	x	x	x			Faible pousse
	XI-A8	1	<i>ppc</i>	Phosphoenolpyruvate carboxylase	4 248 972	4 149 275	275 nt	ATCTGCC	x	x	x	x	x	x	x			
Métabolisme général	XXXIX-C6	4	<i>purA</i>	Adenylosuccinate synthetase	4 503 781	4 403 419	639 nt	GATAAAG	x	x	x	x	x	x	x			
	XXII-G10	2	<i>purA</i>	Adenylosuccinate synthetase	4 504 323	4 403 961	314 nt	GACAGTA	x	x	x	x	x	x	x			Faible pousse
	XXXII-H10	3	<i>purB</i>	Adenylosuccinate lyase	1 246 221	1 190 786	706 nt	CTTCATCA	x	x	x	x	x	x	x			Faible pousse
	XXXII-G4	3	<i>aspC</i>	Aromatic amino acid/aspartate amino-transferase	1 037 924	893 996	483 nt	GATATGGC	x	x	x	x	x	x	x			Faible pousse
	XL-B10	4	<i>carB</i>	Carbamoyl-phosphate synthase large subunit	31 571	31 571	568 nt	CAGCCCAA	x	x	x	x	x	x	x			Faible pousse
	XIII-D1	5	<i>carA</i>	Carbamoyl phosphate synthetase small subunit	30 662	30 662	554 nt	GTAAGGTA	x	x	x	x	x	x	x			Faible pousse
	XL-F2	4	<i>pyrB</i>	Aspartate carbamoyltransferase, catalytic subunit	4 572 058	4 470 241	539 nt	GGGTACGG	x	x	x	x	x	x	x			Faible pousse
	XXXVIII-G6	4	<i>serC</i>	Phosphoserine/phosphohydroxythreonine aminotransferase	1 011 484	957 556	600 nt	ATCCAGAA	x	x	x	x	x	x	x			Faible pousse
	XVII-E5	1	<i>epfA</i>	Predicted metal dependent hydrolase	4 433 569	4 333 207	145 nt	GCAGCAA	x	x	x	x	x	x	x			Faible pousse
	XXVII-B8	1	<i>hfb</i>	Integration host factor, DNA-binding protein, subunit β	1 017 148	974 557	110 nt	CGGGTAG	x	x	x	x	x	x	x			
	Divers (pas de partenaire)	IV-G6	1	<i>ygaC/mdh</i>	Predicted protein / Predicted glutaredoxin-like protein	2 889 999	2 798 234	734 nt	GATTGACA	x	x	x	x	x	x	x		
IX-H9		1	<i>ybhD</i>	DNA-binding DNA regulator YbhD	853 681	799 753	576 nt	CTGCCAG	x	x	x	x	x	x	x			
XV-F10		2	<i>yjyC/yagQ</i>	Predicted protein / Putative alpha-amylase	2 875 186	2 783 421	83 nt	ATGTAGAA	BLAST dans DH10B et MG1655 seulement									
XXXII-G11		3	<i>ssaA</i>	Stringent starvation protein A	3 473 121	3 375 376	1021 nt	CATCAGGT	x	x	x	x	x	x	x			
XIII-D10		5	<i>yjfn/ppiC</i>	Predicted protein / Peptidyl-prolyl cis-trans isomerase C	4 056 844	3 957 922	338 nt	GGTACATA	BLAST dans DH10B et MG1655 seulement									
XIII-G6		5	<i>yjef/yakE</i>	Conserved protein / Predicted protein	2 098 752	2 007 744	200 nt	TTCCTCGT	x	x	x	x	x	x	x			
IX-H4		1	<i>recD</i>	Exonuclease V	3 042 911	2 949 041	709 nt	GGTGATG	x	x	x	x	x	x	x			PCR <i>dhb/dhbN</i>
XVII-B1		1	<i>ldeA</i>	L,D-carboxypeptidase A	1 281 996	1 241 755	549 nt	GATCGGTA	x	x	x	x	x	x	x			Faible pousse
XXXIII-F7		4	<i>ldeA</i>	L,D-carboxypeptidase A	1 282 264	1 242 023	51 nt	GCCAACAA	x	x	x	x	x	x	x			Faible pousse
XIV-H5		5	<i>rhaA</i>	Sodium-proton antiporter	17 549	17 549	222 nt	GCCGCTAT	x	x	x	x	x	x	x			Faible pousse
XIV-A9		5	<i>pgi</i>	Glucose-phosphate isomerase	4 333 010	4 233 314	577 nt	GTTCCAG	x	x	x	x	x	x	x			Faible pousse

CHAPITRE I : LA PROTÉINE CHAPERON HSP90_{Ec} EST NÉCESSAIRE À LA BIOSYNTHÈSE DE COLIBACTINE ET DE YERSINIABACTINE

Le premier axe de cette thèse s'est intéressé au groupe de gènes codant des protéines chaperons : *htpG* (Hsp90), *dnaJ* (Hsp70) et *groL* (Hsp60). Ces gènes codent des protéines qui appartiennent aux protéines du choc thermique Hsp, dont la synthèse est induite par la bactérie en réponse à différents stress [225]. Cette famille protéique regroupe des chaperons moléculaires, des protéases, des facteurs de transcription et des enzymes du métabolisme général, qui permettent aux bactéries de s'adapter et survivre en conditions de stress. Ces différents acteurs forment un réseau interactif assurant le maintien de l'homéostasie protéique, qui est le résultat d'un équilibre finement régulé entre le repliement conformationnel par les protéines chaperons et la dégradation par les protéases (Figure 35) [226].

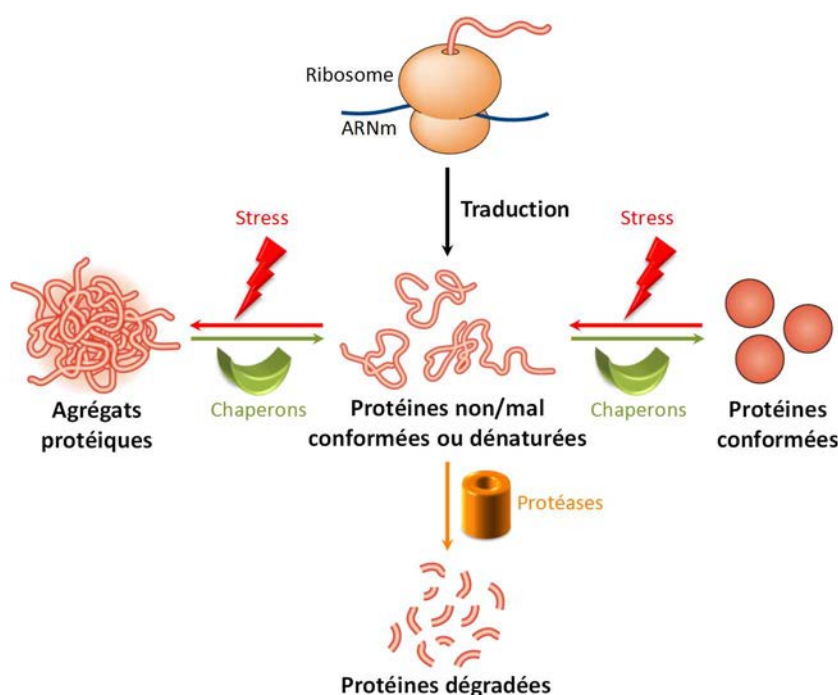


Figure 35. Principes généraux du maintien de l'homéostasie protéique

A l'issue de la traduction d'un ARNm par la machinerie ribosomale, les chaînes d'acides aminés synthétisées ne sont pas conformées. Des protéines chaperons interviennent pour permettre leur repliement conformationnel. Cependant, lors d'un stress thermique par exemple, les protéines peuvent être dénaturées, voire former des agrégats. La mise en place de la réponse au choc thermique induit une production importante de protéines chaperons, qui permettent la désagrégation des protéines et assurent leur repliement. Si les altérations subies sont trop importantes, des protéases dégradent les protéines dénaturées en acides aminés [226].

Nous avons décidé de focaliser notre travail sur l'une des trois protéines chaperons identifiées lors du criblage de la banque de mutants : HtpG, encore appelée Hsp90_{EC}. Cette protéine chaperon semblait particulièrement intéressante dans l'étude de la biosynthèse de la colibactine. L'analogue procaryote de Hsp90 est en effet nécessaire à la synthèse de deux composés de type PK-NRP, de la même famille chimique que la colibactine [177,178]. L'implication de Hsp90_{EC} dans la production de molécules telles que la colibactine soulève également l'hypothèse de son rôle dans l'adaptation et la virulence durant le processus infectieux chez l'hôte, qui a été démontré pour d'autres espèces bactériennes [185,188–190].

Ce projet avait donc pour objectifs d'élucider le rôle de la protéine chaperon Hsp90_{EC} dans la biosynthèse de la colibactine chez *E. coli*, et de déterminer son implication dans la virulence.

Les résultats de ce travail expérimental ont permis de démontrer que la protéine chaperon Hsp90_{EC} est impliquée dans la biosynthèse de deux métabolites secondaires de type PK-NRP produits par *E. coli*, la génotoxine colibactine et le sidérophore yersiniabactine. De plus, la protéase ClpQ intervient de concert avec Hsp90_{EC} dans la production de ces deux composés. Un équilibre entre les quantités de protéine chaperon et de protéase doit être conservé pour permettre la production des deux métabolites, interrompue par une surexpression de ClpQ. Hsp90_{EC} pourrait protéger un substrat requis pour la production de la colibactine de l'activité protéasique de ClpQ. Leur intervention respective constituerait ainsi un mécanisme de régulation post-traductionnelle efficace de la synthèse de colibactine. Enfin, l'analyse des effets de la mutation du gène *htpG* au cours d'une infection systémique chez l'animal, dans des modèles de sepsis et de méningite néonatale chez les rongeurs, démontre le rôle de la protéine de réponse au stress Hsp90_{EC} dans la virulence de souches d'ExPEC.

Cette étude met ainsi en évidence pour la première fois le rôle de la protéine chaperon Hsp90_{EC} dans la production de deux métabolites secondaires impliqués dans la virulence de *E. coli*, la colibactine et la yersiniabactine.

Ces résultats sont rapportés dans l'article « The bacterial stress-responsive Hsp90 chaperone (HtpG) is required for the production of the genotoxin colibactin and the siderophore yersiniabactin in *Escherichia coli* », publié dans *The Journal of Infectious Diseases* (**Article 1**) [227].

Article 1

The Journal of Infectious Diseases

MAJOR ARTICLE



The Bacterial Stress-Responsive Hsp90 Chaperone (HtpG) Is Required for the Production of the Genotoxin Colibactin and the Siderophore Yersiniabactin in *Escherichia coli*

Christophe Garcie,^{1,2} Sophie Tronnet,² Amélie Garénaux,⁴ Alex J. McCarthy,⁵ Alexander O. Brachmann,⁶ Marie Pénary,² Sébastien Houle,⁴ Jean-Philippe Nougayrède,² Jörn Piel,⁶ Peter W. Taylor,⁵ Charles M. Dozois,⁴ Pierre Genevaux,³ Eric Oswald,^{1,2} and Patricia Martin^{1,2}

¹Service de Bactériologie-Hygiène, CHU Toulouse, ²IRSD, Université de Toulouse, INSERM, INRA, ENVT, UPS, and ³Laboratoire de Microbiologie et de Génétique Moléculaires, Centre de Biologie Intégrative, CNRS, Université Toulouse 3 Paul Sabatier, France; ⁴Institut Armand Frappier, INRS, Laval, Canada; ⁵School of Pharmacy, University College London, United Kingdom; and ⁶Institute of Microbiology, Eidgenössische Technische Hochschule, Zurich, Switzerland

The genotoxin colibactin, synthesized by *Escherichia coli*, is a secondary metabolite belonging to the chemical family of hybrid polyketide/nonribosomal peptide compounds. It is produced by a complex biosynthetic assembly line encoded by the *pks* pathogenicity island. The presence of this large cluster of genes in the *E. coli* genome is invariably associated with the high-pathogenicity island, encoding the siderophore yersiniabactin, which belongs to the same chemical family as colibactin. The *E. coli* heat shock protein HtpG (Hsp90_{Ec}) is the bacterial homolog of the eukaryotic molecular chaperone Hsp90, which is involved in the protection of cellular proteins against a variety of environmental stresses. In contrast to eukaryotic Hsp90, the functions and client proteins of Hsp90_{Ec} are poorly known. Here, we demonstrated that production of colibactin and yersiniabactin is abolished in the absence of Hsp90_{Ec}. We further characterized an interplay between the Hsp90_{Ec} molecular chaperone and the ClpQ protease involved in colibactin and yersiniabactin synthesis. Finally, we demonstrated that Hsp90_{Ec} is required for the full in vivo virulence of extraintestinal pathogenic *E. coli*. This is the first report highlighting the role of heat shock protein Hsp90_{Ec} in the production of two secondary metabolites involved in *E. coli* virulence.

Keywords. heat shock proteins; Hsp90; HtpG; stress response; colibactin; yersiniabactin; siderophores; virulence; meningitis; *Escherichia coli*.

Escherichia coli is both a commensal inhabitant of the human gastrointestinal tract and a pathogen associated with a wide range of infections. Certain pathogenic *E. coli* strains (ie, extraintestinal pathogenic *E. coli* [ExPEC]) display an enhanced capacity to cause infection outside the intestinal tract. These strains harbor numerous virulence factors encoded by mobile genetic elements, such as plasmids, transposons, phages, or pathogenicity islands [1].

We previously revealed the presence in the *E. coli* genome of a 54-kb gene cluster, the *pks* genomic island [2]. This highly conserved pathogenicity island is predominately found in *E. coli* strains of the phylogenetic group B2 and in some other pathogenic species of *Enterobacteriaceae*, such as *Citrobacter koseri*, *Klebsiella pneumoniae*, and *Enterobacter aerogenes* [2, 3]. The *pks* island carries genes *clbA* to *clbS* and encodes

modular nonribosomal peptide synthetases, polyketide synthases, and accessory enzymes. This complex biosynthetic machinery synthesizes a microbial secondary metabolite, the genotoxin colibactin, which is a hybrid polyketide/nonribosomal peptide (PK-NRP) compound. Colibactin was demonstrated to generate DNA double-strand breaks in eukaryotic cells both in vitro and in vivo [2–5]. This DNA damage leads to mutations, chromosomal instability, and premature senescence that could ultimately drive tumorigenesis [4, 6, 7]. *E. coli* strains also synthesize other PK-NRP secondary metabolites, including the siderophores enterobactin, salmochelins, and yersiniabactin. Siderophores are low-molecular-weight compounds that facilitate iron uptake by bacteria and are virulence factors [8]. We recently described crosstalk between the biosynthesis of the PK-NRP secondary metabolites colibactin and siderophores [9]. Our recent studies have also demonstrated the importance of the interplay between colibactin and siderophores in the virulence of ExPEC in rodent models of sepsis and neonatal meningitis [5, 9, 10].

Entry of a pathogen into a warm-blooded host is usually accompanied by a temperature upshift. In bacteria, such change triggers a global stress response, named the heat shock response, which is orchestrated by heat shock proteins (HSPs) [11]. HSPs include highly conserved molecular chaperones, proteases,

Received 6 April 2016; accepted 6 July 2016; published online 13 July 2016.
Presented in part: French Society for Microbiology National Meeting, Paris, France, 1 April 2014.

Correspondence: E. Oswald, Institut de Recherche en Santé Digestive, INSERM-INRA-ENVT-Université de Toulouse, CHU Toulouse Purpan, Place du Docteur Baylac, CS 60039, 31024 Toulouse Cedex 3, France (eric.oswald@inserm.fr).

The Journal of Infectious Diseases® 2016;214:916–24

© The Author 2016. Published by Oxford University Press for the Infectious Diseases Society of America. All rights reserved. For permissions, e-mail journals.permissions@oup.com.
DOI: 10.1093/infdis/jiv294

transcription factors, and key metabolic enzymes that facilitate adaptation and survival in response to exogenous stressors [12]. Heat shock genes are also induced at later stages of infection, reflecting their role in the protection of the bacterium against a variety of other stresses and host defense mechanisms, such as oxidative stress, low pH, defensins, or bactericidal serum activity [13]. These processes are essential for the survival of the pathogen within the host.

Molecular chaperones are ubiquitous and highly conserved proteins that maintain intracellular protein homeostasis. They assist with the folding, translocation, quality control, targeting to proteases, assembly, and disassembly of protein complexes [14, 15]. One of these chaperones, high-temperature protein G (HtpG), is the bacterial homolog of eukaryotic heat shock protein 90 (Hsp90). The Hsp90 machinery is involved in diverse cellular processes, including protein folding or repairing and signal transduction [16, 17], and has been demonstrated as a key therapeutic target in cancers or neurodegenerative diseases [18, 19]. However, in contrast to the essential nature of eukaryotic Hsp90, deletion of *htpG* is not lethal to bacterial cells but results in impaired growth at high temperatures [20, 21]. Furthermore, cellular functions and client proteins of the Hsp90 bacterial homolog remain enigmatic, despite its high conservation among bacteria and high abundance in the cell [22, 23]. In this work, we demonstrate that the *E. coli* molecular chaperone HtpG (Hsp90_{Ec}) is required for the synthesis of both colibactin and yersiniabactin and is involved in the extraintestinal virulence of *E. coli*. These results confirm the role of the stress response during infection and indicate that bacterial Hsp90 should be considered a potential target for antimicrobial treatments.

MATERIAL AND METHODS

N-Myristoyl-D-Asparagine (Colibactin Prodrug Motif) Quantification by Liquid Chromatography–Mass Spectrometry (LC-MS)

Strains were grown in Dulbecco's modified Eagle's medium at 37°C for 18 hours (Supplementary Methods). Supernatants of cultures were obtained following centrifugation of bacterial cells at 3200 × *g* for 15 minutes and were filtered on 0.2- μ m membranes. Aliquots of 1 mL of supernatant were prepared. Each strain was cultured in triplicate (deriving from 3 independent clones), and each supernatant was analyzed by LC-MS.

Quantification experiments were conducted with ultra-performance LC–high-resolution/heated electrospray ionization MS. The data were recorded on a Thermo Scientific Q Exactive Hybrid Quadrupole–Orbitrap mass spectrometer coupled to a Dionex Ultimate 3000 ultra-performance LC system. The following solvent gradient (A = H₂O + 0.1% formic acid and B = acetonitrile + 0.1% formic acid, with B at 30% from 0–1 minute, at 30%–95% from 1–6 minutes, and at 95% from 6–7 minutes at a flow rate of 0.5 mL/minute) was used on a Phenomenex Kinetex 5- μ m EVO C18 (50 × 2.1 mm) column at 30°C. The mass spectrometer was operated in positive

ionization mode at a scan range of 200–500 *m/z* and a resolution of 35 000. The spray voltage was set to 3.5 kV, the S-lens to 35, the auxiliary gas heater temperature to 438°C, and the capillary temperature to 270°C. Absolute quantification was achieved by using Schotten-Baumann reaction-derived *N*-myristoyl-L-asparagine (isomer of the *N*-myristoyl-D-asparagine colibactin cleavage product) as a standard. The standard curve was recorded with methanol-diluted concentrations of 10 ng/mL, 100 ng/mL, 500 ng/mL, 1 μ g/mL, and 10 μ g/mL from a 1-mg/mL methanol stock solution. Data were obtained from undiluted cell-free sample supernatants and analyzed for *N*-myristoyl-D-asparagine, and concentrations were calculated using the Thermo Xcalibur 2.2 Quan Browser.

Siderophore Quantification by LC-MS

Culture supernatants were obtained following centrifugation of bacterial cells at 3200 × *g* for 15 minutes and were filtered on 0.2- μ m membranes (Supplementary Methods). Aliquots of 1 mL of supernatant were prepared, and 0.12 ng/mL of 5,6,7,8-tetradeutero-3,4-dihydroxy-2-heptylquinoline was added as an internal control. Each strain was cultured in triplicate, and each culture supernatant was analyzed by LC-MS.

Multiple-reaction-monitoring analyses were performed using a Waters 2795 Alliance HT high-performance LC system coupled to a Micromass Quattro Premier XE mass spectrometer (Micromass MS Technologies). Samples were injected onto a Phenomenex Kinetex 2.6- μ m C8, 100- Å , 150 × 4.6-mm column at a flow rate of 400 μ L/min and with a linear gradient of water-acetonitrile with 1% acetic acid. The transition for yersiniabactin was *m/z* 482 > 295. The specific transitions from pseudomolecular to daughter ions of salmochelins, enterobactin, and aerobactin are described elsewhere [24]. These transitions were used for relative quantification.

Mouse Sepsis Model

The procedure has been described in detail previously [5]. Briefly, 9-week-old female C57BL/6J mice (Janvier) were injected into the footpad with 10⁸ SP15 wild-type and Δ *htpG* mutant strains (Supplementary Methods). Twenty hours after injection, mice were treated with 100 μ L of 1 mg/mL gentamicin injected intraperitoneally, together with ringer solution injected subcutaneously (2 × 500 μ L) for rehydration.

Rat Neonatal Meningitis Model

This procedure has been described in detail previously [25]. Briefly, all 12 two-day-old members of a litter of Wistar rat pups (Harlan, United Kingdom) were fed 20 μ L of mid-logarithmic-phase *E. coli* (6 × 10⁶ colony-forming units) from an Eppendorf micropipette to induce gastrointestinal colonization (Supplementary Methods). Disease progression was determined by daily evaluation of all rat pups for symptoms of systemic infection and scored on a scale of rising severity, from 0 to 3. Pups scoring 3 were culled. Systemic infection was confirmed by quantifying *E. coli* K1 in blood samples cultured on MacConkey

agar, and the expression of the K1 capsule was confirmed by testing the susceptibility of colonies to bacteriophage K1E.

Ethics Statement

Animal experiments were performed in accordance with the European directive for the protection of animals used for scientific purposes. Mouse experiments were approved by the local ethics committee on animal experimentation (the Comité d'éthique pour la protection de l'animal de laboratoire Midi-Pyrénées [C2EA-22]) and were conducted under protocol

MP/03/63/07/12. Rat experiments were approved by the Ethical Committee of the UCL School of Pharmacy and the United Kingdom Home Office (HO) and were conducted under the HO Project License PPL 70/7773.

Statistical Analysis

Statistical analyses were conducted using GraphPad Prism 6.0c. P values were calculated using the unpaired t test or by 1-way analysis of variance test followed by a Bonferroni post hoc test. For in vivo experiments, survival curves were analyzed

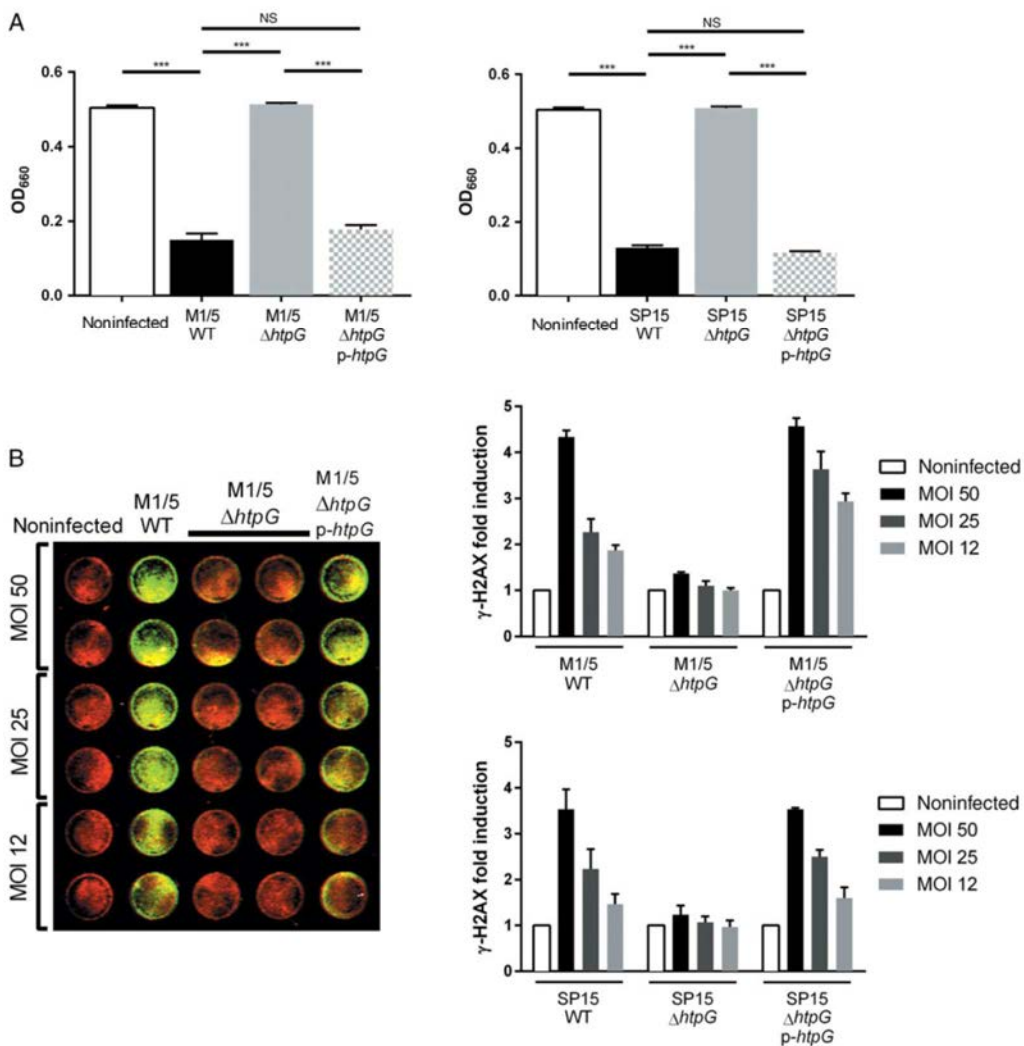


Figure 1. The *Escherichia coli* heat shock protein HtpG (Hsp90_{Ec}) is required for *E. coli* pks⁺ strains genotoxicity. The production of colibactin by *E. coli* strains M1/5 and SP15 and their derivatives was determined by quantification of megalocytosis (A) and of H2AX phosphorylation (B). Wild-type (WT) *E. coli*, Δ htpG mutants, and complemented derivatives were cocultivated with HeLa cells for 4 hours and then washed as previously described [2]. At the end of infection, bacterial growth was similar for all strains. A, After infection, the cells were incubated for 72 hours with appropriate antibiotics before protein staining with methylene blue. The staining was quantified by acid extraction of methylene blue and measurement of absorbance at OD₆₀₀. The multiplicity of infection (MOI) was 200. ***P < .001, by 1-way analysis of variance. B, After infection, the cells were incubated 4 hours in Dulbecco's modified Eagle's medium supplemented with antibiotics before fixation, then permeabilized and labeled for DNA (pseudo-colored red) and phosphorylated H2AX histone (γ -H2AX, pseudo-colored green) using an In-Cell Western method [26]. The MOIs ranged from 12 to 50. All bar graphs show mean values \pm standard errors of the mean. Abbreviation: NS, not significant.

using a log-rank test. A *P* value of <.05 was considered statistically significant.

RESULTS

Hsp90_{Ec} Is Required for the Genotoxicity of *pks*⁺ *E. coli*

The *pks* island was initially identified as the genomic determinant for the synthesis of colibactin through the screening of a transposon mutant library [2]. Interestingly, screening of this library had revealed that several transposon mutants had insertions in *htpG*, suggesting that *htpG* was potentially involved in colibactin biosynthesis (unpublished data). To investigate the impact of Hsp90_{Ec} in colibactin production, deletion of *htpG* was constructed (Supplementary Methods) in commensal and pathogenic *pks*⁺ *E. coli* strains M1/5 and SP15, respectively (Supplementary Table 1). The resulting Δ *htpG* mutant strains were subsequently complemented with plasmid *p-htpG* harboring *htpG* under the control of an IPTG-inducible promoter (Supplementary Table 1). The production of colibactin was measured in the mutant and complemented strains through bacteria–host cell interactions and subsequent quantification of megalocytosis (Figure 1A) and histone H2AX phosphorylation (Figure 1B), which correlate with DNA double-strand breaks resulting from the genotoxic effect of colibactin [2, 4].

The megalocytosis assay [2] and the quantification of H2AX phosphorylation, using an In-Cell Western assay [26], revealed that the inactivation of *htpG* in both the M1/5 and SP15 strains abrogated the genotoxic effect induced by the colibactin (Figure 1). Transformation of the Δ *htpG* mutants with *p-htpG* carrying functional wild-type *htpG* resulted in a complete restoration of the genotoxicity (Figure 1). Together, these data demonstrated that Hsp90_{Ec} is required for colibactin-mediated genotoxicity.

Hsp90_{Ec} Is Directly Involved in Colibactin Biosynthesis

The mature colibactin genotoxin has not been fully characterized yet. However, a biosynthetic byproduct derived from the colibactin assembly line (ie, the *N*-myristoyl-D-asparagine moiety) has been recently characterized [27, 28]. This moiety, generated by the ClbN enzyme, is a prodrug motif cleaved by the ClbP peptidase in the late activation step. We developed a LC-MS assay to quantitatively measure the amount of *N*-myristoyl-D-asparagine as a means to indirectly quantify the production of the genotoxin by wild-type and Δ *htpG* mutant *E. coli* M1/5 (Figure 2). This revealed that the amount of colibactin prodrug motif was not detectable in the Δ *htpG* mutant, compared with the wild-type strain (Figure 2). These results indicated that Hsp90_{Ec} was directly or indirectly required for the synthesis of the biosynthetic intermediate *N*-myristoyl-D-asparagine and, therefore, for colibactin biosynthesis.

Hsp90_{Ec} Does Not Regulate Transcription of the *pks* Island Genes

To investigate whether Hsp90_{Ec} was involved in the regulation of colibactin production via transcriptional regulation of colibactin biosynthesis genes, we quantified the transcription of *clb* genes

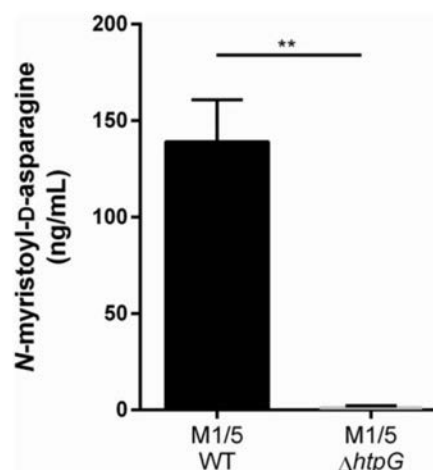


Figure 2. Hsp90_{Ec} is required for the synthesis of the colibactin prodrug motif. The colibactin prodrug motif *N*-myristoyl-D-asparagine produced by wild-type (WT) *E. coli* strain M1/5 and the Δ *htpG* mutant was quantified by liquid chromatography–mass spectrometry (LC-MS). Bacteria were cultivated at 37°C for 18 hours in Dulbecco’s modified Eagle’s medium, and *N*-myristoyl-D-asparagine was quantified in culture supernatants by LC-MS, using *N*-myristoyl-L-asparagine (an isomer of the *N*-myristoyl-D-asparagine colibactin prodrug motif) as a standard. The results were normalized to the bacterial biomass and are presented as the quantity of *N*-myristoyl-D-asparagine. Data represented in the graph were obtained from 3 biological replicates. Similar results were observed for 2 independent experiments. Data are mean values \pm standard errors of the mean. ***P* < .01, by the unpaired *t* test.

encoded on the *pks* island, in relation to *htpG* expression (Supplementary Methods and Figure 3). Wild-type M1/5 and its Δ *htpG* derivative were grown at 37°C, after which the temperature was shifted to 45°C to provoke a heat shock stress, known to induce Hsp90_{Ec} production [29]. Total RNA was isolated and used for a transcriptional analysis of *htpG* and *clbA*, *clbC*, *clbJ*, *clbP*, and *clpQ*, which were selected to monitor the expression of the different transcriptional units identified in the *pks* island [30]. This revealed that a transient heat shock resulted in an increased transcription of *htpG* (Figure 3), as previously shown [29]. Nonetheless, the transcription of the *clb* genes was not altered in response to a heat shock stress and was unchanged whether *htpG* was functional or inactivated (Figure 3). These data suggested that Hsp90_{Ec} was not involved in the transcription of the *pks* island genes.

Hsp90_{Ec} Is Involved in Yersiniabactin Production

E. coli possesses biosynthetic pathways that yield compounds belonging to the same chemical family as colibactin (ie, siderophores), which mediate bacterial uptake of iron and other metals [8]. *E. coli* strains are known to synthesize up to 4 different types of siderophores: aerobactin, enterobactin, salmochelins, and yersiniabactin, which are PK-NRP compounds except for aerobactin. To determine whether Hsp90_{Ec} is also involved in siderophore production, we performed LC-MS to quantify the amount of each siderophore [31] synthesized by wild-type or Δ *htpG*-mutant *E. coli* strains M1/5 and SP15 (Figure 4). This

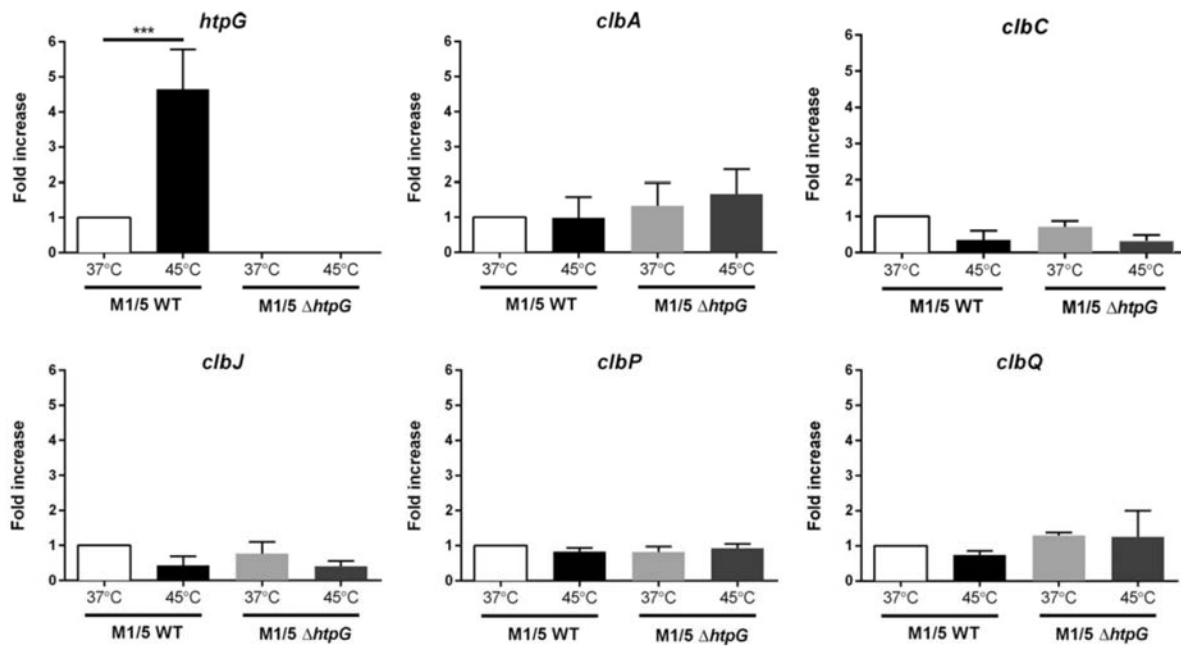


Figure 3. Hsp90_{Ec} does not regulate the transcription of genes located on the *pks* island. Gene expression level of *htpG* and 5 *clb* genes of the *pks* island during heat shock was measured by quantitative reverse transcription–polymerase chain reaction (qRT-PCR). Wild-type (WT) *E. coli* strain M1/5 or its Δ *htpG* mutant were cultivated at 37°C for 3 hours. A fraction of the cultures was transferred at 45°C during 30 minutes to induce heat shock. After total RNA extraction, transcription levels of *htpG*, *clbA*, *clbC*, *clbJ*, *clbP*, and *clbQ* were determined by qRT-PCR. Results were normalized to the level of *hcaT* reference gene expression and are presented as increases (n-fold) in expression level, compared with M1/5 WT cultivated at 37°C. Data are mean values \pm standard errors of the mean. *** $P < .001$, by 1-way analysis of variance.

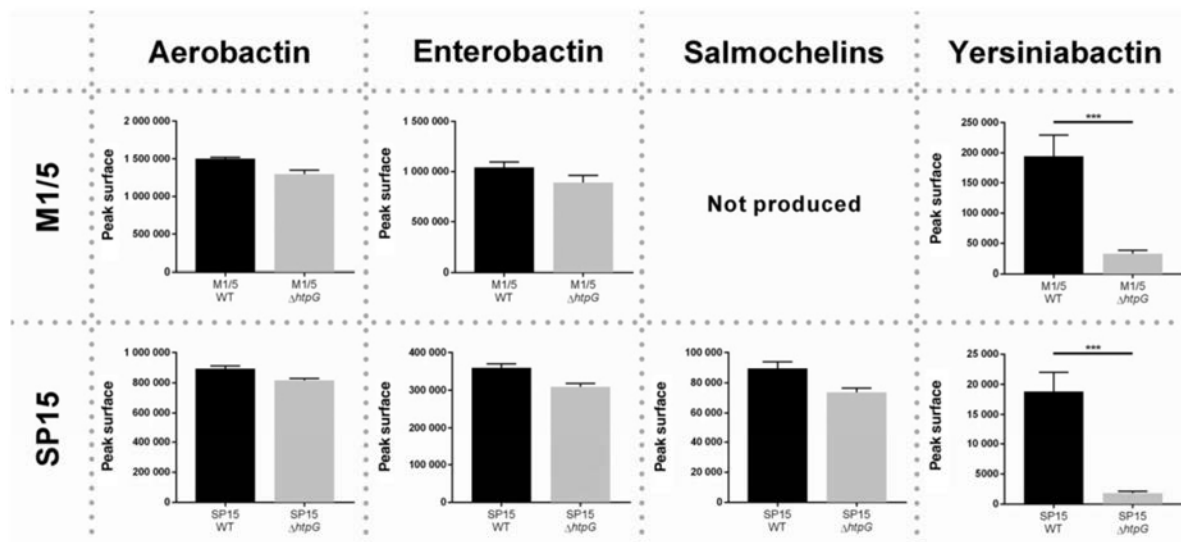


Figure 4. Hsp90_{Ec} is also involved in yersiniabactin production. Siderophore production by *E. coli* strains M1/5, SP15, and derivatives was quantified by liquid chromatography–mass spectrometry (LC-MS). Wild-type (WT) *E. coli* strains M1/5 and SP15 and Δ *htpG* mutants were cultivated at 37°C for 18 hours in Dulbecco’s modified Eagle’s medium. Siderophore production was quantified by LC-MS, as described previously [31]. The results were normalized to the bacterial biomass and are presented as peak surfaces. Data represented in the graph were obtained from 3 biological replicates. Similar results were observed for 2 independent experiments. Data are mean values \pm standard errors of the mean. *** $P < .001$, by the unpaired *t* test.

revealed that the synthesis of yersiniabactin was decreased in both M1/5 $\Delta htpG$ and SP15 $\Delta htpG$ mutant strains, compared with the wild-type strains. The amount of the other siderophores remained unchanged whether *htpG* was disrupted or not (Figure 4). These data indicated that Hsp90_{Ec} also contributed to the biosynthesis or accumulation of yersiniabactin.

Interplay Between Hsp90_{Ec} and the ClpQ Protease Modulates Colibactin and Yersiniabactin Biosynthesis

Specific interplays between Hsp90_{Ec} and the chaperone/protease network were previously reported [32, 33]. To assess whether the absence of Hsp90_{Ec} would abolish competition with proteases for colibactin-involved client binding, we tested the effect of the disruption of the 3 major *E. coli* cytosolic

stress proteases Lon, ClpP, and ClpQ (also called HslV) on the $\Delta htpG$ mutant phenotype. *lon*, *clpP*, or *clpQ* (or *hslV*) were individually inactivated in the M1/5 $\Delta htpG$ mutant strain (Supplementary Methods). The resulting double mutants (Supplementary Table 1) were analyzed for the production of colibactin by quantification of the megalocytosis phenotype (Figure 5A), quantification of colibactin prodrug motif synthesis (Figure 5B), and for the production of yersiniabactin (Figure 5C).

The megalocytosis assay revealed that inactivation of *lon* or *clpP* did not restore the genotoxic effect of colibactin in the $\Delta htpG$ mutant (Figure 5A). On the contrary, infection of HeLa cells with the $\Delta htpG \Delta clpQ$ double mutant resulted in a

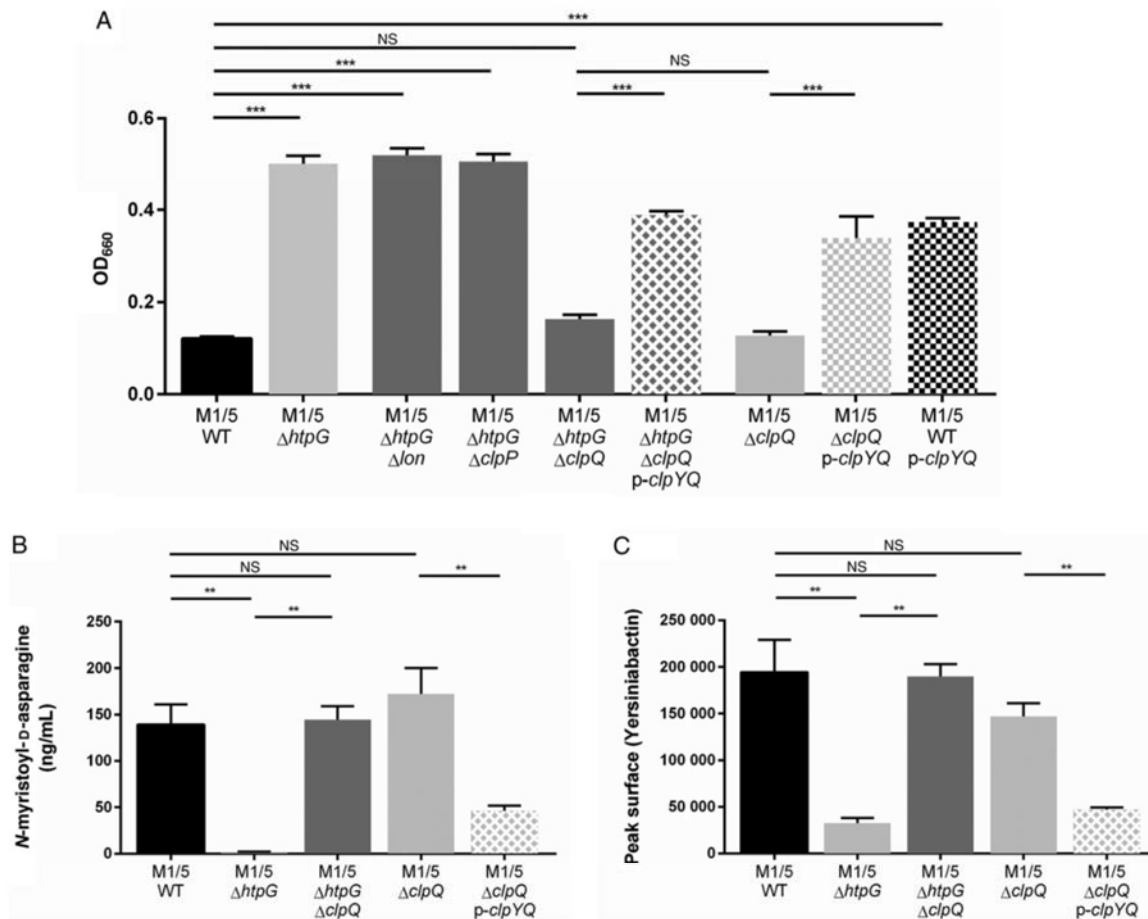


Figure 5. The protease ClpQ is involved in colibactin and yersiniabactin production in combination with Hsp90_{Ec}. **A**, Colibactin-mediated genotoxicity was determined by infection of HeLa cells and quantification of megalocytosis for *E. coli* strain M1/5 and derivatives, as described in Figure 1A. The multiplicity of infection was 200. ****P* < .001, by 1-way analysis of variance (ANOVA). **B**, Colibactin prodrug motif *N*-myristoyl-D-asparagine production by *E. coli* strain M1/5 and derivatives was quantified by liquid chromatography–mass spectrometry (LC-MS), as described in Figure 2. The results were normalized to the bacterial biomass and are presented as the quantity of *N*-myristoyl-D-asparagine. ***P* < .01, by 1-way ANOVA. **C**, Yersiniabactin production by *E. coli* strain M1/5 and derivatives was quantified by LC-MS, as described in Figure 4. The results were normalized to bacterial biomass and are presented as peak surfaces. All bar graphs show mean values ± standard errors of the mean. ***P* < .01, by 1-way ANOVA. Abbreviations: NS, nonsignificant; WT, wild-type.

restored megalocytosis phenotype, indicating the production of colibactin (Figure 5A). Transformation of this double mutant with a plasmid carrying the functional *clpYQ* operon (*p-clpYQ*; Supplementary Table 1) resulted in decreased colibactin activity (Figure 5A). We then analyzed a $\Delta clpQ$ mutant in *E. coli* strain M1/5 (Supplementary Table 1) for colibactin production in the megalocytosis assay (Figure 5A). Inactivation of *clpQ* resulted in a colibactin production level similar to that of wild-type strain (Figure 5A). When the $\Delta clpQ$ mutant was transformed with an inducible and high-copy-number plasmid carrying the functional *clpYQ* operon (*p-clpYQ*; Supplementary Table 1), the resulting complemented derivative lost the ability to induce megalocytosis (Figure 5A). Moreover, introduction of the same plasmid in the wild-type strain also resulted in the loss of the megalocytosis phenotype.

We quantified by LC-MS the amount of *N*-myristoyl-D-asparagine (Figure 5B) and the production of yersiniabactin (Figure 5C) synthesized by $\Delta htpG$ $\Delta clpQ$ and $\Delta clpQ$ mutants and complemented derivatives of *E. coli* M1/5. Patterns similar to those obtained in the megalocytosis experiments (Figure 5A) were observed.

Together, these results indicated that disruption of the ClpQ protease allowed restoration of colibactin and yersiniabactin production in the absence of Hsp90_{Ec}. These data suggested that both the Hsp90_{Ec} molecular chaperone and the ClpQ protease might share specific substrate(s) of the colibactin and the yersiniabactin biosynthesis pathways and that Hsp90_{Ec} could protect the substrate(s) from ClpQ-mediated degradation.

Hsp90_{Ec} Is Required for Full Virulence of ExPEC in a Meningitis Model

Colibactin and yersiniabactin are bona fide virulence factors [5, 8–10]. To address the biological relevance of the chaperone protein Hsp90_{Ec} on *E. coli* virulence in vivo, we analyzed the effects of *htpG* disruption during systemic infection in animals, using a mouse model of sepsis and a rat model of neonatal meningitis (Figure 6).

Wild-type and $\Delta htpG$ -mutant *E. coli* SP15 strains were injected into mice footpads to induce sepsis, as previously described [5]. Infected mice were then treated with antibiotics and hydration 20 hours after infection. Monitoring of animal survival revealed that the mortality of mice inoculated with wild-type *E. coli* SP15 appeared higher than that in the *E. coli* SP15 $\Delta htpG$ group (Figure 6A); however, the finding of log-rank test was not statistically significant.

The second infection model tested the ability of wild-type and $\Delta htpG$ -mutant *E. coli* A192PP strains to induce systemic infection in neonatal rats following oral feeding of live A192PP bacteria [34]. The wild-type *E. coli* A192PP strain produced lethal infection in all colonized pups. On the contrary, although a proportion of pups colonized with the A192PP $\Delta htpG$ mutant did not survive, the overall lethal effect of this mutation was

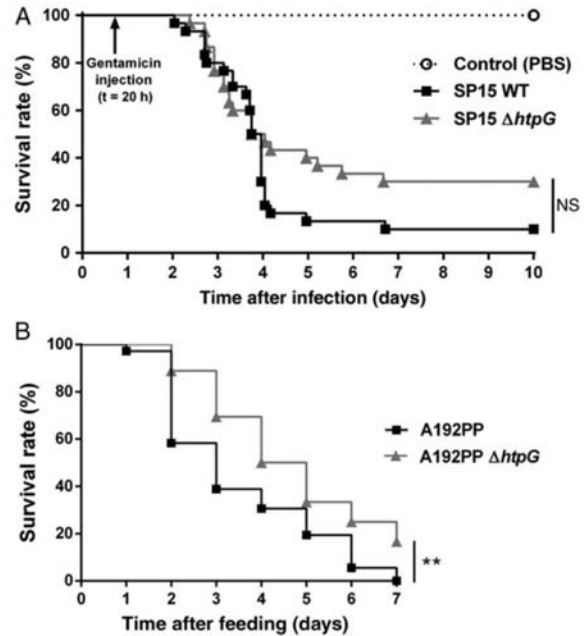


Figure 6. Effect of *htpG* inactivation in in vivo models of systemic infection. *A*, Virulence of wild-type (WT) *Escherichia coli* strain SP15 and the $\Delta htpG$ mutant was evaluated in a murine model of sepsis with antibiotic rescue [5]. Mice received a footpad injection with phosphate-buffered saline (PBS) or 10^8 colony-forming units (CFU) of WT *E. coli* strain SP15 or the $\Delta htpG$ mutant. Mice were then treated with gentamicin (100 μ g per mouse) 20 hours after injection. The percentage of surviving mice was monitored for the different groups (15 mice/group). Data are pooled from 2 independent experiments. $P = .117$, by the log-rank test. *B*, Virulence of *E. coli* strain A192PP and the $\Delta htpG$ mutant was evaluated in a rat model of neonatal meningitis [10]. Two-day-old rats received 2×10^6 – 6×10^6 CFU of WT *E. coli* strain A192PP or the $\Delta htpG$ mutant, orally. The percentage of surviving rats was monitored for the different groups (12 rats/group). Data are pooled from 3 independent experiments. $**P < .005$, by the log-rank test ($P = .0031$). Abbreviation: NS, not significant.

significantly attenuated ($P = .0032$), with an increased survival of about 20% (Figure 6B). This demonstrated that Hsp90_{Ec} chaperone protein was required to maintain full virulence of ExPEC during systemic infection.

DISCUSSION

In this study, we demonstrated that Hsp90_{Ec} is mandatory for the production of two virulence factors produced by *E. coli*, the genotoxin colibactin and the siderophore yersiniabactin. Thus, our work provides new insights into the role of the Hsp90_{Ec} molecular chaperone, together with the recent studies that highlighted Hsp90_{Ec} functions in *E. coli* [21, 32, 35, 36]. Other bacterial Hsp90 molecular chaperones were previously shown to be required for the biosynthesis of PK-NRP compounds, such as albicidin, an antibiotic and phytotoxin produced by *Xanthomonas albilineans* [37], or arthrobaactin, a biosurfactant produced by *Pseudomonas* strains [38]. Based on these studies, we could hypothesize that the molecular

chaperone Hsp90_{Ec} either facilitates the folding or prevents a rapid degradation of a colibactin-synthesis enzyme(s), as proposed for the biosynthesis of arthrobactin in *Pseudomonas* species [38]. Moreover, disruption of *clpQ* allowed restoration of colibactin and yersiniabactin biosynthesis in the absence of Hsp90_{Ec}, which suggests that the chaperone and the protease could share common client protein(s) in the biosynthetic pathways. The fact that overexpression of the *clpYQ* operon induced a decrease of colibactin-mediated genotoxicity is in agreement with the substrate overlap and suggests that Hsp90_{Ec} could protect a substrate required for colibactin production from degradation by the ClpQ protease. The interplay between Hsp90_{Ec} and ClpQ might be an efficient way to posttranslationally control colibactin synthesis. Unfortunately, we have not yet identified a specific client protein involved in colibactin and yersiniabactin biosynthesis directly interacting with Hsp90_{Ec}. So far, only two characterized client proteins for Hsp90_{Ec} have been reported in *E. coli*, the ribosomal protein L2 [39] and the DNA-replication initiator DnaA [40]. The incapacity to identify a specific partner in these pathways could also be explained by a more general function of the chaperone. We hypothesize that Hsp90_{Ec} participates in the assembly or stabilization of the biosynthetic machinery complex, as has been proposed for the biosynthesis of albicidin in *X. albilineans*, where no HtpG-specific client protein has been identified [37].

Colibactin and yersiniabactin have been characterized as virulence factors in vivo [5, 8–10]. Here, we demonstrated that Hsp90_{Ec} is required for full virulence of ExPEC in a model of neonatal meningitis, thus linking a major stress-induced molecular chaperone involved in protein homeostasis to a successful infection. Our study is the first to highlight the role of Hsp90_{Ec} in *E. coli* virulence. In other bacterial species, HtpG has been demonstrated to be involved in the virulence of *Edwardsiella tarda* [41], *Francisella tularensis* [42], *Leptospira interrogans* [43], and *Salmonella enterica* serovar Typhimurium [44]. Here, we showed that Hsp90_{Ec} is required for successful infection in a rat model of neonatal meningitis, which supports the fact that bacterial stress adaptation through the global heat shock response is essential for a successful infectious process.

The involvement of HtpG in the virulence of various bacterial species led us to propose bacterial Hsp90 as a potential antimicrobial therapeutic target, as described for eukaryotic Hsp90 in the treatment of various human cancers [18], but also protozoan [45] and fungal [46] infections. Geldanamycin and radicicol are two natural products that have both been shown to inhibit the ATPase activity and function of the eukaryotic Hsp90 chaperone [18]. Both compounds have also been demonstrated to inhibit bacterial Hsp90 [47, 48]. The use of these two inhibitors in our animal models would provide us with invaluable information. The major challenge of this promising strategy would be to develop Hsp90 inhibitors specific for the prokaryotic

isoform of the molecular chaperone to avoid side effects of antimicrobial treatment by eukaryotic Hsp90 inhibition. Additional knowledge about Hsp90_{Ec} will be required to reach that objective.

Supplementary Data

Supplementary materials are available at <http://jid.oxfordjournals.org>. Consisting of data provided by the author to benefit the reader, the posted materials are not copyedited and are the sole responsibility of the author, so questions or comments should be addressed to the author.

Notes

Acknowledgments. We thank Michèle Boury, for technical assistance; Marie-Pierre Castanié-Cornet, for discussion; and the ANINFIMIP (an EQUIPEX [EQUIPEment d'EXcellence] and state-of-the-art, high-security animal facility that is part of the French funding programme Investissements d'Avenir), for facilitating this work.

Disclaimer. The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Financial support. This work was supported by the French Agence Nationale de la Recherche (grants ANR-13-BSV3-0015-02 and ANR-13-BSV1-0028-01), the French Institut National du Cancer (grant INCA-PLBIO13-123), the United Kingdom Medical Research Council (grant MR/K018396/1 to A. J. M.), and the Canadian Natural Sciences and Engineering Research Council (Discovery grant to C. M. D.).

Potential conflicts of interest. All authors: No reported conflicts. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References

- Kaper JB, Nataro JP, Mobley HL. Pathogenic *Escherichia coli*. *Nat Rev Microbiol* 2004; 2:123–40.
- Nougayrède J-P, Homburg S, Taieb F, et al. *Escherichia coli* induces DNA double-strand breaks in eukaryotic cells. *Science* 2006; 313:848–51.
- Putze J, Hennequin C, Nougayrède J-P, et al. Genetic structure and distribution of the colibactin genomic island among members of the family *Enterobacteriaceae*. *Infect Immun* 2009; 77:4696–703.
- Cuevas-Ramos G, Petit CR, Marcq I, Boury M, Oswald E, Nougayrède J-P. *Escherichia coli* induces DNA damage in vivo and triggers genomic instability in mammalian cells. *Proc Natl Acad Sci U S A* 2010; 107:11537–42.
- Marcq I, Martin P, Payros D, et al. The genotoxin colibactin exacerbates lymphopenia and decreases survival rate in mice infected with septicemic *Escherichia coli*. *J Infect Dis* 2014; 210:285–94.
- Secher T, Samba-Louaka A, Oswald E, Nougayrède J-P. *Escherichia coli* producing colibactin triggers premature and transmissible senescence in mammalian cells. *PLoS One* 2013; 8:e77157.
- Cougnoux A, Dalmasso G, Martinez R, et al. Bacterial genotoxin colibactin promotes colon tumour growth by inducing a senescence-associated secretory phenotype. *Gut* 2014; 63:1932–42.
- Garénaux A, Caza M, Dozois CM. The ins and outs of siderophore mediated iron uptake by extra-intestinal pathogenic *Escherichia coli*. *Vet Microbiol* 2011; 153:89–98.
- Martin P, Marcq I, Magistro G, et al. Interplay between siderophores and colibactin genotoxin biosynthetic pathways in *Escherichia coli*. *PLoS Pathog* 2013; 9:e1003437.
- McCarthy AJ, Martin P, Cloup E, Stabler RA, Oswald E, Taylor PW. The genotoxin Colibactin is a determinant of virulence in *Escherichia coli* K1 experimental neonatal systemic infection. *Infect Immun* 2015; 83:3704–11.
- Richter K, Haslbeck M, Buchner J. The heat shock response: life on the verge of death. *Mol Cell* 2010; 40:253–66.
- Henderson B, Allan E, Coates ARM. Stress wars: the direct role of host and bacterial molecular chaperones in bacterial infection. *Infect Immun* 2006; 74:3693–706.
- Guisbert E, Yura T, Rhodius VA, Gross CA. Convergence of molecular, modeling, and systems approaches for an understanding of the *Escherichia coli* heat shock response. *Microbiol Mol Biol Rev MMBR* 2008; 72:545–54.
- Doyle SM, Genest O, Wickner S. Protein rescue from aggregates by powerful molecular chaperone machines. *Nat Rev Mol Cell Biol* 2013; 14:617–29.

15. Mogk A, Huber D, Bukau B. Integrating protein homeostasis strategies in prokaryotes. *Cold Spring Harb Perspect Biol* **2011**; 3:a004366.
16. Picard D. Heat-shock protein 90, a chaperone for folding and regulation. *Cell Mol Life Sci CMLS* **2002**; 59:1640–8.
17. Wandinger SK, Richter K, Buchner J. The Hsp90 chaperone machinery. *J Biol Chem* **2008**; 283:18473–77.
18. Jhaveri K, Taldone T, Modi S, Chiosis G. Advances in the clinical development of heat shock protein 90 (Hsp90) inhibitors in cancers. *Biochim Biophys Acta* **2012**; 1823:742–55.
19. Luheshi LM, Dobson CM. Bridging the gap: from protein misfolding to protein misfolding diseases. *FEBS Lett* **2009**; 583:2581–6.
20. Bardwell JC, Craig EA. Ancient heat shock gene is dispensable. *J Bacteriol* **1988**; 170:2977–83.
21. Grudniak AM, Pawlak K, Bartosik K, Wolska KI. Physiological consequences of mutations in the *htpG* heat shock gene of *Escherichia coli*. *Mutat Res* **2013**; 745–746:1–5.
22. Chen B, Zhong D, Monteiro A. Comparative genomics and evolution of the HSP90 family of genes across all kingdoms of organisms. *BMC Genomics* **2006**; 7:156.
23. Buchner J. Bacterial Hsp90—desperately seeking clients. *Mol Microbiol* **2010**; 76:540–4.
24. Caza M, Lépine F, Milot S, Dozois CM. Specific roles of the *iroBCDEN* genes in virulence of an avian pathogenic *Escherichia coli* O78 strain and in production of salmochelins. *Infect Immun* **2008**; 76:3539–49.
25. Dalgakiran F, Witcomb LA, McCarthy AJ, Birchenough GMH, Taylor PW. Non-invasive model of neuropathogenic *Escherichia coli* infection in the neonatal rat. *J Vis Exp JoVE* **2014**; 92:e52018.
26. Audebert M, Riu A, Jacques C, et al. Use of the γ H2AX assay for assessing the genotoxicity of polycyclic aromatic hydrocarbons in human cell lines. *Toxicol Lett* **2010**; 199:182–92.
27. Bian X, Fu J, Plaza A, et al. In vivo evidence for a prodrug activation mechanism during colibactin maturation. *Chembiochem Eur J Chem Biol* **2013**; 14:1194–7.
28. Brotherton CA, Balskus EP. A prodrug resistance mechanism is involved in colibactin biosynthesis and cytotoxicity. *J Am Chem Soc* **2013**; 135:3359–62.
29. Mason CA, Dünner J, Indra P, Colangelo T. Heat-induced expression and chemically induced expression of the *Escherichia coli* stress protein HtpG are affected by the growth environment. *Appl Environ Microbiol* **1999**; 65:3433–40.
30. Homburg S, Oswald E, Hacker J, Dobrindt U. Expression analysis of the colibactin gene cluster coding for a novel polyketide in *Escherichia coli*. *FEMS Microbiol Lett* **2007**; 275:255–62.
31. Bachman MA, Oyler JE, Burns SH, et al. *Klebsiella pneumoniae* yersiniabactin promotes respiratory tract infection through evasion of lipocalin 2. *Infect Immun* **2011**; 79:3309–16.
32. Genest O, Hoskins JR, Camberg JL, Doyle SM, Wickner S. Heat shock protein 90 from *Escherichia coli* collaborates with the DnaK chaperone system in client protein remodeling. *Proc Natl Acad Sci U S A* **2011**; 108:8206–11.
33. Kumar M, Sourjik V. Physical map and dynamics of the chaperone network in *Escherichia coli*. *Mol Microbiol* **2012**; 84:736–47.
34. Mushtaq N, Redpath MB, Luzio JP, Taylor PW. Prevention and cure of systemic *Escherichia coli* K1 infection by modification of the bacterial phenotype. *Antimicrob Agents Chemother* **2004**; 48:1503–8.
35. Li H, Sourjik V. Assembly and stability of flagellar motor in *Escherichia coli*. *Mol Microbiol* **2011**; 80:886–99.
36. Yosef I, Goren MG, Kiro R, Edgar R, Qimron U. High-temperature protein G is essential for activity of the *Escherichia coli* clustered regularly interspaced short palindromic repeats (CRISPR)/Cas system. *Proc Natl Acad Sci U S A* **2011**; 108:20136–41.
37. Vivien E, Megessier S, Pieretti I, et al. *Xanthomonas albilineans* HtpG is required for biosynthesis of the antibiotic and phytoalexin albicidin. *FEMS Microbiol Lett* **2005**; 251:81–9.
38. Washio K, Lim SP, Roongsawang N, Morikawa M. Identification and characterization of the genes responsible for the production of the cyclic lipopeptide arthrofactin by *Pseudomonas* sp. MIS38. *Biosci Biotechnol Biochem* **2010**; 74:992–9.
39. Motojima-Miyazaki Y, Yoshida M, Motojima F. Ribosomal protein L2 associates with *E. coli* HtpG and activates its ATPase activity. *Biochem Biophys Res Commun* **2010**; 400:241–5.
40. Grudniak AM, Markowska K, Wolska KI. Interactions of *Escherichia coli* molecular chaperone HtpG with DnaA replication initiator DNA. *Cell Stress Chaperones* **2015**; 20:951–7.
41. Dang W, Hu Y, Sun L. HtpG is involved in the pathogenesis of *Edwardsiella tarda*. *Vet Microbiol* **2011**; 152:394–400.
42. Weiss DS, Brotcke A, Henry T, Margolis JJ, Chan K, Monack DM. In vivo negative selection screen identifies genes required for *Francisella* virulence. *Proc Natl Acad Sci U S A* **2007**; 104:6037–42.
43. King AM, Pretre G, Bartpho T, et al. High-temperature protein G is an essential virulence factor of *Leptospira interrogans*. *Infect Immun* **2014**; 82:1123–31.
44. Verbrugghe E, Van Parys A, Leyman B, Boyen F, Haesebrouck F, Pasmans F. HtpG contributes to *Salmonella* Typhimurium intestinal persistence in pigs. *Vet Res* **2015**; 46:118.
45. Angel SO, Matrajt M, Echeverria PC. A review of recent patents on the protozoan parasite HSP90 as a drug target. *Recent Pat Biotechnol* **2013**; 7:2–8.
46. Veri A, Cowen LE. Progress and prospects for targeting Hsp90 to treat fungal infections. *Parasitology* **2014**; 141:1127–37.
47. Panaretou B, Prodromou C, Roe SM, et al. ATP binding and hydrolysis are essential to the function of the Hsp90 molecular chaperone in vivo. *EMBO J* **1998**; 17:4829–36.
48. Schulte TW, Akinaga S, Murakata T, et al. Interaction of radicicol with members of the heat shock protein 90 family of molecular chaperones. *Mol Endocrinol Baltim Md* **1999**; 13:1435–48.

SUPPLEMENTARY METHODS

Bacterial strains, mutagenesis and growth conditions

Bacterial strains used in this study are listed in [Supplementary Table 1](#). Strain M1/5 is a human commensal *E. coli* strain belonging to the B2 phylogenetic group isolated from the feces of a healthy adult [49]. Strains SP15 and A192 are ExPEC strains of serotype O18:K1, isolated from patients with, respectively, neonatal meningitis [50] and septicemia [51]. A192PP was generated by serial passage through neonatal rat pups as described previously [34], and was demonstrated to be virulent in a rat model of neonatal meningitis. Strains M1/5, SP15 and A192PP harbor the *pks* island and were previously shown to produce colibactin [9,10]. Genes for aerobactin, enterobactin and yersiniabactin synthesis are present in strains M1/5, SP15 and A192PP. Strains SP15 and A192PP also possess genes for salmochelins production.

Gene inactivation of *htpG*, *lon*, *clpP* and *clpQ* was engineered by using the lambda Red recombinase method [52] using pairs of primers presented in Supplementary Table 2. For the Δ *htpG* mutant, gene inactivation was followed by excision of the kanamycin resistance cassette as previously described [52]. For complementation, the *htpG* gene was cloned into pUFR043 (p-*htpG*, Supplementary Table 1, [37]), and the *clpYQ* operon was cloned into pSE380 (p-*clpYQ*, Supplementary Table 1, [53]).

For heat shock experiments, two identical cultures were inoculated from overnight cultures of *E. coli* in lysogeny broth (LB) medium (1:100 dilution). Cultures were grown 3 h at 37°C with shaking. One culture was then transferred to 45°C for 30 min, whereas the second culture was kept at 37°C [29]. The two cultures were immediately processed for RNA isolation.

For siderophore quantification, all *E. coli* strains were grown in LB medium supplemented with appropriate antibiotics for 6 h at 37°C with shaking. These cultures were then diluted 1:100 in Dulbecco's modified Eagle's medium (DMEM Glutamax, Invitrogen) or M9 glucose 0.2% medium supplemented with appropriate antibiotics and inducers, and grown for 18 h at 37°C with shaking. After centrifugation, bacterial pellets were resuspended in fresh DMEM or M9 glucose medium and grown for 2 h at 37°C with shaking. After centrifugation, culture supernatants were then immediately processed for siderophore quantification by LC-MS.

Before injection of mice, all *E. coli* strains were grown overnight in LB, at 37°C with shaking. These cultures were diluted 1:100 in LB and grown for 3 h at 37°C with shaking. After centrifugation, bacterial pellets were resuspended in sterile PBS to the appropriate concentration. All the strains were shown to display similar growth kinetics *in vitro* in LB broth.

Eukaryotic cell culture and infection

HeLa cells (ATCC) were maintained by serial passage in DMEM supplemented with 10% Fetal Calf Serum (FCS, Eurobio) and 1% Non Essential Amino Acid (Invitrogen) at 37°C in a 5% CO₂ atmosphere. HeLa cells were infected as described previously [2].

RNA isolation and reverse transcription-PCR

Total RNAs from *E. coli* grown in LB were isolated at OD_{600nm} = 0.6, using the RNeasy Mini kit (Qiagen, Hilden, Germany). The iScript cDNA Synthesis Kit (Bio-Rad) was used for RT-PCR experiments according to the manufacturer's instructions. After reverse transcription, cDNA samples were used for PCR amplification with primers indicated in Supplementary Table 2. The *hcaT* gene was used as a reference gene [54]. Negative controls were also performed to check for genomic DNA contamination.

SUPPLEMENTARY REFERENCES

49. Clermont O, Christenson JK, Denamur E, Gordon DM. The Clermont *Escherichia coli* phylo-typing method revisited: improvement of specificity and detection of new phylo-groups. *Environ Microbiol Rep* **2013**; 5:58–65.
50. Johnson JR, Oswald E, O'Bryan TT, Kuskowski MA, Spanjaard L. Phylogenetic distribution of virulence-associated genes among *Escherichia coli* isolates associated with neonatal bacterial meningitis in the Netherlands. *J Infect Dis* **2002**; 185:774–784.
51. Achtman M, Mercer A, Kusecek B, *et al.* Six widespread bacterial clones among *Escherichia coli* K1 isolates. *Infect Immun* **1983**; 39:315–335.
52. Datsenko KA, Wanner BL. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* **2000**; 97:6640–6645.
53. Genevaux P, Keppel F, Schwager F, Langendijk-Genevaux PS, Hartl FU, Georgopoulos C. *In vivo* analysis of the overlapping functions of DnaK and trigger factor. *EMBO Rep* **2004**; 5:195–200.
54. Zhou K, Zhou L, Lim Q 'En, Zou R, Stephanopoulos G, Too H-P. Novel reference genes for quantifying transcriptional responses of *Escherichia coli* to protein overexpression by quantitative PCR. *BMC Mol Biol* **2011**; 12:18.

Supplementary Table 1. Strains and plasmids used in this study

Strain or plasmid	Genotype or phenotype	Source
<i>E. coli</i> strain		
M1/5	Commensal <i>E. coli</i> strain isolated from feces of a healthy adult; B2 phylogenetic group; colibactin genotoxin producer; aerobactin, enterobactin and yersiniabactin siderophores producer	Gift from U. Dobrindt
M1/5 Δ <i>htpG</i>	<i>htpG</i> mutant of strain M1/5	This study
M1/5 Δ <i>htpG</i> p- <i>htpG</i>	<i>htpG</i> mutant of strain M1/5 carrying p- <i>htpG</i> , Gen ^r	This study
M1/5 Δ <i>htpG</i> Δ <i>lon</i>	<i>htpG lon</i> mutant of strain M1/5, Kan ^r	This study
M1/5 Δ <i>htpG</i> Δ <i>clpP</i>	<i>htpG clpP</i> mutant of strain M1/5, Kan ^r	This study
M1/5 Δ <i>htpG</i> Δ <i>clpQ</i>	<i>htpG clpQ</i> mutant of strain M1/5, Kan ^r	This study
M1/5 Δ <i>htpG</i> Δ <i>clpQ</i> p- <i>clpYQ</i>	<i>htpG clpQ</i> mutant of strain M1/5 carrying p- <i>clpYQ</i> , Kan ^r Amp ^r	This study
M1/5 Δ <i>clpQ</i>	<i>clpQ</i> mutant of strain M1/5, Kan ^r	This study
M1/5 Δ <i>clpQ</i> p- <i>clpYQ</i>	<i>clpQ</i> mutant of strain M1/5 carrying p- <i>clpYQ</i> , Kan ^r Amp ^r	This study
M1/5 p- <i>clpYQ</i>	Strain M1/5 carrying p- <i>clpYQ</i> , Amp ^r	This study
SP15	ExPEC strain isolated from spinal fluid of a neonate with meningitis; O18:K1 serotype; colibactin genotoxin producer; aerobactin, enterobactin, salmochelins and yersiniabactin siderophores producer; virulent in a mouse model of sepsis	[5,9,34]
SP15 Δ <i>htpG</i>	<i>htpG</i> mutant of strain SP15	This study
SP15 Δ <i>htpG</i> p- <i>htpG</i>	<i>htpG</i> mutant of strain SP15 carrying p- <i>htpG</i> , Gen ^r	This study
A192PP	ExPEC strain isolated from blood culture of a septicemic patient; O18:K1 serotype; colibactin genotoxin producer; genes for aerobactin, enterobactin, salmochelins and yersiniabactin production; virulent in a rat model of neonatal	[34]
A192PP Δ <i>htpG</i>	<i>htpG</i> mutant of strain A192PP, Cm ^r	This study
Plasmid		
p- <i>htpG</i>	Low copy number pUFR043 cosmid carrying <i>htpG</i> gene, Gen ^r	[37]
p- <i>clpYQ</i>	High copy number pSE380 plasmid carrying <i>clpYQ</i> operon, Amp ^r	[53]

Supplementary Table 2. Oligonucleotides used in this study

Primer	Sequence
<i>htpG</i> -F	TGAGGTAGACTACATGAAAGGACAAGAACTCGTGGTTTTTCAGTCAGAAGTGTAGGCTGGAGCTGCTTC
<i>htpG</i> -R	GACTGGAACGTTTTACCGTCGAACTCAGTCAGATAGTTTCATCATCCACTCATGGGAATTAGCCATGGTCC
<i>lon</i> -F	CGTCGGGACCGTGGCCTCTA
<i>lon</i> -R	ACGCCAGCCGAGCTTTAG
<i>clpP</i> -F	AAAGCCTCTTTTCGGTGTAGC
<i>clpP</i> -R	GCCGCCCTGGATAAGTATAG
<i>clpQ</i> -F	GCCAGCGTTACGCAGAGCA
<i>clpQ</i> -R	CGCGGAACAGGCAGAGACGG
RT- <i>hcaT</i> -F	GCTGGCACTGCTGACA
RT- <i>hcaT</i> -R	CGCCGAGCCAATGACA
RT- <i>htpG</i> -F	GAACAGTTCATGCCGAA
RT- <i>htpG</i> -R	GCCAAACTGTTGCCAGA
RT- <i>clbA</i> -F	TGGACGCTCAAAGAA
RT- <i>clbA</i> -R	TGGGACTGCATAGGAAA
RT- <i>clbC</i> -F	CCTACCATGCGCTGGA
RT- <i>clbC</i> -R	CACCGGACCGTTGAGA
RT- <i>clbJ</i> -F	GCCTCTGCTGTCGACA
RT- <i>clbJ</i> -R	TGGCATGCCTATCACCA
RT- <i>clbP</i> -F	ATAGACGCCTGGTTGCA
RT- <i>clbP</i> -R	GCTGATCCGGTCGCAA
RT- <i>clbQ</i> -F	CCTTCGAACTGGTGCA
RT- <i>clbQ</i> -R	GGGGGTGAAAATTGTCA

CHAPITRE II : LA SPERMIDINE MODULE L'ACTIVITÉ GÉNOTOXIQUE DE LA COLIBACTINE

Le second axe de cette thèse s'est intéressé au groupe de gènes impliqués dans le métabolisme des polyamines, identifiés par le criblage de la banque de mutants : *speB*, *speC* et *puuR*. Les gènes *speB* et *speC* codent deux enzymes permettant la synthèse de putrescine par des voies de biosynthèse alternatives, l'agmatinase SpeB et l'ornithine décarboxylase SpeC, respectivement. Le gène *puuR* code une protéine régulant négativement la dégradation de putrescine par la voie PuuABCD. Ce gène a été exclu de notre étude, car il n'est pas présent dans le génome des souches de *E. coli* du phylogroupe B2, et donc des souches *pks*⁺.

Les polyamines sont des molécules cationiques comportant plusieurs fonctions amines, qui se lient aux acides nucléiques au sein d'une cellule, formant majoritairement des complexes [ARN–polyamine] [200,203]. Grâce à ces interactions, elles sont largement impliquées dans la synthèse protéique en régulant la traduction à différents niveaux [219]. Les polyamines peuvent également être intégrées dans la structure de métabolites secondaires, tels que des sidérophores appartenant à la famille de composés de type PK-NRP [220,221].

Ce projet, pour le moment inachevé, a donc pour objectif d'élucider le rôle du métabolisme des polyamines dans la biosynthèse de la colibactine chez *E. coli*.

Une approche de microbiologie moléculaire de l'étude du métabolisme des polyamines a démontré que l'interruption du gène *speE* codant la spermidine synthase altère l'activité génotoxique et la production de colibactine des souches *pks*⁺. Ces phénotypes sont restaurés par complémentation plasmidique, mais aussi par supplémentation en spermidine exogène, indiquant que cette triamine est la polyamine impliquée dans la production de colibactine. Des résultats préliminaires suggèrent que la spermidine participerait à la régulation de l'expression de certains gènes de l'îlot génomique *pks*, et de fait modulerait la production de colibactine. Des études complémentaires sont en cours pour élucider les mécanismes impliqués.

Ces résultats sont rapportés dans une ébauche de manuscrit, intitulé « Connection between primary and secondary metabolism in *Escherichia coli*: the polyamine spermidine modulates the production of the genotoxin colibactin » (**Article 2**). La discussion, qui sera sujette à modifications en fonction des résultats complémentaires, est rédigée en français.

Article 2

Connection between primary and secondary metabolism in *Escherichia coli*: the polyamine spermidine modulates the production of the genotoxin colibactin

Christophe Garcie,^{1,2} Nadège Greif,² Sophie Tronnet,² Alexander O. Brachmann,³ Jörn Piel,³
Jean-Philippe Nougayrède,² Patricia Martin² and Eric Oswald^{1,2,*}

¹CHU Toulouse, Service de Bactériologie-Hygiène, Toulouse, France; ²IRSD, Université de Toulouse, INSERM, INRA, ENVT, UPS, Toulouse, France; ³Institute of Microbiology, Eidgenössische Technische Hochschule (ETH), Zurich, Switzerland.

* Corresponding author: Pr. Eric Oswald; eric.oswald@inserm.fr; +33 (0) 5 62 74 45 43.

INTRODUCTION

The genotoxin colibactin, synthesized by *Escherichia coli*, is a microbial secondary metabolite that was demonstrated to generate DNA double-strand breaks in eukaryotic cells both *in vitro* and *in vivo* (1–3). This DNA damage leads to gene mutations, chromosomal instability and premature senescence that could ultimately drive tumorigenesis (2,4,5). It was recently shown in different mouse models that colibactin-producing *E. coli* strains promote intestinal tumor progression (5–7) and modulate gut homeostasis (8–10). Colibactin has also been recently demonstrated as a virulence factor of extraintestinal pathogenic *E. coli* (ExPEC) in rodent models of sepsis and neonatal meningitis (3,11,12).

The genetic determinant of colibactin is a 54-kb gene cluster, the *pks* genomic island (1). This highly conserved pathogenicity island is predominately found in *E. coli* strains of the phylogenetic group B2 and in some other pathogenic species of *Enterobacteriaceae* (1,13). The

pks island carries genes *clbA* to *clbS*, and encodes modular nonribosomal peptide synthetases (NRPSs), polyketide synthases (PKSs) and accessory enzymes (1). This complex biosynthetic assembly line is responsible for the synthesis of the yet uncharacterized colibactin that belongs to the chemical family of hybrid polyketide/non-ribosomal peptide (PK-NRP) compound. In the past three years, new insights in the colibactin biosynthesis have been published, but if many intermediates have been identified (14–23), the complete structure of the genotoxic colibactin remains elusive.

Although many effects of colibactin on host cells have been described in the past decade, and numerous *pks* assembly line by-products have been characterized recently, details on the regulation of colibactin expression are largely unknown. We recently made a major step in highlighting that the synthesis of this secondary metabolite is closely related to the primary metabolism. We showed that the bacterial molecular chaperone Hsp90_{Ec}, together with the ClpQ protease, were involved in the production of colibactin (24). Hsp90_{Ec} could protect a substrate required for colibactin production from degradation by the ClpQ protease; this interplay may be an efficient way to posttranslationally control colibactin synthesis. We thus wanted to assess whether the primary metabolism was widely involved in the regulation of this secondary metabolite synthesis.

Polyamines are small polycationic molecules associated, in eukaryotic and prokaryotic cells, with a broad range of biological functions: translation, gene regulation, cell proliferation and viability (25–28). In bacteria, the diamine putrescine and the triamine spermidine are the predominant polyamines (25). Synthesis of bacterial spermidine and putrescine depends on the decarboxylation of precursor amino acids which are then converted into functional polyamines (**Figure 1A**) (25,29). In addition to the *de novo* synthesis pathways, specific transport systems allow the uptake of polyamines from the external environment (30). Polyamine intracellular content is tightly regulated by concerted biosynthesis and uptake mechanisms, as well as by degradation and efflux processes (25).

These small aliphatic cations bind to nucleic acids in the cell, existing mostly as [polyamine–RNA] complexes in bacteria (31). Through these interactions, polyamines are involved in the biosynthesis of proteins and thus affect translation at various steps (32). Furthermore, polyamines can be incorporated into the chemical structure of specialized secondary metabolites, such as siderophores, small iron-scavenging PK-NRP compounds produced by bacteria to take up iron (33,34).

In this study, we demonstrated that the polyamine spermidine is involved in colibactin synthesis, and could modulate its production through the regulation of the *pks* island gene expression. These results provide a new insight into the regulation process of colibactin production, and highlight an additional interplay between the primary metabolism and the synthesis of the secondary metabolite colibactin.

MATERIAL AND METHODS

Bacterial strains, mutagenesis and growth conditions

Bacterial strains used in this study are listed in **Table 1**. For genetic manipulations, *E. coli* strains were routinely grown at 37°C under shaking in 5mL of lysogeny broth (LB, Invitrogen). Appropriate antibiotics were added to the medium when required (chloramphenicol 25µg/mL, kanamycin 50µg/mL, ampicillin 50µg/mL). Inactivation of genes *speB*, *speC*, *speE*, and *speG* was performed by using the lambda Red recombinase method (35) using pairs of primers presented in **Table 2**. Allelic exchanges were confirmed by PCR. For complementation, the *speE* gene was PCR amplified using primers *speE_CompF* and *speE_CompR* and cloned into pSC-A-amp/kan using the StrataClone PCR kit (Agilent) (p-*speE*, **Table 1**).

For megalocytosis assay, genotoxicity quantification, DNA crosslinking assay, *N*-myristoyl-D-asparagine quantification and bioluminescence measurements, *E. coli* strains were pre-grown overnight at 37°C with shaking in DMEM-Hepes (Gibco), a spermidine-free medium. Overnight cultures were then diluted 1:50 in DMEM-Hepes and grown until OD_{600nm} = 0.6 and then processed for experiments. For enumeration experiments, each strain was pre-grown in

LB to reach exponential growth ($OD_{600nm} = 0.4$) and then inoculated at 2×10^6 bacteria/mL in LB and grown overnight (17h). For colony-forming unit (CFU) count, cultures were serially diluted in PBS, plated on LB agar plates and incubated overnight at 37°C.

Determination of the megalocytosis and genotoxic effect induced by colibactin

The megalocytosis and genotoxic effect induced by colibactin were determined as previously described (11). Briefly, HeLa cells were dispensed in 96-well cell culture plate (5 to 7.5×10^3 cells/well) and incubated for 24 hours. For bacterial infections, pre-cultures of *E. coli* were diluted in interaction medium (DMEM, 5% FCS, 25mM HEPES) and cell cultures were infected with a multiplicity of infection (MOI, number of bacteria per HeLa cell at the onset of infection) of 100 to 400. Four hours post-inoculation, cells were washed 3 times with HBSS and incubated in cell culture medium with 200µg/mL gentamicin until analysis.

For megalocytosis quantification, cells were incubated for 72h before protein staining with methylene blue (1% w/v in Tris-HCl 0.01M). The methylene blue was extracted with HCl 0.1N. The quantification of staining was measured at OD_{660nm} .

For H2AX phosphorylation quantification, cells were incubated for 3h before fixation with paraformaldehyde, permeabilization and blocking, as previously described (36). Cells were then incubated for 2h at room temperature with rabbit monoclonal anti-γ-H2AX antibody #9718 (Cell Signaling Technology, 1:200). An infrared fluorescent IRDyeTM800CW-conjugated goat anti-rabbit secondary antibody (Rockland, 1:200) was used to detect γ-H2AX. DNA was counterstained with RedDot2 (Biotium, 1:500). DNA and γ-H2AX were visualized simultaneously with an Odyssey Infrared Imaging Scanner (Li-Cor Biosciences) using 680nm and 800nm channels for RedDot2 and IRDyeTM800, respectively. Relative fluorescence units for γ-H2AX per cell (as determined by γ-H2AX divided by DNA content) were divided by untreated controls to determine percentage change in phosphorylation of γ-H2AX levels relative to control.

***N*-myristoyl-D-asparagine (colibactin prodrug motif) quantification by liquid chromatography/mass spectrometry (LC/MS)**

Quantification of colibactin prodrug motif was performed as previously described (24). Briefly, pre-cultivated strains were grown in DMEM-Hepes at 37°C for 18h under shaking (240rpm). Supernatants of cultures were obtained by centrifugation of bacterial cells at 3,200 × g for 15min and were filtered on 0.2µm membranes. Each strain was cultured in triplicate (deriving from three independent clones) and each supernatant was analyzed by LC-MS/MS.

Quantification experiments were conducted with ultra-performance liquid chromatography-high resolution/heated electrospray ionization mass spectrometry (UPLC-HR/HESI-MS). The data were recorded on a Thermo Scientific Q Exactive Hybrid Quadrupole-Orbitrap Mass Spectrometer coupled to a Dionex Ultimate 3000 UPLC. The following solvent gradient (A = H₂O + 0.1% formic acid, B = acetonitrile + 0.1% formic acid with B at 30% from 0-1min, 30-95% from 1-6min and 95% from 6-7min at a flowrate of 0.5mL/min) was used on a Phenomenex Kinetex 5µm EVO C18 (50 × 2.1mm) column at 30°C. The MS was operated in positive ionization mode at a scan range of 200-500 m/z and a resolution of 35,000. The spray voltage was set to 3.5kV, the S-lens to 35, the auxiliary gas heater temperature to 438°C and the capillary temperature to 270°C. Absolute quantification was achieved by using a Schotten-Baumann reaction-derived *N*-myristoyl-L-asparagine (isomer of the *N*-myristoyl-D-asparagine colibactin cleavage product) as a standard. Data were obtained from undiluted cell free sample supernatants and analyzed for *N*-myristoyl-D-asparagine and concentrations were calculated using Thermo Xcalibur 2.2 Quan Browser.

DNA crosslinking assay

Linearized DNA was obtained by digesting pUC19 plasmid with *Bam*H1 restriction enzyme (NEB). Purified linearized DNA was quantified and diluted to obtain a 200ng/µL stock solution. For bacteria-DNA interactions, 1 × 10⁶, 2 × 10⁶, or 4 × 10⁶ bacteria were cocultivated, in a 96-

well plate, with 200ng of linearized DNA during 4h at 37°C without shaking. Wells content was then centrifuged 5min at $5000 \times g$ to pellet bacteria, and the DNA present in supernatants was purified using the PCR Purification Kit (Qiagen) following the manufacturer recommendations.

The agarose gel was prepared by dissolving 1.0g of agarose in 100mL of an aqueous 100mM NaCl and 2mM EDTA solution (pH 8.0). The prepared gel was then soaked (2h) in an aqueous alkaline running buffer solution (40mM NaOH and 1mM EDTA, pH ~12.0). 7.5 μ L of each DNA sample were loaded on the agarose gel. The gel was run at 25V (45min) and then at 50V (3h). The gel was neutralized for a total of 45min in an aqueous 100 mM Tris pH 7.4 buffer solution containing 150mM NaCl, refreshing the neutralizing solution every 15min. The gel was stained with an aqueous solution containing GelRed for 20min and revealed with UV exposure using ChemiDoc Imaging System (BioRad).

Luciferase measurements

Promoter activities of genes *clbA*, *clbB*, *clbQ*, and *clbR* were determined by time-course quantification of luciferase, as previously described (37). *E. coli* strains were grown overnight in DMEM-Hepes, and subcultured into DMEM-Hepes up to $OD_{600nm} = 0.6$. Samples of 100 μ L were then used to inoculate black 96-well plate (Greiner Bio-one) and grown without shaking at 37°C in a luminometer (Tecan Infinite Pro microplate reader). Both OD_{600nm} and light emission (relative luminescence units, RLU, 6000ms aperture/sample) were recorded every 30min, simultaneously.

Statistical analysis

Statistical analyses were conducted using GraphPad Prism 6.01. The mean and the standard error of the mean (SEM) are shown in figures, unless otherwise stated. *P* values were calculated by 1-way analysis of variance (ANOVA) test followed by a Bonferroni post-hoc test. A *P* value of $< .05$ was considered statistically significant.

RESULTS

The spermidine synthase SpeE is required for colibactin-associated megalocytosis phenotype

The *pks* island was initially identified as the genomic determinant for the synthesis of colibactin through the screening of a transposon mutant library (1). Screening of this library had revealed that a subset of transposon mutants had insertions in the *speB* and *speC* genes, suggesting that the polyamines metabolism was potentially involved in colibactin biosynthesis (unpublished data). To investigate the impact of polyamines in colibactin production, deletion of the *speB*, *speC*, *speE*, and *speG* genes were constructed in the *E. coli* laboratory strain DH10B harboring the *pks* island on a bacterial artificial chromosome (pBAC*pks*), *i.e.* strain DH10B *pks*⁺ (**Table 1**). The production of colibactin was analyzed in the mutant strains through bacteria-host cells interactions and subsequent observation of megalocytosis (**Figure 1**), which correlates with DNA double strand breaks resulting from the genotoxic effect of colibactin (1,2).

The megalocytosis assay revealed that the inactivation of *speB*, *speC*, and *speE* genes in DH10B *pks*⁺ strain abrogated the genotoxic effect induced by the colibactin (**Figure 1A**). On the contrary, disruption of *speG* gene did not affect the megalocytosis phenotype. Thus, we focused our analysis on *speE*, which encodes the spermidine synthase SpeE, an enzyme catalyzing the final step in the biosynthesis of the polyamine spermidine (38).

The loss of the megalocytosis phenotype with the Δ *speE* mutant was then confirmed by semi-quantification with methylene blue staining in the laboratory *E. coli* strain DH10B *pks*⁺ (**Figure 1B**), but also in the M1/5 commensal *E. coli* strain and the SP15 ExPEC strain (**Table 1**, **Figure 1C**). Inactivation of *speE* in DH10B *pks*⁺, M1/5, and SP15 strains resulted in decreased colibactin activity for the three tested genetic contexts (**Figure 1B** and **C**). Transformation of the Δ *speE* mutant with p-*speE* carrying the functional wild-type *speE* gene in strain DH10B *pks*⁺ resulted in a complete restoration of the genotoxicity (**Figure 1B**). Altogether, these data demonstrated that SpeE was required for colibactin-associated megalocytosis phenotype.

The spermidine synthase SpeE is required for full genotoxicity of colibactin-producing *E. coli*

We then used a quantitative assay to measure the genotoxic activity of the wild-type, $\Delta speE$ mutant, and its complemented derivative of *E. coli* DH10B *pks*⁺ strains through bacteria-host cells interactions and subsequent quantification of histone H2AX phosphorylation, a sensitive marker of DNA double-strand breaks resulting from the genotoxic effect of colibactin (1,2). A $\Delta clbA$ mutant, in which production of colibactin is completely abolished, was used as a negative control.

The quantification of H2AX phosphorylation using an In-Cell Western assay (36) revealed that the inactivation of the *speE* gene in the DH10B *pks*⁺ strain drastically decreased the genotoxic effect induced by the colibactin (**Figure 2**). Noteworthy, the $\Delta speE$ mutant presented a residual genotoxic activity compared to the $\Delta clbA$ mutant (**Figure 2**). The complemented mutant displayed a restoration of the genotoxic activity (**Figure 2**). These results indicated that SpeE was required for full genotoxicity of colibactin-producing *E. coli* strains, but a $\Delta speE$ mutant still showed a residual genotoxic activity.

Spermidine supply restores the colibactin-mediated genotoxicity in a *speE* mutant

In order to determine whether the role of *speE* in colibactin production was associated to the enzymatic activity of the enzyme SpeE or to its product, the polyamine spermidine, we supplemented the interaction medium with increasing concentrations of spermidine during host cell infection (**Figure 3**). The production of colibactin in the mutant strain supplemented or not with exogenous spermidine was measured through bacteria-host cells interactions and subsequent quantification of histone H2AX phosphorylation (**Figure 3**). This revealed that spermidine supplementation of the $\Delta speE$ mutant in the DH10B *pks*⁺ strain restored the genotoxic effect induced by the colibactin in a dose-dependent manner (**Figure 3**). A concentration of at least 0.5 $\mu\text{g/mL}$ was required to observe partial complementation, whereas maximum restoration seemed to be reached with a concentration of 2 $\mu\text{g/mL}$. Furthermore, spermidine supplementation of the wild-type strain did not enhance its genotoxic activity (**Figure 3**). These data demonstrated that the spermidine polyamine was involved in colibactin production.

Spermidine is not mandatory for colibactin-associated DNA crosslink

To test the capacity of the genotoxin colibactin to generate crosslinks with DNA (19), we developed an assay to monitor colibactin-associated DNA crosslinking after a 4-hours coculture of live bacteria with linearized DNA (Greif *et al.*, unpublished data). After interaction with wild-type *E. coli* strain DH10B *pks*⁺, Δ *clbA* mutant, Δ *speE* mutant and its complemented derivative in DMEM-Hepes, DNA was purified and analyzed in alkaline agarose gel (**Figure 4**). This revealed the presence of two major bands for the wild-type strain: the upper one corresponded to crosslinked DNA, whereas the lower one matched with linearized single-stranded DNA (**Figure 4, WT**). The formation of DNA interstrand crosslink occurred in a dose-dependent manner (**Figure 4, WT**). This experiment also showed that *speE* inactivation deeply attenuated the colibactin DNA crosslinking activity, which was observed only for the highest bacterial dose (**Figure 4, Δ *speE***). Either exogenous addition of spermidine (+ **spermidine**) or *speE* plasmidic complementation (+ **p-*speE***) restored the ability of the Δ *speE* mutant to induce DNA crosslink (**Figure 4**). Together, these data suggested that spermidine was not mandatory for direct interaction of colibactin with DNA neither for crosslink formation, but potentiated colibactin activity or enhanced its production.

Spermidine is not required for colibactin autotoxicity

ClbS, a resistance protein blocking the genotoxicity of colibactin, participates in the self-protection of colibactin-producing *E. coli* (39). A Δ *clbS* mutant displays an autotoxicity phenotype characterized by a marked decrease of viable bacterial cells in overnight cultures compared to the wild-type strain. In addition, the Δ *clbS* mutant grown on agar plates exhibits a distinct colony morphology with flattened or depressed centers. In order to determine whether spermidine was required for the colibactin-associated autotoxicity, we constructed a Δ *speE* Δ *clbS* double-mutant in the *E. coli* strain DH10B *pks*⁺ (**Table 1**).

Wild-type strain, Δ *speE* mutant, Δ *clbS* mutant, and Δ *speE* Δ *clbS* double-mutant were grown overnight on LB agar plates, then observed for colony morphology (**Figure 5A**). The Δ *clbS*

mutant colonies presented depressed centers, as previously described (39). The $\Delta speE \Delta clbS$ double-mutant exhibited a similar phenotype, whereas wild-type and $\Delta speE$ mutant colonies displayed a normal convex appearance (**Figure 5A**). The same strains were cultivated in LB during 17 hours to reach the stationary growth phase, then plated on LB agar plates to determine CFU counts (**Figure 5B**). As expected, the $\Delta clbS$ mutant presented a decreased CFU number, compared to the wild-type strain (**Figure 5B**). A drop in CFU counts was also observed for the $\Delta speE \Delta clbS$ double-mutant, compared to the $\Delta speE$ mutant, but to a lesser extent (**Figure 5B**), suggesting that *speE* disruption slightly decreases colibactin autotoxicity in a $\Delta clbS$ mutant. However, this experiment will be repeated in spermidine-free medium to confirm these data.

Spermidine is involved in colibactin biosynthesis

The mature colibactin genotoxin has not been fully characterized yet. Another biosynthetic by-product derived from the colibactin assembly line has been characterized, *i.e.* the *N*-myristoyl-D-asparagine moiety (14,15). This moiety, generated by the ClbN enzyme, is a prodrug motif cleaved by the ClbP peptidase in the late activation step of inactive pre-colibactin to the active compound. We developed a LC-MS assay to quantitatively measure the amount of *N*-myristoyl-D-asparagine (24) as a means to indirectly quantify the production of the genotoxin by wild-type, $\Delta clbA$ mutant, $\Delta speE$ mutant and its complemented derivative in *E. coli* DH10B *pks*⁺ (**Figure 6**). This revealed that the amount of colibactin prodrug motif was drastically decreased in the $\Delta speE$ mutant, compared to the wild-type strain (**Figure 6**). Complementation with *p-speE* or supplementation with exogenous spermidine partially restored the production of *N*-myristoyl-D-asparagine by the $\Delta speE$ mutant (**Figure 6**). These results indicated that spermidine was required for the synthesis of the biosynthetic intermediate *N*-myristoyl-D-asparagine, and therefore for colibactin biosynthesis.

Spermidine could regulate the expression of the *pks* island genes

[This paragraph presents preliminary results that will require further investigations]

To investigate whether the polyamine spermidine was involved in the regulation of colibactin production *via* transcriptional or translational regulation of the colibactin biosynthetic genes, *E. coli* strain Nissle 1917 mutated or not for *speE*, and harboring luciferase fusions with the genes *clbA*, *clbB*, *clbQ*, and *clbR*, were studied (**Table 1** and **Figure 7**) (40). These strains were grown in chemically defined and spermidine-free DMEM-Hepes. The expression rates of genes *clbA*, *clbB*, *clbQ*, and *clbR* were determined as OD_{600nm} standardized-relative luminescence units (RLU) (**Figure 7**). A continuous increase of bioluminescence emission was observed for the four genes during the bacterial growth kinetics, to reach a maximal value at early stationary phase of growth, followed by a decrease (**Figure 7**). *clbA* was the most highly expressed in DMEM-Hepes, as previously reported (**Figure 7A**, compared to **Figure 7B, C** and **D**) (40).

We observed that the expression of the *clbA* gene was repressed in the Δ *speE* mutant compared to the wild-type strain, but was not restored with spermidine supplementation (20 μ g/mL) (**Figure 7A**). Addition of spermidine (20 μ g/mL) in the culture medium of the wild-type strain resulted in a decrease of *clbA* expression (**Figure 7A**). This result was not in agreement with those presented above, as spermidine supplementation for the wild-type strain did not modify genotoxic activity (**Figure 3**). We could hypothesize that the expression of the *clbA* gene is induced by spermidine at physiological concentrations, but repressed in spermidine-depleted or in high-polyamine conditions. This hypothesis will be investigated by testing various conditions of spermidine-supplementation (increasing concentrations from 0.1 to 100 μ g/mL).

In contrast, when the same scale axis as for *clbA* was applied to the three other genes, the expression of *clbB*, *clbQ*, and *clbR* did not seem to be significantly modified by the *speE* disruption (**Figure 7B, C** and **D**, respectively), neither when the culture medium was supplemented in spermidine. However, visualization of the same results in a different RLU

scale showed that expression of *clbQ* and *clbR*, but not *clbB*, was significantly attenuated in the Δ *speE* mutant, and totally restored with spermidine supplementation (**Figure 7F, G and E**, respectively). Finally, polyamine supply with 20 μ g/mL of spermidine in the culture medium of the wild-type strain did not modify gene expression of *clbB*, *clbQ*, and *clbR* (**Figure 7E, F and G**, respectively). These results could suggest that spermidine induce *clbQ* and *clbR* expression, but in a different manner than *clbA*.

Together, these data could support the hypothesis that spermidine modulates the expression of genes of the *pks* island. However, further investigations will be performed to elucidate the regulation mechanism(s). Proteomic and transcriptomic approaches in conditions of spermidine-depletion or spermidine-supplementation might contribute to characterize the role of spermidine in the regulation of the colibactin production.

DISCUSSION

[Discussion préliminaire en français qui évoluera en fonction des résultats complémentaires]

Ce travail démontre que la spermidine est impliquée dans la production de colibactine. Les polyamines étant principalement associées aux ARN au sein de la cellule, nous avons testé l'hypothèse de l'implication de la spermidine dans la régulation de l'expression des gènes de l'îlot *pks*. Les résultats préliminaires obtenus suggèrent que la spermidine modulerait l'expression du gène *clbA*, et dans une moindre mesure des gènes *clbQ* et *clbR*. Le mécanisme de régulation exercé par la spermidine semble être cependant différent pour ces trois gènes, puisqu'une supplémentation en spermidine pour une souche mutée pour *speE* restaure l'expression physiologique des gènes *clbQ* et *clbR*, alors que l'ajout de spermidine dans le milieu de culture de la souche sauvage inhibe l'expression de *clbA*. Les mécanismes de régulation semblent donc complexes, et leur élucidation nécessitera des investigations complémentaires, comme par exemple l'utilisation d'approches protéomiques et transcriptomiques en l'absence ou en présence de spermidine. Ces résultats n'excluent pas pour

autant que la spermidine soit incorporée dans la structure de la colibactine, potentialisant son activité génotoxique, tout en exerçant un rôle régulateur dans l'expression des gènes de l'îlot *pks*. Cette hypothèse est d'ailleurs explorée en collaboration avec une équipe de chimistes spécialistes de la synthèse des composés PK-NRP (Dr. Alexander Brachmann and Pr. Jörn Piel, Institut de Microbiologie, ETH Zurich, Suisse).

Si le rôle de la spermidine dans la modulation de la synthèse de colibactine n'est pas clairement élucidé à ce jour, cette étude met en lumière un lien entre le métabolisme primaire avec les polyamines, et le métabolisme secondaire avec les composés PK-NRP. La compréhension des interconnexions entre ces voies de biosynthèse est essentielle pour mieux appréhender les mécanismes de régulation de la production de facteurs d'adaptation ou de virulence bactériens de type métabolites secondaires, et pour identifier de nouvelles cibles thérapeutiques antimicrobiennes.

REFERENCES

1. Nougayrède J-P, Homburg S, Taieb F, *et al.* *Escherichia coli* induces DNA double-strand breaks in eukaryotic cells. *Science*. **2006**; 313(5788):848–851.
2. Cuevas-Ramos G, Petit CR, Marcq I, Boury M, Oswald E, Nougayrède J-P. *Escherichia coli* induces DNA damage *in vivo* and triggers genomic instability in mammalian cells. *Proc Natl Acad Sci U S A*. **2010**; 107(25):11537–11542.
3. Marcq I, Martin P, Payros D, *et al.* The genotoxin colibactin exacerbates lymphopenia and decreases survival rate in mice infected with septicemic *Escherichia coli*. *J Infect Dis*. **2014**; 210(2):285–294.
4. Secher T, Samba-Louaka A, Oswald E, Nougayrède J-P. *Escherichia coli* producing colibactin triggers premature and transmissible senescence in mammalian cells. *PLoS One*. **2013**; 8(10):e77157.
5. Cougnoux A, Dalmaso G, Martinez R, *et al.* Bacterial genotoxin colibactin promotes colon tumour growth by inducing a senescence-associated secretory phenotype. *Gut*. **2014**; 63(12):1932–1942.

6. Arthur JC, Perez-Chanona E, Mühlbauer M, *et al.* Intestinal inflammation targets cancer-inducing activity of the microbiota. *Science*. **2012**; 338(6103):120–123.
7. Dalmasso G, Cougnoux A, Delmas J, Darfeuille-Michaud A, Bonnet R. The bacterial genotoxin colibactin promotes colon tumor growth by modifying the tumor microenvironment. *Gut Microbes*. **2014**; 5(5):675–680.
8. Olier M, Marcq I, Salvador-Cartier C, *et al.* Genotoxicity of *Escherichia coli* Nissle 1917 strain cannot be dissociated from its probiotic activity. *Gut Microbes*. **2012**; 3(6):501–509.
9. Payros D, Secher T, Boury M, *et al.* Maternally acquired genotoxic *Escherichia coli* alters offspring's intestinal homeostasis. *Gut Microbes*. **2014**; 5(3):313–325.
10. Secher T, Payros D, Brehin C, *et al.* Oral tolerance failure upon neonatal gut colonization with *Escherichia coli* producing the genotoxin colibactin. *Infect Immun*. **2015**; 83(6):2420–2429.
11. Martin P, Marcq I, Magistro G, *et al.* Interplay between siderophores and colibactin genotoxin biosynthetic pathways in *Escherichia coli*. *PLoS Pathog*. **2013**; 9(7):e1003437.
12. McCarthy AJ, Martin P, Cloup E, Stabler RA, Oswald E, Taylor PW. The genotoxin Colibactin is a determinant of virulence in *Escherichia coli* K1 experimental neonatal systemic infection. *Infect Immun*. **2015**; 83(9):3704–3711.
13. Putze J, Hennequin C, Nougayrède J-P, *et al.* Genetic structure and distribution of the colibactin genomic island among members of the family *Enterobacteriaceae*. *Infect Immun*. **2009**; 77(11):4696–4703.
14. Brotherton CA, Balskus EP. A prodrug resistance mechanism is involved in colibactin biosynthesis and cytotoxicity. *J Am Chem Soc*. **2013**; 135(9):3359–3362.
15. Bian X, Fu J, Plaza A, *et al.* *In vivo* evidence for a prodrug activation mechanism during colibactin maturation. *Chembiochem Eur J Chem Biol*. **2013**; 14(10):1194–1197.
16. Vizcaino MI, Engel P, Trautman E, Crawford JM. Comparative metabolomics and structural characterizations illuminate colibactin pathway-dependent small molecules. *J Am Chem Soc*. **2014**; 136(26):9244–9247.
17. Brotherton CA, Wilson M, Byrd G, Balskus EP. Isolation of a metabolite from the *pks* island provides insights into colibactin biosynthesis and activity. *Org Lett*. **2015**; 17(6):1545–1548.

18. Bian X, Plaza A, Zhang Y, Müller R. Two more pieces of the colibactin genotoxin puzzle from *Escherichia coli* show incorporation of an unusual 1-aminocyclopropanecarboxylic acid moiety. *Chem Sci.* **2015**; 6(5):3154–3160.
19. Vizcaino MI, Crawford JM. The colibactin warhead crosslinks DNA. *Nat Chem.* **2015**; 7(5):411–417.
20. Li Z-R, Li Y, Lai JYH, *et al.* Critical Intermediates Reveal New Biosynthetic Events in the Enigmatic Colibactin Pathway. *Chembiochem Eur J Chem Biol.* **2015**; 16(12):1715–1719.
21. Zha L, Wilson MR, Brotherton CA, Balskus EP. Characterization of Polyketide Synthase Machinery from the *pks* Island Facilitates Isolation of a Candidate Precolibactin. *ACS Chem Biol.* **2016**; 11(5):1287–1295.
22. Healy AR, Vizcaino MI, Crawford JM, Herzon SB. Convergent and Modular Synthesis of Candidate Precolibactins. Structural Revision of Precolibactin A. *J Am Chem Soc.* **2016**; 138(16):5426–5432.
23. Li Z-R, Li J, Gu J-P, *et al.* Divergent biosynthesis yields a cytotoxic aminomalonate-containing precolibactin. *Nat Chem Biol.* **2016**; 12(10):773–775.
24. Garcie C, Tronnet S, Garénaux A, *et al.* The Bacterial Stress-Responsive Hsp90 Chaperone (HtpG) Is Required for the Production of the Genotoxin Colibactin and the Siderophore Yersiniabactin in *Escherichia coli*. *J Infect Dis.* **2016**; 214(6):916–924.
25. Tabor CW, Tabor H. Polyamines in microorganisms. *Microbiol Rev.* **1985**; 49(1):81–99.
26. Igarashi K, Kashiwagi K. Modulation of cellular function by polyamines. *Int J Biochem Cell Biol.* **2010**; 42(1):39–51.
27. Miller-Fleming L, Olin-Sandoval V, Campbell K, Ralser M. Remaining Mysteries of Molecular Biology: The Role of Polyamines in the Cell. *J Mol Biol.* **2015**; 427(21):3389–3406.
28. Michael AJ. Polyamines in Eukaryotes, Bacteria and Archaea. *J Biol Chem.* **2016**; DOI:10.1074/jbc.R116.734780.
29. Charlier D, Glansdorff N. Biosynthesis of Arginine and Polyamines. *EcoSal Plus.* **2004**; 3(6):1–55.
30. Igarashi K, Kashiwagi K. Characteristics of cellular polyamine transport in prokaryotes and eukaryotes. *Plant Physiol Biochem PPB Société Fr Physiol Végétale.* **2010**; 48(7):506–512.

31. Miyamoto S, Kashiwagi K, Ito K, Watanabe S, Igarashi K. Estimation of polyamine distribution and polyamine stimulation of protein synthesis in *Escherichia coli*. Arch Biochem Biophys. **1993**; 300(1):63–68.
32. Igarashi K, Kashiwagi K. Modulation of protein synthesis by polyamines. IUBMB Life. **2015**; 67(3):160–169.
33. Keating TA, Marshall CG, Walsh CT. Vibriobactin biosynthesis in *Vibrio cholerae*: VibH is an amide synthase homologous to nonribosomal peptide synthetase condensation domains. Biochemistry (Mosc). **2000**; 39(50):15513–15521.
34. Tan W, Verma V, Jeong K, *et al*. Molecular characterization of vulnibactin biosynthesis in *Vibrio vulnificus* indicates the existence of an alternative siderophore. Aquat Microbiol. **2014**; 5:1.
35. Datsenko KA, Wanner BL. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc Natl Acad Sci U S A. **2000**; 97(12):6640–6645.
36. Audebert M, Riu A, Jacques C, *et al*. Use of the γ H2AX assay for assessing the genotoxicity of polycyclic aromatic hydrocarbons in human cell lines. Toxicol Lett. **2010**; 199(2):182–192.
37. Tronnet S, Garcie C, Rehm N, Dobrindt U, Oswald E, Martin P. Iron homeostasis regulates the genotoxicity of *Escherichia coli* producing colibactin. Infect Immun. **2016**; DOI:10.1128/IAI.00659-16.
38. Tabor CW, Tabor H, Xie QW. Spermidine synthase of *Escherichia coli*: localization of the *speE* gene. Proc Natl Acad Sci U S A. **1986**; 83(16):6040–6044.
39. Bossuet-Greif N, Dubois D, Petit C, *et al*. *Escherichia coli* ClbS is a colibactin resistance protein. Mol Microbiol. **2016**; 99(5):897–908.
40. Homburg S, Oswald E, Hacker J, Dobrindt U. Expression analysis of the colibactin gene cluster coding for a novel polyketide in *Escherichia coli*. FEMS Microbiol Lett. **2007**; 275(2):255–262.
41. Johnson JR, Oswald E, O’Bryan TT, Kuskowski MA, Spanjaard L. Phylogenetic distribution of virulence-associated genes among *Escherichia coli* isolates associated with neonatal bacterial meningitis in the Netherlands. J Infect Dis. **2002**; 185(6):774–784.

TABLES

Table 1. Strains and plasmids used in this study.

Strain or plasmid	Genotype or phenotype	Source
<i>E. coli</i> strain		
DH10B <i>pks</i> ⁺	K-12 <i>E. coli</i> laboratory strain carrying pBAC <i>pks</i> , Cm ^r	(1)
DH10B <i>pks</i> ⁺ Δ <i>speB</i>	<i>speB</i> mutant of strain DH10B pBAC <i>pks</i> , Cm ^r Kan ^r	This study
DH10B <i>pks</i> ⁺ Δ <i>speC</i>	<i>speC</i> mutant of strain DH10B pBAC <i>pks</i> , Cm ^r Kan ^r	This study
DH10B <i>pks</i> ⁺ Δ <i>speE</i>	<i>speE</i> mutant of strain DH10B pBAC <i>pks</i> , Cm ^r Kan ^r	This study
DH10B <i>pks</i> ⁺ Δ <i>speG</i>	<i>speG</i> mutant of strain DH10B pBAC <i>pks</i> , Cm ^r Kan ^r	This study
DH10B <i>pks</i> ⁺ Δ <i>speE</i> + p- <i>speE</i>	<i>speE</i> mutant of strain M1/5 carrying p- <i>speE</i> , Cm ^r Kan ^r Amp ^r	This study
M1/5	Commensal <i>E. coli</i> strain isolated from feces of a healthy adult; B2 phylogenetic group; colibactin genotoxin producer	(11)
M1/5 Δ <i>speE</i>	<i>speE</i> mutant of strain M1/5, Kan ^r	This study
SP15	ExPEC strain isolated from spinal fluid of a neonate with meningitis; O18:K1 serotype; colibactin genotoxin producer	(41)
SP15 Δ <i>speE</i>	<i>speE</i> mutant of strain SP15, Kan ^r	This study
DH10B <i>pks</i> ⁺ Δ <i>clbA</i>	<i>clbA</i> mutant of strain DH10B pBAC <i>pks</i> , Cm ^r	(1)
DH10B <i>pks</i> ⁺ Δ <i>clbS</i>	<i>clbS</i> mutant of strain DH10B pBAC <i>pks</i> , Cm ^r	(39)
DH10B <i>pks</i> ⁺ Δ <i>speE</i> Δ <i>clbS</i>	<i>speE clbS</i> double-mutant of strain DH10B pBAC <i>pks</i> , Cm ^r Kan ^r	This study
EcN <i>clbA-lux</i>	Luciferase fusion of the <i>clbA</i> gene in strain <i>E. coli</i> Nissle 1917, Kan ^r	(40)
EcN <i>clbB-lux</i>	Luciferase fusion of the <i>clbB</i> gene in strain <i>E. coli</i> Nissle 1917, Kan ^r	(40)
EcN <i>clbQ-lux</i>	Luciferase fusion of the <i>clbQ</i> gene in strain <i>E. coli</i> Nissle 1917, Kan ^r	(40)
EcN <i>clbR-lux</i>	Luciferase fusion of the <i>clbR</i> gene in strain <i>E. coli</i> Nissle 1917, Kan ^r	(40)
EcN <i>clbA-lux</i> Δ <i>speE</i>	<i>speE</i> mutant of strain <i>E. coli</i> Nissle 1917 <i>clbA-lux</i> , Kan ^r Cm ^r	This study
EcN <i>clbB-lux</i> Δ <i>speE</i>	<i>speE</i> mutant of strain <i>E. coli</i> Nissle 1917 <i>clbB-lux</i> , Kan ^r Cm ^r	This study
EcN <i>clbQ-lux</i> Δ <i>speE</i>	<i>speE</i> mutant of strain <i>E. coli</i> Nissle 1917 <i>clbQ-lux</i> , Kan ^r Cm ^r	This study
EcN <i>clbR-lux</i> Δ <i>speE</i>	<i>speE</i> mutant of strain <i>E. coli</i> Nissle 1917 <i>clbR-lux</i> , Kan ^r Cm ^r	This study
Plasmid		
p- <i>speE</i>	pSC-A plasmid carrying wild-type <i>speE</i> gene, Amp ^r /Kan ^r	This study

Table 2. Primers used in this study.

Primer	Sequence
<i>speB</i> _P1	TACTGGCGTGCCGTTTCGATATGGCCACTTCTGGTCGTGCGGTGTAGGCTGGAGCTGCTTC
<i>speB</i> _P2	TAATAGCGCGATCGGAGGTCAGGCCGCCAATCACTGGCGTCATATGAATATCCTCCTTAG
<i>speB</i> _F	GCCTAACGACGCGGAAGG
<i>speB</i> _R	GTTTTACCCGTGCGCATCG
<i>speC</i> _F	CGCCATCTCCTTACATTCTCTC
<i>speC</i> _R	GTTGATTTTCGCTGGTTACTCC
<i>speE</i> _P1	GGCATGAAACGCTACACGACCAGTTTGGGCAGTACTTTGGTGTAGGCTGGAGCTGCTTC
<i>speE</i> _P2	GATTGTAATAACGGCATTTCAGGCCAGAGGCGAGAAACATATGAATATCCTCCTTAG
<i>speE</i> _F	AGCTATTATGTTGCGCCCTT
<i>speE</i> _R	AGCCATGCAGTTTCAGTTTTT
<i>speE</i> _CompF	ATCTCGAGCCTTAAGCCTGCTTGCCTACG
<i>speE</i> _CompR	ATCTCGAGGTTTTGGCGTAGCAGATATCG
<i>speG</i> _P1	GAAGCCTTTGTTGAACTCTCTGATCTGTATGATAAGCGTGTAGGCTGGAGCTGCTTC
<i>speG</i> _P2	ATAGAGATTGAGAACGGTAAAGCCATAGTCCATTGCCATATGAATATCCTCCTTAG
<i>speG</i> _F	GCCGCGTTATTACCCCCTAA
<i>speG</i> _R	GGGTTTACACCATCAAAAATACGA

FIGURES

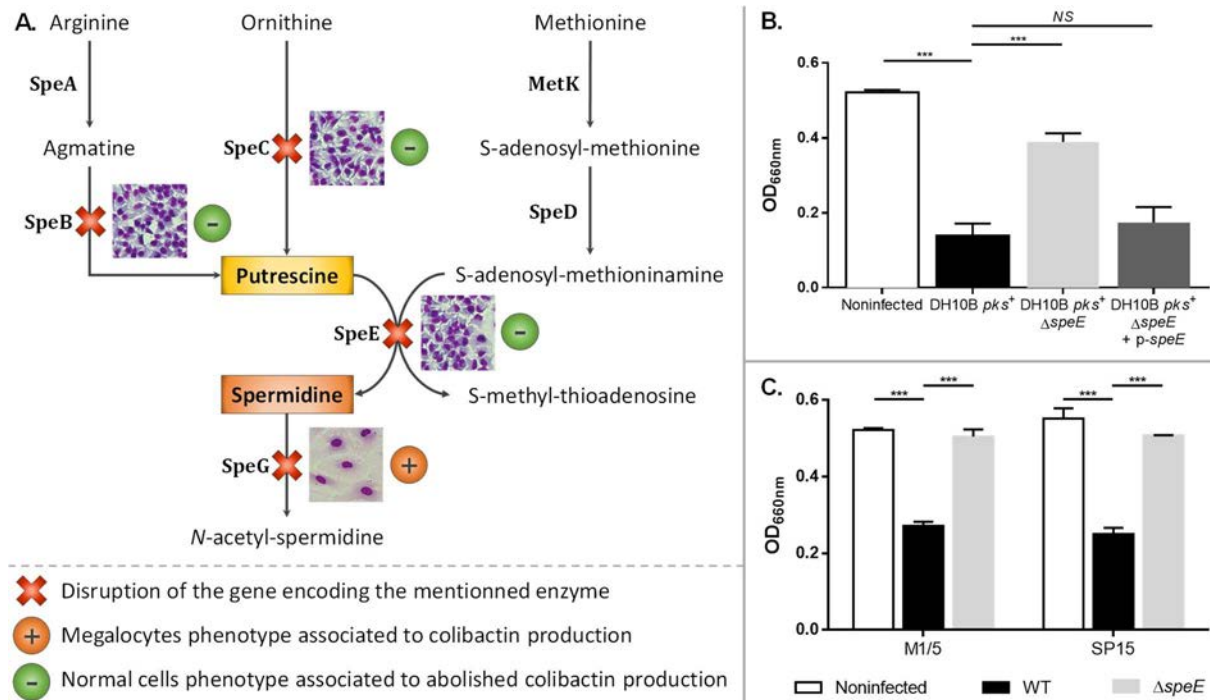


Figure 1. The spermidine synthase SpeE is required for colibactin-associated megalocytosis.

A. Graphic representation of the polyamines biosynthetic pathways. DH10B *pks*⁺ mutants for *speB*, *speC*, *speE*, and *speG* genes were tested for megalocytosis phenotype, as previously described (1). DH10B *pks*⁺ mutants were cocultivated with HeLa cells for 4h. After infection, the cells were incubated for 72h with gentamicin before staining. Pictures of observed phenotype for the different mutants are presented next to the corresponding enzyme encoded by the disrupted gene.

B. and **C.** The production of colibactin by *E. coli* strains DH10B *pks*⁺, M1/5, and SP15 and their derivatives was determined by quantification of megalocytosis. Wild-type (WT) *E. coli*, Δ*speE* mutants, and complemented derivative (in DH10B *pks*⁺ strain) were cocultivated as mentioned above. At the end of infection, bacterial growth was similar for all strains. After 72h of incubation post-infection, the cells were processed for protein staining with methylene blue. The staining was quantified by acid-extraction of methylene blue and measurement of absorbance at an optical density of 660nm (OD_{660nm}). The multiplicity of infection (MOI) was 200. Data are pooled from three independent experiments. ***: $P < .001$, by 1-way analysis of variance (ANOVA). All bar graphs show mean values ± standard errors of the mean (SEM). NS, not significant.

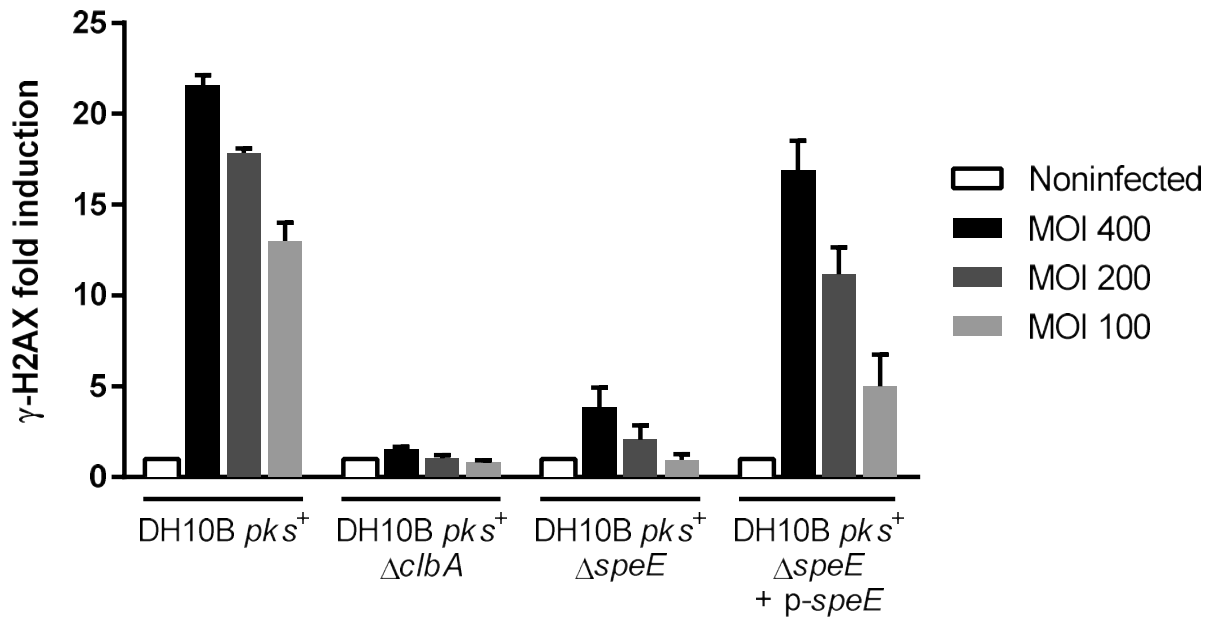


Figure 2. The spermidine synthase SpeE is required for full genotoxicity of colibactin producing *E. coli*.

The production of colibactin by *E. coli* strain DH10B *pks*⁺ and derivatives was determined by quantification of H2AX phosphorylation, which correlates to DNA double-strand breaks resulting from the genotoxic effect of colibactin. WT *E. coli* strain DH10B *pks*⁺, Δ *clbA* mutant, Δ *speE* mutant and its complemented derivative were cocultivated with HeLa cells in DMEM-Hepes for 4h, then washed as previously described (1). The cells were incubated 3h in DMEM supplemented with gentamicin before fixation, then permeabilized and labeled for DNA and phosphorylated H2AX histone (γ -H2AX) using an In-Cell Western method (36). The MOI ranged from 100 to 400. Data shown in the graph are representative from three independent experiments. All bar graphs show mean values \pm SEM.

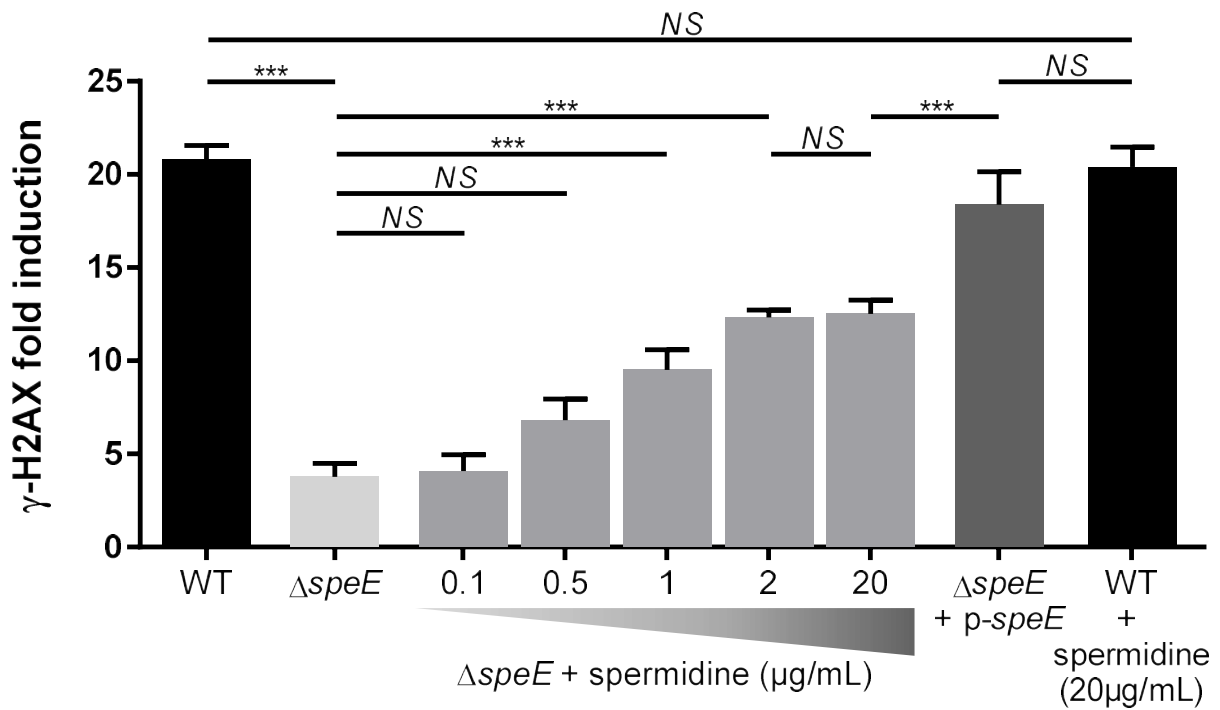


Figure 3. Spermidine supply restores colibactin-associated genotoxicity in a $\Delta speE$ mutant.

Colibactin-mediated genotoxicity was determined by infection of HeLa cells and quantification of H2AX phosphorylation for *E. coli* strain DH10B *pks*⁺ and derivatives, as described in **Figure 2**.

When indicated, spermidine ($\mu\text{g/mL}$) was added to the interaction medium during infection.

The MOI was 400. Data are pooled from three independent experiments. ***: $P < .001$, by

1-way ANOVA. All bar graphs show mean values \pm SEM. NS, not significant.

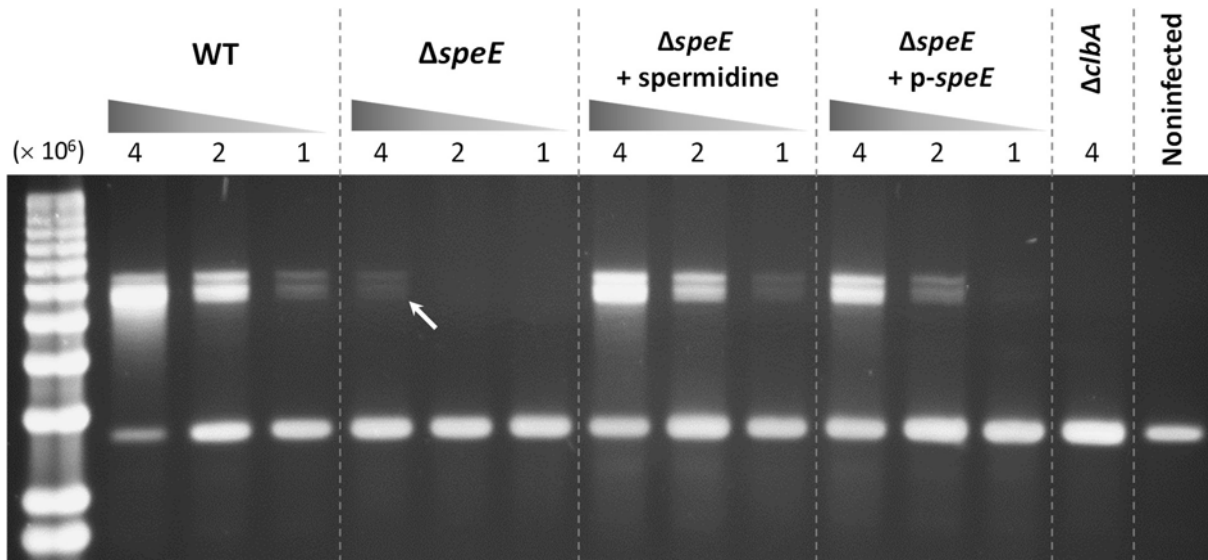


Figure 4. Spermidine is not mandatory for colibactin DNA crosslink activity.

DNA crosslinking was observed by cocultivating colibactin producing *E. coli* strains with linearized plasmidic DNA. In a 96-wells plate, 200ng of *Bam*H1 linearized pUC19 plasmid were added to 1×10^6 , 2×10^6 or 4×10^6 bacteria/well of WT *E. coli* strain DH10B *pks*⁺, $\Delta clbA$ mutant, $\Delta speE$ mutant and its complemented derivative in DMEM-Hepes. When indicated, 20 μ g/mL of spermidine were added to the interaction medium during infection in order to test for exogenous complementation. Each well was performed in duplicates. The plate was then incubated during 4h at 37°C without shaking. Wells content was centrifuged to pellet bacteria. Supernatants were then carefully aspirated, and DNA was purified on silica columns. Each DNA sample was loaded on agarose gel. After migration in alkaline conditions, the gel was neutralized and stained before UV revelation. The picture of the stained agarose gel presented here is representative from two independent experiments.

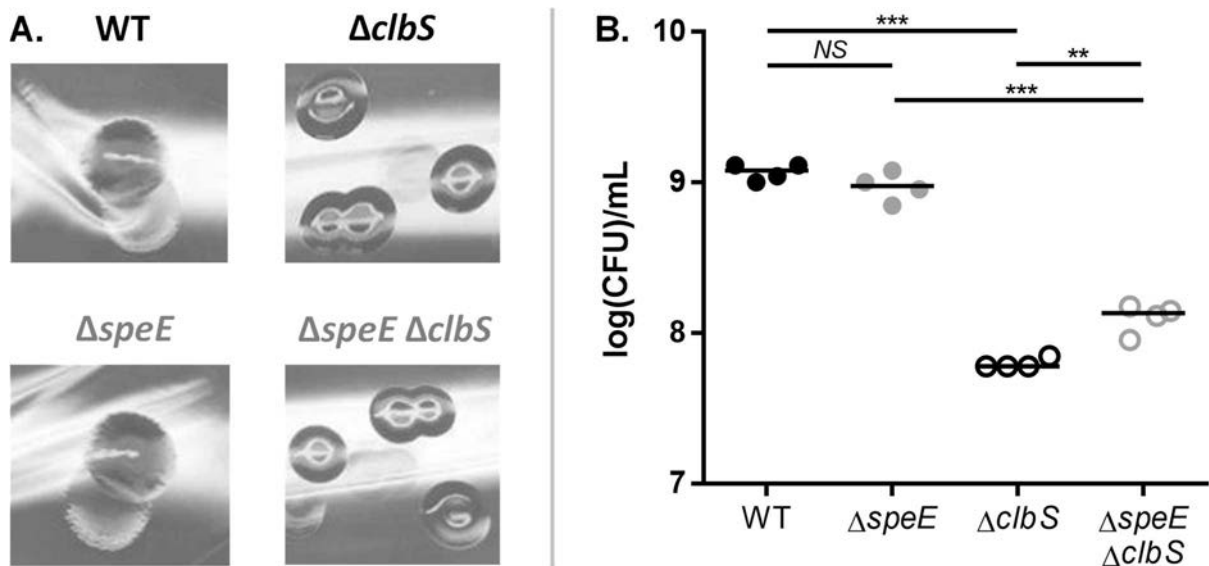


Figure 5. *speE* disruption slightly decreases colibactin autotoxicity in a $\Delta clbS$ mutant.

A. Colony morphology of WT *E. coli* strain DH10B *pks*⁺, $\Delta clbS$ mutant, $\Delta speE$ mutant, and $\Delta speE \Delta clbS$ double-mutant grown on LB agar plates.

B. Enumeration of culturable bacterial cells in the stationary growth phase. WT *E. coli* strain DH10B *pks*⁺, $\Delta clbS$ mutant, $\Delta speE$ mutant, and $\Delta speE \Delta clbS$ double-mutant were pre-grown in LB to reach exponential growth ($\text{OD}_{600\text{nm}} = 0.4$). 2×10^6 bacteria/mL were then inoculated in LB and grown for 17 hours, before plating on LB agar plates to determine colony-forming unit (CFU) numbers. The median and individual results of four independent experiments are shown. ***: $P < .001$, **: $P < .01$, by 1-way ANOVA. NS, not significant.

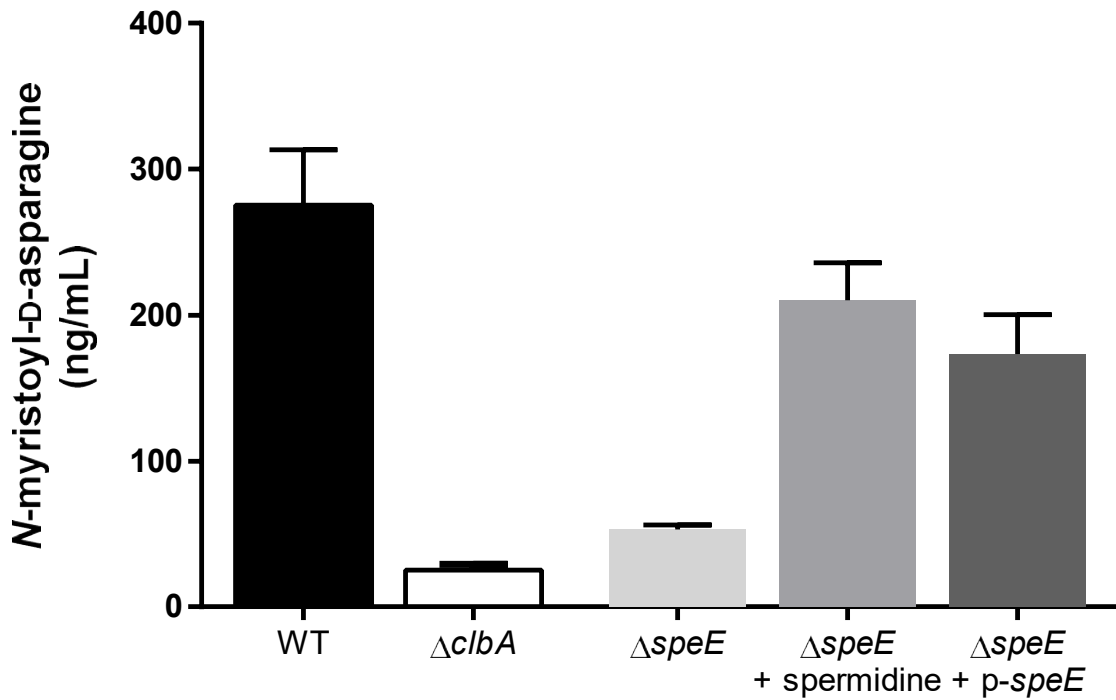


Figure 6. Spermidine is involved in the synthesis of the colibactin prodrug motif.

The colibactin prodrug motif *N*-myristoyl-D-asparagine produced by WT *E. coli* strain DH10B *pks*⁺, $\Delta clbA$ mutant, $\Delta speE$ mutant, and its complemented derivative was quantified by liquid chromatography-mass spectrometry (LC-MS). Bacteria were cultivated at 37°C for 18h in DMEM-Hepes, supplemented when indicated with 20 μ g/mL of spermidine. *N*-myristoyl-D-asparagine was quantified in culture supernatants by LC-MS, using *N*-myristoyl-L-asparagine (isomer of the *N*-myristoyl-D-asparagine colibactin prodrug motif) as a standard. The results were normalized to the bacterial biomass. Data represented in the graph were obtained from three biological replicates. All bar graphs show mean values \pm SEM.

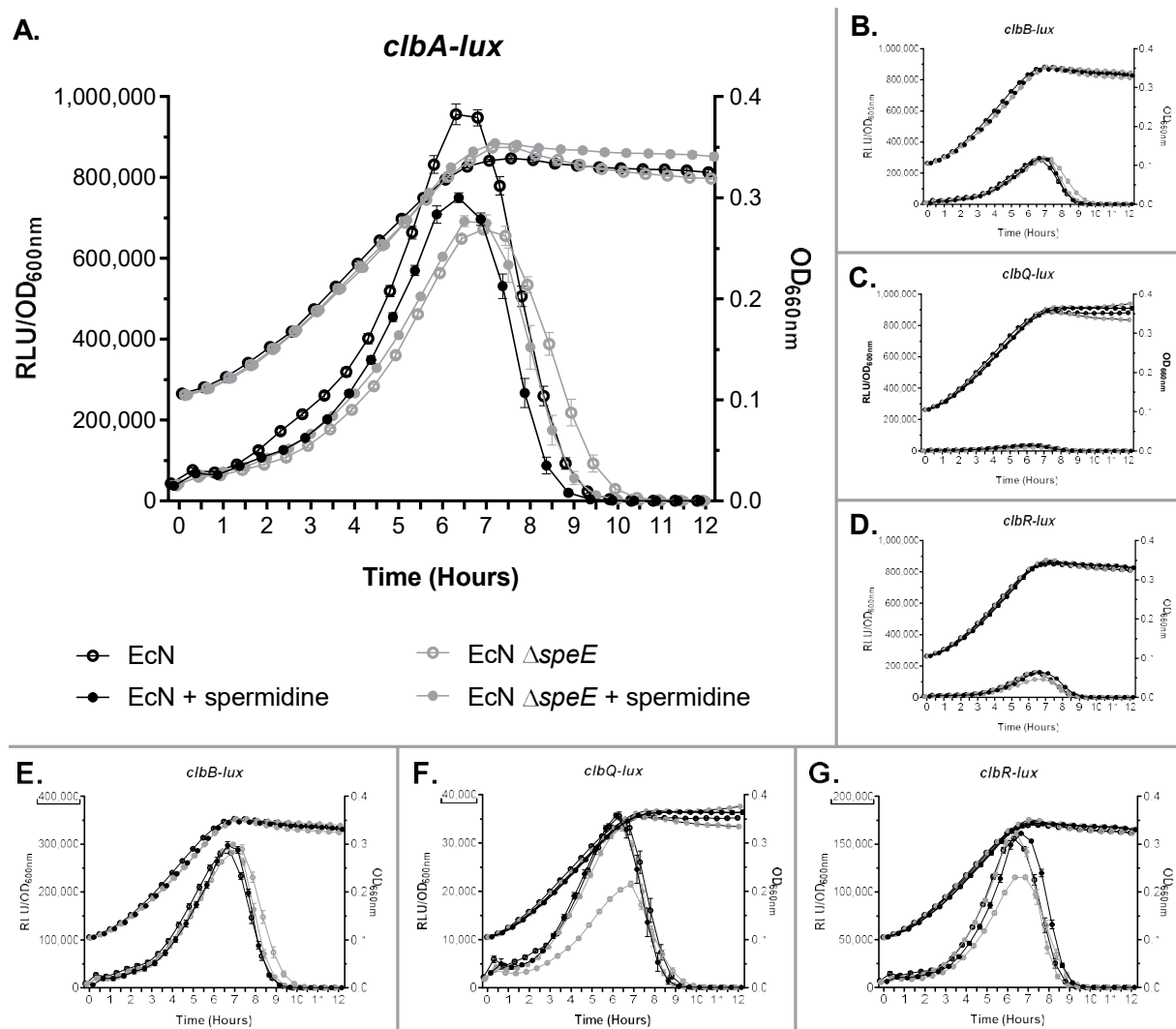


Figure 7. Spermidine modulates the expression of the genes of the *pks* island.

Growth curves (OD_{600nm}) and relative OD_{600nm} standardized-relative luminescence units (RLU/OD_{600nm}) of WT and Δ *speE* mutant of *E. coli* Nissle 1917 (EcN) *clbA-lux* (A), *clbB-lux* (B-E), *clbQ-lux* (C-F), and *clbR-lux* (D-G) fusion strains grown at 37°C in DMEM-Hepes supplemented or not with 20 μ g/mL of spermidine. Data shown in the graph were obtained from three biological replicates, and are pooled from three independent experiments. All bar graphs show mean values \pm SEM.

DISCUSSION
&
PERSPECTIVES

À LA DÉCOUVERTE DE LA RÉGULATION DE L'ÎLOT *pks*

La régulation de l'expression génique en réponse aux facteurs environnementaux est une caractéristique majeure des micro-organismes bactériens, et leur permet de s'adapter très rapidement à des conditions environnementales extrêmement variées [228]. Afin de pleinement bénéficier des avantages apportés par le matériel génétique d'origine exogène nouvellement acquis, il est indispensable de l'intégrer au réseau de régulation préalablement établi au sein de la cellule bactérienne. L'expression des gènes codés au sein d'éléments génétiques mobiles, tels que les îlots génomiques, doit ainsi être régulée de façon coordonnée avec le reste du chromosome, afin que les facteurs d'adaptation ou de virulence qu'ils codent soient produits au bon endroit et au bon moment, en fonction des stimuli provenant du milieu extérieur [228].

L'expression des îlots génomiques est ainsi influencée par l'hôte au sein duquel ils résident [25]. En effet, si ces îlots génomiques comportent fréquemment un ou plusieurs gènes codant des régulateurs, il existe de nombreuses configurations différentes [25,26,228]. Certains régulateurs sont codés au sein de l'îlot et ne régulent que les gènes spécifiques de cet îlot. D'autres systèmes de régulation sont codés par l'îlot mais régulent également des gènes situés en dehors de l'îlot, appartenant à d'autres éléments génétiques mobiles ou au *core genome*. Enfin, l'expression des gènes codés au sein d'un îlot peut être modulée par des régulateurs codés par le *core genome* ou par d'autres éléments génétiques mobiles. Cette régulation est parfois encore plus complexe dans le cas des cascades de régulation, où des régulateurs directs des gènes d'un îlot génomique sont codés au sein de ce même îlot, mais sont eux-mêmes régulés par des systèmes de régulation globaux codés hors de cet îlot [25]. L'îlot génomique *pks* illustre parfaitement cette diversité de mécanismes de régulation.

Classiquement, les systèmes de régulation codés par les îlots génomiques appartiennent à la famille des régulateurs de type AraC [25,229], comme par exemple le régulateur YbtA, codé par le gène *ybtA* présent au sein de l'îlot HPI responsable de la production du sidérophore yersiniabactine et retrouvé chez de nombreuses entérobactéries [230,231]. L'îlot génomique

pks semble comporter en son sein un système de régulation similaire. Le gène *clbR* présent dans l'îlot code en effet une protéine comportant un motif de liaison à l'ADN de type hélice-tour-hélice [2]. ClbR présente certaines homologies avec les régulateurs de la famille LuxR, qui sont impliqués dans la régulation des mécanismes de communication inter-cellulaire, encore appelée détection du quorum (ou *quorum sensing*) [232]. Les mécanismes de régulation assurés par ClbR n'ont pour le moment pas été démontrés.

Dans d'autres cas, un régulateur codé par un îlot peut intervenir dans la modulation de systèmes codés par d'autres îlots génomiques ou au sein du *core genome*. Nous avons par exemple récemment mis en évidence chez *E. coli* une interchangeabilité entre les PPTases, qui assurent l'activation des enzymes de biosynthèse de type PKS et/ou NRPS par modification post-traductionnelle (**cf. Article 3 en annexe**) [92]. La PPTase ClbA, codée par l'îlot *pks* et activant la machinerie enzymatique produisant la colibactine, est également capable d'activer celle de la yersiniabactine codée par l'îlot HPI, et celle de l'entérobactine codée par le *core genome*. Parallèlement, la PPTase EntD codée par l'îlot génomique de l'entérobactine, permet en outre d'activer les enzymes nécessaires à la production de yersiniabactine.

Les îlots génomiques sont par ailleurs fréquemment régulés par des systèmes de régulation à deux composants (ou TCS, pour *two-component system*) codés par le *core genome* [25,229]. Par exemple, les îlots génomiques SPI-1 et SPI-2 retrouvés chez *Salmonella enterica* codent des gènes impliqués dans le processus d'invasion en réponse à des stimuli de l'environnement au cours du processus infectieux chez l'hôte [229]. Ces signaux extérieurs sont captés par plusieurs TCS (BarA/SirA, PhoP/PhoQ, ...), codés en dehors des îlots SPI-1 et SPI-2, qui lorsqu'ils sont activés régulent positivement ou négativement l'expression des gènes de ces deux îlots génomiques [25,229]. Pour le moment, aucune donnée démontrant l'implication d'un TCS dans la régulation de l'îlot *pks* n'a été publiée.

Les autres systèmes de régulation des îlots génomiques codés dans ou hors des îlots classiquement retrouvés incluent les protéines de la famille des H-NS (pour *histone-like nucleoid structure*) ou encore les facteurs sigma alternatifs. Un exemple bien caractérisé impliquant ces deux types de régulateurs est le locus d'effacement des entérocytes (LEE), îlot génomique présent chez les souches d'EPEC ou d'EHEC [25,229]. Le LEE code l'intimine

(protéine d'adhésion aux entérocytes), le système de sécrétion de type III (et les chaperons associés) et des protéines effectrices (dont le récepteur de l'intimine Tir), et est responsable des lésions histopathologiques d'attachement et d'effacement au niveau des entérocytes [1]. Le LEE code également un régulateur transcriptionnel de la famille H-NS, appelé Ler, qui contrôle l'expression des gènes codant le système de sécrétion de type III et le récepteur de l'intimine Tir [25]. Ler contrôle également l'expression de gènes situés hors du LEE, et est lui-même régulé par le *quorum sensing* [25]. Enfin, la transcription des gènes du LEE est en outre régulé par le facteur sigma alternatif RpoS (ou σ^{38}), régulateur central de la réponse au stress induit par divers stimuli environnementaux [25]. À ce jour, aucun mécanisme de régulation de l'îlot *pks* impliquant les protéines H-NS ou les facteurs sigma n'a été identifié.

L'îlot *pks* semble cependant être régulé par d'autres systèmes codés par le chromosome. Grâce à ce travail de thèse, nous avons en effet identifié deux mécanismes potentiels de régulation de l'îlot *pks* codés par le *core genome* de *E. coli*. Nous avons apporté la démonstration de l'implication de la protéine chaperon Hsp90_{Ec} et de la protéase ClpQ dans la modulation de la machinerie enzymatique produisant la colibactine codée par l'îlot *pks* (**cf. Article 1**) [227]. Si le mécanisme n'a pas été totalement élucidé pour le moment, les résultats suggèrent que l'intervention respective de Hsp90_{Ec} et ClpQ constituerait un mécanisme de régulation post-traductionnelle efficace de la production de colibactine. Nous avons également mis en évidence le rôle de la spermidine dans la production de colibactine, qui semble intervenir dans la régulation de l'expression des gènes de l'îlot *pks* (**cf. Article 2**) [Garcie C. *et al.*, en préparation]. Les résultats préliminaires de cette étude doivent être complétés afin de déterminer si la spermidine régule directement ou indirectement l'expression des gènes de l'îlot, et si ce mécanisme intervient au niveau transcriptionnel ou traductionnel.

Enfin, il a été très récemment démontré que la disponibilité en fer dans le milieu extracellulaire module l'expression du gène *clbA* ainsi que la production de colibactine, *via* l'intervention de deux régulateurs de l'homéostasie du fer, le régulateur protéique global Fur (pour *ferric uptake regulatory protein*) et le petit ARN (ou sRNA, pour *small RNA*) non-codant RyhB (**cf. Article 5 en annexe**) [233]. Un autre mécanisme de régulation indépendant de Fur et RyhB semble être également impliqué en conditions d'excès de fer [Tronnet S. *et al.*, soumis dans *Journal of Bacteriology*].

Ces données, synthétisées dans un schéma-bilan ([Figure 36](#)), indiquent que la régulation de l'îlot *pks* est complexe et fait intervenir de multiples systèmes régulateurs, codés dans ou hors de l'îlot, à l'image d'autres îlots génomiques. Ces résultats sont parfois préliminaires, et nécessitent des investigations complémentaires pour être confirmés, ou pour élucider plus en détail le mécanisme impliqué. Cependant, l'intégration de ces informations suggèrent que la production de la colibactine peut être potentiellement modulée par différents signaux : une situation de stress avec Hsp90_{Ec} et ClpQ, le niveau de polyamines, ou encore la disponibilité en fer dépendante ou non des régulateurs Fur et RyhB.

UN ÎLOT *pks*, MAIS DES COLIBACTINES

À ce jour, la structure complète de la colibactine est inconnue, mais plusieurs études publiées ces trois dernières années ont identifié de nombreux intermédiaires de synthèse [82–91]. L'une d'elle a consisté en une analyse métabolomique comparative de deux souches de *E. coli* portant l'îlot *pks*, la souche K-12 DH10B *pks*⁺ (portant l'îlot *pks* issue de la souche d'ExPEC IHE3034 sur un chromosome bactérien artificiel) mutée ou non pour *clbP*, et la souche probiotique Nissle 1917 (EcN) mutée ou non pour *clbP* [84]. Les auteurs ont ainsi identifié un répertoire de 101 métabolites dérivant de la voie de biosynthèse de la colibactine. De façon surprenante, ce répertoire est très différent en fonction de la souche, puisque par exemple seulement dix métabolites sont communs aux souches DH10B *pks*⁺ et EcN [84]. De plus, les derniers travaux publiés sur l'identification des intermédiaires métaboliques issus de la machinerie enzymatique codée par l'îlot *pks* suggèrent l'existence d'au moins deux voies de synthèse alternatives, en relation avec l'activité de la thioestérase non conventionnelle ClbQ, qui produisent des pré-colibactines aux structures radicalement différentes [91]. L'ensemble de ces données suggère que l'îlot *pks* ne synthétiserait pas un unique composé, mais plutôt un ensemble de molécules.

Nous avons récemment obtenu des résultats préliminaires suggérant que l'activité de l'îlot *pks* pourrait varier en fonction du contexte génétique. D'une part, la quantité de *N*-myristoyl-D-asparagine, produit du clivage de la pré-colibactine par ClbP, qui correspond également à l'intermédiaire de synthèse issue de l'activité de ClbN, diffère d'une souche à l'autre [Garcie C. *et al.*, résultats non publiés]. D'autre part, la quantification de l'activité génotoxique est également très variable selon la souche de *E. coli* portant l'îlot *pks* [Taieb F. *et al.*, résultats non publiés]. Nous proposons l'hypothèse que les systèmes de régulation pourraient être différents, ou différemment exprimés, en fonction du contexte génétique, et que la production de métabolites issus de la machinerie enzymatique codée par l'îlot *pks* serait qualitativement ou quantitativement différente selon la souche.

Par ailleurs, une récente étude vient de démontrer que l'îlot *pks* produisait une ou plusieurs molécules à activité antibiotique vis-à-vis de souches de *Staphylococcus aureus*, y compris des résistantes à la méticilline ou multirésistantes [234]. Ces résultats renforcent les données ayant montré une faible activité antibiotique du produit de clivage sur *Bacillus subtilis* [84]. Nous pouvons envisager que la perception, ou non, des stimuli extérieurs spécifiques des différents systèmes impliqués dans la régulation de l'îlot *pks* pourrait avoir des conséquences sur la production relative des différents métabolites issus de cette voie de biosynthèse.

Ainsi, l'îlot *pks* génomique code un machinerie enzymatique capable de synthétiser un répertoire de molécules aux fonctions variées, et est régulé par de multiples systèmes de régulation codés dans ou hors de l'îlot. Les différents régulateurs et les métabolites associés à l'îlot *pks* discutés plus haut sont synthétisés dans le schéma-bilan ci-dessous (Figure 36).

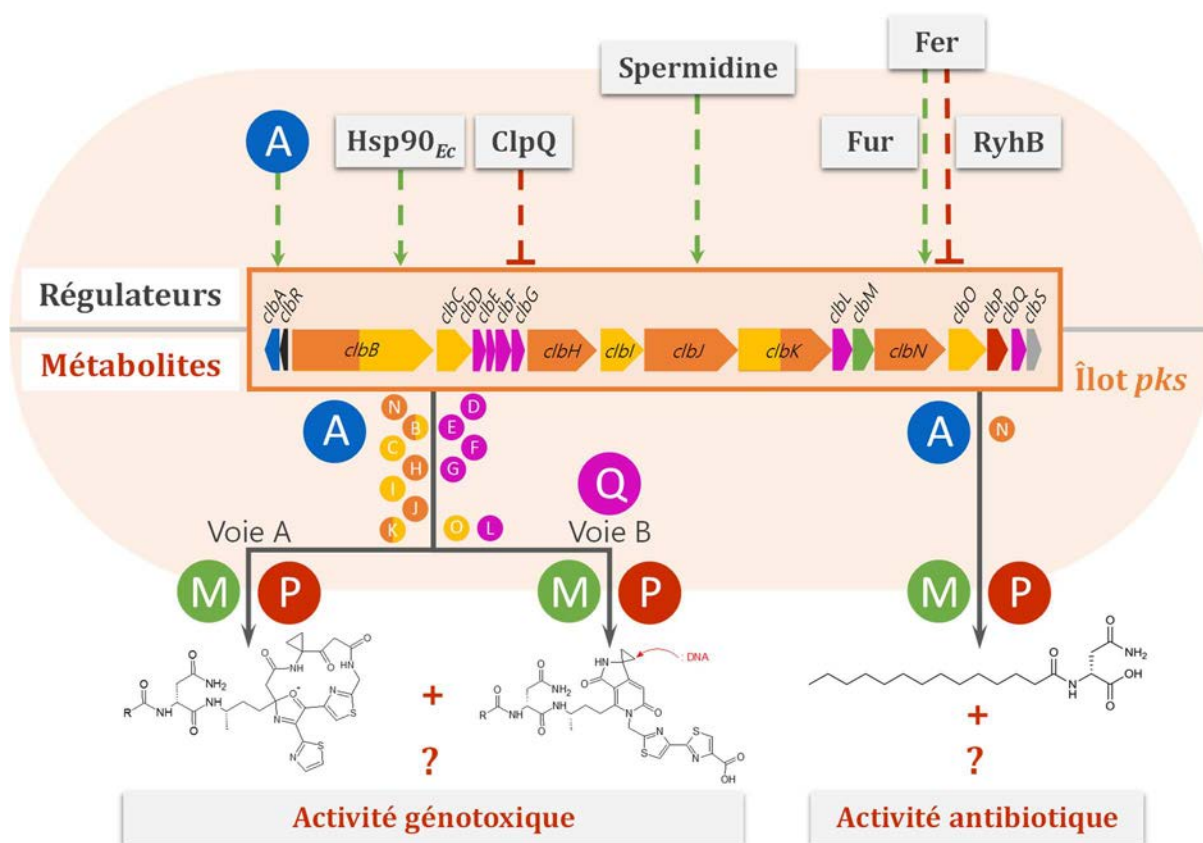


Figure 36. Schéma-bilan : régulateurs et métabolites associés à l'îlot génomique *pks*

● PPTase ; ● PKS ; ● NRPS ; ● enzyme accessoire ; ● pompe à efflux ; ● peptidase ;
 - - - - - régulation positive ; - - - - - régulation négative.

ÉVOLUTION DE DEUX ÎLOTS INTERCONNECTÉS

Ce travail de thèse a permis de démontrer l'implication de la protéine chaperon Hsp90_{Ec} et de la protéase ClpQ dans la modulation de deux voies de biosynthèse permettant la production de métabolites secondaires à fonctions distinctes chez une même bactérie : la colibactine, une génotoxine, et la yersiniabactine, un sidérophore [227]. Cette interconnexion s'ajoute à celle précédemment identifiée dans notre équipe, démontrant que la PPTase ClbA, codée par l'îlot *pks* et principalement impliquée dans la synthèse de colibactine, est capable d'activer les machineries enzymatiques produisant la yersiniabactine, l'entérobactine, et de fait les salmochélines [92]. Il a de plus été très récemment établi que la disponibilité en fer dans le milieu extracellulaire module, outre l'expression des gènes codant la yersiniabactine, l'expression du gène *clbA* ainsi que la production de colibactine, *via* l'intervention des régulateurs de l'homéostasie du fer [233]. Les différentes connexions entre les îlots *pks* et HPI sont synthétisées dans la [Figure 37](#).

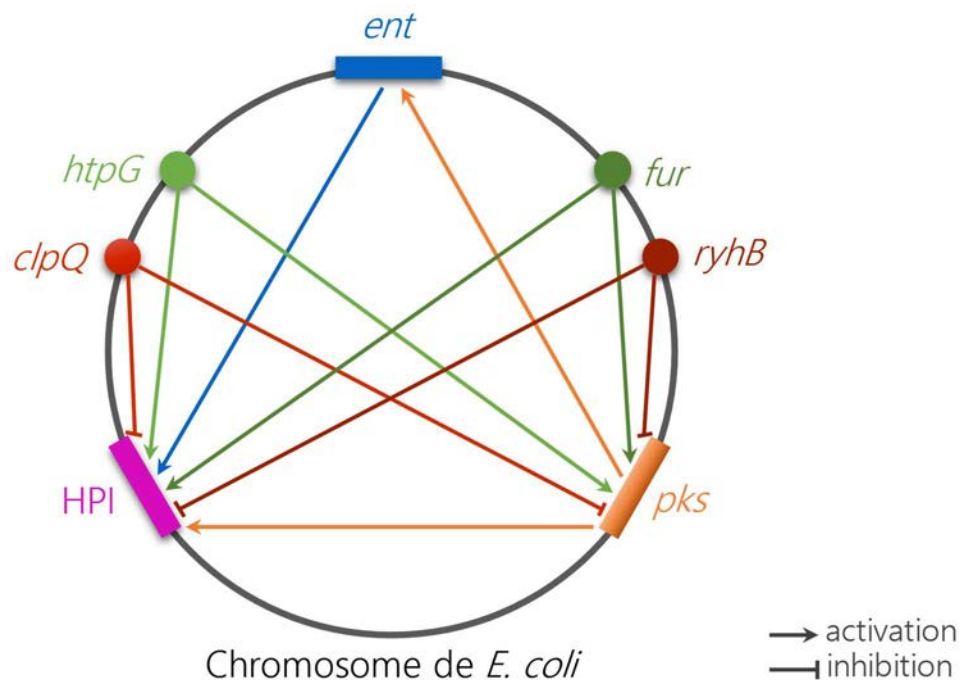


Figure 37. Interconnexions entre les îlots génomiques *pks* et HPI

De façon intéressante, ces interconnexions sont observées entre deux îlots génomiques très conservés et fréquemment associés au sein du groupe phylogénétique B2. Lorsqu'il est présent dans une souche de *E. coli* B2, l'îlot *pks* est en effet systématiquement associé et inséré à proximité de l'îlot génomique HPI codant la yersiniabactine [74]. Ces données suggèrent l'émergence d'un sous-groupe distinct au sein du phylogroupe B2. Cette hypothèse est renforcée par de récentes études épidémiologiques qui ont identifié au sein des souches d'ExPEC une lignée clonale présentant un pouvoir pathogène important [235]. Nous pensons que ce clone, à la fois virulent mais aussi bon colonisateur du tractus digestif de l'Homme, résulterait de l'acquisition séquentielle et sélective de deux îlots génomiques interconnectés. Nous faisons donc l'hypothèse que l'association des îlots *pks* et HPI s'est maintenue dans les souches hypervirulentes du fait des multiples interconnexions entre ces deux îlots : la PPTase ClbA, la régulation par Hsp90_{Ec} et ClpQ, et la régulation par l'homéostasie du fer.

L'acquisition séquentielle de facteurs d'adaptation et/ou de virulence, portés par différents éléments génétiques mobiles, ayant conduit au cours de l'évolution à l'apparition d'un clone hyper-virulent et parfaitement adapté à la colonisation de l'hôte, a précédemment été décrite chez *E. coli*. Par exemple, la souche hyper-virulente d'EHEC de sérotype rare O104:H4, responsable de l'épidémie de SHU survenue en Allemagne en 2011, possédait à la fois le gène *stx2* codant la Shiga-toxine responsable de SHU et caractéristique des STEC (pour *Shiga toxin-producing E. coli*), mais également des gènes caractéristiques des EAEC, formant ainsi un pathotype hybride STEC/EAEC [236]. Plus récemment, une étude a décrit l'émergence en France du nouveau sérotype d'EHEC hypervirulent O80:H2, présentant les caractéristiques génétiques d'un pathotype hybride STEC/ExPEC et pouvant induire un SHU associé à une bactériémie [237]. Ce clone possède à la fois le gène *stx2* codé sur un prophage, le gène *eae-ξ* contenant le LEE, mais aussi le gène *iroN* codant la synthèse des salmochélines ainsi qu'un plasmide similaire au plasmide pS88, classiquement retrouvés chez les NMEC.

PERSPECTIVES

Si ce travail de thèse a permis de lever le voile sur certains acteurs impliqués dans la régulation de la production de la colibactine par *E. coli*, il soulève également de nombreuses questions. Nous proposons ici quelques pistes potentielles à explorer pour poursuivre ce projet.

Le premier axe de ce travail a démontré que la protéine chaperon Hsp90_{Ec} est impliquée dans la biosynthèse de deux métabolites secondaires produits par *E. coli*, et agit de concert avec la protéase ClpQ [227]. Le mécanisme d'action moléculaire et les partenaires potentiels dans la biosynthèse de colibactine et de yersiniabactine n'ont cependant pas été identifiés.

Dans le but d'élucider le mécanisme d'action de Hsp90_{Ec} et ClpQ, il est indispensable d'identifier les protéines clientes dans chaque voie de biosynthèse. En analysant *in silico* la liste des substrats potentiels proposés par les deux études d'interactome de Hsp90_{Ec} [166,167], nous avons identifié un candidat potentiel : EntD, la PPTase impliquée dans la biosynthèse d'entérobactine, de salmocheline et de yersiniabactine, trois sidérophores de *E. coli* [238]. La présence de EntD dans l'interactome de Hsp90_{Ec} [167] suggère l'hypothèse que cette protéine chaperon pourrait participer au repliement conformationnel des PPTases, enzymes indispensables à la synthèse des PK-NRP. Afin de vérifier cette hypothèse, des expériences de co-immunoprécipitation entre la protéine chaperon Hsp90_{Ec} et les deux PPTases de *E. coli*, EntD et ClbA, pourraient être envisagées.

L'implication de Hsp90_{Ec} dans la modulation de deux voies de biosynthèse permettant la production de métabolites secondaires à fonctions distinctes chez une même bactérie, ainsi que la démonstration de son rôle dans la virulence de souches d'ExPEC, constituent deux arguments pour considérer cette protéine chaperon comme cible thérapeutique potentielle, à l'image de ce qui est décrit pour ses analogues eucaryotes dans le traitement de certains cancers [195], de protozooses [198] ou encore d'infections fongiques [199]. Ces arguments sont en outre renforcés par la démonstration de l'implication de HtpG, l'analogue procaryote de Hsp90 dans

la virulence de plusieurs espèces bactériennes [185,188–190]. De nombreux inhibiteurs de l'activité ATPasique de Hsp90 ont été développés depuis la découverte de la geldamycine, la novobiocine ou le radicicol [194]. Les essais cliniques menés actuellement utilisent des molécules dérivées, qui sont mieux tolérées et présentent une meilleure efficacité et des modalités d'administration facilitées [194].

La principale difficulté rencontrée par cette stratégie repose sur l'homologie importante entre les différents analogues de la protéine Hsp90. Avec 40% d'identité de séquences entre Hsp90 humaine et procaryote [149], il est particulièrement difficile de développer des inhibiteurs spécifiques d'un règne du vivant, d'autant plus que la cible d'action moléculaire se situe dans la région très conservée du site ATPasique de la protéine chaperon [197]. Le développement de molécules anti-Hsp90 non spécifiques comme traitement antibiotique, agissant indifféremment sur cellules eucaryotes et procaryotes, et potentiellement responsables d'effets indésirables, n'est pas envisageable à une époque où les stratégies thérapeutiques deviennent ciblées, voire personnalisées [239]. Il est donc indispensable d'approfondir nos connaissances structurales, mécanistiques et fonctionnelles concernant l'analogue procaryote de Hsp90, pour développer une stratégie thérapeutique robuste.

Dans cette optique, une récente étude a présenté les résultats d'un criblage de molécules se liant spécifiquement sur HtpG [240]. Minagawa *et al.* ont notamment mis en évidence que des dérivés de la colistine et de la polymyxine B, deux antibiotiques utilisés en médecine, sont capables de se fixer sur le domaine N-terminal de HtpG, au niveau d'une région distincte du site ATPasique, et inhibent l'activité chaperon [240]. Les données présentées semblent en faveur d'une inhibition spécifique de HtpG, mais doivent être confirmées.

HtpG constitue ainsi un candidat potentiel comme cible thérapeutique de médicaments à action anti-virulence, utilisables en association avec les antibiotiques classiques disponibles. Le développement de stratégies thérapeutiques adjuvantes, telles que les médicaments anti-résistance, anti-virulence ou modulateurs de la réponse de l'hôte, mais aussi la phagothérapie ou les probiotiques, à utiliser conjointement aux antibiotiques traditionnels, sera crucial pour espérer poursuivre la lutte contre les infections bactériennes dans l'ère post-antibiotique [241].

Le deuxième axe de ce travail de thèse a mis en évidence le rôle de la spermidine dans la production de colibactine, qui semble intervenir dans la régulation de l'expression des gènes de l'îlot *pks* [Garcie *et al.*, en préparation]. Le mécanisme de régulation et les gènes cibles n'ont cependant pas été identifiés pour le moment.

Afin de caractériser les cibles de la spermidine, ainsi que son mécanisme d'action, plusieurs stratégies sont envisagées. Les expériences de quantification de l'expression des gènes *clbA*, *clbB*, *clbQ* et *clbR* grâce au système rapporteur de la luciférase doivent être répétées en testant une supplémentation avec une gamme de concentration de spermidine afin de confirmer la régulation différentielle du gène *clbA* par rapport aux gènes *clbQ* et *clbR* par cette polyamine. Par ailleurs, une quantification de l'expression de différents gènes de l'îlot *pks* par RT-PCR quantitative pourrait permettre d'identifier d'autres gènes cibles. Des approches globales comparatives protéomiques et transcriptomiques d'une souche mutée pour *speE* et supplémenteée ou non en spermidine pourrait également apporter des informations complémentaires sur les cibles de la spermidine et les mécanismes de régulation impliqués.

Les résultats obtenus n'excluent pas pour autant que la spermidine soit incorporée dans la structure de la colibactine, potentialisant son activité génotoxique, tout en exerçant un rôle régulateur dans l'expression des gènes de l'îlot *pks*. Cette hypothèse est d'ailleurs explorée en collaboration avec une équipe de chimistes spécialistes de la synthèse des composés PK-NRP (Dr. Alexander Brachmann and Pr. Jörn Piel, Institut de Microbiologie, ETH Zurich, Suisse).

Enfin, l'implication des polyamines dans la virulence a été démontré ces dernières années pour différentes bactéries pathogènes [201,242]. Les polyamines seraient par exemple nécessaires aux mécanismes d'invasion cellulaire chez *Salmonella enterica*, en induisant au moment opportun la transcription des gènes codés au sein des îlots génomique SPI-1 et SPI-2 [243–245]. Il serait donc intéressant d'étudier la relevance biologique de l'implication de la spermidine dans la régulation de l'îlot *pks* dans différents modèles animaux. Des souches d'ExPEC mutées ou non pour *speE* seront testées dans un modèle de méningite néonatale par administration orale à des rats nouveau-nés. L'impact de la colonisation à long terme d'animaux axéniques par une souche commensale mutées pour *speE*, en modulant le contenu du tube digestif en polyamines, pourraient également être évalué.

Grâce à ce travail de thèse, nous avons identifié deux mécanismes potentiels de régulation de l'îlot *pks* codés par le *core genome* de *E. coli*. Nous avons d'une part apporté la démonstration de l'implication de la protéine chaperon Hsp90_{Ec} et de la protéase ClpQ dans la modulation de la machinerie enzymatique produisant la colibactine, et d'autre part mis en évidence le rôle de la spermidine dans la biosynthèse de la génotoxine, probablement en régulant l'expression de certains gènes de l'îlot.

Nous avons de plus caractérisé une interconnexion supplémentaire entre les îlots génomiques *pks* et HPI, qui sont responsables de la synthèse de deux facteurs de virulence de *E. coli* appartenant à la même famille chimique des composés PK-NRP. La mise en évidence de différents niveaux de connexions dans la régulation de la production de ces métabolites secondaires est essentielle pour la compréhension de la pathogénicité de *E. coli*, mais également pour l'identification de nouvelles cibles thérapeutiques et le développement de stratégies anti-infectieuses innovantes.

Les résultats de cette thèse, en identifiant plusieurs connexions bilatérales entre l'îlot génomique *pks* et le reste du génome de *E. coli*, illustrent l'intégration symbiotique d'un élément génétique mobile acquis par transfert horizontal au sein du chromosome bactérien.

Enfin, la banque de mutants pour la production de colibactine, qui a servi de base à ce travail, constitue désormais une véritable mine d'informations à exploiter. La validation des autres gènes candidats potentiels identifiés ouvre de nombreuses perspectives pour approfondir nos connaissances concernant la biosynthèse, la structure, l'export et les mécanismes de régulation de la production de cette insaisissable génotoxine qu'est la colibactine.

RÉFÉRENCES BIBLIOGRAPHIQUES

1. Kaper JB, Nataro JP, Mobley HL. Pathogenic *Escherichia coli*. Nat Rev Microbiol. **2004**; 2(2):123–140.
2. Nougayrède J-P, Homburg S, Taieb F, *et al.* *Escherichia coli* induces DNA double-strand breaks in eukaryotic cells. Science. **2006**; 313(5788):848–851.
3. Cuevas-Ramos G, Petit CR, Marcq I, Boury M, Oswald E, Nougayrède J-P. *Escherichia coli* induces DNA damage *in vivo* and triggers genomic instability in mammalian cells. Proc Natl Acad Sci U S A. **2010**; 107(25):11537–11542.
4. Shulman ST, Friedmann HC, Sims RH. Theodor Escherich: the first pediatric infectious diseases physician? Clin Infect Dis Off Publ Infect Dis Soc Am. **2007**; 45(8):1025–1029.
5. Tenaillon O, Skurnik D, Picard B, Denamur E. The population genetics of commensal *Escherichia coli*. Nat Rev Microbiol. **2010**; 8(3):207–217.
6. van Elsas JD, Semenov AV, Costa R, Trevors JT. Survival of *Escherichia coli* in the environment: fundamental and public health aspects. ISME J. **2011**; 5(2):173–183.
7. Dobrindt U. (Patho-)Genomics of *Escherichia coli*. Int J Med Microbiol. **2005**; 295(6–7):357–371.
8. Croxen MA, Finlay BB. Molecular mechanisms of *Escherichia coli* pathogenicity. Nat Rev Microbiol. **2010**; 8(1):26–38.
9. Orskov F, Orskov I, Evans DJ, Sack RB, Sack DA, Wadström T. Special *Escherichia coli* serotypes among enterotoxigenic strains from diarrhoea in adults and children. Med Microbiol Immunol (Berl). **1976**; 162(2):73–80.
10. Ochman H, Selander RK. Standard reference strains of *Escherichia coli* from natural populations. J Bacteriol. **1984**; 157(2):690–693.
11. Moissenet D, Salauze B, Clermont O, *et al.* Meningitis caused by *Escherichia coli* producing TEM-52 extended-spectrum beta-lactamase within an extensive outbreak in a neonatal ward: epidemiological investigation and characterization of the strain. J Clin Microbiol. **2010**; 48(7):2459–2463.
12. Jaureguy F, Landraud L, Passet V, *et al.* Phylogenetic and genomic diversity of human bacteremic *Escherichia coli* strains. BMC Genomics. **2008**; 9:560.
13. Escobar-Páramo P, Grenet K, Le Menac’h A, *et al.* Large-scale population structure of human commensal *Escherichia coli* isolates. Appl Environ Microbiol. **2004**; 70(9):5698–5700.
14. Touchon M, Hoede C, Tenaillon O, *et al.* Organised genome dynamics in the *Escherichia coli* species results in highly diverse adaptive paths. PLoS Genet. **2009**; 5(1):e1000344.
15. Medini D, Donati C, Tettelin H, Massignani V, Rappuoli R. The microbial pan-genome. Curr Opin Genet Dev. **2005**; 15(6):589–594.
16. Lan R, Reeves PR. Intraspecies variation in bacterial genomes: the need for a species genome concept. Trends Microbiol. **2000**; 8(9):396–401.

17. Tettelin H, Riley D, Cattuto C, Medini D. Comparative genomics: the bacterial pan-genome. *Curr Opin Microbiol.* **2008**; 11(5):472–477.
18. Lukjancenko O, Wassenaar TM, Ussery DW. Comparison of 61 Sequenced *Escherichia coli* Genomes. *Microb Ecol.* **2010**; 60(4):708–720.
19. Kaas RS, Friis C, Ussery DW, Aarestrup FM. Estimating variation within the genes and inferring the phylogeny of 186 sequenced diverse *Escherichia coli* genomes. *BMC Genomics.* **2012**; 13:577.
20. Hendrickson H. Order and disorder during *Escherichia coli* divergence. *PLoS Genet.* **2009**; 5(1):e1000335.
21. Juhas M. Horizontal gene transfer in human pathogens. *Crit Rev Microbiol.* **2015**; 41(1):101–108.
22. Gyles C, Boerlin P. Horizontally Transferred Genetic Elements and Their Role in Pathogenesis of Bacterial Disease. *Vet Pathol Online.* **2014**; 51(2):328–340.
23. Popa O, Dagan T. Trends and barriers to lateral gene transfer in prokaryotes. *Curr Opin Microbiol.* **2011**; 14(5):615–623.
24. Stecher B, Maier L, Hardt W-D. “Blooming” in the gut: how dysbiosis might contribute to pathogen evolution. *Nat Rev Microbiol.* **2013**; 11(4):277–284.
25. Hacker J, Kaper JB. Pathogenicity Islands and the Evolution of Microbes. *Annu Rev Microbiol.* **2000**; 54(1):641–679.
26. Hentschel U, Hacker J. Pathogenicity islands: the tip of the iceberg. *Microbes Infect.* **2001**; 3(7):545–548.
27. Dobrindt U, Hochhut B, Hentschel U, Hacker J. Genomic islands in pathogenic and environmental microorganisms. *Nat Rev Microbiol.* **2004**; 2(5):414–424.
28. Malik VS. Microbial secondary metabolism. *Trends Biochem Sci.* **1980**; 5(3):68–72.
29. Clardy J, Walsh C. Lessons from natural molecules. *Nature.* **2004**; 432(7019):829–837.
30. Nolan EM, Walsh CT. How nature morphs peptide scaffolds into antibiotics. *Chembiochem Eur J Chem Biol.* **2009**; 10(1):34–53.
31. Li JW-H, Vederas JC. Drug discovery and natural products: end of an era or an endless frontier? *Science.* **2009**; 325(5937):161–165.
32. Harvey AL. Natural products in drug discovery. *Drug Discov Today.* **2008**; 13(19–20):894–901.
33. Walsh CT, Fischbach MA. Natural Products Version 2.0: Connecting Genes to Molecules. *J Am Chem Soc.* **2010**; 132(8):2469–2493.
34. Crosa JH, Walsh CT. Genetics and assembly line enzymology of siderophore biosynthesis in bacteria. *Microbiol Mol Biol Rev.* **2002**; 66(2):223–249.

35. Fischbach MA, Walsh CT. Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: logic, machinery, and mechanisms. *Chem Rev.* **2006**; 106(8):3468–3496.
36. Donadio S, Monciardini P, Sosio M. Polyketide synthases and nonribosomal peptide synthetases: the emerging view from bacterial genomics. *Nat Prod Rep.* **2007**; 24(5):1073–1109.
37. Finking R, Marahiel MA. Biosynthesis of nonribosomal peptides. *Annu Rev Microbiol.* **2004**; 58:453–488.
38. Lambalot RH, Gehring AM, Flugel RS, *et al.* A new enzyme superfamily - the phosphopantetheinyl transferases. *Chem Biol.* **1996**; 3(11):923–936.
39. Beld J, Sonnenschein EC, Vickery CR, Noel JP, Burkart MD. The phosphopantetheinyl transferases: catalysis of a post-translational modification crucial for life. *Nat Prod Rep.* **2013**; 31(1):61–108.
40. Meier JL, Burkart MD. The chemical biology of modular biosynthetic enzymes. *Chem Soc Rev.* **2009**; 38(7):2012–2045.
41. Moore BS, Hertweck C. Biosynthesis and attachment of novel bacterial polyketide synthase starter units. *Nat Prod Rep.* **2002**; 19(1):70–99.
42. Schwarzer D, Mootz HD, Linne U, Marahiel MA. Regeneration of misprimed nonribosomal peptide synthetases by type II thioesterases. *Proc Natl Acad Sci U S A.* **2002**; 99(22):14083–14088.
43. Kleinkauf H, Dittmann J, Lawen A. Cell-free biosynthesis of cyclosporin A and analogues. *Biomed Biochim Acta.* **1991**; 50(10–11):S219–224.
44. Crichton RR, Wilmet S, Leggsyer R, Ward RJ. Molecular and cellular mechanisms of iron homeostasis and toxicity in mammalian cells. *J Inorg Biochem.* **2002**; 91(1):9–18.
45. Andrews SC, Robinson AK, Rodríguez-Quiñones F. Bacterial iron homeostasis. *FEMS Microbiol Rev.* **2003**; 27(2–3):215–237.
46. Skaar EP. The battle for iron between bacterial pathogens and their vertebrate hosts. *PLoS Pathog.* **2010**; 6(8):e1000949.
47. Schröder I, Johnson E, de Vries S. Microbial ferric iron reductases. *FEMS Microbiol Rev.* **2003**; 27(2–3):427–447.
48. Miethke M, Marahiel MA. Siderophore-Based Iron Acquisition and Pathogen Control. *Microbiol Mol Biol Rev.* **2007**; 71(3):413–451.
49. Tong Y, Guo M. Bacterial heme-transport proteins and their heme-coordination modes. *Arch Biochem Biophys.* **2009**; 481(1):1–15.
50. Braun V. Iron uptake mechanisms and their regulation in pathogenic bacteria. *Int J Med Microbiol.* **2001**; 291(2):67–79.

51. Garénaux A, Caza M, Dozois CM. The Ins and Outs of siderophore mediated iron uptake by extra-intestinal pathogenic *Escherichia coli*. *Vet Microbiol.* **2011**; 153(1–2):89–98.
52. Ratledge C, Dover LG. Iron metabolism in pathogenic bacteria. *Annu Rev Microbiol.* **2000**; 54:881–941.
53. McHugh JP, Rodríguez-Quinoñes F, Abdul-Tehrani H, *et al.* Global iron-dependent gene regulation in *Escherichia coli*. A new mechanism for iron homeostasis. *J Biol Chem.* **2003**; 278(32):29478–29486.
54. Higgs PI, Larsen RA, Postle K. Quantification of known components of the *Escherichia coli* TonB energy transduction system: TonB, ExbB, ExbD and FepA. *Mol Microbiol.* **2002**; 44(1):271–281.
55. Dozois CM, Daigle F, Curtiss R. Identification of pathogen-specific and conserved genes expressed *in vivo* by an avian pathogenic *Escherichia coli* strain. *Proc Natl Acad Sci U S A.* **2003**; 100(1):247–252.
56. Fischbach MA, Lin H, Zhou L, *et al.* The pathogen-associated *iroA* gene cluster mediates bacterial evasion of lipocalin-2. *Proc Natl Acad Sci U S A.* **2006**; 103(44):16502–16507.
57. Feldmann F, Sorsa LJ, Hildinger K, Schubert S. The salmochelin siderophore receptor IroN contributes to invasion of urothelial cells by extraintestinal pathogenic *Escherichia coli* *in vitro*. *Infect Immun.* **2007**; 75(6):3183–3187.
58. Russo TA, McFadden CD, Carlino-MacDonald UB, Beanan JM, Barnard TJ, Johnson JR. IroN functions as a siderophore receptor and is a urovirulence factor in an extraintestinal pathogenic isolate of *Escherichia coli*. *Infect Immun.* **2002**; 70(12):7156–7160.
59. Nègre VL, Bonacorsi S, Schubert S, Bidet P, Nassif X, Bingen E. The siderophore receptor IroN, but not the high-pathogenicity island or the hemin receptor ChuA, contributes to the bacteremic step of *Escherichia coli* neonatal meningitis. *Infect Immun.* **2004**; 72(2):1216–1220.
60. Peigne C, Bidet P, Mahjoub-Messai F, *et al.* The plasmid of *Escherichia coli* strain S88 (O45:K1:H7) that causes neonatal meningitis is closely related to avian pathogenic *E. coli* plasmids and is associated with high-level bacteremia in a neonatal rat meningitis model. *Infect Immun.* **2009**; 77(6):2272–2284.
61. Schubert S, Rakin A, Karch H, Carniel E, Heesemann J. Prevalence of the “high-pathogenicity island” of *Yersinia* species among *Escherichia coli* strains that are pathogenic to humans. *Infect Immun.* **1998**; 66(2):480–485.
62. Bach S, de Almeida A, Carniel E. The *Yersinia* high-pathogenicity island is present in different members of the family *Enterobacteriaceae*. *FEMS Microbiol Lett.* **2000**; 183(2):289–294.
63. Garcia EC, Brumbaugh AR, Mobley HLT. Redundancy and specificity of *Escherichia coli* iron acquisition systems during urinary tract infection. *Infect Immun.* **2011**; 79(3):1225–1235.

64. Hancock V, Ferrières L, Klemm P. The ferric yersiniabactin uptake receptor FyuA is required for efficient biofilm formation by urinary tract infectious *Escherichia coli* in human urine. *Microbiol Read Engl*. **2008**; 154:167–175.
65. Paauw A, Leverstein-van Hall MA, van Kessel KPM, Verhoef J, Fluit AC. Yersiniabactin reduces the respiratory oxidative stress response of innate immune cells. *PLoS One*. **2009**; 4(12):e8240.
66. Johnson TJ, Siek KE, Johnson SJ, Nolan LK. DNA sequence of a ColV plasmid and prevalence of selected plasmid-encoded virulence genes among avian *Escherichia coli* strains. *J Bacteriol*. **2006**; 188(2):745–758.
67. Valdebenito M, Crumbliss AL, Winkelmann G, Hantke K. Environmental factors influence the production of enterobactin, salmochelin, aerobactin, and yersiniabactin in *Escherichia coli* strain Nissle 1917. *Int J Med Microbiol*. **2006**; 296(8):513–520.
68. Caza M, Lépine F, Dozois CM. Secretion, but not overall synthesis, of catecholate siderophores contributes to virulence of extraintestinal pathogenic *Escherichia coli*. *Mol Microbiol*. **2011**; 80(1):266–282.
69. Taieb F, Petit C, Nougayrède J-P, Oswald E. The Enterobacterial Genotoxins: Cytolethal Distending Toxin and Colibactin. *EcoSal Plus*. **2016**; 7(1):1–21.
70. Mousa JJ, Yang Y, Tomkovich S, *et al*. MATE transport of the *E. coli*-derived genotoxin colibactin. *Nat Microbiol*. **2016**; 1(1):15009.
71. Dubois D, Baron O, Cougnoux A, *et al*. ClbP is a prototype of a peptidase subgroup involved in biosynthesis of nonribosomal peptides. *J Biol Chem*. **2011**; 286(41):35562–35570.
72. Bossuet-Greif N, Dubois D, Petit C, *et al*. *Escherichia coli* ClbS is a colibactin resistance protein. *Mol Microbiol*. **2016**; 99(5):897–908.
73. Johnson JR, Johnston B, Kuskowski MA, Nougayrede J-P, Oswald E. Molecular epidemiology and phylogenetic distribution of the *Escherichia coli* *pks* genomic island. *J Clin Microbiol*. **2008**; 46(12):3906–3911.
74. Putze J, Hennequin C, Nougayrède J-P, *et al*. Genetic structure and distribution of the colibactin genomic island among members of the family *Enterobacteriaceae*. *Infect Immun*. **2009**; 77(11):4696–4703.
75. Bondarev V, Richter M, Romano S, Piel J, Schwedt A, Schulz-Vogt HN. The genus *Pseudovibrio* contains metabolically versatile bacteria adapted for symbiosis. *Environ Microbiol*. **2013**; 15(7):2095–2113.
76. Engel P, Vizcaino MI, Crawford JM. Gut symbionts from distinct hosts exhibit genotoxic activity via divergent colibactin biosynthesis pathways. *Appl Environ Microbiol*. **2015**; 81(4):1502–1512.
77. Dubois D, Delmas J, Cady A, *et al*. Cyclomodulins in urosepsis strains of *Escherichia coli*. *J Clin Microbiol*. **2010**; 48(6):2122–2129.

78. Krieger JN, Dobrindt U, Riley DE, Oswald E. Acute *Escherichia coli* prostatitis in previously health young men: bacterial virulence factors, antimicrobial resistance, and clinical outcomes. *Urology*. **2011**; 77(6):1420–1425.
79. Nowrouzian FL, Oswald E. *Escherichia coli* strains with the capacity for long-term persistence in the bowel microbiota carry the potentially genotoxic *pks* island. *Microb Pathog*. **2012**; 53(3–4):180–182.
80. Payros D, Secher T, Boury M, *et al.* Maternally acquired genotoxic *Escherichia coli* alters offspring's intestinal homeostasis. *Gut Microbes*. **2014**; 5(3):313–325.
81. Hacker J, Hentschel U, Dobrindt U. Prokaryotic chromosomes and disease. *Science*. **2003**; 301(5634):790–793.
82. Brotherton CA, Balskus EP. A prodrug resistance mechanism is involved in colibactin biosynthesis and cytotoxicity. *J Am Chem Soc*. **2013**; 135(9):3359–3362.
83. Bian X, Fu J, Plaza A, *et al.* *In vivo* evidence for a prodrug activation mechanism during colibactin maturation. *Chembiochem Eur J Chem Biol*. **2013**; 14(10):1194–1197.
84. Vizcaino MI, Engel P, Trautman E, Crawford JM. Comparative metabolomics and structural characterizations illuminate colibactin pathway-dependent small molecules. *J Am Chem Soc*. **2014**; 136(26):9244–9247.
85. Brotherton CA, Wilson M, Byrd G, Balskus EP. Isolation of a metabolite from the *pks* island provides insights into colibactin biosynthesis and activity. *Org Lett*. **2015**; 17(6):1545–1548.
86. Bian X, Plaza A, Zhang Y, Müller R. Two more pieces of the colibactin genotoxin puzzle from *Escherichia coli* show incorporation of an unusual 1-aminocyclopropanecarboxylic acid moiety. *Chem Sci*. **2015**; 6(5):3154–3160.
87. Vizcaino MI, Crawford JM. The colibactin warhead crosslinks DNA. *Nat Chem*. **2015**; 7(5):411–417.
88. Li Z-R, Li Y, Lai JYH, *et al.* Critical Intermediates Reveal New Biosynthetic Events in the Enigmatic Colibactin Pathway. *Chembiochem Eur J Chem Biol*. **2015**; 16(12):1715–1719.
89. Zha L, Wilson MR, Brotherton CA, Balskus EP. Characterization of Polyketide Synthase Machinery from the *pks* Island Facilitates Isolation of a Candidate Precolibactin. *ACS Chem Biol*. **2016**; 11(5):1287–1295.
90. Healy AR, Vizcaino MI, Crawford JM, Herzon SB. Convergent and Modular Synthesis of Candidate Precolibactins. Structural Revision of Precolibactin A. *J Am Chem Soc*. **2016**; 138(16):5426–5432.
91. Li Z-R, Li J, Gu J-P, *et al.* Divergent biosynthesis yields a cytotoxic aminomalonate-containing precolibactin. *Nat Chem Biol*. **2016**; 12(10):773–775.
92. Martin P, Marcq I, Magistro G, *et al.* Interplay between siderophores and colibactin genotoxin biosynthetic pathways in *Escherichia coli*. *PLoS Pathog*. **2013**; 9(7):e1003437.

93. Brachmann AO, Garcie C, Wu V, *et al.* Colibactin biosynthesis and biological activity depend on the rare aminomalonyl polyketide precursor. *Chem Commun Camb Engl.* **2015**; 51(66):13138–13141.
94. Cougnoux A, Gibold L, Robin F, *et al.* Analysis of Structure–Function Relationships in the Colibactin-Maturing Enzyme ClbP. *J Mol Biol.* **2012**; 424(3–4):203–214.
95. Kevany BM, Rasko DA, Thomas MG. Characterization of the complete zwittermicin A biosynthesis gene cluster from *Bacillus cereus*. *Appl Environ Microbiol.* **2009**; 75(4):1144–1155.
96. Reimer D, Luxenburger E, Brachmann AO, Bode HB. A new type of pyrrolidine biosynthesis is involved in the late steps of xenocoumarin production in *Xenorhabdus nematophila*. *Chembiochem Eur J Chem Biol.* **2009**; 10(12):1997–2001.
97. Tanasova M, Sturla SJ. Chemistry and Biology of Acylfulvenes: Sesquiterpene-Derived Antitumor Agents. *Chem Rev.* **2012**; 112(6):3578–3610.
98. Secher T, Samba-Louaka A, Oswald E, Nougayrède J-P. *Escherichia coli* producing colibactin triggers premature and transmissible senescence in mammalian cells. *PLoS One.* **2013**; 8(10):e77157.
99. Cougnoux A, Dalmaso G, Martinez R, *et al.* Bacterial genotoxin colibactin promotes colon tumour growth by inducing a senescence-associated secretory phenotype. *Gut.* **2014**; 63(12):1932–1942.
100. Arthur JC, Jobin C. The complex interplay between inflammation, the microbiota and colorectal cancer. *Gut Microbes.* **2013**; 4(3):253–258.
101. Arthur JC, Perez-Chanona E, Mühlbauer M, *et al.* Intestinal inflammation targets cancer-inducing activity of the microbiota. *Science.* **2012**; 338(6103):120–123.
102. Marcq I, Martin P, Payros D, *et al.* The genotoxin colibactin exacerbates lymphopenia and decreases survival rate in mice infected with septicemic *Escherichia coli*. *J Infect Dis.* **2014**; 210(2):285–294.
103. McCarthy AJ, Martin P, Cloup E, Stabler RA, Oswald E, Taylor PW. The genotoxin Colibactin is a determinant of virulence in *Escherichia coli* K1 experimental neonatal systemic infection. *Infect Immun.* **2015**; 83(9):3704–3711.
104. Secher T, Payros D, Brehin C, *et al.* Oral tolerance failure upon neonatal gut colonization with *Escherichia coli* producing the genotoxin colibactin. *Infect Immun.* **2015**; 83(6):2420–2429.
105. Olier M, Marcq I, Salvador-Cartier C, *et al.* Genotoxicity of *Escherichia coli* Nissle 1917 strain cannot be dissociated from its probiotic activity. *Gut Microbes.* **2012**; 3(6):501–509.
106. Jahn TR, Radford SE. Folding versus aggregation: polypeptide conformations on competing pathways. *Arch Biochem Biophys.* **2008**; 469(1):100–117.
107. Ellis RJ, Minton AP. Protein aggregation in crowded environments. *Biol Chem.* **2006**; 387(5):485–497.

108. Mogk A, Huber D, Bukau B. Integrating protein homeostasis strategies in prokaryotes. *Cold Spring Harb Perspect Biol.* **2011**; 3(4):a004366.
109. Doyle SM, Genest O, Wickner S. Protein rescue from aggregates by powerful molecular chaperone machines. *Nat Rev Mol Cell Biol.* **2013**; 14(10):617–629.
110. Komar AA. A pause for thought along the co-translational folding pathway. *Trends Biochem Sci.* **2009**; 34(1):16–24.
111. Castanié-Cornet M-P, Bruel N, Genevaux P. Chaperone networking facilitates protein targeting to the bacterial cytoplasmic membrane. *Biochim Biophys Acta.* **2014**; 1843(8):1442–1456.
112. Valent QA, Kendall DA, High S, Kusters R, Oudega B, Luirink J. Early events in preprotein recognition in *E. coli*: interaction of SRP and trigger factor with nascent polypeptides. *EMBO J.* **1995**; 14(22):5494–5505.
113. Hoffmann A, Bukau B, Kramer G. Structure and function of the molecular chaperone Trigger Factor. *Biochim Biophys Acta.* **2010**; 1803(6):650–661.
114. Hartl FU, Hayer-Hartl M. Converging concepts of protein folding *in vitro* and *in vivo*. *Nat Struct Mol Biol.* **2009**; 16(6):574–581.
115. Kaiser CM, Chang H-C, Agashe VR, *et al.* Real-time observation of trigger factor function on translating ribosomes. *Nature.* **2006**; 444(7118):455–460.
116. Deuerling E, Schulze-Specking A, Tomoyasu T, Mogk A, Bukau B. Trigger factor and DnaK cooperate in folding of newly synthesized proteins. *Nature.* **1999**; 400(6745):693–696.
117. Martinez-Hackert E, Hendrickson WA. Promiscuous substrate recognition in folding and assembly activities of the trigger factor chaperone. *Cell.* **2009**; 138(5):923–934.
118. Genevaux P, Keppel F, Schwager F, Langendijk-Genevaux PS, Hartl FU, Georgopoulos C. *In vivo* analysis of the overlapping functions of DnaK and trigger factor. *EMBO Rep.* **2004**; 5(2):195–200.
119. Genevaux P, Georgopoulos C, Kelley WL. The Hsp70 chaperone machines of *Escherichia coli*: a paradigm for the repartition of chaperone functions. *Mol Microbiol.* **2007**; 66(4):840–857.
120. Swain JF, Dinler G, Sivendran R, Montgomery DL, Stotz M, Gierasch LM. Hsp70 chaperone ligands control domain association via an allosteric mechanism mediated by the interdomain linker. *Mol Cell.* **2007**; 26(1):27–39.
121. Brehmer D, Rüdiger S, Gässler CS, *et al.* Tuning of chaperone activity of Hsp70 proteins by modulation of nucleotide exchange. *Nat Struct Biol.* **2001**; 8(5):427–432.
122. Mayer MP, Bukau B. Hsp70 chaperones: cellular functions and molecular mechanism. *Cell Mol Life Sci CMLS.* **2005**; 62(6):670–684.
123. Rüdiger S, Germeroth L, Schneider-Mergener J, Bukau B. Substrate specificity of the DnaK chaperone determined by screening cellulose-bound peptide libraries. *EMBO J.* **1997**; 16(7):1501–1507.

124. Calloni G, Chen T, Schermann SM, *et al.* DnaK Functions as a Central Hub in the *E. coli* Chaperone Network. *Cell Rep.* **2012**; 1(3):251–264.
125. Fayet O, Ziegelhoffer T, Georgopoulos C. The *groES* and *groEL* heat shock gene products of *Escherichia coli* are essential for bacterial growth at all temperatures. *J Bacteriol.* **1989**; 171(3):1379–1385.
126. Saibil HR, Fenton WA, Clare DK, Horwich AL. Structure and allostery of the chaperonin GroEL. *J Mol Biol.* **2013**; 425(9):1476–1487.
127. Kerner MJ, Naylor DJ, Ishihama Y, *et al.* Proteome-wide analysis of chaperonin-dependent protein folding in *Escherichia coli*. *Cell.* **2005**; 122(2):209–220.
128. Horwich AL, Fenton WA. Chaperonin-mediated protein folding: using a central cavity to kinetically assist polypeptide chain folding. *Q Rev Biophys.* **2009**; 42(2):83–116.
129. Lin Z, Madan D, Rye HS. GroEL stimulates protein folding through forced unfolding. *Nat Struct Mol Biol.* **2008**; 15(3):303–311.
130. Horwich AL, Low KB, Fenton WA, Hirshfield IN, Furtak K. Folding *in vivo* of bacterial cytoplasmic proteins: role of GroEL. *Cell.* **1993**; 74(5):909–917.
131. Guisbert E, Yura T, Rhodius VA, Gross CA. Convergence of molecular, modeling, and systems approaches for an understanding of the *Escherichia coli* heat shock response. *Microbiol Mol Biol Rev.* **2008**; 72(3):545–554.
132. Mogk A, Tomoyasu T, Goloubinoff P, *et al.* Identification of thermolabile *Escherichia coli* proteins: prevention and reversion of aggregation by DnaK and ClpB. *EMBO J.* **1999**; 18(24):6934–6949.
133. Groemping Y, Reinstein J. Folding properties of the nucleotide exchange factor GrpE from *Thermus thermophilus*: GrpE is a thermosensor that mediates heat shock response. *J Mol Biol.* **2001**; 314(1):167–178.
134. Winter J, Linke K, Jatzek A, Jakob U. Severe oxidative stress causes inactivation of DnaK and activation of the redox-regulated chaperone Hsp33. *Mol Cell.* **2005**; 17(3):381–392.
135. Winter J, Ilbert M, Graf PCF, Ozcelik D, Jakob U. Bleach activates a redox-regulated chaperone by oxidative protein unfolding. *Cell.* **2008**; 135(4):691–701.
136. Mogk A, Deuerling E, Vorderwülbecke S, Vierling E, Bukau B. Small heat shock proteins, ClpB and the DnaK system form a functional triade in reversing protein aggregation. *Mol Microbiol.* **2003**; 50(2):585–595.
137. Haslbeck M, Franzmann T, Weinfurtner D, Buchner J. Some like it hot: the structure and function of small heat-shock proteins. *Nat Struct Mol Biol.* **2005**; 12(10):842–846.
138. Weibezahn J, Tessarz P, Schlieker C, *et al.* Thermotolerance requires refolding of aggregated proteins by substrate translocation through the central pore of ClpB. *Cell.* **2004**; 119(5):653–665.

139. Haslberger T, Zdanowicz A, Brand I, *et al.* Protein disaggregation by the AAA+ chaperone ClpB involves partial threading of looped polypeptide segments. *Nat Struct Mol Biol.* **2008**; 15(6):641–650.
140. Sabate R, de Groot NS, Ventura S. Protein folding and aggregation in bacteria. *Cell Mol Life Sci CMLS.* **2010**; 67(16):2695–2715.
141. Leichert LI. Proteomic methods unravel the protein quality control in *Escherichia coli*. *Proteomics.* **2011**; 11(15):3023–3035.
142. Thomas JG, Baneyx F. ClpB and HtpG facilitate *de novo* protein folding in stressed *Escherichia coli* cells. *Mol Microbiol.* **2000**; 36(6):1360–1370.
143. Baneyx F, Mujacic M. Recombinant protein folding and misfolding in *Escherichia coli*. *Nat Biotechnol.* **2004**; 22(11):1399–1408.
144. Sauer RT, Baker TA. AAA+ proteases: ATP-fueled machines of protein destruction. *Annu Rev Biochem.* **2011**; 80:587–612.
145. Maisonneuve E, Fraysse L, Moinier D, Dukan S. Existence of abnormal protein aggregates in healthy *Escherichia coli* cells. *J Bacteriol.* **2008**; 190(3):887–893.
146. Lindner AB, Madden R, Demarez A, Stewart EJ, Taddei F. Asymmetric segregation of protein aggregates is associated with cellular aging and rejuvenation. *Proc Natl Acad Sci U S A.* **2008**; 105(8):3076–3081.
147. Winkler J, Seybert A, König L, *et al.* Quantitative and spatio-temporal features of protein aggregation in *Escherichia coli* and consequences on protein quality control and cellular ageing. *EMBO J.* **2010**; 29(5):910–923.
148. Taipale M, Jarosz DF, Lindquist S. HSP90 at the hub of protein homeostasis: emerging mechanistic insights. *Nat Rev Mol Cell Biol.* **2010**; 11(7):515–528.
149. Gupta RS. Phylogenetic analysis of the 90 kD heat shock family of protein sequences and an examination of the relationship among animals, plants, and fungi species. *Mol Biol Evol.* **1995**; 12(6):1063–1073.
150. Neidhardt FC, VanBogelen RA, Vaughn V. The genetics and regulation of heat-shock proteins. *Annu Rev Genet.* **1984**; 18:295–329.
151. Mason CA, Dünner J, Indra P, Colangelo T. Heat-induced expression and chemically induced expression of the *Escherichia coli* stress protein HtpG are affected by the growth environment. *Appl Environ Microbiol.* **1999**; 65(8):3433–3440.
152. Bardwell JC, Craig EA. Ancient heat shock gene is dispensable. *J Bacteriol.* **1988**; 170(7):2977–2983.
153. Huai Q, Wang H, Liu Y, Kim H-Y, Toft D, Ke H. Structures of the N-terminal and middle domains of *E. coli* Hsp90 and conformation changes upon ADP binding. *Struct Lond Engl* 1993. **2005**; 13(4):579–590.

154. Harris SF, Shiau AK, Agard DA. The crystal structure of the carboxy-terminal dimerization domain of HtpG, the *Escherichia coli* Hsp90, reveals a potential substrate binding site. *Struct Lond Engl* 1993. **2004**; 12(6):1087–1097.
155. Shiau AK, Harris SF, Southworth DR, Agard DA. Structural Analysis of *E. coli* Hsp90 reveals dramatic nucleotide-dependent conformational rearrangements. *Cell*. **2006**; 127(2):329–340.
156. Pearl LH, Prodromou C. Structure and mechanism of the Hsp90 molecular chaperone machinery. *Annu Rev Biochem*. **2006**; 75:271–294.
157. Genest O, Reidy M, Street TO, *et al.* Uncovering a region of heat shock protein 90 important for client binding in *E. coli* and chaperone function in yeast. *Mol Cell*. **2013**; 49(3):464–473.
158. Graf C, Stankiewicz M, Kramer G, Mayer MP. Spatially and kinetically resolved changes in the conformational dynamics of the Hsp90 chaperone machine. *EMBO J*. **2009**; 28(5):602–613.
159. Krukenberg KA, Förster F, Rice LM, Sali A, Agard DA. Multiple conformations of *E. coli* Hsp90 in solution: insights into the conformational dynamics of Hsp90. *Struct Lond Engl* 1993. **2008**; 16(5):755–765.
160. Krukenberg KA, Southworth DR, Street TO, Agard DA. pH-dependent conformational changes in bacterial Hsp90 reveal a Grp94-like conformation at pH 6 that is highly active in suppression of citrate synthase aggregation. *J Mol Biol*. **2009**; 390(2):278–291.
161. Southworth DR, Agard DA. Species-dependent ensembles of conserved conformational states define the Hsp90 chaperone ATPase cycle. *Mol Cell*. **2008**; 32(5):631–640.
162. Genest O, Hoskins JR, Camberg JL, Doyle SM, Wickner S. Heat shock protein 90 from *Escherichia coli* collaborates with the DnaK chaperone system in client protein remodeling. *Proc Natl Acad Sci U S A*. **2011**; 108(20):8206–8211.
163. Richter K, Buchner J. Hsp90: twist and fold. *Cell*. **2006**; 127(2):251–253.
164. Street TO, Lavery LA, Agard DA. Substrate binding drives large-scale conformational changes in the Hsp90 molecular chaperone. *Mol Cell*. **2011**; 42(1):96–105.
165. Wandinger SK, Richter K, Buchner J. The Hsp90 chaperone machinery. *J Biol Chem*. **2008**; 283(27):18473–18477.
166. Butland G, Peregrín-Alvarez JM, Li J, *et al.* Interaction network containing conserved and essential protein complexes in *Escherichia coli*. *Nature*. **2005**; 433(7025):531–537.
167. Arifuzzaman M, Maeda M, Itoh A, *et al.* Large-scale identification of protein-protein interaction of *Escherichia coli* K-12. *Genome Res*. **2006**; 16(5):686–691.
168. Motojima-Miyazaki Y, Yoshida M, Motojima F. Ribosomal protein L2 associates with *E. coli* HtpG and activates its ATPase activity. *Biochem Biophys Res Commun*. **2010**; 400(2):241–245.

169. Grudniak AM, Markowska K, Wolska KI. Interactions of *Escherichia coli* molecular chaperone HtpG with DnaA replication initiator DNA. *Cell Stress Chaperones*. **2015**; 20(6):951–957.
170. Yosef I, Goren MG, Kiro R, Edgar R, Qimron U. High-temperature protein G is essential for activity of the *Escherichia coli* clustered regularly interspaced short palindromic repeats (CRISPR)/Cas system. *Proc Natl Acad Sci U S A*. **2011**; 108(50):20136–20141.
171. Grudniak AM, Pawlak K, Bartosik K, Wolska KI. Physiological consequences of mutations in the htpG heat shock gene of *Escherichia coli*. *Mutat Res*. **2013**; 745–746:1–5.
172. Inoue T, Shingaki R, Hirose S, Waki K, Mori H, Fukui K. Genome-wide screening of genes required for swarming motility in *Escherichia coli* K-12. *J Bacteriol*. **2007**; 189(3):950–957.
173. Press MO, Li H, Creanza N, *et al*. Genome-scale co-evolutionary inference identifies functions and clients of bacterial Hsp90. *PLoS Genet*. **2013**; 9(7):e1003631.
174. Street TO, Zeng X, Pellarin R, *et al*. Elucidating the mechanism of substrate recognition by the bacterial Hsp90 molecular chaperone. *J Mol Biol*. **2014**; 426(12):2393–2404.
175. Li H, Sourjik V. Assembly and stability of flagellar motor in *Escherichia coli*. *Mol Microbiol*. **2011**; 80(4):886–899.
176. Sato T, Minagawa S, Kojima E, Okamoto N, Nakamoto H. HtpG, the prokaryotic homologue of Hsp90, stabilizes a phycobilisome protein in the cyanobacterium *Synechococcus elongatus* PCC 7942. *Mol Microbiol*. **2010**; 76(3):576–589.
177. Vivien E, Megessier S, Pieretti I, *et al*. *Xanthomonas albilineans* HtpG is required for biosynthesis of the antibiotic and phytotoxin albicidin. *FEMS Microbiol Lett*. **2005**; 251(1):81–89.
178. Washio K, Lim SP, Roongsawang N, Morikawa M. Identification and characterization of the genes responsible for the production of the cyclic lipopeptide arthrofactin by *Pseudomonas* sp. MIS38. *Biosci Biotechnol Biochem*. **2010**; 74(5):992–999.
179. Roongsawang N, Washio K, Morikawa M. Diversity of nonribosomal peptide synthetases involved in the biosynthesis of lipopeptide biosurfactants. *Int J Mol Sci*. **2010**; 12(1):141–172.
180. Saito M, Watanabe S, Yoshikawa H, Nakamoto H. Interaction of the molecular chaperone HtpG with uroporphyrinogen decarboxylase in the cyanobacterium *Synechococcus elongatus* PCC 7942. *Biosci Biotechnol Biochem*. **2008**; 72(5):1394–1397.
181. Tanaka N, Nakamoto H. HtpG is essential for the thermal stress management in cyanobacteria. *FEBS Lett*. **1999**; 458(2):117–123.
182. Hossain MM, Nakamoto H. Role for the cyanobacterial HtpG in protection from oxidative stress. *Curr Microbiol*. **2003**; 46(1):70–76.
183. Versteeg S, Mogk A, Schumann W. The *Bacillus subtilis* htpG gene is not involved in thermal stress management. *Mol Gen Genet*. **1999**; 261(3):582–588.

184. Mann MS, Dragovic Z, Schirrmacher G, Lütke-Eversloh T. Over-expression of stress protein-encoding genes helps *Clostridium acetobutylicum* to rapidly adapt to butanol stress. *Biotechnol Lett.* **2012**; 34(9):1643–1649.
185. Dang W, Hu Y, Sun L. HtpG is involved in the pathogenesis of *Edwardsiella tarda*. *Vet Microbiol.* **2011**; 152(3–4):394–400.
186. Choi S, Jang K, Choi S, Yun H-J, Kang D-H. Identification of the *Vibrio vulnificus htpG* gene and its influence on cold shock recovery. *J Microbiol Seoul Korea.* **2012**; 50(4):707–711.
187. García-Descalzo L, Alcazar A, Baquero F, Cid C. Identification of *in vivo* HSP90-interacting proteins reveals modularity of HSP90 complexes is dependent on the environment in psychrophilic bacteria. *Cell Stress Chaperones.* **2011**; 16(2):203–218.
188. Weiss DS, Brotcke A, Henry T, Margolis JJ, Chan K, Monack DM. *In vivo* negative selection screen identifies genes required for *Francisella* virulence. *Proc Natl Acad Sci U S A.* **2007**; 104(14):6037–6042.
189. King AM, Pretre G, Bartpho T, *et al.* High-temperature protein G is an essential virulence factor of *Leptospira interrogans*. *Infect Immun.* **2014**; 82(3):1123–1131.
190. Verbrugghe E, Van Parys A, Leyman B, Boyen F, Haesebrouck F, Pasmans F. HtpG contributes to *Salmonella* Typhimurium intestinal persistence in pigs. *Vet Res.* **2015**; 46:118.
191. Genest O, Hoskins JR, Kravats AN, Doyle SM, Wickner S. Hsp70 and Hsp90 of *E. coli* Directly Interact for Collaboration in Protein Remodeling. *J Mol Biol.* **2015**; 427(24):3877–3889.
192. Kumar M, Sourjik V. Physical map and dynamics of the chaperone network in *Escherichia coli*. *Mol Microbiol.* **2012**; 84(4):736–747.
193. Nakamoto H, Fujita K, Ohtaki A, *et al.* Physical interaction between bacterial heat shock protein (Hsp) 90 and Hsp70 chaperones mediates their cooperative action to refold denatured proteins. *J Biol Chem.* **2014**; 289(9):6110–6119.
194. Jackson SE. Hsp90: structure and function. *Top Curr Chem.* **2013**; 328:155–240.
195. Jhaveri K, Taldone T, Modi S, Chiosis G. Advances in the clinical development of heat shock protein 90 (Hsp90) inhibitors in cancers. *Biochim Biophys Acta.* **2012**; 1823(3):742–755.
196. Luheshi LM, Dobson CM. Bridging the gap: from protein misfolding to protein misfolding diseases. *FEBS Lett.* **2009**; 583(16):2581–2586.
197. Neckers L, Tatu U. Molecular chaperones in pathogen virulence: emerging new targets for therapy. *Cell Host Microbe.* **2008**; 4(6):519–527.
198. Angel SO, Matrajt M, Echeverria PC. A review of recent patents on the protozoan parasite HSP90 as a drug target. *Recent Pat Biotechnol.* **2013**; 7(1):2–8.
199. Veri A, Cowen LE. Progress and prospects for targeting Hsp90 to treat fungal infections. *Parasitology.* **2014**; 141(9):1127–1137.

200. Tabor CW, Tabor H. Polyamines in microorganisms. *Microbiol Rev.* **1985**; 49(1):81–99.
201. Shah P, Swiatlo E. A multifaceted role for polyamines in bacterial pathogens. *Mol Microbiol.* **2008**; 68(1):4–16.
202. Michael AJ. Biosynthesis of polyamines and polyamine-containing molecules. *Biochem J.* **2016**; 473(15):2315–2329.
203. Miyamoto S, Kashiwagi K, Ito K, Watanabe S, Igarashi K. Estimation of polyamine distribution and polyamine stimulation of protein synthesis in *Escherichia coli*. *Arch Biochem Biophys.* **1993**; 300(1):63–68.
204. Cohen SS. *A Guide to the Polyamines*. Oxford University Press; **1998**.
205. Igarashi K, Kashiwagi K. Modulation of cellular function by polyamines. *Int J Biochem Cell Biol.* **2010**; 42(1):39–51.
206. Dubin DT, Rosenthal SM. The acetylation of polyamines in *Escherichia coli*. *J Biol Chem.* **1960**; 235:776–782.
207. Charlier D, Glansdorff N. Biosynthesis of Arginine and Polyamines. *EcoSal Plus.* **2004**; 3(6):1–55.
208. Fukuchi J, Kashiwagi K, Takio K, Igarashi K. Properties and structure of spermidine acetyltransferase in *Escherichia coli*. *J Biol Chem.* **1994**; 269(36):22581–22585.
209. Watson N, Duniak DS, Rosey EL, Slonczewski JL, Olson ER. Identification of elements involved in transcriptional regulation of the *Escherichia coli* cad operon by external pH. *J Bacteriol.* **1992**; 174(2):530–540.
210. Barbagallo M, Di Martino ML, Marcocci L, *et al.* A new piece of the Shigella Pathogenicity puzzle: spermidine accumulation by silencing of the *speG* gene [corrected]. *PLoS One.* **2011**; 6(11):e27226.
211. Igarashi K, Kashiwagi K. Characteristics of cellular polyamine transport in prokaryotes and eukaryotes. *Plant Physiol Biochem.* **2010**; 48(7):506–512.
212. Furuchi T, Kashiwagi K, Kobayashi H, Igarashi K. Characteristics of the gene for a spermidine and putrescine transport system that maps at 15 min on the *Escherichia coli* chromosome. *J Biol Chem.* **1991**; 266(31):20928–20933.
213. Vassilyev DG, Tomitori H, Kashiwagi K, Morikawa K, Igarashi K. Crystal structure and mutational analysis of the *Escherichia coli* putrescine receptor. Structural basis for substrate specificity. *J Biol Chem.* **1998**; 273(28):17604–17609.
214. Higashi K, Ishigure H, Demizu R, *et al.* Identification of a Spermidine Excretion Protein Complex (MdtJI) in *Escherichia coli*. *J Bacteriol.* **2008**; 190(3):872–878.
215. Reitzer L. Catabolism of Amino Acids and Related Compounds. *EcoSal Plus.* **2005**; 3(4):1–56.
216. Antognoni F, Del Duca S, Kuraishi A, *et al.* Transcriptional inhibition of the operon for the spermidine uptake system by the substrate-binding protein PotD. *J Biol Chem.* **1999**; 274(4):1942–1948.

217. Cunningham-Rundles S, Maas WK. Isolation, characterization, and mapping of *Escherichia coli* mutants blocked in the synthesis of ornithine decarboxylase. *J Bacteriol.* **1975**; 124(2):791–799.
218. Igarashi K, Ito K, Kashiwagi K. Polyamine uptake systems in *Escherichia coli*. *Res Microbiol.* **2001**; 152(3):271–278.
219. Igarashi K, Kashiwagi K. Modulation of protein synthesis by polyamines. *IUBMB Life.* **2015**; 67(3):160–169.
220. Keating TA, Marshall CG, Walsh CT. Vibriobactin biosynthesis in *Vibrio cholerae*: VibH is an amide synthase homologous to nonribosomal peptide synthetase condensation domains. *Biochemistry (Mosc).* **2000**; 39(50):15513–15521.
221. Tan W, Verma V, Jeong K, *et al.* Molecular characterization of vulnibactin biosynthesis in *Vibrio vulnificus* indicates the existence of an alternative siderophore. *Front Microbiol.* **2014**; 5:1.
222. Nusca TD, Kim Y, Maltseva N, *et al.* Functional and structural analysis of the siderophore synthetase AsbB through reconstitution of the petrobactin biosynthetic pathway from *Bacillus anthracis*. *J Biol Chem.* **2012**; 287(19):16058–16072.
223. Aouida M, Poulin R, Ramotar D. The human carnitine transporter SLC22A16 mediates high affinity uptake of the anticancer polyamine analogue bleomycin-A5. *J Biol Chem.* **2010**; 285(9):6275–6284.
224. Koepsell H. The SLC22 family with transporters of organic cations, anions and zwitterions. *Mol Aspects Med.* **2013**; 34(2–3):413–435.
225. Henderson B, Allan E, Coates ARM. Stress wars: the direct role of host and bacterial molecular chaperones in bacterial infection. *Infect Immun.* **2006**; 74(7):3693–3706.
226. Richter K, Haslbeck M, Buchner J. The heat shock response: life on the verge of death. *Mol Cell.* **2010**; 40(2):253–266.
227. Garcie C, Tronnet S, Garénaux A, *et al.* The Bacterial Stress-Responsive Hsp90 Chaperone (HtpG) Is Required for the Production of the Genotoxin Colibactin and the Siderophore Yersiniabactin in *Escherichia coli*. *J Infect Dis.* **2016**; 214(6):916–924.
228. Dobrindt U, Chowdary MG, Krumbholz G, Hacker J. Genome dynamics and its impact on evolution of *Escherichia coli*. *Med Microbiol Immunol (Berl).* **2010**; 199(3):145–154.
229. Schmidt H, Hensel M. Pathogenicity Islands in Bacterial Pathogenesis. *Clin Microbiol Rev.* **2004**; 17(1):14–56.
230. Carniel E. The *Yersinia* high-pathogenicity island: an iron-uptake island. *Microbes Infect.* **2001**; 3(7):561–569.
231. Schubert S, Rakin A, Heesemann J. The *Yersinia* high-pathogenicity island (HPI): evolutionary and functional aspects. *Int J Med Microbiol.* **2004**; 294(2–3):83–94.
232. Chen J, Xie J. Role and regulation of bacterial LuxR-like regulators. *J Cell Biochem.* **2011**; 112(10):2694–2702.

233. Tronnet S, Garcie C, Rehm N, Dobrindt U, Oswald E, Martin P. Iron homeostasis regulates the genotoxicity of *Escherichia coli* producing colibactin. *Infect Immun*. **2016**; 84(12):3358-3368.
234. Faïs T, Cougnoux A, Dalmaso G, Laurent F, Delmas J, Bonnet R. Antibiotic activity of *Escherichia coli* against Multiresistant *Staphylococcus aureus*. *Antimicrob Agents Chemother*. **2016**; DOI:10.1128/AAC.00130-16.
235. Köhler C-D, Dobrindt U. What defines extraintestinal pathogenic *Escherichia coli*? *Int J Med Microbiol*. **2011**; 301(8):642–647.
236. Frank C, Werber D, Cramer JP, *et al*. Epidemic Profile of Shiga Toxin–Producing *Escherichia coli* O104:H4 Outbreak in Germany. *N Engl J Med*. **2011**; 365(19):1771–1780.
237. Soysal N, Mariani-Kurkdjian P, Smail Y, *et al*. Enterohemorrhagic *Escherichia coli* Hybrid Pathotype O80:H2 as a New Therapeutic Challenge. *Emerg Infect Dis*. **2016**; 22(9):1604–1612.
238. Gehring AM, Bradley KA, Walsh CT. Enterobactin biosynthesis in *Escherichia coli*: isochorismate lyase (EntB) is a bifunctional enzyme that is phosphopantetheinylated by EntD and then acylated by EntE using ATP and 2,3-dihydroxybenzoate. *Biochemistry (Mosc)*. **1997**; 36(28):8495–8503.
239. Hudson TJ. Genome variation and personalized cancer medicine. *J Intern Med*. **2013**; 274(5):440–450.
240. Minagawa S, Kondoh Y, Sueoka K, Osada H, Nakamoto H. Cyclic lipopeptide antibiotics bind to the N-terminal domain of the prokaryotic Hsp90 to inhibit the chaperone activity. *Biochem J*. **2011**; 435(1):237–246.
241. Gill EE, Franco OL, Hancock REW. Antibiotic Adjuvants: Diverse Strategies for Controlling Drug-Resistant Pathogens. *Chem Biol Drug Des*. **2015**; 85(1):56–78.
242. Di Martino ML, Campilongo R, Casalino M, Micheli G, Colonna B, Prosseda G. Polyamines: Emerging players in bacteria–host interactions. *Int J Med Microbiol*. **2013**; 303(8):484–491.
243. Jelsbak L, Thomsen LE, Wallrodt I, Jensen PR, Olsen JE. Polyamines Are Required for Virulence in *Salmonella enterica* serovar Typhimurium. *PLoS One*. **2012**; 7(4):e36149.
244. Schroll C, Christensen JP, Christensen H, *et al*. Polyamines are essential for virulence in *Salmonella enterica* serovar Gallinarum despite evolutionary decay of polyamine biosynthesis genes. *Vet Microbiol*. **2014**; 170(1–2):144–150.
245. Espinel IC, Guerra PR, Jelsbak L. Multiple roles of putrescine and spermidine in stress resistance and virulence of *Salmonella enterica* serovar Typhimurium. *Microb Pathog*. **2016**; 95:117–123.

ANNEXES

Article 3

OPEN ACCESS Freely available online

PLOS PATHOGENS

Interplay between Siderophores and Colibactin Genotoxin Biosynthetic Pathways in *Escherichia coli*

Patricia Martin^{1,2,3,4}, Ingrid Marcq^{1,2,3,4,5}, Giuseppe Magistro⁶, Marie Penary^{1,2,3,4}, Christophe Garcia^{1,2,3,4,7}, Delphine Payros^{1,2,3,4}, Michèle Boury^{1,2,3,4}, Maiwenn Olier^{1,2,3,4,8}, Jean-Philippe Nougayrède^{1,2,3,4}, Marc Audebert⁹, Christian Chalut¹⁰, Sören Schubert⁵, Eric Oswald^{1,2,3,4,7*}

1Inserm, UMR1043, Toulouse, France, **2**INRA, USC 1360, Toulouse, France, **3**CNRS, UMR5282, Toulouse, France, **4**Université de Toulouse, UPS, Centre de Physiopathologie de Toulouse Purpan (CPTP), Toulouse, France, **5**Jules Verne Picardie University, Medical school, Amiens, France, **6**Max von Pettenkofer-Institut für Hygiene und Medizinische Mikrobiologie, München, Germany, **7**CHU Toulouse, Hôpital Purpan, Service de bactériologie-Hygiène, Toulouse, France, **8**Neuro-gastroentérologie et Nutrition, UMR INRA/ENVIT 1331, Toulouse, France, **9**INRA, UMR1331, Toxalim, Research Centre in Food Toxicology, Toulouse, France, **10**Université de Toulouse, UPS, IPBS, Toulouse, France

Abstract

In *Escherichia coli*, the biosynthetic pathways of several small iron-scavenging molecules known as siderophores (enterobactin, salmochelins and yersiniabactin) and of a genotoxin (colibactin) are known to require a 4'-phosphopantetheinyl transferase (PPTase). Only two PPTases have been clearly identified: EntD and ClbA. The gene coding for EntD is part of the core genome of *E. coli*, whereas ClbA is encoded on the *pks* pathogenicity island which codes for colibactin. Interestingly, the *pks* island is physically associated with the high pathogenicity island (HPI) in a subset of highly virulent *E. coli* strains. The HPI carries the gene cluster required for yersiniabactin synthesis except for a gene coding its cognate PPTase. Here we investigated a potential interplay between the synthesis pathways leading to the production of siderophores and colibactin, through a functional interchangeability between EntD and ClbA. We demonstrated that ClbA could contribute to siderophores synthesis. Inactivation of both *entD* and *clbA* abolished the virulence of extra-intestinal pathogenic *E. coli* (ExPEC) in a mouse sepsis model, and the presence of either functional EntD or ClbA was required for the survival of ExPEC *in vivo*. This is the first report demonstrating a connection between multiple phosphopantetheinyl-requiring pathways leading to the biosynthesis of functionally distinct secondary metabolites in a given microorganism. Therefore, we hypothesize that the strict association of the *pks* island with HPI has been selected in highly virulent *E. coli* because ClbA is a promiscuous PPTase that can contribute to the synthesis of both the genotoxin and siderophores. The data highlight the complex regulatory interaction of various virulence features with different functions. The identification of key points of these networks is not only essential to the understanding of ExPEC virulence but also an attractive and promising target for the development of anti-virulence therapy strategies.

Citation: Martin P, Marcq I, Magistro G, Penary M, Garcia C, et al. (2013) Interplay between Siderophores and Colibactin Genotoxin Biosynthetic Pathways in *Escherichia coli*. PLoS Pathog 9(7): e1003437. doi:10.1371/journal.ppat.1003437

Editor: Xavier Nassif, Faculté de Médecine Paris Descartes, site Necker, France

Received: February 19, 2013; **Accepted:** May 2, 2013; **Published:** July 11, 2013

Copyright: © 2013 Martin et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This research was supported by grants from FEDER (MYCA), from the Centre de Physiopathologie de Toulouse Purpan (CPTP, Sepsis: the critical role of iron) and from the French National Research Agency (ANR). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: eric.oswald@inserm.fr

Introduction

Escherichia coli is a normal resident of the lower-gut of humans and animals. Although usually a commensal, *E. coli* can be also a pathogen, associated with diarrheal disease and extra-intestinal infections [1,2]. The majority of *E. coli* strains can be assigned to one of five main phylogenetic groups: A, B1, B2, D and E [3]. Strains of the distinct phylogenetic groups differ in their phenotypic and genotypic characteristics [4–6]. Extra-intestinal pathogenic *E. coli* (ExPEC), which display enhanced ability to cause infection outside the intestinal tract, carry specific genetic determinants or virulence factors that are clustered on different pathogenicity islands [7]. These virulence factors associated with extra-intestinal infections are nonrandomly distributed, and strains of the *E. coli* phylogenetic group B2 harbor the greatest frequency and diversity of virulence traits [8,9].

As iron bioavailability is limited in the host, ExPEC are known to synthesize up to four types of siderophores involved in iron uptake: enterobactin, salmochelins, yersiniabactin and aerobactin [10,11]. The biosynthesis of the first three requires a 4'-phosphopantetheinyl transferase (PPTase). These enzymes activate polyketide synthases (PKSs) and nonribosomal peptide synthetases (NRPSs) by catalyzing the transfer of a phosphopantetheinyl (P-pant) moiety from coenzyme A to conserved serine residues on PKSs and NRPSs [12,13]. In organisms containing multiple P-pant-requiring pathways, each pathway generally involves a dedicated cognate PPTase [12]. In *E. coli*, the EntD PPTase is involved in the synthesis of enterobactin [14] and salmochelins, which are glycosylated forms of enterobactin [15]. The *IroA* locus responsible for salmochelins production is located either on a chromosomal pathogenicity island or on a transmissible plasmid [16]. Contrary to enterobactin, salmochelins are able to evade the

Author Summary

The synthesis of numerous molecules involved in the virulence potential and fitness of pathogenic bacteria requires a particular enzyme family, *i.e.* phosphopantetheinyl transferases (PPTases). To date, the synthesis of a given bioactive metabolite was thought to require a specific PPTase. As PPTases are being investigated as promising targets for antibacterial development, we addressed the question of a possible functional interchangeability between PPTases in *Escherichia coli*. PPTases are known to be involved in the synthesis of low-molecular weight iron chelators (siderophores), and of a genotoxin named colibactin. Here we demonstrated interplay between the synthesis pathways leading to the production of siderophores and of colibactin. We showed that inactivation of both PPTases abolished the virulence of extra-intestinal pathogenic *E. coli* (ExPEC) in a mouse sepsis model. To our knowledge, this is the first demonstration of interplay between multiple PPTases-requiring pathways leading to the biosynthesis of functionally distinctive virulence factors, in a given microorganism. The extensive substrate specificity of PPTase ClbA could account for the co-selection and co-evolution of genomic islands encoding colibactin and yersiniabactin siderophore.

mammalian innate immune response protein lipocalin 2 (siderocalin) and are therefore more potent virulence factors [17]. The other siderophore necessitating a PPTase is yersiniabactin. This siderophore is encoded by the high-pathogenicity island (HPI) that was acquired through horizontal transfer [18]. The HPI core region was detected in more than 70% of ExPEC isolated from blood cultures, urine samples and cerebrospinal fluid [19]. While yersiniabactin production in *Yersinia* requires the YbtD PPTase encoded outside the HPI [20], no gene homologous to *ybtD* has been identified in the genome of *E. coli* strains producing yersiniabactin. The PPTase committed to the synthesis of yersiniabactin in *E. coli* remains unknown.

We have shown that a number of *E. coli* strains from phylogenetic group B2 display also the *pks* island, which codes for the production of colibactin, a polyketide-non ribosomal peptide genotoxin [21]. Colibactin is known to induce DNA double-strand breaks, cell cycle arrest in G2-phase and megakaryocytosis in infected eukaryotic cells [21]. *E. coli* strains harboring the *pks* island can induce DNA damage in enterocytes *in vivo* and trigger genomic instability in mammalian cells [22]. In a rodent model of colon inflammation, colibactin was demonstrated to potentiate the development of colon cancer [23]. Surprisingly, colibactin is also required for the colonic anti-inflammatory properties of the probiotic *E. coli* strain Nissle 1917 [24]. The synthesis of colibactin requires a PPTase encoded by the *clbA* gene located on the *pks* island [21]. Epidemiological studies revealed that the majority (73.1%) of the colibactin-positive *E. coli* strains was clinical ExPEC and that the *pks* island was significantly associated with a highly virulent subset of ExPEC isolates [25]. Strikingly, an analysis of the prevalence of the colibactin island among Enterobacteriaceae revealed that the *pks* island was constantly associated with the yersiniabactin gene cluster [26].

In this work we investigated a potential interplay between the biosynthetic pathways leading to the production of siderophores and of the colibactin genotoxin, through a possible functional interchangeability between PPTases in *E. coli*. We demonstrated that ClbA can contribute to the synthesis of siderophores both *in vitro* and *in vivo*. We proved in a mouse model of sepsis that the

presence of either functional EntD or ClbA is required to maintain full virulence of ExPEC. This evidenced the interconnection between pathways leading to the synthesis of distinct secondary metabolites, via the PPTase ClbA. Therefore, the strict association of the *pks* island with HPI could have been selected in highly virulent *E. coli* isolates because ClbA can contribute to the synthesis of both the genotoxin and yersiniabactin.

Results

The *pks* island does not code for the biosynthesis of a siderophore *in vitro*

Because colibactin and siderophores belong to the same family of chemical compounds, we investigated first whether the *pks* island could not only allow the production of a genotoxin, but also of a siderophore. The *entE* gene, that encodes the ligase component of synthase multienzyme complex necessary for the enterobactin biosynthesis, was inactivated in the enterobactin producer *E. coli* strain MG1655. The resulting MG1655 *entE* mutant strain was shown not to produce any siderophore, as detected on CAS plate (Fig. 1A). The wild type (WT) and *entE* derivative of strain MG1655 were transformed with the bacterial artificial chromosome (BAC) harboring the entire *pks* island (BAC *pks*⁺). Both strains MG1655+BAC *pks*⁺ and MG1655 *entE*+BAC *pks*⁺ were shown to produce the genotoxin, as evidenced by the induction of double-strand breaks in eukaryotic cells (data not shown). The production of siderophore was qualitatively investigated in the resulting strains by plating on CAS plates (Fig. 1A). A yellow halo was not observed surrounding the bacterial colonies of

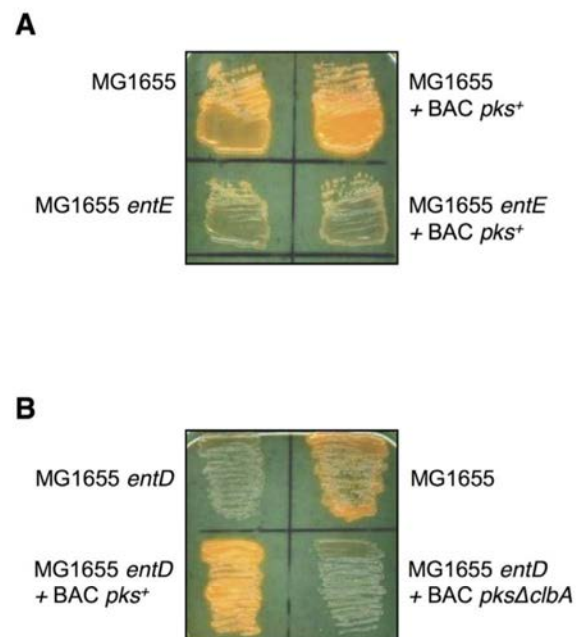


Figure 1. Siderophore production by *Escherichia coli* strain MG1655 and derivatives. Chrome azurol 5 (CAS) plates upon which the *E. coli* strain MG1655 and derivatives have been grown overnight. **A.** Wild type and *entE* derivatives of strain MG1655. **B.** Wild type and *entD* derivatives of strain MG1655. BAC *pks*⁺ is a bacterial artificial chromosome (BAC) harboring the entire *pks* island. BAC *pks* Δ *clbA* is a BAC harboring the entire *pks* island where the *clbA* gene was deleted. A yellow halo is produced around siderophore secreting bacteria. doi:10.1371/journal.ppat.1003437.g001

strain MG1655 *entE*+BAC *pks*⁺. This showed that the *pks* island did not code for the biosynthesis of a siderophore.

ClbA, the PPTase encoded on the *pks* island, can support the enterobactin siderophore synthesis *in vitro*

In order to test whether the ClbA PPTase was functionally capable of participating to the biosynthesis of enterobactin, the *entD* gene was disrupted in *E. coli* strain MG1655. The resulting MG1655 *entD* mutant strain was subsequently transformed with BAC *pks*⁺ and with the BAC harboring the entire *pks* island where the *clbA* gene was deleted (BAC *pksΔclbA*). The production of siderophore was investigated by plating the resulting strains on CAS medium (Fig. 1B). This revealed that disruption of the *entD* gene in strain MG1655 resulted in the abrogation of the production of enterobactin (Fig. 1B). The introduction of the intact *pks* island in strain MG1655 *entD* restored the production of yellow pigmentation surrounding the colonies. This was not observed upon the introduction of the *pks* island disrupted for the *clbA* gene (Fig. 1B). Introduction of a functional plasmidic *clbA* gene in strain MG1655 *entD*+BAC *pksΔclbA* and in strain MG1655 *entD* restored the production of enterobactin (data not shown).

These data evidenced that the ClbA PPTase can contribute to the enterobactin siderophore synthesis *in vitro*.

Both EntD and ClbA can support the yersiniabactin siderophore synthesis *in vitro*

Yersiniabactin is a siderophore the biosynthesis of which requires the PPTase YbtD in *Yersinia pestis* [20]. Although numerous *E. coli* strains were shown to produce yersiniabactin, an *in silico* analysis of the genome of all the *E. coli* strains available to date did not reveal any gene homologous to the *ybtD* gene.

In order to test whether the ClbA PPTase was functionally proficient to participate to the biosynthesis of yersiniabactin, we analyzed the enterobactin and yersiniabactin producer *E. coli* strain SE15. The *entD* gene was disrupted in *E. coli* strain SE15. The resulting SE15 *entD* mutant was subsequently transformed with plasmids carrying wild type *entD* gene or *clbA* gene. The production of total siderophores was qualitatively (Fig. 2A) and quantitatively (Fig. 2B) investigated using the CAS assay. This revealed that disruption of the *entD* gene resulted in the abrogation of the production of siderophores in strain SE15 (Fig. 2A and 2B). As expected, complementation with *entD* gene restored the production of siderophores. Remarkably, complementation with *clbA* gene also resulted in the synthesis of siderophores (Fig. 2A and 2B). The synthesis of yersiniabactin was specifically quantified in the different SE15 derivatives (Fig. 2C). This revealed that in the *entD* mutant, the yersiniabactin biosynthesis was abolished. The introduction of *entD* or *clbA* genes in SE15 *entD* mutant strain resulted in the restoration of yersiniabactin production.

These data showed that in *E. coli* strain SE15, EntD is the PPTase dedicated to the synthesis of yersiniabactin. Moreover, the EntD function can be substituted by ClbA. This suggests that both EntD and ClbA are involved in the synthesis of yersiniabactin in *E. coli* strains producing endogenously EntD and ClbA.

Colibactin synthesis cannot be sustained by EntD *in vitro*

As our data demonstrated that ClbA could complement EntD for the synthesis of enterobactin and yersiniabactin, we investigated whether EntD could rescue a *clbA* mutant for the production of colibactin. The *entD* gene was disrupted alone or in combination with the *clbA* gene in the colibactin producing *E. coli* strain M1/5.

The M1/5 *entD clbA* double mutant was transformed with multicopy plasmids harboring wild type *entD* or *clbA* genes. The production of colibactin was quantified in the resulting strains through the quantification of megalocytic cells (Fig. 3A) and phosphorylation of H2AX histone (Fig. 3B) which correlate with DNA double strand breaks resulting from the genotoxic effect of colibactin [21,22].

HeLa cells were infected with the different strains for 4 hours, fixed and stained with methylene blue in order to quantify the megalocytosis effect, as previously described [21]. This revealed that the megalocytosis effect observed with the M1/5 *entD* mutant strain was similar to the effect measured with the wild type M1/5 strain (Fig. 3A). Inactivation of the *clbA* gene in the M1/5 *entD* mutant abrogated the colibactin effect (Fig. 3A). Transformation of the M1/5 *entD clbA* mutant strain with plasmids carrying the functional wild type *clbA* gene resulted in the restoration of the megalocytosis. A partial complementation of the double mutation was observed with plasmid p-*clbA* (1) whereas the double mutant was fully complemented with p-*clbA* (2). The different copy number of the plasmids can account for the quantitative differences observed below. A complementation was not observed when the wild type *entD* gene was expressed from a multicopy plasmid in the double mutant (Fig. 3A).

Genotoxicity of colibactin [21] was also examined in HeLa cells using H2AX assay based on indirect DNA double strand break detection using In Cell Western (ICW) with infrared fluorescence for H2AX phosphorylation (γ -H2AX) quantification [27]. HeLa cells were infected with strains M1/5, M1/5 *entD*, M1/5 *entD clbA* or M1/5 *entD clbA* complemented with *entD*. Following the quantification of the γ -H2AX (green) and the DNA (red) signals (Fig. 3B), respectively, the fold induction of γ -H2AX per cell was calculated. This revealed a genotoxic dose-response depending on the multiplicity of infection (MOI, Fig. 3B). No difference of γ -H2AX per cell was observed between WT and *entD* mutant strains. Infection of HeLa cells with mutant M1/5 *entD clbA* did not induce phosphorylation of H2AX. Moreover, the introduction of the functional *entD* gene did not result in the generation of DNA double strand breaks in strain M1/5 *entD clbA* (Fig. 3B).

Altogether, these data evidenced that EntD does not contribute to the colibactin synthesis, even when highly expressed on a multicopy plasmid.

Colibactin synthesis can be sustained by exogenous PPTases *in vitro*

We then investigated whether other PPTases, originated from other bacterial species, could rescue a *clbA* mutant for the production of colibactin. The *clbA* gene was disrupted in *E. coli* strain M1/5. The M1/5 *clbA* mutant was transformed with plasmids harboring wild type *ybtD* gene that encodes the YbtD PPTase in *Yersinia pestis*, *pptT* gene the PptT PPTase in *Mycobacterium tuberculosis*, *sfp* gene the Sfp PPTase in *Bacillus subtilis*, and *clbA* gene. PptT is involved in biosynthesis of the mycobactin siderophore [28] and is essential for mycobacterial viability [29]. Sfp is required for production of the peptide antibiotic surfactin [30]. The production of colibactin was quantified in the resulting strains through the quantification of megalocytic cells (Fig. 4A) and phosphorylation of H2AX histone (Fig. 4B). This revealed that both the megalocytosis and the H2AX phosphorylation were restored in the *clbA* mutant upon introduction of *ybtD*, *pptT* and *sfp* genes.

These data evidenced that ClbA can be xeno-complemented for the colibactin synthesis.

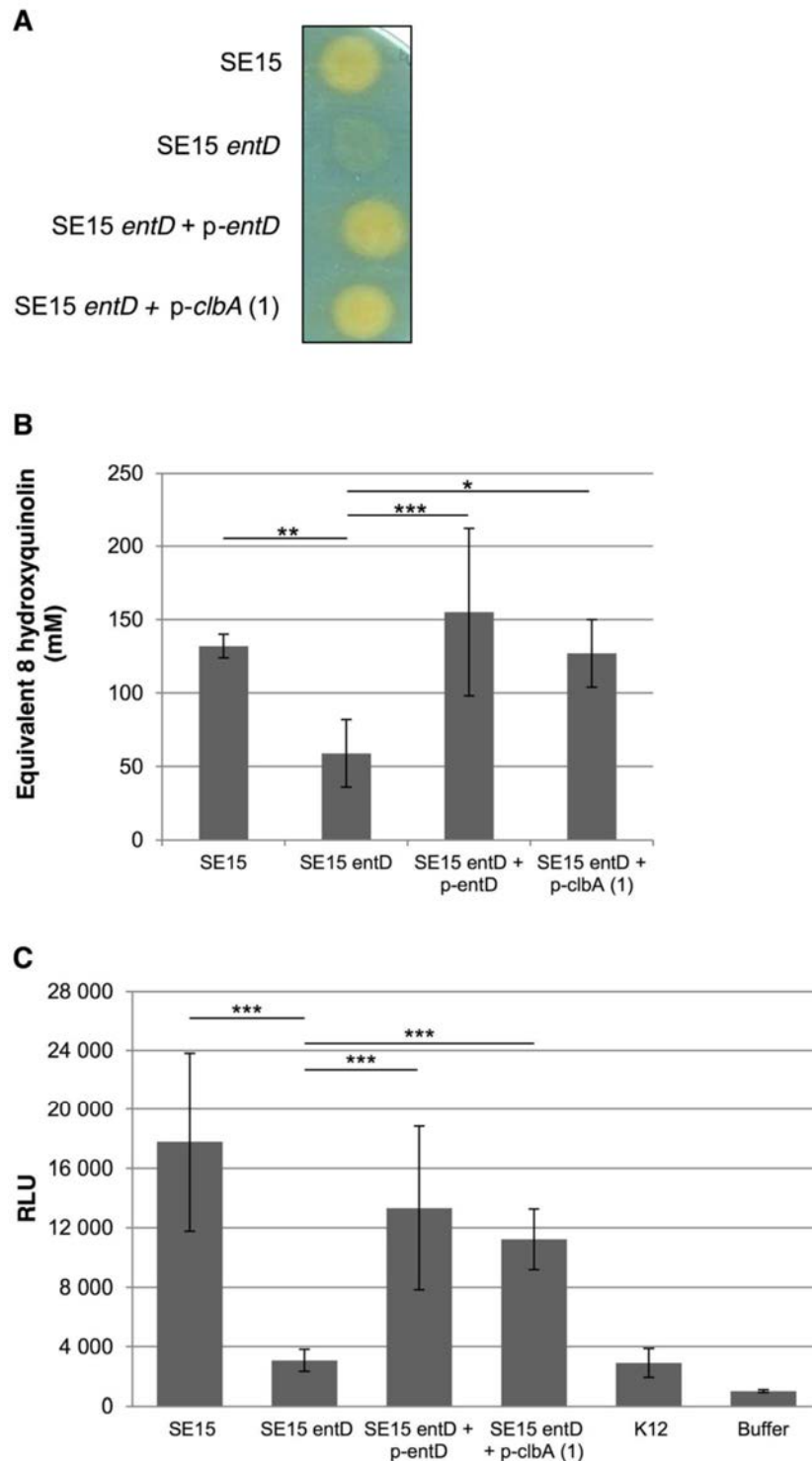


Figure 2. Both EntD and ClbA can support the yersiniabactin siderophore synthesis *in vitro*. Siderophore production by the enterobactin and yersiniabactin siderophores producer *Escherichia coli* strain SE15 and derivatives. **A.** Chrome azurol S (CAS) plate upon which *E. coli* strain SE15 and derivatives have been streaked for overnight growth. **B.** Quantification of total siderophore production in supernatants of *E. coli* strain SE15 and derivatives determined by the CAS assay. The data are the means and standard deviations of 5 independent experiments. **C.** Quantification of the yersiniabactin siderophore production in *E. coli* strain SE15 and derivatives. *E. coli* strains HB101, MG1655 and DH5 α were used as negative controls (K12). RLU: relative light units. ***: $P < 0.001$, **: $P < 0.01$, *: $P < 0.05$. doi:10.1371/journal.ppat.1003437.g002

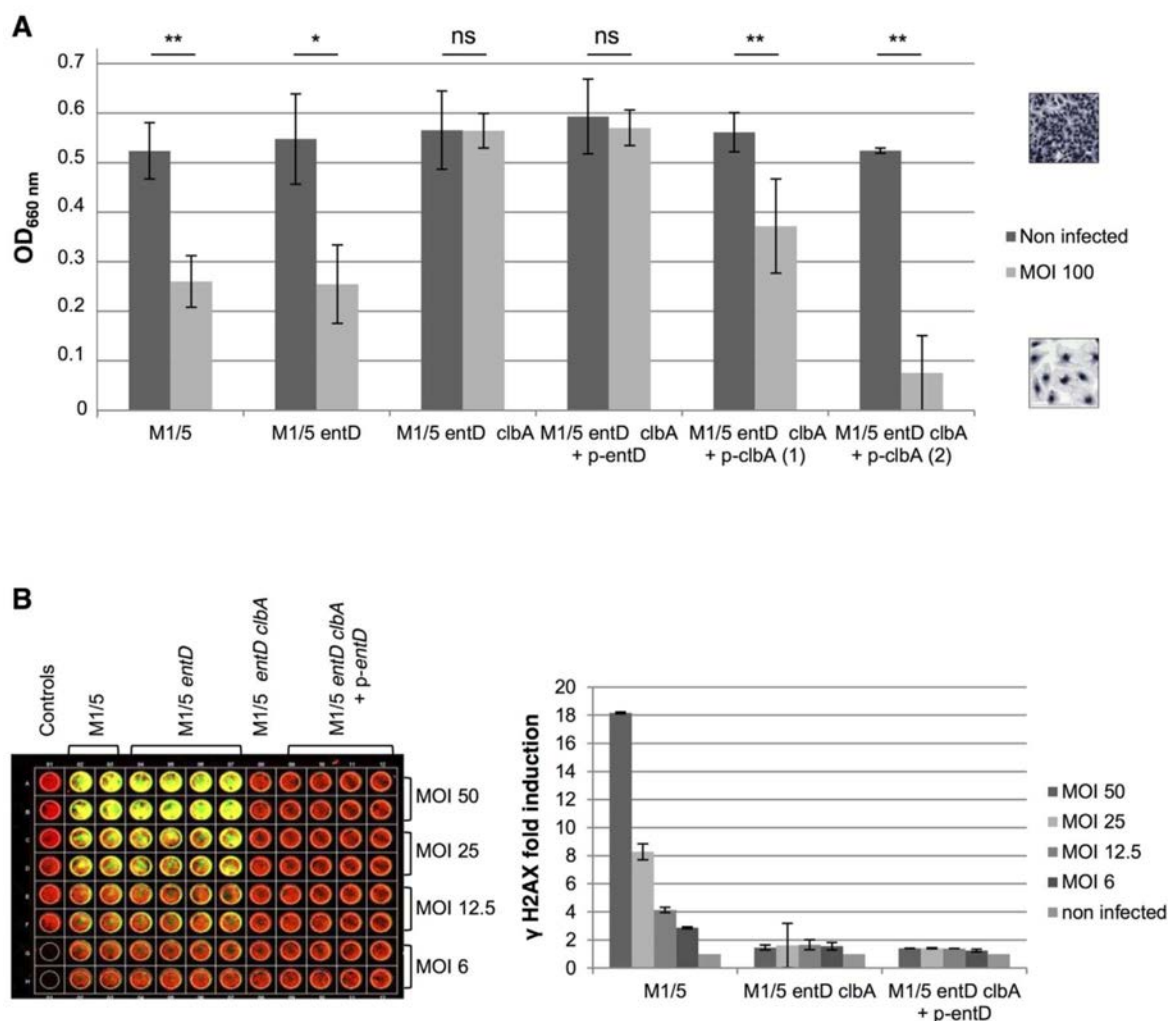


Figure 3. Colibactin synthesis cannot be sustained by EntD *in vitro*. Colibactin production by *Escherichia coli* strain M1/5 and derivatives determined by megalocytosis (**A**) and by quantification of DNA double strand breaks (**B**) in infected HeLa cells. **A.** Live *E. coli* wild type strain M1/5, mutants and complemented derivatives were added directly onto HeLa cells [multiplicity of infection (MOI) = 100], cocultivated for 4 h, then washed as described in Nougayrède et al. [21]. The cells were incubated for 72 h with gentamicin before protein staining with methylene blue. The quantification of staining was measured at OD 660 nm. **: $P < 0.01$, *: $P < 0.05$, ns: not significant. **B.** Quantification of DNA double strand breaks through the quantification of phosphorylated H2AX (γ -H2AX) using In Cell Western method [27]. HeLa cells were infected 4 h with strain M1/5 and derivatives [MOI = 50 to 6] fixed, and examined 8 h post infection for quantification of γ -H2AX. doi:10.1371/journal.ppat.1003437.g003

ClbA is more promiscuous in its substrate specificity than EntD

In order to confirm that EntD and ClbA have narrow and broad substrate-specificity, respectively, we investigated whether EntD and ClbA had the capacity to activate the carrier protein involved in a reporter biosynthetic pathway. When activated by a PPTase, the single-module non-ribosomal peptide synthetase BpsA from *Streptomyces lavendulae* synthesizes a colored product (indigoidine), from a single substrate (L-glutamine) [31]. Plasmid p-BpsA that encodes BpsA was transformed into strain MG1655 *entD*. The resulting MG1655 Δ entD+p-BpsA strain was subsequently transformed with plasmids carrying *ybtD*, *pptT*, *sfp*, *clbA*, or *entD* genes. In addition, *E. coli* strain MG1655 BAC *pks*⁺ and MG1655 BAC *pks* Δ clbA were transformed with p-BpsA. The resulting strains that carry both the NRPS and a functional PPTase were grown in

auto-induction medium, as previously described [32]. A blue coloration was detectable in cultures after overnight incubation for all strains but strain MG1655 Δ entD+p-BpsA+p-entD (Fig. 5A). A quantification of the indigoidine production was determined for all the strains (Fig. 5B). This confirmed that contrary to EntD, the PPTases YbtD, PptT, Sfp and ClbA were able to participate to the synthesis of the blue pigment.

This strengthens the fact that ClbA is more promiscuous in its substrate specificity than EntD in *E. coli*.

Both EntD and ClbA must be inactivated to abolish virulence of ExPEC in a mouse model of sepsis

In order to address the consequences, on the virulence of *E. coli*, of the cross talk between the synthesis pathways of colibactin and

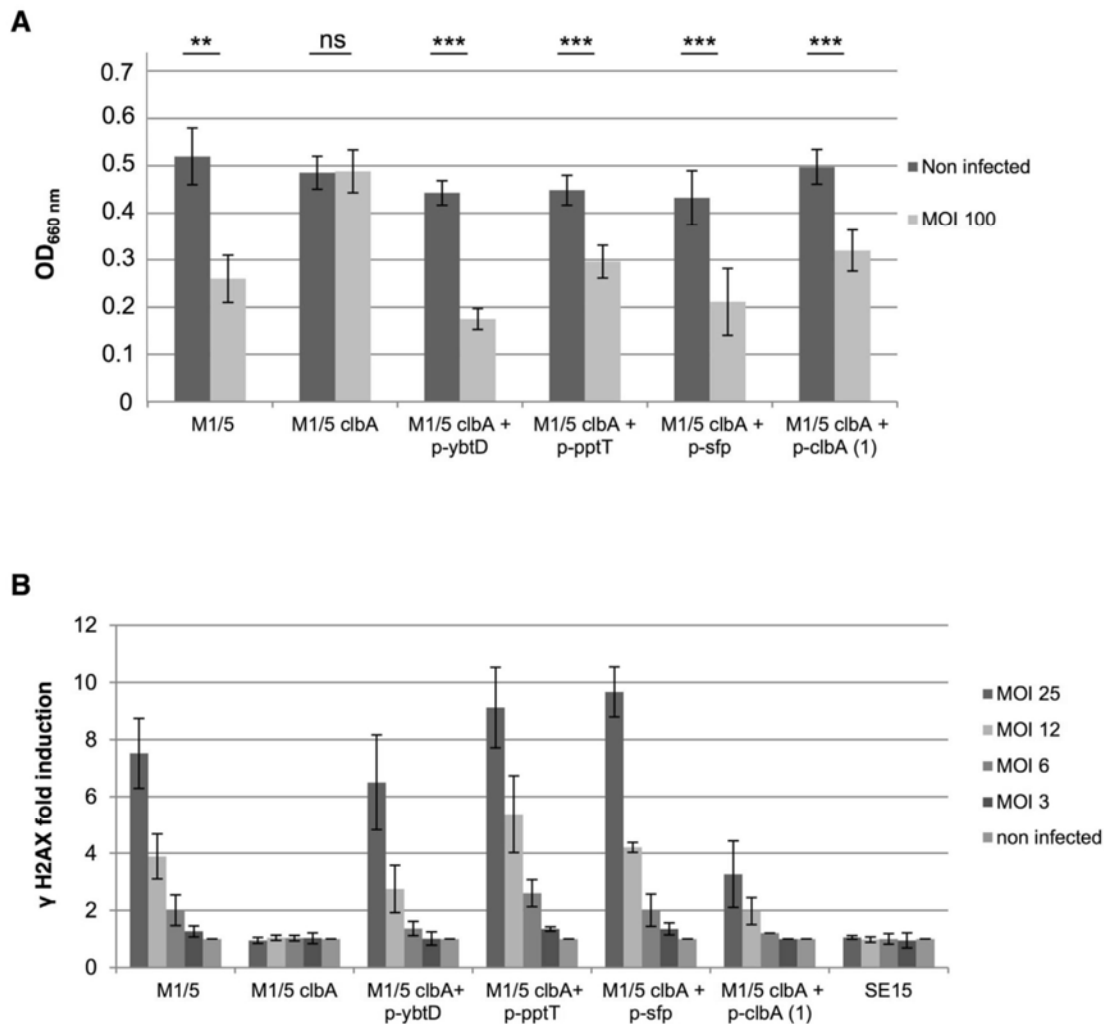


Figure 4. Colibactin synthesis can be sustained by exogenous PPTases *in vitro*. Colibactin production by *Escherichia coli* strain M1/5 and derivatives determined by megalocytosis (A) and by quantification of DNA double strand breaks (B), as in Fig. 3. ***: $P < 0.001$, **: $P < 0.01$, ns: not significant. *E. coli* strain SE15, which is devoid of colibactin locus, was used as a negative control. The *ybtD* gene encodes the YbtD PPTase in *Yersinia pestis*, the *pptT* gene the PptT PPTase in *Mycobacterium tuberculosis*, and the *sfp* gene the Sfp PPTase in *Bacillus subtilis*. doi:10.1371/journal.ppat.1003437.g004

siderophores demonstrated *in vitro*, we investigated *E. coli* strain SP15, an extra-intestinal pathogenic *E. coli* strain (ExPEC) of serotype O18:K1:H7 isolated from neonatal meningitis, in a mouse model of sepsis. *E. coli* strain SP15 produces colibactin and four different siderophores (aerobactin, yersiniabactin, enterobactin and salmochelin). The *entD* or *clbA* genes were disrupted individually and in combination. The strains were injected individually into the mice footpad; and the mice survival was monitored (Fig. 6A). This revealed that all the strains but SP15 *entD clbA* induced 70% mortality within 40 hours after injection. In contrast, virulence of strain SP15 *entD clbA* was completely attenuated in this mouse model of sepsis (Fig. 6A).

The bacterial dissemination in the mice was analyzed (Fig. 6B). Mice were sacrificed 18 hours post injection with PBS, WT strain, single or double mutants. Spleens and blood samples were collected, and bacteria were quantified by plating on selective medium (Fig. 6B). We observed that in both spleen and blood of

infected animals the bacterial loads were similar with all the strains, but strain SP15 *entD clbA*. No bacteria were recovered from spleen or blood of mice injected with the double mutant SP15 *entD clbA* (Fig. 6B).

This demonstrated that both EntD and ClbA must be inactivated to abolish virulence of ExPEC in a mouse model of sepsis.

The presence of either functional EntD or ClbA is required to maintain full virulence of ExPEC in a mouse model of sepsis

In order to investigate the relative importance of EntD and ClbA in the virulence of *E. coli*, the SP15 *entD clbA* mutant strain was transformed with plasmids harboring *clbA* or *entD* functional genes. The resulting complemented strains were injected in mice (Fig. 7). This showed that complementation of strain SP15 *entD clbA* with either *clbA* or *entD* totally restored the virulence of the

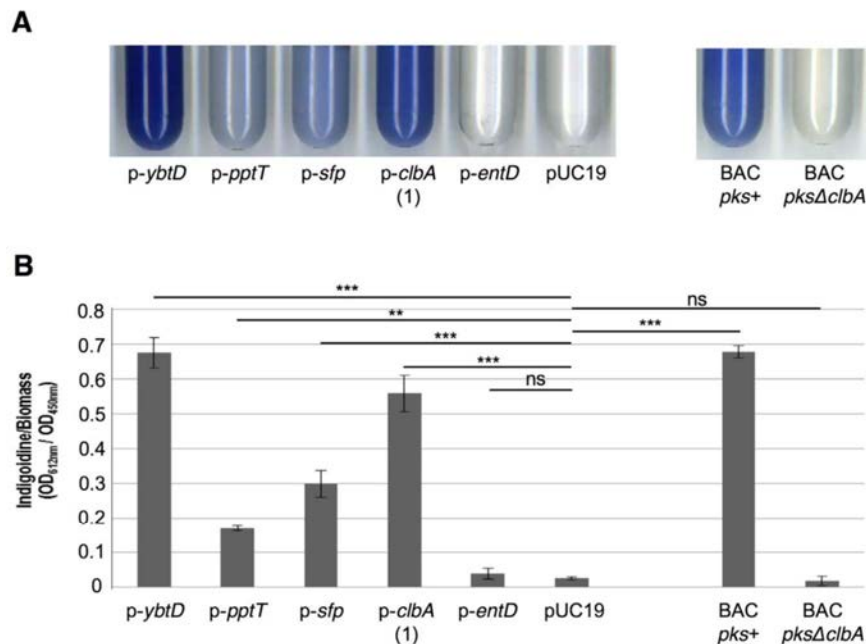


Figure 5. ClbA is more promiscuous in its substrate specificity than EntD. The synthesis of the single-module non-ribosomal peptide synthetase BpsA from *Streptomyces lavendulae* resulting in the production of indigoidine was qualitatively (A) and quantitatively (B) assessed as previously described [31,51]. *E. coli* strain MG1655 Δ entD+p-BpsA was transformed with the plasmids p-ybtD, p-pptT, p-sfp, p-clbA (1), p-entD and pUC19 (left). *E. coli* strain MG1655 BAC pks⁺ and MG1655 BAC pks Δ clbA were transformed with p-BpsA (right). **A.** All the strains were grown overnight at 18 °C in an auto-inducer medium. The cultures were left at room temperature so that the blue pigmentation (indigoidine) is produced. **B.** Quantification of indigoidine in the different strains. The data are the means and standard deviations of 3 independent experiments. ***: $P < 0.001$, **: $P < 0.01$, ns: not significant.

doi:10.1371/journal.ppat.1003437.g005

strain (Fig. 7A). A slight but statistically significant delay in survival kinetics was observed when strain SP15 *entD clbA* complemented with the *clbA* gene was used for the injections (Fig. 7A). The quantification of bacteria in spleen and blood of the infected animals was determined (Fig. 7B). This revealed that complementation with *clbA* or *entD* allowed the survival of strain SP15 *entD clbA* *in vivo*, in a statistically significant manner at least in blood (Fig. 7B).

This evidenced that the presence of either functional EntD or ClbA is required to maintain full virulence of ExPEC in a mouse model of sepsis.

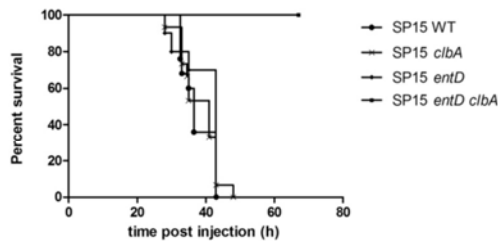
Discussion

Our work demonstrates the interplay between the biosynthetic pathways of a genotoxin and multiple siderophores. We have shown that ClbA, encoded by the *pks* island, is a promiscuous PPTase which promotes the synthesis of colibactin, yersiniabactin, enterobactin and consequently salmochelins. Although we demonstrated that ClbA could substitute for an *entD* mutation, the reciprocity was not observed. EntD seems to be specific for the synthesis of siderophores, which is consistent with other published reports [32]. In contrast, YbtD, the PPTase involved in yersiniabactin production in *Yersinia* was shown to substitute for a *clbA* mutation and allowed the production of colibactin. Attempts to relate conserved motifs of the group II subfamily of PPTases [12] with substrate specificity did not allow us to understand the functional promiscuity evidenced among certain PPTases, since type II PPTases usually have very remote primary sequences. Unfortunately, it is not possible to compare either the 3D structure

of these PPTases because only the structure of Sfp is available [33]. Type II PPTases are predicted to have a similar folding and very similar secondary structures [33]. However it is difficult to draw conclusions on the folding of proteins and to correlate it with substrate specificity. Only the comparison of 3D high-resolution structures would provide information about the structure/function relationship of PPTases. Our work provides novel evidence that make PPTases promising targets for antibacterial development [34], because these enzymes are crucial for the biosynthesis of a multitude of a pathogen's collection of essential metabolites and virulence factors [35].

Iron is an essential element for survival of *E. coli*. Therefore, *E. coli* strains have evolved a strategy for iron acquisition which uses multiple siderophores with high-affinity for ferric iron. These include enterobactin, salmochelins, aerobactin and yersiniabactin [11]. Each siderophore has specific affinity for iron and may be differentially regulated to provide different advantages, potentially allowing extra-intestinal pathogenic *E. coli* (ExPEC) to adapt to different environmental conditions or to overcome host innate immunity [10,36,37]. In our model of sepsis, the ExPEC mutant that produced only aerobactin as a siderophore (strain SP15 *entD clbA*) was completely attenuated. This suggests that aerobactin plays a minor role in the iron uptake in this sepsis model; but the importance of each siderophore can be host and strain dependent [38]. Interestingly, either ClbA or EntD were able to restore the virulence of strain SP15 *entD clbA*. However, we have shown that colibactin synthesis cannot be sustained by EntD. This suggests that not colibactin, but the siderophore systems (alone or in combination) are critical during the first step of the infection in this

A



B

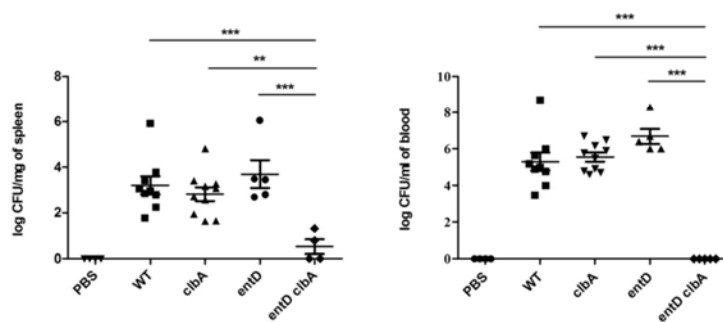


Figure 6. Both EntD and ClbA must be inactivated to abolish virulence of ExPEC. Mice underwent footpad injection with 10^8 CFU of *E. coli* SP15 wild type strain or derivatives. **A.** The percentage of mice survival was monitored. 10 to 25 mice were used per group. **B.** 18 h post infection 4 to 10 mice per group were sacrificed. Bacteria were quantified in spleen and blood collected from each animal. For statistical analysis, two-factor ANOVA and Bonferroni's multiple comparison test was performed. ***: $P < 0.001$, **: $P < 0.01$. doi:10.1371/journal.ppat.1003437.g006

mouse model of sepsis. Indeed, the bacterial loads in both spleen and blood were similar in animal infected with SP15 *entD clbA* mutant complemented either with ClbA or EntD. Analysis of bacteria present in the popliteal lymph node confirmed this analysis (data not shown). Since the carriage of the *pks* island is correlated with successful long-term gut colonization in humans [39], colibactin could be important for the commensal lifestyle of ExPEC. Moreover, our unpublished data suggest that the genotoxin colibactin could also play a role in natural sepsis since lymphocytes are susceptible to the genotoxin.

Phylogroup B2, which includes the majority of ExPEC isolates, is considered to represent the evolutionary eldest lineage within the species [40]. Interestingly, the *pks* island found in B2 isolates is highly conserved, and is physically associated to a highly conserved High-Pathogenicity Island. This might even point towards a recent emergence of a distinct subgroup within phylogroup B2. In fact, epidemiological knowledge allows defining specific clonal lineages with high ExPEC virulence potential [41]. We believe that the most virulent and also the best colonizer of human gut resulted from a step-by-step acquisition and selection of different mobile elements. We propose here a scenario with the sequential integration of at least two pathogenicity islands and the cross talk via two PPTases (Fig. 8). At first, all *E. coli* strains produce at least one siderophore *i.e.* enterobactin. The *entD* gene and the other genes of the enterobactin system are part of the core genome and have been identified in all the *E. coli* strains isolated so far [42]. In contrast, the HPI encoding the yersiniabactin siderophore system devoid of any PPTase gene was acquired by horizontal gene

transfer. Almost all *E. coli* HPIs appear to result from a single ancestor, which entered the *E. coli* species rather recently [43]. All strains of the phylogenetic group B2 and almost all of group D carry the HPI, whereas strains of groups A and B1 were found to be only occasionally HPI positive (Fig. 8, [19]). The spread of the HPI must have occurred in a dramatically fast fashion, which may indicate a strong selective pressure. We have shown in this study that EntD is actually the PPTase that mediates the synthesis of a functional yersiniabactin. *E. coli* strains that contain the HPI were demonstrated to be more virulent than isolates that lack the island [18]. Moreover, yersiniabactin is frequently associated with urinary tract infections [44,45]. The *pks* island is known to be confined to the phylogenetic group B2. Besides, the *pks* island is highly represented within an especially highly virulent subset of B2 strains that exhibit extremely elevated virulence scores and an increased likelihood of causing bacteremia [25]. It has been previously demonstrated that all the *E. coli* strains that acquired the *pks* island encoding the colibactin through horizontal transfer, also displayed the HPI locus, with an integration site in tRNA *asnW* gene and *asnT* gene, respectively (Fig. 8; [26]). The *pks* island appears to be highly conserved (or even identical) in terms of nucleotide sequence in different *E. coli* isolates [26]. This may be a hint to a more recent acquisition of the *pks* island, compared to the HPI, which displays about 1–2% sequence divergence among the *E. coli* isolates (Schubert, unpublished data). We hypothesize that the association of the *pks* island with HPI has been selected in the highly virulent *E. coli* isolates because ClbA can contribute to the synthesis of both the genotoxin and yersiniabactin (and also

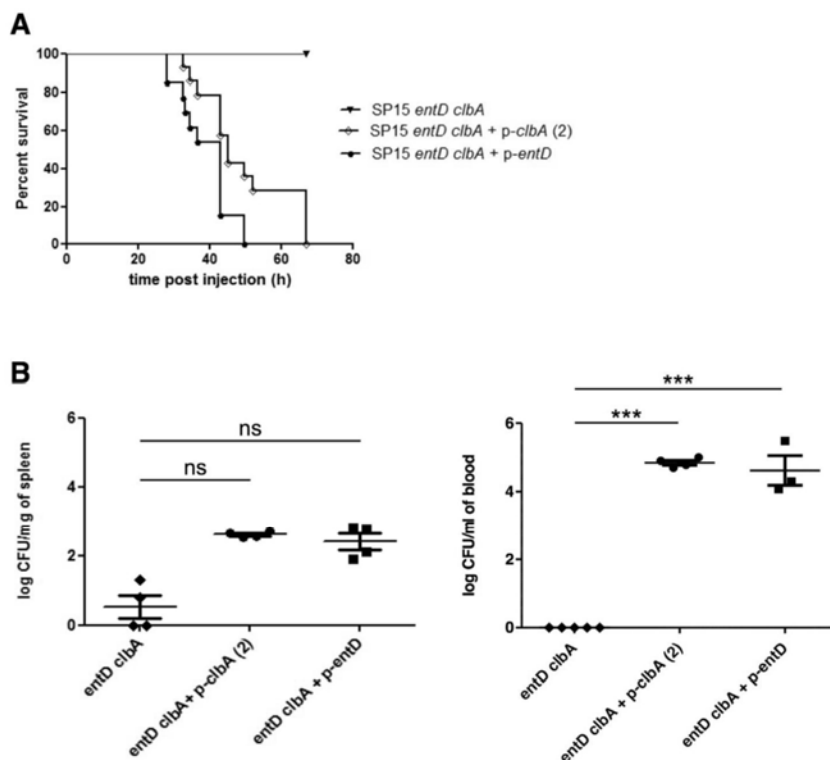


Figure 7. The presence of either EntD or ClbA is required to maintain full virulence of ExPEC. Mice underwent footpad injection with 10^8 CFU of *E. coli* SP15 *entD clbA* strain and complemented derivatives. **A.** The percentage of mice survival was monitored. 10 to 25 mice were used per group. **B.** 18 h post infection 3 to 5 mice per group were sacrificed. Bacteria were quantified in spleen and blood collected from each animal. For statistical analysis, two-factor ANOVA and Bonferroni's multiple comparison test was performed. ***: $P < 0.001$, ns: not significant. doi:10.1371/journal.ppat.1003437.g007

enterobactin and consequently salmochelins). This deadly association is not confined in *E. coli*. Similar events also occurred in other pathogenic *Enterobacteriaceae* since the *pks* island was also detected in *Klebsiella pneumoniae*, *Enterobacter aerogenes*, and *Citrobacter koseri* isolates where the island is also physically associated on the chromosome with the HPI locus [26].

Materials and Methods

Bacterial strains, mutagenesis procedures and growth conditions

Bacterial strains used in this study are listed in Table 1. *E. coli* SE15 (O150:H5) is a human commensal bacterium isolated from feces of a healthy adult and classified into *E. coli* phylogenetic group B2 [46]. Strain SE15 is devoid of the *pks* island. *E. coli* M1/5 is a human commensal bacterium isolated from feces of a healthy adult and classified into *E. coli* phylogenetic group B2. Strain M1/5 harbors of the *pks* island. Strain SP15 is an extra-intestinal pathogenic *E. coli* strain (ExPEC) of serotype O18:K1:H7 isolated from neonatal meningitis. Strain SP15 harbors the *pks* island. The repertoire of siderophores the *E. coli* strains possess is indicated in Table 1. Gene inactivations were engineered by using the lambda Red recombinase method [47] using primers listed in Table 2. For complementation, the *clbA* gene was cloned into plasmid pASK75, a cloning vector that harbors a pBR322 origin of replication and therefore is low copy number plasmid (p-*clbA* (1),

table 1) or PCR-Script, a cloning vector that harbors a ColE1 origin of replication and therefore is high copy number plasmid (p-*clbA* (2), table 1). For complementation, the *entD* gene was cloned into PCR-Script (p-*entD*, table 1).

Before injection to mice, all *E. coli* strains were grown overnight in LB broth supplemented with antibiotics if required, at 37°C with shaking. These cultures were diluted 1:100 in LB broth with antibiotics when necessary and grown for 3 h at 37°C with shaking. Bacterial cells were resuspended in sterile PBS to the appropriate concentration (2×10^9 CFU/mL). All the strains were shown to display similar growth kinetics *in vitro* in LB broth (data not shown).

Detection and quantification of total siderophores

Chrome azurol S (CAS) assay was used to detect siderophores produced by *E. coli*. The CAS solution was prepared according to Schwyn and Neilands [48]. *E. coli* strains were grown on CAS agar plates and incubated at 37°C overnight in the dark. The colonies with orange zones were siderophore-producing strains [48].

To quantify siderophore synthesis, 500 μ L of CAS indicator solution containing 4 mM sulfosalicylic acid was mixed with the same volume of supernatant. The reaction mixtures were incubated for 60 min at room temperature to allow complex formation, and the siderophore-dependent color change was determined at OD_{630 nm}. For quantification, the iron chelating agent 8 hydroxyquinoline (8HQ, sigma-aldrich) was used as the standard.

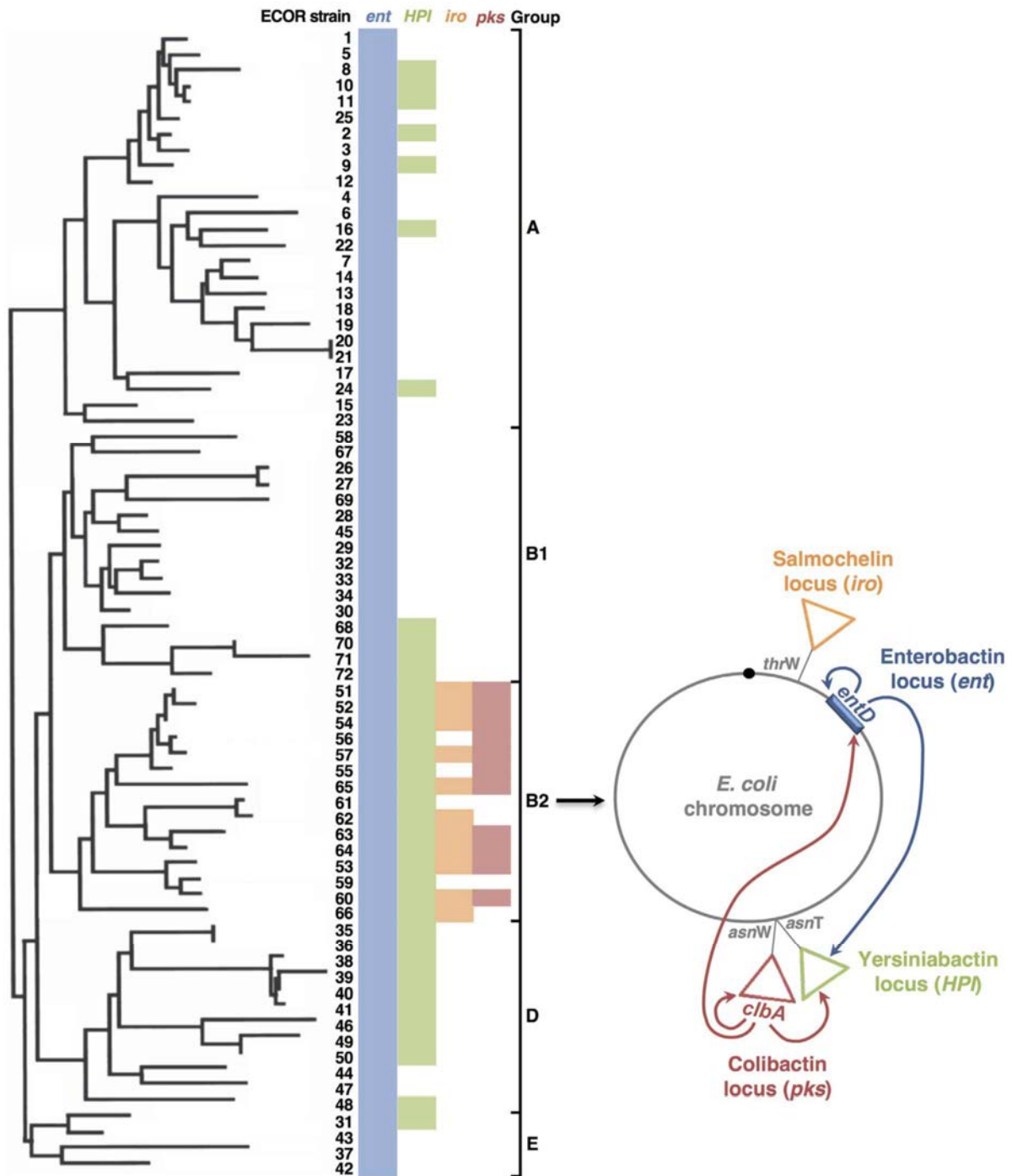


Figure 8. Model for the co-evolution of the *pks* and high pathogenicity islands in *E. coli*. **Left.** Phylogenetic relationships amongst the *E. coli* reference strains (ECOR, [3,52]). The phylogeny was based on MLST of back-bone genes [3,43]. Ent locus positive strains are indicated in blue, HPI island positive strains are indicated in green [43], *iro* locus positive strains are indicated in orange and *pks* island positive strains are indicated in pink [21]. The presence of the *iro* locus was determined only in B2 strains. **Right.** The archetypal chromosome of phylogroup B2 *E. coli* strains. The loci encoding enterobactin (*ent*), yersiniabactin (HPI), salmochelin (*iro*) and colibactin (*pks*) are represented. The arrows originating from PPTases EntD and ClbA and pointing towards other loci illustrate the capacity of the PPTase to contribute to the synthesis of metabolites from other loci. doi:10.1371/journal.ppat.1003437.g008

Table 1. Strains and plasmids used in the study.

Strain or plasmid	Genotype or phenotype	Source or reference
<i>E. coli</i> strains		
DH10B	Enterobactin siderophore producer	
DH5 α	Enterobactin siderophore producer	
HB101	Enterobactin siderophore producer	
MG1655	Enterobactin siderophore producer	
WR1542+pACYC5.3L	Tc ^r , Ap ^r , Kan ^r , Cm ^r ; <i>fepA::Tn10dTc</i> , <i>iroN::pGP704</i> , <i>cir::MudJ</i> carrying pACYC5.3L plasmid	Gift from W. Rabsch
MG1655 <i>entE</i>	<i>entE</i> mutant of strain MG1655; Kan ^r	This study
MG1655 <i>entE</i> +BAC <i>pk</i> s+	<i>entE</i> mutant of strain MG1655 carrying BAC <i>pk</i> s+; Kan ^r , Cm ^r	This study
MG1655 <i>entD</i>	<i>entD</i> mutant of strain MG1655; Kan ^r	This study
MG1655 <i>entD</i> +BAC <i>pk</i> s+	<i>entD</i> mutant of strain MG1655 carrying BAC <i>pk</i> s+; Kan ^r , Cm ^r	This study
MG1655 <i>entD</i> +BAC <i>pk</i> s Δ <i>clbA</i>	<i>entD</i> mutant of strain MG1655 carrying BAC <i>pk</i> s Δ <i>clbA</i> ; Kan ^r , Cm ^r	This study
MG1655 <i>entD</i> +p- <i>clbA</i> (1)	<i>entD</i> mutant of strain MG1655 carrying p- <i>clbA</i> (1); Kan ^r Amp ^r	This study
MG1655 Δ <i>entD</i>	<i>entD</i> mutant of strain MG1655	This study
MG1655 Δ <i>entD</i> +p- <i>bpsA</i>	<i>entD</i> mutant of strain MG1655 carrying p- <i>bpsA</i> ; Kan ^r	This study
MG1655 Δ <i>entD</i> +p- <i>bpsA</i> +p- <i>entD</i>	<i>entD</i> mutant of strain MG1655 carrying p- <i>bpsA</i> and p- <i>entD</i> ; Kan ^r Amp ^r	This study
MG1655 Δ <i>entD</i> +p- <i>bpsA</i> —+p- <i>clbA</i> (1)	<i>entD</i> mutant of strain MG1655 carrying p- <i>bpsA</i> and p- <i>clbA</i> (1); Kan ^r Amp ^r	This study
MG1655 Δ <i>entD</i> +p- <i>bpsA</i> +p- <i>sfp</i>	<i>entD</i> mutant of strain MG1655 carrying p- <i>bpsA</i> and p- <i>sfp</i> ; Kan ^r Amp ^r	This study
MG1655 Δ <i>entD</i> +p- <i>bpsA</i> +p- <i>pptT</i>	<i>entD</i> mutant of strain MG1655 carrying p- <i>bpsA</i> and p- <i>pptT</i> ; Kan ^r Amp ^r	This study
MG1655+BAC <i>pk</i> s+	strain MG1655 carrying BAC <i>pk</i> s+; Cm ^r	This study
MG1655+BAC <i>pk</i> s Δ <i>clbA</i>	strain MG1655 carrying BAC <i>pk</i> s Δ <i>clbA</i> ; Cm ^r	This study
MG1655+BAC <i>pk</i> s++p- <i>bpsA</i>	strain MG1655 carrying BAC <i>pk</i> s+ and p- <i>bpsA</i> ; Kan ^r Cm ^r	This study
MG1655+BAC <i>pk</i> s Δ <i>clbA</i> +p- <i>bpsA</i>	strain MG1655 carrying BAC <i>pk</i> s Δ <i>clbA</i> and p- <i>bpsA</i> ; Kan ^r Cm ^r	This study
SE15	Enterobactin and yersiniabactin siderophores producer	[46]
SE15 <i>entD</i>	<i>entD</i> mutant of strain SE15; Kan ^r	This study
SE15 <i>entD</i> +p- <i>entD</i>	<i>entD</i> mutant of strain SE15 carrying p- <i>entD</i> plasmid; Kan ^r Amp ^r	This study
SE15 <i>entD</i> +p- <i>clbA</i> (1)	<i>entD</i> mutant of strain SE15 carrying p- <i>clbA</i> (1) plasmid; Kan ^r Amp ^r	This study
M1/5	Enterobactin, aerobactin and yersiniabactin siderophores producer	Gift from U. Dobrindt
M1/5 <i>entD</i>	<i>entD</i> mutant of strain M1/5; Kan ^r	This study
M1/5 <i>clbA</i>	<i>clbA</i> mutant of strain M1/5; Kan ^r	This study
M1/5 <i>entD clbA</i>	<i>entD clbA</i> mutant of strain M1/5; Kan ^r	This study
M1/5 <i>entD clbA</i> +p- <i>entD</i>	<i>entD clbA</i> mutant of strain M1/5 carrying p- <i>entD</i> ; Kan ^r Amp ^r	This study
M1/5 <i>entD clbA</i> +p- <i>clbA</i> (1)	<i>entD clbA</i> mutant of strain M1/5 carrying p- <i>clbA</i> (1); Kan ^r Amp ^r	This study
M1/5 <i>entD clbA</i> +p- <i>clbA</i> (2)	<i>entD clbA</i> mutant of strain M1/5 carrying p- <i>clbA</i> (2); Kan ^r Amp ^r	This study
M1/5 <i>clbA</i> +p- <i>clbA</i> (1)	<i>clbA</i> mutant of strain M1/5 carrying p- <i>clbA</i> (1); Kan ^r Amp ^r	This study
M1/5 <i>clbA</i> +p- <i>sfp</i>	<i>clbA</i> mutant of strain M1/5 carrying p- <i>sfp</i> plasmid; Kan ^r Amp ^r	This study
M1/5 <i>clbA</i> +p- <i>pptT</i>	<i>clbA</i> mutant of strain M1/5 carrying p- <i>pptT</i> plasmid; Kan ^r Amp ^r	This study
SP15	Enterobactin, salmochelin, aerobactin and yersiniabactin siderophores producer	[53]
SP15 <i>clbA</i>	<i>clbA</i> mutant of strain SP15; Kan ^r	This study
SP15 <i>entD</i>	<i>entD</i> mutant of strain SP15; Kan ^r	This study
SP15 <i>entD clbA</i>	<i>entD clbA</i> mutant of strain SP15; Kan ^r	This study
SP15 <i>entD clbA</i> +p- <i>clbA</i> (2)	<i>clbA</i> mutant of strain SP15 carrying p- <i>clbA</i> (2); Kan ^r Amp ^r	This study
SP15 <i>entD clbA</i> +p- <i>entD</i>	<i>entD clbA</i> mutant of strain SP15 carrying p- <i>entD</i> ; Kan ^r Amp ^r	This study
Plasmids		
pACYC5.3L	<i>fyuA</i> -, <i>ybtA</i> -, <i>fyuA-luc</i> -, <i>irp6-8</i> , Cm ^r	Gift from W. Rabsch
p- <i>entD</i>	High copy number PCR-Script plasmid carrying <i>entD</i> gene; Amp ^r	This study
p- <i>clbA</i> (1)	Low copy number pASK75 plasmid carrying <i>clbA</i> gene; Amp ^r	Gift from U. Dobrindt
p- <i>clbA</i> (2)	pMB808, high copy number PCR-Script plasmid carrying <i>clbA</i> gene; Amp ^r	[22]
BAC <i>pk</i> s+	Bacterial artificial chromosome carrying the entire <i>pk</i> s island; Cm ^r	[21]
BAC <i>pk</i> s Δ <i>clbA</i>	Bacterial artificial chromosome carrying the entire <i>pk</i> s island with deleted <i>clbA</i> gene; Cm ^r	[22]

Table 1. Cont.

Strain or plasmid	Genotype or phenotype	Source or reference
p- <i>sfp</i>	Low copy number pET26b plasmid carrying <i>sfp</i> gene from <i>Bacillus subtilis</i> ; Amp ^r	Gift from C. Chalut
p- <i>pptT</i>	Low copy number pET28a plasmid carrying <i>pptT</i> gene from <i>Mycobacterium tuberculosis</i> ; Amp ^r	Gift from C. Chalut
p- <i>bpsA</i>	Low copy number pET26b plasmid carrying gene <i>bpsA</i> from <i>Streptomyces lavendulae</i> ; Kan ^r	Gift from C. Chalut

doi:10.1371/journal.ppat.1003437.t001

Quantification of yersiniabactin

The expression of the *fyuA* gene encoding the yersiniabactin receptor (FyuA) is known to be up-regulated in the presence of extracellular yersiniabactin [49]. Thus, yersiniabactin-dependent up-regulation of *fyuA* expression can be monitored by means of a *fyuA*-reporter fusion in the indicator strain [50].

Bacterial strains were cultivated in NBD medium, *i.e.* Nutrient Broth (NB) medium supplemented with 200 μ M α,α' -dipyridyl (Sigma), for 24 h at 37°C. Bacteria were pelleted by centrifugation and the supernatant was added to the indicator strain WR1542 carrying plasmid pACYC5.3L (kind gift of W. Rabsch, Wernigerode). The plasmid encodes all genes necessary for yersiniabactin uptake; *i.e.* *irp6*, *irp7*, *irp8*, *fyuA* and *yblA*. Additionally, the *fyuA* promoter region fused to the luciferase reporter gene is included on pACYC5.3L. After further 24 h of incubation at 37°C the indicator strain was pelleted and resuspended in bacterial lysis buffer (100 mM potassium phosphate buffer [pH 7.8], 2 mM EDTA, 1% [wt/vol] Triton X-100, 5 mg/ml bovine serum albumin, 1 mM dithiothreitol, 5 mg/ml lysozyme). Complete lysis was performed by incubation at room temperature for 20 min and repeated mixing. The samples were centrifuged and supernatants of lysates were analyzed by addition of luciferase reagent (20 mM Tricine-HCl (pH 7.8), 1.07 mM (MgCO₃)₄ Mg(OH)₂, 100 μ M EDTA, 470 μ M D(-) luciferin, 33.3 mM dithiothreitol, 270 μ M Li₃ coenzyme A, 530 μ M Mg-ATP). Luciferase activities were determined in triplicates using the multimode reader Berthold Tristar LB 941. Values were corrected by relating luciferase activity to the OD₆₀₀ of bacterial cultures grown 24 h in NBD medium. K12 *E. coli* strain DH5 α served as negative control. The experiments were repeated at least three times.

Detection and quantification of indigoidine

After overnight cultures in LB broth supplemented with the appropriate antibiotics, bacteria were diluted 1:10 in M9 minimal medium supplemented with 100 mM L-glutamine and 1 mM IPTG, and cultivated 16 h at 18–20°C under shaking [31,51].

Bacteria were then collected by centrifugation at 900 \times g for 5 min. At this speed, bacterial cells were pelleted while indigoidine still remained in the supernatant [51]. Indigoidine production was quantified by measuring the absorbance of blue-colored supernatant (OD_{612 nm}). The bacterial pellet was resuspended in PBS, and biomass was quantified by measuring the absorbance (OD_{450 nm}). Finally, the indigoidine production was normalized with the ratio Indigoidine/Biomass (*e.g.* ratio OD_{612 nm}/OD_{450 nm}).

Determination of the megalocytosis induced by colibactin

HeLa cells were maintained by serial passage in DMEM supplemented with 10% FCS, non-essential amino acids and 50 μ g/mL gentamicin. HeLa cells were dispensed in 96-well cell culture plate (5×10^3 cells/well). For bacterial infections, overnight LB broth cultures of *E. coli* were diluted in interaction medium (DMEM, 5% FCS, 25 mM HEPES) and cell cultures (~70% confluent) were infected with a multiplicity of infection (number of bacteria per HeLa cell at the onset of infection) of 3 to 400. Four hours post-inoculation, cells were washed 3 times with HBSS and incubated in cell culture medium 72 h with 200 μ g/mL gentamicin before protein staining with methylene blue (1% w/v in Tris-HCl 0.01M). The methylene blue was extracted with HCl 0.1N. The quantification of staining was measured at OD_{660 nm}.

Determination of the genotoxic effect induced by colibactin

The In Cell Western procedure was performed as described previously [27]. Briefly, HeLa cells were dispensed in 96-well cell culture plate (1.5×10^5 cells/200 μ L/wells). Twenty four hours later, cells were infected with *E. coli* strains for 4 h. Eight hours post-infection the cells were directly fixed in the plate with 4% paraformaldehyde. Paraformaldehyde was neutralized, and cells were permeabilized as previously described [27]. Cells were blocked with MAXblock Blocking medium (Active Motif,

Table 2. Oligonucleotides used in the study.

Primers	Sequences
entD-P1	GGGCGGATCGCTGCAATTTATGCGCTGCGGGAATATGGCTATAAATGTGTGCGTGTAGGCTGGAGCTGCTTC
entD-P2	TCACCTTGCCCTTAAATGCGCTCTCTTTGGCGGAAATGCCAGTGTACAGGCCATATGAATATCCTCCTTAG
entD-Up	CCCCCGGGGGGACGTACGTGGTATATGAGC
entD-Down	AACTGCAGAAGCACCTGCTTTACACTTTCG
entE-P1	TATCGACGGCGAGCGACAGTTGAGTTATCGGGAGCTGAATCAGCGCGCTGTAGGCTGGAGCTGCTTC
entE-P2	AAGCGCAGCTTTTTTCGCCATCAGCTCATCTTCCATGCTCACCAGTGCCATATGAATATCCTCCTTAG
JPN42	CAG ATA CAC AGA TAC CAT TCA
JPN46	CTA GAT TAT CCG TGG CGA TTC

doi:10.1371/journal.ppat.1003437.t002

Belgium) supplemented with phosphatase inhibitor PHOSTOP (Roche), followed by overnight incubation with rabbit monoclonal anti γ -H2AX (Cell Signaling) (1:200). An infrared fluorescent secondary antibody absorbing at 800 nm (IRDye™ 800CW, Rockland) was then applied (dilution 1:500). For DNA labeling, RedDot2 (Biotium) was used (dilution 1:500) together with the secondary antibody. The DNA and the γ -H2AX were simultaneously visualized using an Odyssey Infrared Imaging Scanner (Li-Cor ScienceTec, Les Ulis, France) with the 680 nm fluorophore (red color) and the 800 nm fluorophore (green dye). Relative fluorescence units from the scanning allowed a quantitative analysis. Relative fluorescent units for γ -H2AX per cell (as determined by γ -H2AX divided by DNA content) were divided by vehicle controls to determine percent change in phosphorylation of H2AX levels relative to control. All experiments were carried out in triplicate.

Mouse sepsis model

Animal experimentations were carried out in accordance with the European directive for the protection of animals used for scientific purposes. The protocols were validated by the local ethics committee on animal experiment "Comité d'éthique Midi Pyrénées pour l'expérimentation animale" which is affiliated to "Comité National de Réflexion Ethique sur l'Expérimentation Animale" and linked to the french ministry of research (Referenced protocols: PX-ANI-A2-94, 95, 96, 99, 100, and 101). Nine week old female C57BL/6J mice (JANVIER) were injected into the footpad with 10^8 ExPEC WT, *clbA* mutants, *entD* mutant, *entD clbA* mutant and *entD clbA* mutant complemented with *clbA* (*entD clbA*+p-*clbA*(2)) and *entD* (*entD clbA*+p-*entD*), together with intraperitoneal injection of 100 μ L of carbenicillin (1.6 mg/mL) or PBS. When required, mice were sacrificed by lethal anaesthesia (rompun/ketamine in 0.9% NaCl) 18 h post injection. The

abdominal cavity of anesthetized mouse was opened. The widest part of the posterior vena cava was localized and sectioned. Blood was collected by aspiration from the abdominal cavity. Spleens were surgically removed. Bacteria located in spleen cells were isolated from the mechanical dissociation of the splenic tissue using Precellys tissue homogenizer. Bacteria were quantified by plating of serial dilutions of blood and dissociated spleen on appropriate selective MacConkey agar. The antibiotics used to supplement the medium correspond to the resistance displayed by the different strains and are indicated in Table 1.

Statistical analysis

Statistical analyses were conducted using GraphPad Prism 5.0d. The mean \pm standard deviation (SD) is shown in figures, and *P* values were calculated using a one-way or two-way ANOVA followed by a Bonferroni post-test unless otherwise stated. For bacterial quantification, CFU by mg of spleen or mL of blood were log transformed for the analysis. A *P* value of less than 0.05 was considered statistically significant and is denoted by *. *P*<0.01 is denoted by ** and *P*<0.001 by ***.

Acknowledgments

We thank Dr. Damien Dubois for the *in silico* analysis of conserved domains of PPTases and for critical reading of the manuscript. We thank Claude Watrin for excellent technical help.

Author Contributions

Conceived and designed the experiments: PM IM JPN EO. Performed the experiments: PM IM GM MP CG DP MB MO. Analyzed the data: PM GM MO MA JPN SS EO. Contributed reagents/materials/analysis tools: MA CC. Wrote the paper: PM EO.

References

- Kaper JB, Nataro JP, Mobley HL (2004) Pathogenic *Escherichia coli*. *Nat Rev Microbiol* 2: 123–140.
- Russo TA, Johnson JR (2000) Proposal for a new inclusive designation for extraintestinal pathogenic isolates of *Escherichia coli*: ExPEC. *J Infect Dis* 181: 1753–1754.
- Chaudhuri RR, Henderson IR (2012) The evolution of the *Escherichia coli* phylogeny. *Infect Genet Evol* 12: 214–226.
- Gordon DM, Cowling A (2003) The distribution and genetic structure of *Escherichia coli* in Australian vertebrates: host and geographic effects. *Microbiology* 149: 3575–3586.
- Escobar-Paramo P, Grenet K, Le Menac'h A, Rode L, Salgado E, et al. (2004) Large-scale population structure of human commensal *Escherichia coli* isolates. *Appl Environ Microbiol* 70: 5698–5700.
- Escobar-Paramo P, Le Menac'h A, Le Gall T, Amorin C, Gouriou S, et al. (2006) Identification of forces shaping the commensal *Escherichia coli* genetic structure by comparing animal and human isolates. *Environ Microbiol* 8: 1975–1984.
- Hacker J, Blum-Oehler G, Muhldorfer I, Tschape H (1997) Pathogenicity islands of virulent bacteria: structure, function and impact on microbial evolution. *Mol Microbiol* 23: 1089–1097.
- Johnson JR, Delavari P, Kuskowski M, Stell AL (2001) Phylogenetic distribution of extraintestinal virulence-associated traits in *Escherichia coli*. *J Infect Dis* 183: 78–88.
- Bingen E, Picard B, Brahimi N, Mathy S, Desjardins P, et al. (1998) Phylogenetic analysis of *Escherichia coli* strains causing neonatal meningitis suggests horizontal gene transfer from a predominant pool of highly virulent B2 group strains. *J Infect Dis* 177: 642–650.
- Gao Q, Wang X, Xu H, Xu Y, Ling J, et al. (2012) Roles of iron acquisition systems in virulence of extraintestinal pathogenic *Escherichia coli*: salmochelin and aerobactin contribute more to virulence than heme in a chicken infection model. *BMC Microbiol* 12: 143.
- Garenaux A, Caza M, Dozois CM (2011) The Ins and Outs of siderophore mediated iron uptake by extra-intestinal pathogenic *Escherichia coli*. *Vet Microbiol* 153: 89–98.
- Copp JN, Neilan BA (2006) The phosphopantetheinyl transferase superfamily: phylogenetic analysis and functional implications in cyanobacteria. *Appl Environ Microbiol* 72: 2298–2305.
- Lambalot RH, Gehring AM, Flugel RS, Zuber P, LaCelle M, et al. (1996) A new enzyme superfamily - the phosphopantetheinyl transferases. *Chem Biol* 3: 923–936.
- Gehring AM, Bradley KA, Walsh CT (1997) Enterobactin biosynthesis in *Escherichia coli*: isochorismate lyase (EntB) is a bifunctional enzyme that is phosphopantetheinylated by EntD and then acylated by EntE using ATP and 2,3-dihydroxybenzoate. *Biochemistry* 36: 8495–8503.
- Fischbach MA, Lin H, Liu DR, Walsh CT (2005) *In vitro* characterization of IroB, a pathogen-associated C-glycosyltransferase. *Proc Natl Acad Sci U S A* 102: 571–576.
- Sorsa IJ, Dufke S, Heesemann J, Schubert S (2003) Characterization of an *iroBCDEV* gene cluster on a transmissible plasmid of uropathogenic *Escherichia coli*: evidence for horizontal transfer of a chromosomal virulence factor. *Infect Immun* 71: 3285–3293.
- Goetz DH, Holmes MA, Borregaard N, Bluhm ME, Raymond KN, et al. (2002) The neutrophil lipocalin NGAL is a bacteriostatic agent that interferes with siderophore-mediated iron acquisition. *Mol Cell* 10: 1033–1043.
- Schubert S, Picard B, Gouriou S, Heesemann J, Denamur E (2002) Yersinia high-pathogenicity island contributes to virulence in *Escherichia coli* causing extraintestinal infections. *Infect Immun* 70: 5335–5337.
- Schubert S, Rakin A, Heesemann J (2004) The Yersinia high-pathogenicity island (HPI): evolutionary and functional aspects. *Int J Med Microbiol* 294: 83–94.
- Bobrov AG, Geoffroy VA, Perry RD (2002) Yersiniabactin production requires the thioesterase domain of HMWP2 and YbtD, a putative phosphopantetheinylate transferase. *Infect Immun* 70: 4204–4214.
- Nougayrede JP, Homburg S, Taieb F, Boury M, Brzuszkiewicz E, et al. (2006) *Escherichia coli* induces DNA double-strand breaks in eukaryotic cells. *Science* 313: 848–851.
- Cuevas-Ramos G, Petit CR, Marq I, Boury M, Oswald E, et al. (2010) *Escherichia coli* induces DNA damage *in vivo* and triggers genomic instability in mammalian cells. *Proc Natl Acad Sci U S A* 107: 11537–11542.
- Arthur JC, Perez-Chanona E, Muhlbauer M, Tomkovich S, Uronis JM, et al. (2012) Intestinal inflammation targets cancer-inducing activity of the microbiota. *Science* 338: 120–123.
- Olier M, Marq I, Salvador-Cartier C, Secher T, Dobrindt U, et al. (2012) Genotoxicity of *Escherichia coli* Nissle 1917 strain cannot be dissociated from its probiotic activity. *Gut Microbes* 3: 501–509.

25. Johnson JR, Johnston B, Kuskowski MA, Nougayrede JP, Oswald E (2008) Molecular epidemiology and phylogenetic distribution of the *Escherichia coli* *pks* genomic island. *J Clin Microbiol* 46: 3906–3911.
26. Putze J, Hennequin C, Nougayrede JP, Zhang W, Homburg S, et al. (2009) Genetic structure and distribution of the colibactin genomic island among members of the family Enterobacteriaceae. *Infect Immun* 77: 4696–4703.
27. Audebert M, Riu A, Jacques C, Hillenweck A, Jamin EL, et al. (2010) Use of the gamma H2AX assay for assessing the genotoxicity of polycyclic aromatic hydrocarbons in human cell lines. *Toxicology Letters* 199: 182–192.
28. Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, et al. (1998) Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* 393: 537–544.
29. Chalut C, Botella L, de Sousa-D'Auria C, Houssin C, Guilhot C (2006) The nonredundant roles of two 4'-phosphopantetheinyl transferases in vital processes of *Mycobacterium*. *Proc Natl Acad Sci U S A* 103: 8511–8516.
30. Quadri LE, Weinreb PH, Lei M, Nakano MM, Zuber P, et al. (1998) Characterization of Sfp, a *Bacillus subtilis* phosphopantetheinyl transferase for peptidyl carrier protein domains in peptide synthetases. *Biochemistry* 37: 1585–1595.
31. Muller M, Auslander S, Auslander D, Kemmer C, Fussenegger M (2012) A novel reporter system for bacterial and mammalian cells based on the non-ribosomal peptide indigoidine. *Metab Eng* 14: 325–335.
32. Owen JG, Robins KJ, Parachin NS, Ackerley DF (2012) A functional screen for recovery of 4'-phosphopantetheinyl transferase and associated natural product biosynthesis genes from metagenome libraries. *Environ Microbiol* 14: 1198–1209.
33. Reuter K, Mofid MR, Marahiel MA, Ficner R (1999) Crystal structure of the surfactin synthetase-activating enzyme sfp: a prototype of the 4'-phosphopantetheinyl transferase superfamily. *EMBO J* 18: 6823–6831.
34. Yasgar A, Foley TL, Jadhav A, Inglesse J, Burkart MD, et al. (2010) A strategy to discover inhibitors of *Bacillus subtilis* surfactin-type phosphopantetheinyl transferase. *Mol Biosyst* 6: 365–375.
35. Duckworth BP, Aldrich CC (2010) Development of a high-throughput fluorescence polarization assay for the discovery of phosphopantetheinyl transferase inhibitors. *Anal Biochem* 403: 13–19.
36. Fischbach MA, Lin H, Zhou L, Yu Y, Abergel RJ, et al. (2006) The pathogen-associated *iroA* gene cluster mediates bacterial evasion of lipocalin 2. *Proc Natl Acad Sci U S A* 103: 16502–16507.
37. Dozois CM, Daigle F, Curtiss R, 3rd (2003) Identification of pathogen-specific and conserved genes expressed *in vivo* by an avian pathogenic *Escherichia coli* strain. *Proc Natl Acad Sci U S A* 100: 247–252.
38. Caza M, Lepine F, Dozois CM (2011) Secretion, but not overall synthesis, of catecholate siderophores contributes to virulence of extraintestinal pathogenic *Escherichia coli*. *Mol Microbiol* 80: 266–282.
39. Nowrouzian FL, Oswald E (2012) *Escherichia coli* strains with the capacity for long-term persistence in the bowel microbiota carry the potentially genotoxic *pks* island. *Microb Pathog* 53: 180–182.
40. Tenaillon O, Skurnik D, Picard B, Denamur E (2010) The population genetics of commensal *Escherichia coli*. *Nat Rev Microbiol* 8: 207–217.
41. Kohler CD, Dobrindt U (2011) What defines extraintestinal pathogenic *Escherichia coli*? *Int J Med Microbiol* 301: 642–647.
42. Dobrindt U, Hochhut B, Hentschel U, Hacker J (2004) Genomic islands in pathogenic and environmental microorganisms. *Nat Rev Microbiol* 2: 414–424.
43. Schubert S, Darlu P, Clermont O, Wieser A, Magistro G, et al. (2009) Role of intraspecies recombination in the spread of pathogenicity islands within the *Escherichia coli* species. *PLoS Pathog* 5: e1000257.
44. Henderson JP, Crowley JR, Pinkner JS, Walker JN, Tsukayama P, et al. (2009) Quantitative metabolomics reveals an epigenetic blueprint for iron acquisition in uropathogenic *Escherichia coli*. *PLoS Pathog* 5: e1000305.
45. Johnson JR, Kuskowski MA, Gajewski A, Soto S, Horcajada JP, et al. (2005) Extended virulence genotypes and phylogenetic background of *Escherichia coli* isolates from patients with cystitis, pyelonephritis, or prostatitis. *J Infect Dis* 191: 46–50.
46. Toh H, Oshima K, Toyoda A, Ogura Y, Ooka T, et al. (2010) Complete genome sequence of the wild-type commensal *Escherichia coli* strain SE15, belonging to phylogenetic group B2. *J Bacteriol* 192: 1165–1166.
47. Datsenko KA, Wanner BL (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* 97: 6640–6645.
48. Schwyn B, Neilands JB (1987) Universal chemical assay for the detection and determination of siderophores. *Anal Biochem* 160: 47–56.
49. Pelludat C, Rakin A, Jacobi CA, Schubert S, Hecsemann J (1998) The yersiniabactin biosynthetic gene cluster of *Yersinia enterocolitica*: organization and siderophore-dependent regulation. *J Bacteriol* 180: 538–546.
50. Schubert S, Cuenca S, Fischer D, Hecsemann J (2000) High-pathogenicity island of *Yersinia pestis* in enterobacteriaceae isolated from blood cultures and urine samples: prevalence and functional expression. *J Infect Dis* 182: 1268–1271.
51. Yu D, Xu F, Valiente J, Wang S, Zhan J (2012) An indigoidine biosynthetic gene cluster from *Streptomyces chromofuscus* ATCC 49982 contains an unusual IndB homologue. *J Ind Microbiol Biotechnol*.
52. Ochman H, Selander RK (1984) Standard reference strains of *Escherichia coli* from natural populations. *J Bacteriol* 157: 690–693.
53. Johnson JR, Oswald E, O'Bryan TT, Kuskowski MA, Spanjaard L (2002) Phylogenetic distribution of virulence-associated genes among *Escherichia coli* isolates associated with neonatal bacterial meningitis in the Netherlands. *J Infect Dis* 185: 774–784.

Article 4

ChemComm



COMMUNICATION

View Article Online
View Journal

Cite this: DOI: 10.1039/c5cc02718g

Received 1st April 2015,
Accepted 7th July 2015

DOI: 10.1039/c5cc02718g

www.rsc.org/chemcomm

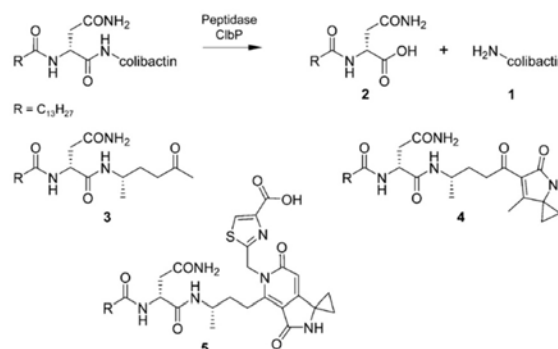
Colibactin biosynthesis and biological activity depend on the rare aminomalonyl polyketide precursor†

A. O. Brachmann,^a C. Garcie,^{bcd} V. Wu,^a P. Martin,^{bcd} R. Ueoka,^a E. Oswald^{bcd} and J. Piet^{*a}

The as-yet unidentified *E. coli* metabolite colibactin induces DNA damage in eukaryotic cells and promotes tumorigenesis. Its wide distribution in pathogenic and probiotic strains has raised great interest in its structure and biosynthesis. Here we show that colibactin formation involves a rare aminomalonyl unit used as a building block.

The cryptic natural product colibactin (**1**) has received much attention due to its association with human diseases, widespread occurrence, and elusive identity. In various pathogenic *E. coli* strains, the colibactin biosynthetic genes (termed *clb* or *pks*) are positively correlated with DNA double-strand breaks in host cells,¹ colon inflammation and damage in mammalian epithelial tissues.² They are also involved in cellular senescence accompanied by tumor promotion in colitis-associated colorectal cancer.³ The *clb* gene cluster encodes a hybrid nonribosomal peptide synthetase–polyketide synthase (NRPS–PKS) system and is widely distributed among different enterobacteria,⁴ including coral⁵ and honeybee symbionts.⁶ Considering the toxic effect in pathogens, it is remarkable that the probiotic *E. coli* Nissle 1917 strain, which also carries the *clb* genes, is being used extensively for the treatment of inflammatory bowel disease.⁷

Despite substantial efforts by several laboratories to identify **1**,^{1b,5,7,8} its structure remains unknown. First insights into the biosynthesis were provided by the identification of ClbP as a membrane bound peptidase that is essential for the maturation of a hypothetical pre-colibactin to induce genotoxicity.^{8a} Based on precedence from other pathways,⁹ biochemical experiments,^{8c} and *in vivo* studies,^{8b} the prodrug activation mechanism should involve removal of an *N*-fatty acyl-*D*-asparaginyl moiety by the peptidase, releasing the active compound (Scheme 1). Recently, MS-guided identification of intermediate-derived metabolites



Scheme 1 Peptidase activation mechanism for colibactin (top) and detected metabolites **2–5** from *E. coli* cultures.

2–5 revealed further structural features.^{8d–h} **4** and **5** contain an unusual 1-aminocyclopropane-1-carboxylic acid (ACC) unit that contributes to the instability of **1** and also the mechanism of action.^{8e,g} Comparison of **4** and thiazol-containing **5** with the NRPS–PKS architecture (Fig. S1, ESI†) suggests that two further polyketide synthase modules (ClbO and first module of ClbK) are involved in the incorporation of uncharacterized polyketide building blocks.

Here we present biochemical data on eight proteins of the biosynthetic pathway. The results suggest that the colibactin pathway employs, in addition to the amino acids glycine and cysteine, a rare 2-aminomalonyl (AM) precursor. Based on MS studies and bioactivity profiles of mutant strains, the AM building block likely belongs to an as yet unknown portion of the colibactin structure.

Initial experiments were motivated by the bioinformatics analysis of the *clb* cluster in the meningitis-associated strain *E. coli* IHE3034 (Fig. S1, ESI†) that suggested the homology of ClbD and ClbF to dehydrogenases (Table S1, ESI†). Close relatives are ZmaG and ZmaI from zwittermixin (*zma*) biosynthesis, which were shown to jointly catalyze the formation of the unusual PKS extender AM-ACP from seryl-ACP, indicating related functions

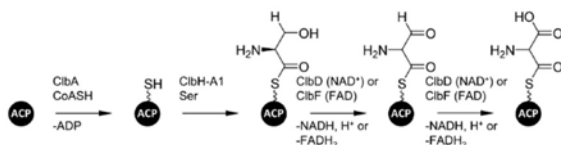
^a Institute of Microbiology, Eigenössische Technische Hochschule (ETH) Zurich, Vladimir-Prelog-Weg 1-5/10, 8093 Zurich, Switzerland. E-mail: jpiel@ethz.ch

^b Inserm, UMR1043, INRA, USC 1360, CNRS, UMR5282, Toulouse, F-31000, France

^c Université de Toulouse III, Toulouse, F-31000, France

^d CHU Toulouse, Service de Bactériologie-Hygiène, Toulouse, F-31000, France

† Electronic supplementary information (ESI) available: Experimental procedures and additional figures and tables. See DOI: 10.1039/c5cc02718g



Scheme 2 Proposed formation of aminomalonyl-ACP and role of colibactin biosynthetic enzymes.

in the *clb* pathway (Scheme 2).¹⁰ Interestingly, zwittermicin biosynthesis also features a similar prodrug activation mechanism as proposed for **1**.^{10a} In the *zma* pathway, seryl-ACP is generated by an isolated adenylation enzyme that loads a cognate, free-standing ACP. A close *clb* homolog of this ACP was identified as ClbE. The only matching adenylation enzyme with predicted¹¹ serine specificity, however, was an A domain (ClbH-A1) at the N terminus of the unusual NRPS protein ClbH. In contrast to the canonical NRPS architecture, ClbH lacks a peptidyl carrier protein (PCP) domain downstream of the A domain, but is directly connected to a condensation (C) domain (overall architecture A-C-A-PCP). Resequencing of the NRPS region confirmed the correct architecture of ClbH.

To biochemically determine the amino acid substrate of ClbH-A1, we expressed and characterized the domain using a previously described γ -[¹⁸O₄]-ATP pyrophosphate exchange assay.¹² In addition to proteinogenic amino acids, we also tested several non-standard amino acids including ACC and *S*-adenosyl-L-methionine (SAM). These were previously considered together with L-Met as possible substrates for either ClbH-A1 or ClbH-A2, the second A domain on ClbH, to account for the presence of ACC in **4** and **5**.^{8f-g} However, among all tested amino acids, significant conversion was observed only for L-Ser (Fig. 1), in agreement with the prediction. We measured low exchange rates for ACC, but previous feeding experiments suggest that this compound is not directly incorporated^{18g} and should instead be derived from NRPS-activated L-Met or SAM.^{8f-g} Our data show that these two amino acids are unlikely substrates for ClbH-A1. We also expressed and assayed ClbH-A2, but did not detect significant exchange rates with any test substrate, including L-Met, SAM, and ACC (Fig. S2c, ESI[†]). In conclusion, our data support the role of ClbH-A1 in the provision of a seryl rather than an ACC unit.

We next investigated whether ClbH-A1 in combination with other enzymes is also able to provide AM-ACP. As additional components, we individually expressed the ACP ClbE, predicted

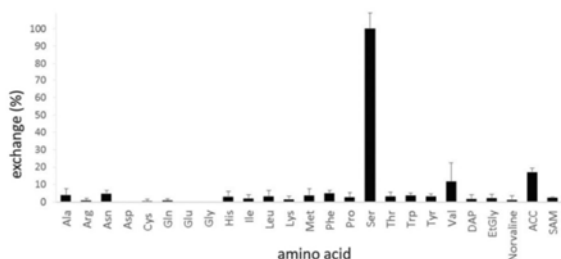


Fig. 1 Exchange rates for the A domain ClbH-A1 and various amino acids. EtGly, ethylglycine; DAP, 2,3-diaminopropionic acid.

dehydrogenases ClbD and ClbF, and the cluster-encoded PPTase ClbA. The PPTase was suspected to transform the *apo*-ACP into the active *holo*-ACP similar to other thiotemplate pathways.¹³ The PPTase activity was assayed in mixtures containing ClbA, ClbE, and the substrate coenzyme A. Analysis by liquid chromatography-high-resolution MS (LC-HRMS) showed the appearance of a new species with a characteristic mass shift of +340 Da as compared to unmodified ClbA, consistent with a phosphopantetheinyl moiety attached to the ACP (Fig. S3, ESI[†]). ClbA was previously shown to replace the siderophore-associated NRPS PPTase EntD in enterobactin biosynthesis.¹⁴ The activation of ClbE suggests that ClbA belongs to the promiscuous PPTases that can activate PCP as well as ACP domains.

With the *holo*-ACP in hand, we tested whether it is loaded with L-Ser after addition of ClbH-A1. The formation of seryl-ACP was supported by a mass shift of +87 Da for one of the major ACP species (Fig. 2a and Fig. S4, ESI[†]). This interpretation was also corroborated by detecting an additional mass shift of +3 Da after using ¹³C₃-labeled L-Ser (Fig. S5, ESI[†]). Further oxidation to AM-ACP was tested by adding both dehydrogenases ClbF and ClbD and their respective cofactors NAD⁺ and FAD (Fig. 2). Here, additional ions were only observed when both dehydrogenases were added (Fig. 2e and Table S2, ESI[†]), whereas neither ClbF nor ClbD alone was able to yield new ACP intermediates (Fig. 2c and d). For assays containing both dehydrogenases, MS analysis revealed a major new species with a detected mass corresponding to glycyl-ACP. This product was previously shown to be indicative of AM-ACP due to spontaneous decarboxylation.^{10a}

The data raised the question of whether the AM unit serves as a precursor to an alternative ACC pathway or is incorporated into an as yet unknown portion of **1**. Based on the position of *clbH* in the gene cluster, which corresponds to the timing of ACC incorporation, the latter hypothesis is counterintuitive. To nevertheless test both possibilities and to further test the relevance of AM for colibactin production, we generated two

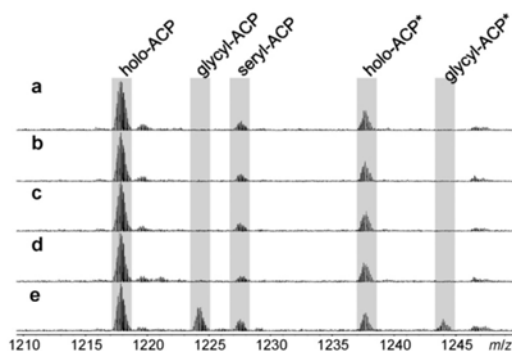


Fig. 2 Isotopically resolved spectra of ACP species [M + 9H]⁹⁺. (a) ACP species after addition of PPTase ClbA, ACP ClbE, adenylation domain ClbH A1 and their respective cofactors. The further spectra show the same experiment with the addition of (b) boiled dehydrogenases ClbD and ClbF, (c) only ClbD, (d) only ClbF, and (e) both dehydrogenases ClbD and ClbF. ACPs labeled with an asterisk correspond to *N*-glycosylated species. Table S2 (ESI[†]) shows the identified ACP species in (e), their detected mass after deconvolution, and relative abundance.

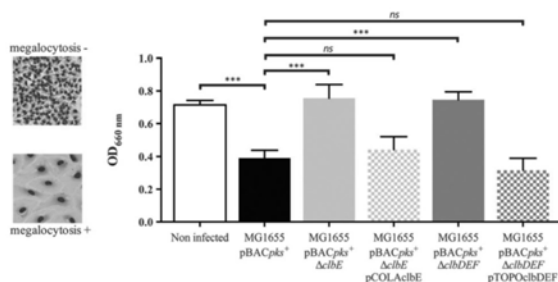


Fig. 3 Megalocytosis quantification for *clbE* and *clbDEF* mutants, and complemented derivatives (right). HeLa cells infected with the *pks*⁺ strain harbor a megalocytosis phenotype (left, bottom), while those infected with *clbE* or *clbDEF* mutants show a normal phenotype (left, top). Complementation of these mutants by functional *clbE* or *clbDEF* genes restores the megalocytosis effect. Multiplicity of infection: MOI = 400. Statistical analysis: one-way Anova. ***: $p < 0.001$, ns: not significant.

deletion mutants in the producer *E. coli* MG1655 pBACpks⁺. The mutant *pks*⁺ Δ*clbE* lacked the ACP gene, whereas the mutant *pks*⁺ Δ*clbDEF* carried additional deletions of both dehydrogenase genes. The production of **1** was assessed in the resulting strains through quantification of megalocytosis, which correlates with DNA double strand breaks resulting from the genotoxic effect of colibactin.¹ HeLa cells were infected with the different strains for 4 hours, fixed, and stained with methylene blue in order to quantify the megalocytosis effect. These assays revealed that the inactivation of the *clbE* gene or the *clbDEF* operon abrogated the colibactin effect (Fig. 3). The transformation of these mutants with plasmids carrying the functional wild type *clbE* or *clbDEF* genes resulted in a complete restoration of megalocytosis (Fig. 3).

The further high-resolution mass spectrometric analysis of the deletion mutants revealed that ACC-containing compounds **4** and **5** are still produced with comparable intensities to colibactin-positive strains (Fig. S6, ESI[†]). These results suggest that *clbDEF* and thus AM are not involved in ACC formation, but on the other hand are indispensable for colibactin production as assessed by the megalocytosis assay. It is therefore likely that AM is incorporated into the unknown part of the colibactin structure. This is also supported by Ser feeding experiments of the Müller and Crawford groups that resulted in poor labeling of the known intermediates.^{8f,g} Candidates for PKS modules that could process the AM unit are present in ClbK and ClbO (Fig. S1, ESI[†]). Both contain highly aberrant domains resembling acyltransferases (ATs) but lacking the conserved active site serine residue within a GSHxG motif that is normally needed for building block transfer to the ACP (Fig. S7, ESI[†]).¹⁵ In addition to these integrated AT-like domains, the cluster encodes a candidate for a free-standing AT, ClbG. The protein features the GHSxG motif, but a HAFH motif, which in malonyl-specific ATs correlates with substrate specificity,¹⁶ is changed to VPYH. A related motif (GPFH) is also present in the stand-alone AT ZmaF that transfers an aminomalonyl unit in zwittermixin biosynthesis.¹⁰ Thus, ClbG is a good candidate for aminomalonyl transfer in colibactin biosynthesis. Aminomalonyl building blocks

rarely occur in polyketide biosynthesis and to our knowledge have been reported only for zwittermixin¹⁰ and guadinomine.¹⁷

Finally, we also wished to obtain biochemical data for other uncharacterized colibactin enzymes and therefore expressed the A domains of NRPS modules ClbJ-A1, ClbJ-A2, and ClbK-A2 (the numbers refer to the module number within the protein). While ClbJ-A2 yielded only insoluble protein under various conditions, the other A domains were obtained in soluble, active form. The adenylation assays showed clear preference for Gly by ClbJ-A1 and for Cys by ClbK-A2, but other amino acids were not or poorly accepted (Fig. S2, ESI[†]). Gly is a component of intermediate **5**, thus assigning ClbJ to this moiety. ClbK corresponds to an as-yet uncharacterized portion.

In summary, our functional study of eight protein components of the *clb* pathway provides insights into the missing regions of colibactin. These are predicted to contain aminomalonyl as well as Cys units. The NRPS architecture further suggests that the Cys residue is converted into a thiazol moiety. This information could aid in the targeted isolation and characterization of colibactin as a basis to study the pharmacology of this elusive and intriguing metabolite.

Notes and references

- (a) J. P. Nougayrède, S. Homburg, F. Taieb, M. Boury, E. Brzuszkiewicz, G. Gottschalk, C. Buchrieser, J. Hacker, U. Dobrindt and E. Oswald, *Science*, 2006, **313**, 848–851; (b) G. Cuevas-Ramos, C. R. Petit, I. Marceq, M. Boury, E. Oswald and J. P. Nougayrède, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 11537–11542.
- J. Raisch, E. Buc, M. Bonnet, P. Sauvanet, E. Vazeille, A. de Vallee, P. Dechelotte, C. Darcha, D. Pezet, R. Bonnet, M. A. Bringer and A. Darfeuille-Michaud, *World J. Gastroenterol.*, 2014, **20**, 6560–6572.
- (a) T. Secher, A. Samba-Louaka, E. Oswald and J. P. Nougayrède, *PLoS One*, 2013, **8**, e77157; (b) A. Cougnoux, G. Dalmasso, R. Martinez, E. Buc, J. Delmas, L. Gibold, P. Sauvanet, C. Darcha, P. Dechelotte, M. Bonnet, D. Pezet, H. Wodrich, A. Darfeuille-Michaud and R. Bonnet, *Gut*, 2014, **63**, 1932–1942.
- (a) J. Putze, C. Hennequin, J. P. Nougayrède, W. Zhang, S. Homburg, H. Karch, M. A. Bringer, C. Fayolle, E. Carniel, W. Rabsch, T. A. Oelschlaeger, E. Oswald, C. Forestier, J. Hacker and U. Dobrindt, *Infect. Immun.*, 2009, **77**, 4696–4703; (b) J. R. Johnson, B. Johnston, M. A. Kuskowski, J. P. Nougayrède and E. Oswald, *J. Clin. Microbiol.*, 2008, **46**, 3906–3911.
- V. Bondarev, M. Richter, S. Romano, J. Piel, A. Schwedt and H. N. Schulz-Vogt, *Environ. Microbiol.*, 2013, **15**, 2095–2113.
- P. Engel, M. I. Vizcaino and J. M. Crawford, *Appl. Environ. Microbiol.*, 2014, **81**, 1502–1512.
- S. Homburg, E. Oswald, J. Hacker and U. Dobrindt, *FEMS Microbiol. Lett.*, 2007, **275**, 255–262.
- (a) A. Cougnoux, L. Gibold, F. Robin, D. Dubois, N. Pradel, A. Darfeuille-Michaud, G. Dalmasso, J. Delmas and R. Bonnet, *J. Mol. Biol.*, 2012, **424**, 203–214; (b) X. Bian, J. Fu, A. Plaza, J. Herrmann, D. Pistorius, A. F. Stewart, Y. Zhang and R. Müller, *ChemBioChem*, 2013, **14**, 1194–1197; (c) C. A. Brotherton and E. P. Balskus, *J. Am. Chem. Soc.*, 2013, **135**, 3359–3362; (d) M. I. Vizcaino, P. Engel, E. Trautman and J. M. Crawford, *J. Am. Chem. Soc.*, 2014, **136**, 9244–9247; (e) C. A. Brotherton, M. Wilson, G. Byrd and E. P. Balskus, *Org. Lett.*, 2015, **17**, 2294; (f) X. Bian, A. Plaza, Y. Zhang and R. Müller, *Chem. Sci.*, 2015, **6**, 3154–3160; (g) M. I. Vizcaino and J. M. Crawford, *Nat. Chem.*, 2015, **7**, 411–417; (h) Z. R. Li, Y. Li, J. Y. Lai, J. Tang, B. Wang, L. Lu, G. Zhu, X. Wu, Y. Xu and P. Y. Qian, *ChemBioChem*, 2015, DOI: 10.1002/cbic.201500239.
- D. Reimer, K. M. Pos, M. Thines, P. Grun and H. B. Bode, *Nat. Chem. Biol.*, 2011, **7**, 888–890.
- (a) Y. A. Chan and M. G. Thomas, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 14349–14354; (b) B. M. Kevany, D. A. Rasko and M. G. Thomas,

Communication

ChemComm

- Appl. Environ. Microbiol.*, 2009, **75**, 1144–1155; (c) Y. A. Chan and M. G. Thomas, *Biochemistry*, 2010, **49**, 3667–3677.
- 11 (a) T. Stachelhaus, H. D. Mootz and M. A. Marahiel, *Chem. Biol.*, 1999, **6**, 493–505; (b) G. L. Challis, J. Ravel and C. A. Townsend, *Chem. Biol.*, 2000, **7**, 211–224.
- 12 V. V. Phelan, Y. Du, J. A. McLean and B. O. Bachmann, *Chem. Biol.*, 2009, **16**, 473–478.
- 13 R. H. Lambalot, A. M. Gehring, R. S. Flugel, P. Zuber, M. LaCelle, M. A. Marahiel, R. Reid, C. Khosla and C. T. Walsh, *Chem. Biol.*, 1996, **3**, 923–936.
- 14 P. Martin, I. Marcq, G. Magistro, M. Penary, C. Garcie, D. Payros, M. Boury, M. Olier, J. P. Nougayrede, M. Audebert, C. Chalut, S. Schubert and E. Oswald, *PLoS Pathog.*, 2013, **9**, e1003437.
- 15 S. Smith and S. C. Tsai, *Nat. Prod. Rep.*, 2007, **24**, 1041–1072.
- 16 F. Del Vecchio, H. Petkovic, S. G. Kendrew, L. Low, B. Wilkinson, R. Lill, J. Cortes, B. A. Rudd, J. Staunton and P. F. Leadlay, *J. Ind. Microbiol. Biotechnol.*, 2003, **30**, 489–494.
- 17 T. C. Holmes, A. E. May, K. Zaleta-Rivera, J. G. Ruby, P. Skewes-Cox, M. A. Fischbach, J. L. DeRisi, M. Iwatsuki, S. Omura and C. Khosla, *J. Am. Chem. Soc.*, 2012, **134**, 17797–17806.

Article 5

AMERICAN
SOCIETY FOR
MICROBIOLOGYInfection and
ImmunityCrossMark
click for updatesIron Homeostasis Regulates the Genotoxicity of *Escherichia coli* That Produces ColibactinSophie Tronnet,^{a,b} Christophe Garcia,^{a,b,c} Nadine Rehm,^{d,e} Ulrich Dobrindt,^{d,e} Eric Oswald,^{a,b,c} Patricia Martin^{a,b,c}IRSD, Université de Toulouse, INSERM, INRA, ENVT, UPS, Toulouse, France^a; Université Toulouse III Paul Sabatier, Toulouse, France^b; CHU Toulouse, Service de Bactériologie-Hygiène, Toulouse, France^c; Institute of Hygiene, University of Münster, Münster, Germany^d; Interdisciplinary Center for Clinical Research (IZKF), University of Münster, Münster, Germany^e

The genotoxin colibactin is a secondary metabolite produced by a variety of pathogenic enterobacteria. Its biosynthesis requires the enzymatic activity of the phosphopantetheinyl transferase (PPTase) ClbA. We previously showed that ClbA can also contribute to the production of siderophores. Because the biosynthesis of siderophores is regulated by iron availability, we hypothesized that iron could also modulate the production of colibactin through the transcriptional regulation of *clbA*. This study revealed an increased transcription of *clbA* under iron-limiting conditions and a decrease of *clbA* expression in iron-rich media. We demonstrate that *clbA* transcription is regulated by both the ferric uptake regulator (Fur) and the small regulatory noncoding RNA RyhB. We evidenced that the regulation of the transcription of *clbA* by Fur and RyhB leads to the regulation of colibactin production. This work highlights the complex mechanism of regulation of an important virulence factor by the two major regulators of bacterial iron homeostasis, making iron a key environmental factor contributing to bacterial virulence and carcinogenesis.

In the human body, most of the iron is intracellular or sequestered by proteins, which makes it unavailable to invading microbes. This process, i.e., nutritional immunity, is a critical host defense strategy against bacterial pathogens that perceive the scarcity of iron within the vertebrate host as a means to sense that they are in their host (1). The global response to iron deprivation in bacteria relies on the ferric uptake regulator (Fur) protein and the small regulatory noncoding RNA (sRNA) RyhB (2–4). In the classical Fur regulation paradigm, Fur binds ferrous ions and the dimeric Fe²⁺-Fur complex (holo-Fur) recognizes target sequences upstream of iron-regulated genes and represses their transcription. However, nowadays numerous reports support four modes of Fur regulation, i.e., apo- and holo-Fur activation and repression, establishing a significant deviation from the classical model of Fur regulation (5). In bacteria, archaea, and eukaryotes, small riboregulators have been shown to mediate posttranscriptional mechanisms of gene regulation (6). In many cases, sRNAs form base pairs with and sequester mRNA ribosome-binding sites, resulting in translational repression and accelerated transcript decay. In contrast, a growing number of examples of translational activation and mRNA stabilization by sRNAs have now been documented. A given sRNA often employs a conserved region to interact with and regulate both repressed and activated targets (6, 7).

To counteract iron deprivation, bacterial pathogens synthesize small iron-scavenging molecules, i.e., siderophores, that are crucial for their survival and play a significant role in virulence (8–10). Siderophores are nonribosomal peptides (NRP) or polyketide (PK)-NRP hybrids (11). A prerequisite for the synthesis of all NRPs, PKs, and PK-NRP hybrids is the posttranslational attachment of P-Pant arms from coenzyme A (CoA) to a conserved serine residue of a carrier protein, converting inactive apo-synthases to active holo-synthases. This reaction is catalyzed by members of the type II family of phosphopantetheinyl transferases (PPTases) (12). The core genome of *Escherichia coli* codes for two type II PPTases, i.e., EntD and YieE (*ECK3705*, *b3712*, *JW3690*, *JW3690*) (13). Whereas the role of YieE remains elusive, EntD is

the PPTase involved in the biosynthesis of the siderophores enterobactin, salmochelin, and yersiniabactin.

An additional type II PPTase, i.e., ClbA, is encoded on the *pks* gene cluster, responsible for the synthesis of the PK-NRP hybrid colibactin in diverse pathogenic enterobacteria, including approximately 50% of the *E. coli* strains that belong to phylogenetic group B2 (14). Colibactin is a genotoxin that induces DNA double-strand breaks, senescence, and chromosomal abnormalities in eukaryotic cells (15–17) and was demonstrated to be a bona fide virulence factor in a mouse model of sepsis (18) and in a rat model of neonatal meningitis (19). We have recently highlighted the existence of cross talk between colibactin and siderophore biosynthesis (20). Indeed, ClbA can replace EntD and contribute to siderophore biosynthesis. Therefore, we speculated that this interplay could constitute a fitness advantage for *E. coli* and that iron could play a key role in this connection.

Here we show that the expression of *clbA* is controlled by iron bioavailability and that the production of colibactin is regulated by Fur and RyhB. Therefore, iron could constitute a key environmental factor contributing to the virulence and carcinogenicity of *E. coli* strains producing colibactin.

Received 30 July 2016 Returned for modification 24 August 2016

Accepted 3 September 2016

Accepted manuscript posted online 12 September 2016

Citation Tronnet S, Garcia C, Rehm N, Dobrindt U, Oswald E, Martin P. 2016. Iron homeostasis regulates the genotoxicity of *Escherichia coli* that produces colibactin. *Infect Immun* 84:000–000. doi:10.1128/IAI.00659-16.

Editor: A. J. Bäuml, University of California, Davis

Address correspondence to Eric Oswald, eric.oswald@inserm.fr, or Patricia Martin, patricia.martin@inserm.fr.

Supplemental material for this article may be found at <http://dx.doi.org/10.1128/IAI.00659-16>.

Copyright © 2016, American Society for Microbiology. All Rights Reserved.

Tronnet et al.

TABLE 1 Strains, plasmids, and primers used in this study

Strain, plasmid, or primer	Description or sequence ^a	Reference or source
<i>E. coli</i> strains		
Nissle 1917 <i>clbA-lux</i>	Luciferase fusion of the gene <i>clbA</i> , Kan ^r	22
Nissle 1917 <i>clbR-lux</i>	Luciferase fusion of the gene <i>clbR</i> , Kan ^r	22
Nissle 1917 <i>clbB-lux</i>	Luciferase fusion of the gene <i>clbB</i> , Kan ^r	22
Nissle 1917 <i>clbQ-lux</i>	Luciferase fusion of the gene <i>clbQ</i> , Kan ^r	22
Nissle 1917 <i>clbA-lux ΔryhB</i>	<i>ryhB</i> mutant strain, Kan ^r Cm ^r	This study
Nissle 1917 <i>clbA-lux ΔryhB/p-ryhB</i>	<i>ryhB</i> mutant strain carrying p- <i>ryhB</i> , Kan ^r Cm ^r Amp ^r	This study
Nissle 1917 <i>clbA-lux Δfur</i>	<i>fur</i> mutant strain, Kan ^r Cm ^r	This study
Nissle 1917 <i>clbA-lux Δfur/p-fur</i>	<i>fur</i> mutant strain carrying p- <i>fur</i> , Kan ^r Cm ^r Amp ^r	This study
M1/5	Colibactin producer, commensal strain	20
M1/5 <i>ΔryhB</i>	<i>ryhB</i> mutant strain, Cm ^r	This study
M1/5 <i>ΔryhB/p-ryhB</i>	<i>ryhB</i> mutant strain carrying p- <i>ryhB</i> , Cm ^r Amp ^r	This study
M1/5 <i>Δfur</i>	<i>fur</i> mutant strain, Kan ^r	This study
M1/5 <i>Δfur/p-fur</i>	<i>fur</i> mutant strain carrying p- <i>fur</i> , Kan ^r Amp ^r	This study
Nissle 1917	Colibactin producer, probiotic strain	16
Nissle 1917 <i>ΔclbA/pBAD33-clbA</i>		This study
Nissle 1917 <i>ΔclbA/pBAD33-clbA</i>	<i>clbA</i> mutant strain carrying pBAD33- <i>clbA</i> , Kan ^r	This study
Nissle 1917 <i>ΔclbA/pBAD33-clbAΔclbR</i>	<i>clbA</i> mutant strain carrying pBAD33- <i>clbAΔclbR</i> , Kan ^r	This study
Nissle 1917 <i>ΔryhB</i>	<i>ryhB</i> mutant strain, Cm ^r	This study
Nissle 1917 <i>ΔryhB/p-ryhB</i>	<i>ryhB</i> mutant strain carrying p- <i>ryhB</i> , Cm ^r Amp ^r	This study
Nissle 1917 <i>Δfur</i>	<i>fur</i> mutant strain, Kan ^r	This study
Nissle 1917 <i>Δfur/p-fur</i>	<i>fur</i> mutant strain carrying p- <i>fur</i> , Kan ^r Amp ^r	This study
SP15	Colibactin producer, pathogenic strain	35
SP15 <i>ΔryhB</i>	<i>ryhB</i> mutant strain, Cm ^r	This study
SP15 <i>ΔryhB/p-ryhB</i>	<i>ryhB</i> mutant strain carrying p- <i>ryhB</i> , Cm ^r Amp ^r	This study
SP15 <i>Δfur</i>	<i>fur</i> mutant strain, Kan ^r	This study
SP15 <i>Δfur/p-fur</i>	<i>fur</i> mutant strain carrying p- <i>fur</i> , Kan ^r Amp ^r	This study
MG1655 <i>ΔryhB</i>	<i>ryhB</i> mutant strain, Cm ^r	Gift from C. M. Dozois
Bw25113 <i>Δfur</i>	<i>fur</i> mutant strain, Kan ^r	Gift from C. M. Dozois
Plasmids		
p- <i>ryhB</i>	High-copy-no. pSC-A plasmid carrying the <i>ryhB</i> gene; Amp ^r Kan ^r	This study
p- <i>fur</i>	High-copy-no. pSC-A plasmid carrying the <i>fur</i> gene; Amp ^r Kan ^r	This study
pBAD33-T7-Venus	High-copy-no. pBAD33 plasmid containing the <i>Venus</i> gene coding for a variant of the yellow fluorescent protein of <i>Aequorea victoria</i>	This study
pBAD33-T7- <i>clbA</i>	High-copy-no. pBAD33-T7 plasmid carrying the <i>Venus</i> fusion of the <i>clbA</i> gene	This study
pBAD33-T7- <i>clbAΔclbR</i>	High-copy-no. pBAD33-T7 plasmid carrying the <i>Venus</i> fusion of the <i>clbA</i> gene and mutated <i>clbR</i>	This study
Primers		
Fur-P1	AAAATCCTGGAAGTTCTTCAGGAGCCGGACAACCATCACGTCAGTGCCGAGTGTAGG CTGGAGCTGCTTC	This study
Fur-P2	TTCGCGGCAATCGCCTTCGGCACAGTGACCGTAAAGATAGAGACTGTGGTCATATGA ATATCCTCCTTAG	This study
Fur_F	CGCCCTAAAGAAAGCTGGCC	This study
Fur_R	CCTTCGTGCGCATGTTTCATC	This study
Fur_F_Compl	CTGTAAGCTGTGCCAGTTC	This study
Fur_R_compl	CTGAGAGCTGTAACCTCCGCTTTTC	This study
CMD1171_ST (<i>ryhB::cat</i>)	TTTGGGGTAAATGTCCCTTC	29
CMD1172_ST (<i>ryhB::cat</i>)	GTGCGCATAACGAACACAAG	29
RyhB_F_Compl	CCTCTCGAGAAAGCGGACGTGGTTCCTAC	This study
RyhB_R_Compl	GACTCGAGTGTTCCTGCGTGGCGTATTAC	29
Venus-XbaI	GCGCCTAGATGGTGTCTACTCACTAAAGATCAAATC	This study
Venus-fus-as	GCGCGTCCGACTTACTTGTACAGCTCGTCCATG	This study
T7-term-fw	AAGGGGATCCGGCTGCTAAC	This study
T7-term-rev	GCGCAATTGCGGATATAGTTCCTCCTTTTCAG	This study
pBAD33-Venus-Bam-rev	GCGCGGATCCGGCTGCTAAACAAAAG	This study
pBAD33-Venus-fw	P-TGGTGTCTATCACTAAAGATCAAATC	This study
<i>clbA</i> p-BamHI-fw	GCGCGGATCCCAACCATCACCTTATTATCGG	This study
<i>clbA</i> p-rev1-1	TTAGATAATCTCATTCCCTGTTAGC	This study

(Continued on following page)

TABLE 1 (Continued)

Strain, plasmid, or primer	Description or sequence ^a	Reference or source
<i>clbAp-DelclbR-fw2</i>	CCGTTATCTAAGACAAGTATTGCGCATG	This study
<i>clbAp-DelclbR-rev2</i>	ACTTGTCTTAGATAACGGGTTTTTTCTTTG	This study
<i>fyuA-up</i>	CAACGCGCAGGCCTTTAC	This study
<i>fyuA-rev</i>	GCGTGCTTTTCGTCTTGCTG	This study
<i>clbC-up</i>	CGAGTCAAATGCGCCATCAC	This study
<i>clbC-rev</i>	GGACCGCCATACCAATAATG	This study
JPN41	CAGATACACAGATACCATTC	14
JPN42	TCAATGAGGAAGAAATAAAAC	14

^a Restriction sites are underlined. A P in front of the sequence indicates a phosphate residue.

MATERIALS AND METHODS

Bacterial strains, mutagenesis, and growth conditions. The bacterial strains used in this study are listed in Table 1. For genetic manipulations, *E. coli* strains were routinely grown at 37°C with shaking at 240 rpm in 3 ml of Lennox L broth (LB; Invitrogen). Ampicillin (50 µg/ml), kanamycin (50 µg/ml), or chloramphenicol (25 µg/ml) was added to the medium when required. Gene inactivations were performed using the bacteriophage lambda Red recombinase method (21) and the primers listed in Table 1. Briefly, the *ryhB::cat* and *fur::kan* alleles were PCR amplified from the chromosomal DNA of strain MG1655 Δ *ryhB* or Bw25113 Δ *fur*, respectively, using primers CMD1171_ST/CMD1172_ST or Fur_F/Fur_R, respectively. The purified PCR product was used to transform the strains of interest. The allelic exchanges were confirmed by PCR using the same pairs of primers. The resulting mutants where *ryhB* or *fur* was disrupted by a resistance cassette were named the Δ *ryhB* and Δ *fur* mutants, respectively. To generate the *fur::cat* mutation in strain Nissle 1917 *clbA-lux*, the chloramphenicol cassette was PCR amplified using primers Fur_P1/Fur_P2. The purified PCR product was used to transform *E. coli* Nissle 1917 *clbA-lux*. The transformants were tested by PCR amplification using primers Fur_F/Fur_R. For cloning of the wild-type *ryhB* and *fur* genes in cloning vectors, the *ryhB* and *fur* genes were PCR amplified using primer pairs RyhB_F_Compl/RyhB_R_Compl and Fur_F_Compl/Fur_R_Compl, respectively. The resulting PCR products were then cloned into the pSC-A-amp/kan vector using a StrataClone PCR cloning kit (Agilent Technologies). The presence of the appropriate insert in the resulting plasmids was checked by PCR amplification, and the plasmids were transformed into the mutant strains.

For the megalocytosis assay and assays for bioluminescence and fluorescence measurements, the *E. coli* strains were grown overnight in Dulbecco's modified Eagle's medium (DMEM) containing HEPES (DMEM-HEPES; Gibco), supplemented with antibiotics when required, at 37°C with shaking (240 rpm). The overnight cultures were then diluted 1:50 in DMEM-HEPES supplemented or not with iron (FeCl₃, 100 µM; FeSO₄, 100 µM), magnesium (MgSO₄, 100 µM), copper (CuCl₂, 100 µM), or iron chelator (8-hydroxyquinoline [8-HQ; 40 µM] or desferrioxamine [DFO; 0.2 µM]) and grown until the optical density at 600 nm (OD₆₀₀) was 0.6. The ranges of concentrations of the chelator and the iron supplement were analyzed, and the highest concentrations of substances which did not alter bacterial growth were chosen.

Construction of Venus fusions. A low-copy-number vector system was used to generate *clbA* promoter fusions with the *Venus* gene as a reporter. For this purpose, a derivative of vector pBAD33, which was devoid of the elements required for l-arabinose-dependent expression, was employed. An overview about the construction of the plasmids is given in Fig. S3 in the supplemental material.

In more detail, the translational fusion of the first 24 bp of *rplL* and the *Venus* gene, coding for a variant of the yellow fluorescent protein of *Aequorea victoria*, was amplified from pMB*rplL-Venus* using the primers Venus-XbaI and Venus-fus-as. The PCR product was digested with XbaI and Sall and cloned into pBAD33 (1). The resulting plasmid was digested

with EcoRV and SacI, treated with mung bean nuclease (MBN), and religated as plasmid pBAD33-*Venus* in order to delete the *araBAD* promoter and the *araC* gene (2). Since the aim was to clone the whole *clbR* intergenic region together with the *clbR* gene, which harbors the *clbA* promoter, the transcription into the opposite direction of *clbA*, facing *clbB*, had to be blocked. To achieve this, the T7 terminator sequence of pTXB1 (NEB) was amplified with primers T7-term-for and T7-term-rev, subsequently digested with BamHI and MfeI, and cloned upstream of *Venus* into pBAD33-*Venus*, resulting in pBAD33-T7-*Venus* (3). The plasmid was verified by sequencing.

The whole pBAD33-T7-*Venus* sequence was amplified via PCR as a precursor for the following plasmids by using primers pBAD33-*Venus*-Bam-rev and pBAD33-*Venus*-fw and subsequently digested with BamHI (4). To yield pBAD33-T7-*clbA*, the 648-bp DNA sequence upstream of *clbA* in *E. coli* strain M1/5 (20), comprising the putative *clbA* promoter, the *clbR* gene, as well as the whole *clbBR* intergenic region (*clbBR* in. reg.), was amplified with primers *clbAp*-BamHI-fw and *clbAp*-rev1-1, digested with BamHI, and then ligated with the digested pBAD33-T7-*Venus* amplicon (5a). To generate pBAD33-T7-*clbA* Δ *clbR*, a derivative of pBAD33-T7-*clbA* with a deletion of *clbR* nucleotides 33 to 40, a two-step-PCR was carried out using *E. coli* M1/5 as a template. Amplicons were generated with the primer pairs *clbAp*-BamHI-fw/*clbAp*-DelclbR-rev2 (PCR 1) and *clbAp*-DelclbR-fw2/*clbAp*-rev1-1 (PCR 2). PCR products 1 and 2 were fused using primers *clbAp*-BamHI-fw and *clbAp*-rev1-1 (PCR 3) and then digested with BamHI and ligated with PCR-amplified and BamHI-digested pBAD33-T7-*Venus*, resulting in pBAD33-T7-*clbA* Δ *clbR*. The plasmids were verified by sequencing.

Luciferase and Venus measurements. The promoter activities of genes *clbA*, *clbB*, *clbQ*, and *clbR* in different media were determined by time course quantification of luciferase or Venus expression. When an OD₆₀₀ of 0.6 was reached, bacterial subcultures were diluted to an OD₆₀₀ of 0.1 in the appropriate medium. Samples of 100 µl were then used to inoculate a black 96-well plate (Greiner Bio-One), and the bacteria were grown without shaking at 37°C in a luminometer (Tecan Infinite Pro microplate reader). For luminescence measurements, light emission (relative light units [RLU]) was recorded (6,000-ms aperture per sample) every 30 min in parallel with the OD₆₀₀. For Venus fluorescence measurements, fluorescence was recorded at 535 nm (bandwidth, 25 nm) using an excitation wavelength of 485 nm (bandwidth, 20 nm) with the luminometer. The OD₆₀₀ was measured in parallel at each time point. The area under the curve (AUC), which quantifies the cumulative luminescence or fluorescence, was calculated with GraphPad Prism (version 6.0) software (baseline $Y = 0$, with the units being the number of RLU per OD₆₀₀ times time [in hours]).

Quantification of colibactin-associated genotoxic effect by megalocytosis assay. Quantification of the colibactin-associated genotoxic effect by megalocytosis assay was performed as previously described (14). Briefly, HeLa cells were grown in DMEM (GlutaMAX; Invitrogen) supplemented with 10% (vol/vol) fetal calf serum (FCS; Eurobio) and 1% (vol/vol) nonessential amino acids (Invitrogen) at 37°C in a 5% CO₂

Tronnet et al.

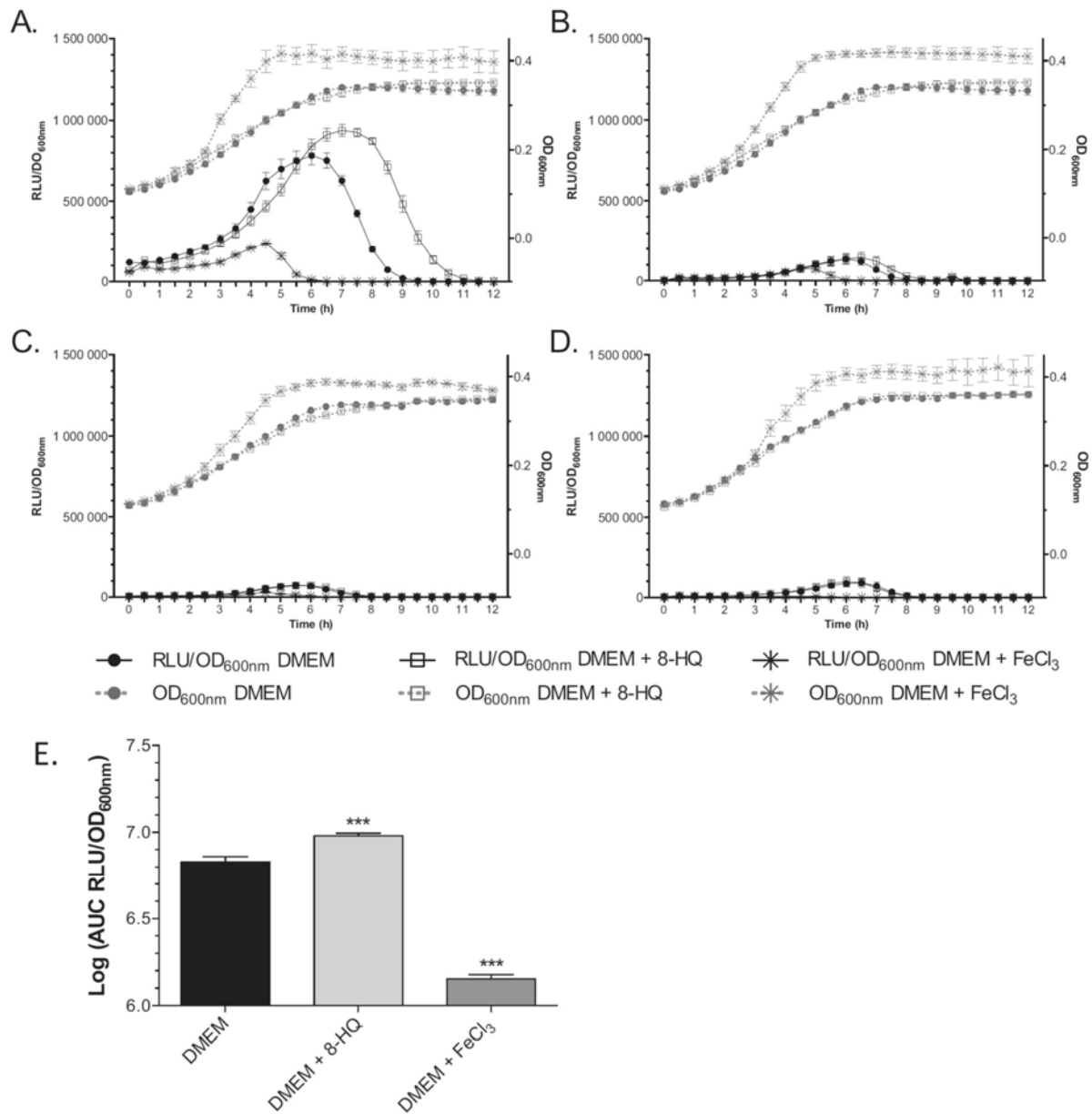


FIG 1 Iron availability modulates the transcription of *clbA*. Growth curves (OD_{600nm}) and the RLU/OD_{600nm} of Nissle 1917 *clbA-lux* (A), *clbB-lux* (B), *clbC-lux* (C), and *clbR-lux* (D) fusion strains grown at 37°C in DMEM-HEPES, DMEM-HEPES supplemented with 40 μM 8-HQ, or DMEM-HEPES supplemented with 100 μM FeCl₃. (E) AUC, determined with GraphPad Prism (version 6.0) software, of the RLU/OD_{600nm} of *clbA* under the three different conditions. The given average values are the mean number of RLU and SEMs from five independent experiments. Statistical analysis was performed using two-way ANOVA and the Bonferroni posttest. ***, $P < 0.001$ compared to the mean values obtained in DMEM-HEPES.

atmosphere. HeLa cells were dispensed in a 96-well cell culture plate (5×10^3 cells/well) for 24 h and then infected at a multiplicity of infection (MOI; number of bacteria per HeLa cell at the onset of infection) of 100 with the *E. coli* strains, which had previously been pregrown in the appropriate medium. At 4 h postinoculation, the cells were washed 3 times with Hanks' balanced salt solution (HBSS; Gibco) and incubated in cell culture medium for 72 h with 200 mg/ml gentamicin before fixation (4% form-

aldehyde) and protein staining with methylene blue (1% [wt/vol] in 0.01 M Tris-HCl). The methylene blue was extracted with 0.1 N HCl. Staining was quantified by measurement of the OD_{660nm}.

Electrophoretic mobility shift assay (EMSA). PCR products were generated by using genomic DNA from *E. coli* strain M1/5 as the template to yield the probes and the primers JPN42/JPN41 (*clbA*, 343 bp), *clbC*-up/*clbC*-rev (*clbC*, 219 bp), and *fyuA*-up/*fyuA*-rev (*fyuA*, 648 bp) (Table 1).

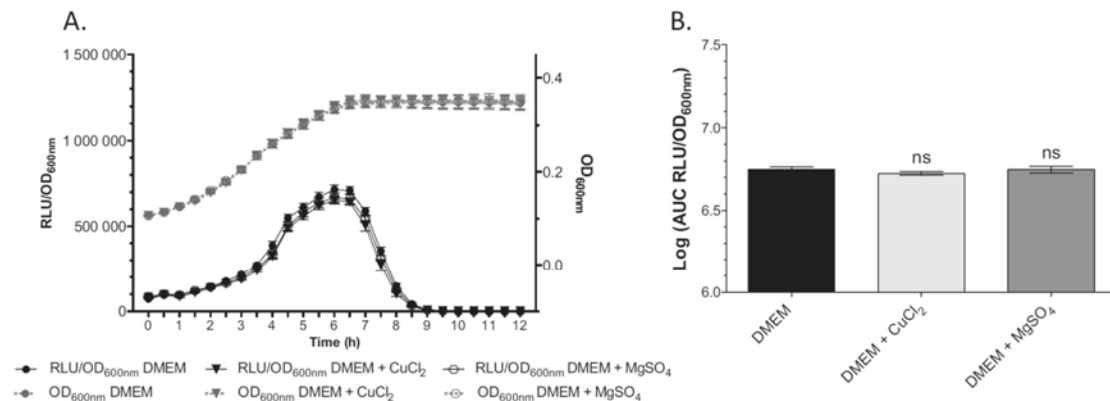


FIG 2 The transcription of *clbA* is not modulated by Cu^{2+} or Mg^{2+} . Growth curves (OD_{600}) and the $\text{RLU}/\text{OD}_{600}$ of *E. coli* Nissle 1917 *clbA-lux* grown at 37°C in DMEM-HEPES and DMEM-HEPES supplemented with either $100\ \mu\text{M}$ MgSO_4 or $100\ \mu\text{M}$ CuCl_2 (A) and the associated AUC (B). The given average values are the mean number of RLU and SEMs from at least three independent experiments. Statistical analysis was performed by comparison of the values to the mean values obtained in DMEM-HEPES using two-way ANOVA and the Bonferroni posttest. ***, $P < 0.001$; ns, not significant.

The PCR products were gel purified using a Wizard SV gel and PCR cleanup system (Promega). A $20\text{-}\mu\text{l}$ binding reaction mixture containing recombinant Fur protein (MyBioSource, USA), $50\ \mu\text{g}\ \text{ml}^{-1}$ of poly(dI-dC), $4\ \mu\text{l}$ of $5\times$ binding buffer (50 mM Tris-HCl, pH 7.5, 200 mM KCl, 5

mM MgCl_2 , 0.625 mM MnCl_2 , 25% [vol/vol] glycerol, 2.5 mM dithiothreitol, $500\ \mu\text{g}\ \text{ml}^{-1}$ bovine serum albumin), and the DNA probe (10 nM) was incubated at room temperature for 30 min, loaded onto a 4% Tris-borate polyacrylamide gel, and electrophoresed in $0.5\times$ Tris-borate,

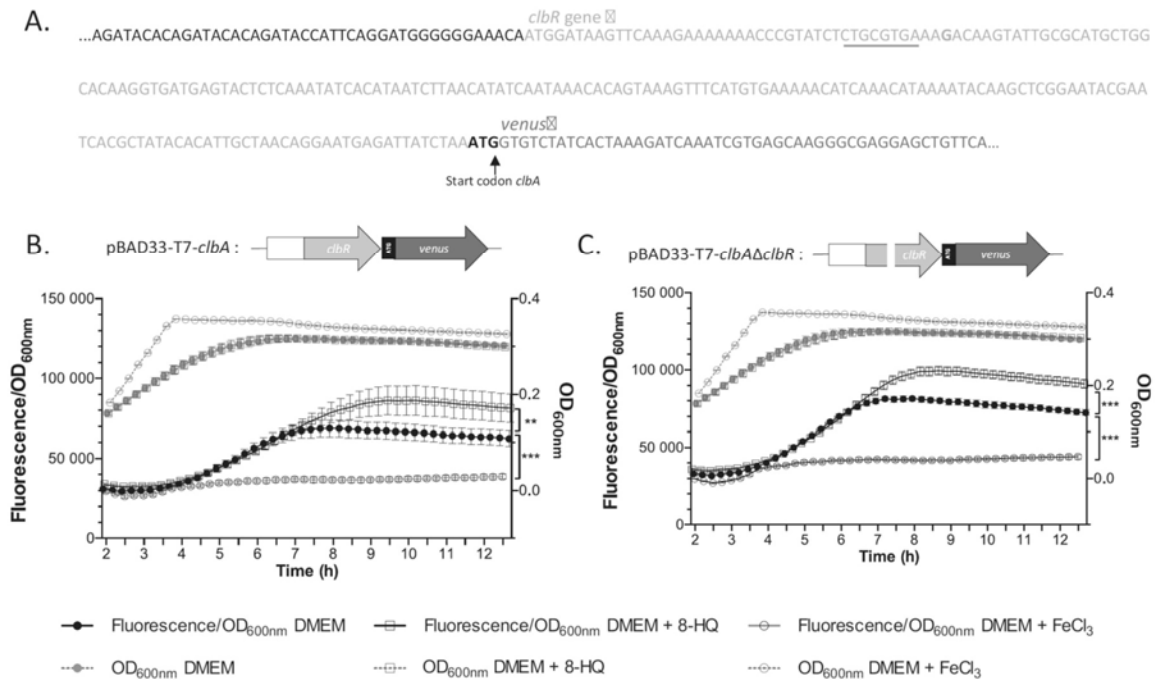


FIG 3 The transcription of *clbA* is dependent on the iron concentration in a ClbR-independent manner. (A) Sequence of the construction of the translational *Venus* fusion of *clbA* (*clbA-Venus* fusion). Two different constructs, one containing the wild-type *clbR* gene and one containing an inactive *clbR* gene, were generated. The underlined sequence was mutated to inactivate *clbR*. The wild-type and the mutated constructs were cloned into plasmid pBAD33-T7, and the plasmid was transformed into strain Nissle 1917 in which *clbA* was mutated. (B and C) Schematic representation of the plasmid construct obtained and results of the quantification of fluorescence of strain Nissle 1917 in which *clbA* was mutated and transformed with the plasmid containing the wild-type construct (B) or the mutated one (C). The fluorescence was measured in DMEM-HEPES or DMEM-HEPES supplemented with $100\ \mu\text{M}$ FeCl_3 or $40\ \mu\text{M}$ 8-HQ. The given average values are the mean fluorescence/OD₆₀₀ and SEMs three independent experiments. Statistical analysis was performed by comparison of the values to the mean values obtained in DMEM-HEPES using two-way ANOVA and the Bonferroni posttest. ***, $P < 0.001$; **, $P < 0.01$.

Tronnet et al.

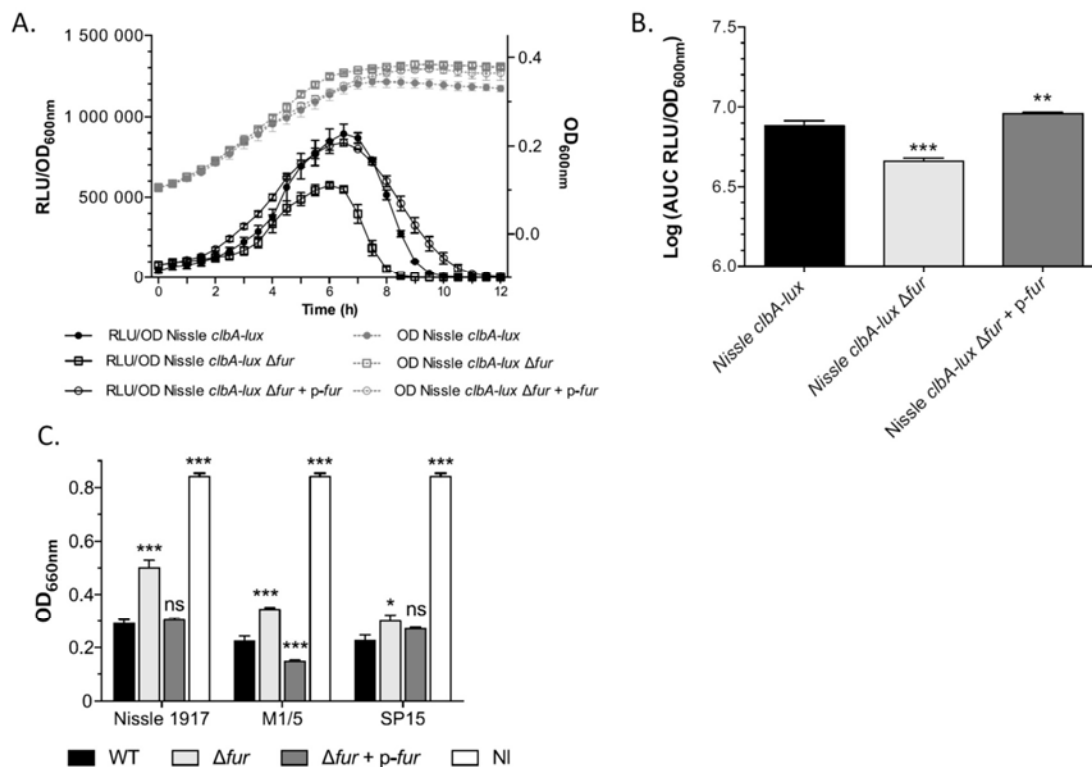


FIG 4 Fur regulates the transcription of *clbA* and colibactin biosynthesis. (A) Growth curves (OD₆₀₀) and the RLU/OD₆₀₀ of strains Nissle 1917 *clbA-lux*, Nissle 1917 *clbA-lux Δfur*, and Nissle 1917 *clbA-lux Δfur/p-fur* grown at 37°C in DMEM-HEPES. The given average values of the number of RLU and SEMs resulted from three independent experiments. (B) AUC of the RLU/OD₆₀₀ of *clbA* in the three different strains. (C) Quantification of colibactin production. The production of colibactin by the *E. coli* strains and derivatives was determined by quantification of megalocytosis as previously described (15). The multiplicity of infection was 100 bacteria per cell. Three distinct genetic contexts of *E. coli* were investigated: Nissle 1917 (a probiotic strain), M1/5 (a commensal strain), and SP15 (a strain isolated from a patient with neonatal meningitis). Both cells that were not infected (NI) and cells that were infected with *E. coli* strains that do not produce colibactin gave the same quantitative results (data not shown). The given quantification average values are represented as the mean values and SEMs from three independent experiments. Statistical analysis was performed by comparison of the values to the mean values obtained for wild-type (WT) strains using two-way ANOVA and the Bonferroni posttest. ***, $P < 0.001$; **, $P < 0.01$; *, $P < 0.05$; ns, not significant.

pH 7.5, containing 0.2 mM MnCl₂ at 100 V. The gels were stained with ethidium bromide.

Statistical analysis. Statistical analyses were conducted using GraphPad Prism (version 6.0) software. The mean with the standard error of the mean (SEM) is shown in the figures, and P values were calculated using a one-way or two-way analysis of variance (ANOVA) followed by a Bonferroni posttest, unless otherwise stated. P values of less than 0.05 were considered statistically significant.

RESULTS

Iron availability modulates *clbA* transcription. In order to test whether the promoter activities of the genes located in the *pks* island (see Fig. S1A in the supplemental material) were dependent on iron availability, *E. coli* strain Nissle 1917 harboring transcriptional luciferase fusions with the genes *clbA*, *clbB*, *clbQ*, and *clbR*, four genes involved in colibactin biosynthesis, were studied (Table 1; see also Fig. S1A and B) (22). These strains were grown in chemically defined DMEM-HEPES, DMEM-HEPES supplemented with 100 μM FeCl₃, and DMEM-HEPES supplemented with an iron chelator, 8-hydroxyquinoline (8-HQ; 40 μM). The ranges of concentrations of the chelator and the iron supplement were an-

alyzed, and the highest concentrations of substances which did not alter bacterial growth were chosen. The transcription rates of the genes *clbA*, *clbB*, *clbQ*, and *clbR* were determined as the number of OD₆₀₀-standardized relative luminescence units (RLU) (Fig. 1).

This analysis revealed that only *clbA* was significantly transcribed in DMEM-HEPES, as previously reported (22) (compare Fig. 1A to Fig. 1B to D). During the bacterial growth kinetics, a continuous increase in bioluminescence emission was observed to reach a maximal value at early stationary phase of growth, followed by a decrease of transcription (Fig. 1A).

The transcription of the genes *clbB*, *clbQ*, and *clbR* was not significantly modified when the culture medium was iron depleted (DMEM-HEPES supplemented with 40 μM 8-HQ) or iron replete (DMEM-HEPES supplemented with 100 μM FeCl₃) (Fig. 1B, C, and D, respectively). In contrast, under iron-depleted conditions, the expression of *clbA* was enhanced and maintained longer than that in DMEM-HEPES (Fig. 1A). Under iron-supplemented conditions, *clbA* transcription was strongly repressed (Fig. 1A). Calculation of the area under the curve (AUC) of the relative number of OD₆₀₀-standardized relative luminescence units (RLU/

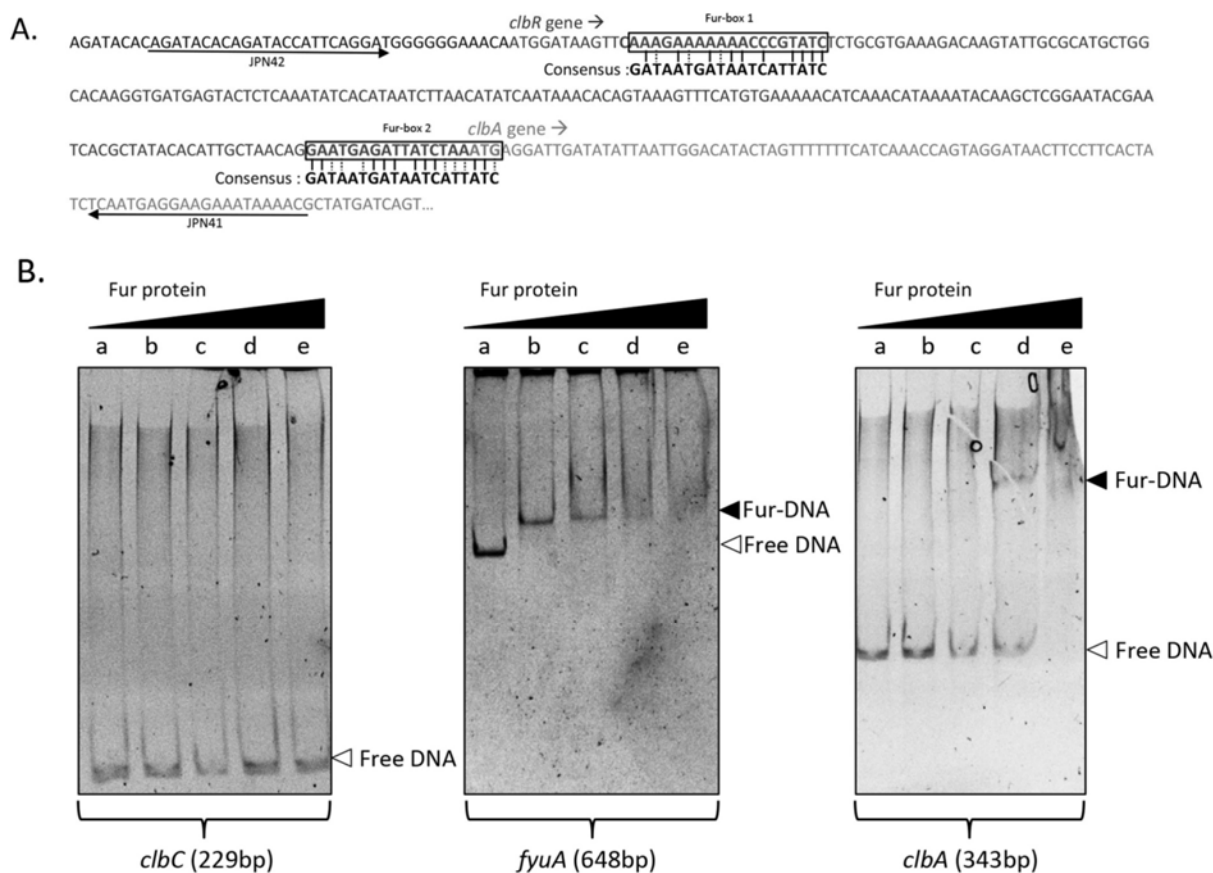


FIG 5 Fur binds to the *clbA* promoter. (A) Identification and localization of the two putative Fur boxes located upstream of *clbA*. (B) EMSA was performed in the presence of 0 μM (lanes a), 0.25 μM (lanes b), 0.5 μM (lanes c), 1 μM (lanes d), and 4 μM (lanes e) Fur protein and 10 ng of the PCR product amplified with primers JPN41/JPN42. Three PCR products were tested: a negative control (*clbC*, 229 bp), a positive control (*fyuA*, 648 bp), and the promoter region upstream of *clbA* (343 bp).

OD₆₀₀) confirmed the iron-dependent differential *clbA* expression (Fig. 1E). Supplementation with an alternative iron chelator (desferrioxamine [DFO], 0.2 μM) or an alternative iron source (FeSO₄, 100 μM) resulted in similar patterns (see Fig. S2 in the supplemental material).

In order to test whether other divalent metals could influence *clbA* transcription, we monitored the expression of *clbA* in strain Nissle 1917 *clbA-lux* grown in DMEM-HEPES supplemented with copper (CuCl₂, 100 μM) or magnesium (MgSO₄, 100 μM) (Fig. 2). This revealed that the transcription of *clbA* was not altered in the presence of Mg²⁺ or Cu²⁺.

Altogether, these results show that *clbA* transcription is regulated by iron availability in the medium.

The iron-dependent transcription of *clbA* is independent of the LuxR-like protein ClbR. The gene *clbR* encodes a LuxR-like protein that exhibits a helix-turn-helix DNA-binding motif, which is suspected to be a regulator of colibactin biosynthesis (22). Moreover, cotranscription was previously demonstrated for the genes *clbR* and *clbA* in DMEM-HEPES medium with 5% fetal bovine serum (22). Our results described above showed that only *clbA* transcription and not *clbR* transcription is regulated by iron. To test whether ClbR is

involved in the iron-dependent regulation of *clbA*, the *clbA* gene was transcriptionally fused with an alternative reporter gene, i.e., the *Venus* gene (coding for a variant of yellow fluorescent protein of *Aequorea victoria*) in plasmid pBAD33-T7-*Venus* to produce pBAD33-T7-*clbA*. The *clbR* gene was inactivated to produce plasmid pBAD33-T7-*clbA* Δ *clbR*. The resulting plasmids, pBAD33-T7-*clbA* and pBAD33-T7-*clbA* Δ *clbR*, were transformed into strain Nissle 1917 Δ *clbA*. Monitoring of the fluorescence of the Venus protein transcriptionally fused with the *clbA* gene in strain Nissle 1917 Δ *clbA*/pBAD33-T7-*clbA* confirmed the iron-dependent regulation of *clbA* (Fig. 3A and B). Monitoring of the fluorescence of strain Nissle 1917 Δ *clbA*/pBAD33-T7-*clbA* Δ *clbR* under iron-supplemented and iron-chelated conditions revealed that *clbA* transcription was regulated by the iron concentration in the absence of functional ClbR (Fig. 3A and C). These results show that ClbR is not involved in the iron-dependent regulation of *clbA*.

Fur positively regulates the production of colibactin. Because the global response to iron availability in bacteria relies on the ferric uptake regulator (Fur) protein, we tested whether the iron-dependent regulation of *clbA* transcription was mediated by Fur. We mutated the *fur* gene in strain Nissle 1917 *clbA-lux*. The com-

Tronnet et al.

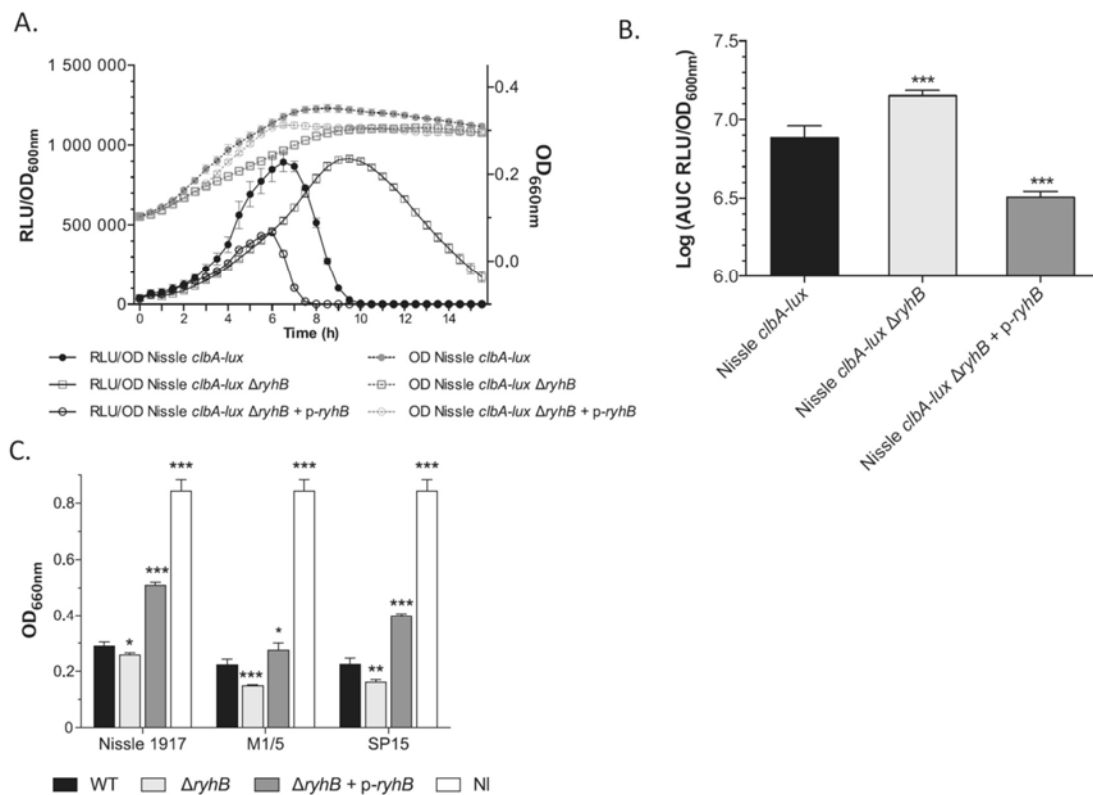


FIG 6 RyhB regulates the transcription of *clbA* and colibactin biosynthesis. (A) Growth curves (OD₆₀₀) and the RLU/OD₆₀₀ of strains Nissle 1917 *clbA-lux*, Nissle 1917 *clbA-lux ΔryhB*, and Nissle 1917 *clbA-lux ΔryhB/p-ryhB* grown at 37°C in DMEM-HEPES. The given average values are the mean number of RLU and SEMs from three independent experiments. (B) AUC of the RLU/OD₆₀₀ of *clbA* in the three different strains. (C) Quantification of colibactin production. The production of colibactin by *E. coli* strains and its derivatives was determined by quantification of megalocytosis as previously described (14). The multiplicity of infection was 100 bacteria per cell. Three distinct genetic contexts of *E. coli* were investigated: Nissle 1917 (a probiotic strain), M1/5 (a commensal strain), and SP15 (a strain isolated from a patient with neonatal meningitis). Both cells that were not infected (NI) and cells that were infected with *E. coli* strains that do not produce colibactin gave the same quantitative results (data not shown). The given quantification average values are represented as mean values and SEMs from three independent experiments. Statistical analysis was performed by comparison of the values to the mean values obtained for wild-type strains using two-way ANOVA and the Bonferroni posttest. ***, $P < 0.001$; **, $P < 0.01$; *, $P < 0.05$.

plemented strain Nissle 1917 *clbA-lux Δfur/p-fur*, where the plasmid-borne wild-type *fur* gene was transformed into strain Nissle 1917 *clbA-lux Δfur*, was also constructed (Table 1). Monitoring of the expression of *clbA* in the resulting strains (Fig. 4A and B) revealed that the transcription of *clbA* was significantly decreased in the strain in which *fur* was mutated (Fig. 4A and B). Transformation of the strain in which *fur* was mutated with a plasmid carrying the functional wild-type *fur* gene (Fig. 4A and B) totally restored the expression of *clbA*. This indicates that the transcription of *clbA* is positively regulated by Fur.

Since ClbA is required for the production of colibactin, the role of Fur in the synthesis of colibactin was then investigated. Eukaryotic cells were infected with *pks*⁺ *E. coli* strains in which *fur* was mutated or not. Three different genetic contexts were analyzed: strains Nissle 1917 (a probiotic strain), M1/5 (a commensal strain), and SP15 (a pathogenic strain isolated from a patient with neonatal meningitis) (Table 1). The genotoxic effect of colibactin was monitored by quantification of megalocytosis, as previously described (14). This revealed that all the Δfur mutants displayed a genotoxicity significantly reduced compared to that of the wild-

type strains (Fig. 4C). Transformation of the strains in which *fur* was mutated with a plasmid carrying the functional wild-type *fur* gene resulted in restoration of the megalocytosis phenotype (Fig. 4C). Altogether, these results evidence that Fur positively regulates colibactin production.

To determine whether Fur could regulate *clbA* directly and/or indirectly, we investigated a putative direct binding of Fur to *clbA*. We detected two putative Fur-binding sites in the DNA sequence of the *clbA* promoter (Fig. 5A). We tested whether purified Fur protein bound to its putative binding sites in the *clbA* promoter region (Fig. 5B). EMSA showed that migration of the 343-bp DNA fragment was retarded when it was incubated with the Fur protein. This suggests that Fur can regulate *clbA* expression through direct binding.

RyhB negatively regulates the production of colibactin. Because the second major regulator of bacterial iron metabolism is the small regulatory noncoding RNA RyhB, we investigated whether RyhB was involved in the iron-mediated regulation of *clbA* transcription. The gene *ryhB* was inactivated in strain Nissle 1917 *clbA-lux*, and the complemented strain was also engineered

A.

```

clbA 5'  A   UUUU  U A   AAAA  AUUUUUUU  ACCUUUCC  AUA  UA   C   A   ACU  GCUU  U  GAUA  A  3'
           UAGU  CC C  GUUGGC  AAG   UA   AUACU  GA  CAGU  AGC  GU  UGCU  UUAG  UUCUC  GCGA  GG  GUC  UUGA
           GUCG  GG G  CGACCG  UUC   AU   UAUGA  CU  GUUA  UCG  CA  ACGA  AGUC  AAGAG  CGCU  CC  CAG  GACU
RyhB 3'  UUUUCG  U   C   A
  
```

B.

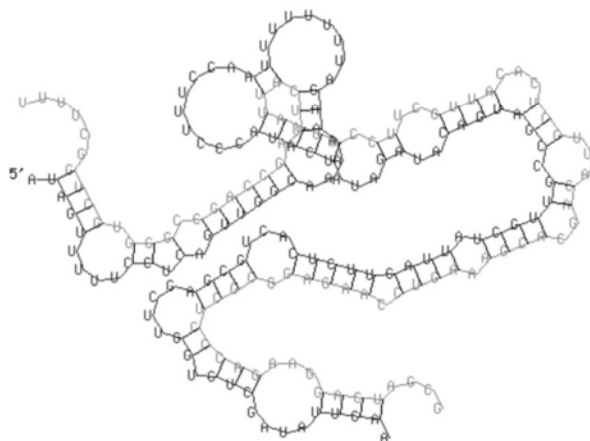


FIG 7 RyhB seems to interact directly with the *clbA* gene. The RNA Hybrid computational program predicted a region of potential interaction between RyhB and *clbA*. The nucleotide sequence involved in the interaction between *clbA* (gray) and RyhB (black) (A) and the two-dimensional structure of RyhB (green) bound to *clbA* (red) (B) are shown. The predicted pairing region with RyhB began at nucleotide 238 in the *clbA* open reading frame (minimal free energy, -51.3 kcal/mol).

(Table 1). Monitoring of the bioluminescence in the resulting strain, Nissle 1917 *clbA-lux* Δ *ryhB*, revealed that the level of transcription of *clbA* was enhanced and maintained longer than it was in strain Nissle 1917 *clbA-lux* (Fig. 6A and B). Overexpression of the wild-type *ryhB* gene expressed from a high-copy-number plasmid in strain Nissle 1917 *clbA-lux* Δ *ryhB* resulted in the repression of *clbA* transcription (Nissle 1917 *clbA-lux* Δ *ryhB*/p-*ryhB*) (Fig. 6A and B). These data evidence that RyhB plays a critical role in the transcription of *clbA*.

The role of RyhB in the synthesis of colibactin was then investigated. Eukaryotic cells were infected with *pks*⁺ *E. coli* strains in which *ryhB* was mutated or not. The genotoxic effect of colibactin, quantified by megalocytosis, revealed that all the Δ *ryhB* mutants displayed significantly increased genotoxicity compared to the wild-type strains (Fig. 6C). Overexpression of *ryhB* in strains in which the *ryhB* gene was mutated resulted in a decrease in the level of the colibactin-associated genotoxic effect (Fig. 6C). Altogether, these results evidence that RyhB negatively regulates colibactin production under these conditions.

To determine whether RyhB could regulate *clbA* directly and/or indirectly, we investigated a putative direct pairing of RyhB to the *clbA* promoter. The RNA Hybrid computational program (23) predicted a region of potential pairing between RyhB and *clbA* (Fig. 7). The region of pairing with *clbA* was located 238 to 344 nucleotides downstream from the translational start (minimal free energy, -51.3 kcal/mol) (Fig. 7) in the reading frame of *clbA*. This suggests that *ryhB* could regulate *clbA* through direct binding.

DISCUSSION

This work highlights that the transcription of *clbA*, the gene encoding the PPTase involved in colibactin production, is regulated

by iron availability (see our model in Fig. 8). Therefore, this work provides new insights into the transcriptional regulation of *clbA*, as only one study on its regulation has been published previously (22). In addition to carbon source and growth phase (22), our work demonstrates that *clbA* is regulated by iron availability. Moreover, this regulation appears to be independent of ClbR, the LuxR-like protein predicted to regulate colibactin biosynthesis (22), although the *clbR* gene was demonstrated to be coexpressed with *clbA* under particular culture conditions (22). This suggests that *clbA* could be expressed from a second promoter specifically activated under specific conditions of iron availability. This iron-dependent regulation of *clbA* transcription could constitute a fitness advantage for the bacteria, as we previously demonstrated that ClbA can also sustain siderophore biosynthesis (20).

This work highlights that Fur is positively involved in the production of the virulence factor colibactin through the regulation of *clbA*. Fur is a global transcriptional regulator that controls the transcription of over 90 genes involved in iron uptake, storage, and metabolism. Only a few studies have reported the involvement of Fur in the regulation of the biosynthesis of toxins in *E. coli*. The *stx* genes, encoding Shiga toxin in enterohemorrhagic *E. coli* (24, 25), and *hly* plasmids, encoding hemolysin (26), are directly regulated by Fur. In addition to its well-known role as a repressor, Fur was characterized as a positive regulator of gene expression, such as the expression of *acnA*, *fumA*, *ftnA*, *bfr*, and *sodB* (27).

Our work also evidenced that RyhB regulates *clbA* transcription, as we observed that *clbA* transcription was upregulated when the *ryhB* gene was inactivated, demonstrating that *clbA* is down-regulated by RyhB. RyhB is a small RNA that recruits RNase E and facilitates the degradation of mRNA targets (27, 28). Whereas only a few studies have demonstrated the impact of RyhB on virulence

Tronnet et al.

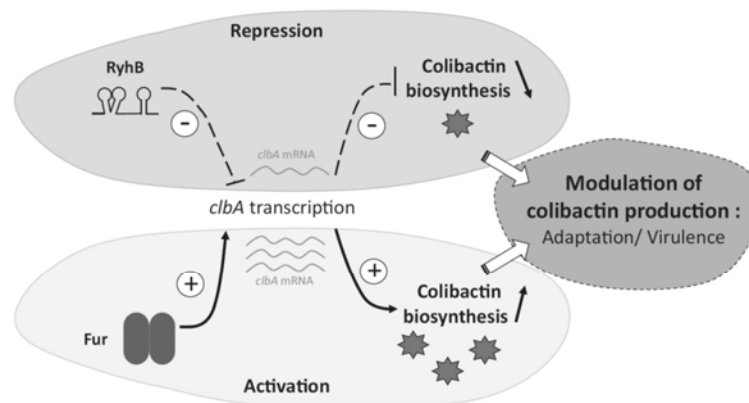


FIG 8 Model for the regulation of the transcription of *clbA* by Fur and RyhB. The biosynthesis of the genotoxin colibactin requires the enzymatic activity of the PPTase ClbA. The transcription of *clbA* is dependent on both the ferric uptake regulator (Fur) and the small regulatory noncoding RNA RyhB. This makes iron a key environmental factor contributing to the regulation of colibactin production.

in pathogenic *E. coli*, for instance, through siderophore production (29), several studies have demonstrated that this sRNA is implicated in the virulence-associated processes of other pathogenic bacteria, such as *Shigella flexneri* (2) or *Shigella dysenteriae* (30). This study provides new evidence that RyhB is involved in the production of a virulence factor in pathogenic *E. coli* through the regulation of *clbA* transcription, which leads to the modulation of colibactin production.

Here we demonstrated that modulation of the expression of *clbA* via regulators of iron homeostasis leads to the modulation of colibactin production. The mutation of *ryhB* in a colibactin-producing *E. coli* strain leads to an increase in the level of colibactin production, whereas the mutation of *fur* results in a decrease in the level of colibactin production in three different genetic contexts of *E. coli*, in which the strains belonged to the B2 phylogenetic group. We confirmed that this regulation occurred in three different strains; however, the quantitative differences observed could be explained by the genetic diversity between strains. Colibactin was demonstrated to be a virulence factor, for example, in neonatal systemic infections (19), but it is also associated with the development of colorectal cancer (31–33). The fact that colibactin is regulated by the two major regulators of iron homeostasis could link iron availability in the gut and blood and *E. coli*-mediated carcinogenesis and systemic infections, respectively. Moreover, the main reservoir of *E. coli* is the intestine and, more precisely, the colon. A recent publication confirmed the emergence of the B2 group of *E. coli* in developed countries (34). Given the existence of a gradient of iron concentration from the lumen to the intestinal epithelial cell, it is conceivable that the fine-tuning of *clbA* expression allows the production of colibactin when the pathogenic *E. coli* isolate is located in an appropriate site in the gut. This iron-dependent regulation could be one reason for the emergence of the B2 group *E. coli* strains, as we know that there is cross talk between colibactin and siderophores. Integration of the regulation of virulence factors, such as siderophores and colibactin, into networks that respond to specific environmental signals, such as the local iron concentration and the balance between Fur and RyhB, could result in the accurate production of colibactin and siderophores, so that the bacteria can adapt to the competitive environment in the gut.

Further understanding of how *E. coli* senses environments to coordinate the expression of *clbA* and the mechanism of regulation of the biosynthesis of these virulence factors may uncover novel pathways for the development of potential targets against this pathogenic enterobacterium.

ACKNOWLEDGMENTS

We have no conflict of interest to declare.

We sincerely thank Charles M. Dozois for providing us with strain MG1655 Δ *ryhB*.

FUNDING INFORMATION

This work, including the efforts of Sophie Tronnet, Christophe Garcia, Eric Oswald, and Patricia Martin, was funded by ANR (ANR-13-BSV3-0015-02 and ANR-13-BSV1-0028-01). This work, including the efforts of Sophie Tronnet, Christophe Garcia, Eric Oswald, and Patricia Martin, was funded by INCA (INCA-PLBIO13-123). This work, including the efforts of Nadine Rehm and Ulrich Dobrindt, was funded by IZKF (Dob2/013/12).

REFERENCES

- Hood MI, Skaar EP. 2012. Nutritional immunity: transition metals at the pathogen-host interface. *Nat Rev Microbiol* 10:525–537. <http://dx.doi.org/10.1038/nrmicro2836>.
- Oglesby AG, Murphy ER, Iyer VR, Payne SM. 2005. Fur regulates acid resistance in *Shigella flexneri* via RyhB and *ydeP*. *Mol Microbiol* 58:1354–1367. <http://dx.doi.org/10.1111/j.1365-2958.2005.04920.x>.
- Troxell B, Hassan HM. 2013. Transcriptional regulation by ferric uptake regulator (Fur) in pathogenic bacteria. *Front Cell Infect Microbiol* 3:59. <http://dx.doi.org/10.3389/fcimb.2013.00059>.
- Porcheron G, Dozois CM. 2015. Interplay between iron homeostasis and virulence: Fur and RyhB as major regulators of bacterial pathogenicity. *Vet Microbiol* 179:2–14. <http://dx.doi.org/10.1016/j.vetmic.2015.03.024>.
- Butcher J, Sarvan S, Brunzelle JS, Couture J-F, Stintzi A. 2012. Structure and regulon of *Campylobacter jejuni* ferric uptake regulator Fur define apo-Fur regulation. *Proc Natl Acad Sci U S A* 109:10047–10052. <http://dx.doi.org/10.1073/pnas.1118321109>.
- Papenfort K, Vanderpool CK. 2015. Target activation by regulatory RNAs in bacteria. *FEMS Microbiol Rev* 39:362–378. <http://dx.doi.org/10.1093/femsre/fuv016>.
- Bobrovskyy M, Vanderpool CK, Richards GR. 2015. Small RNAs regulate primary and secondary metabolism in Gram-negative bacteria. *Microbiol Spectr* 3:MBP-0009-2014. <http://dx.doi.org/10.1128/microbiolspec.MBP-0009-2014>.
- Miethke M, Marahiel MA. 2007. Siderophore-based iron acquisition and

- pathogen control. *Microbiol Mol Biol Rev* 71:413–451. <http://dx.doi.org/10.1128/MMBR.00012-07>.
9. Garénaux A, Caza M, Dozois CM. 2011. The ins and outs of siderophore mediated iron uptake by extra-intestinal pathogenic *Escherichia coli*. *Vet Microbiol* 153:89–98. <http://dx.doi.org/10.1016/j.vetmic.2011.05.023>.
 10. Brumbaugh AR, Smith SN, Subashchandrabose S, Himpf SD, Hazen TH, Rasko DA, Mobley HLT. 2015. Blocking yersiniabactin import attenuates extraintestinal pathogenic *Escherichia coli* in cystitis and pyelonephritis and represents a novel target to prevent urinary tract infection. *Infect Immun* 83:1443–1450. <http://dx.doi.org/10.1128/IAI.02904-14>.
 11. Walsh CT. 2008. The chemical versatility of natural-product assembly lines. *Acc Chem Res* 41:4–10. <http://dx.doi.org/10.1021/ar7000414>.
 12. Beld J, Sonnenschein EC, Vickery CR, Noel JP, Burkart MD. 2014. The phosphopantetheinyl transferases: catalysis of a post-translational modification crucial for life. *Nat Prod Rep* 31:61–108. <http://dx.doi.org/10.1039/C3NP70054B>.
 13. Blattner FR, Plunkett G, Bloch CA, Perna NT, Burland V, Riley M, Collado-Vides J, Glasner JD, Rode CK, Mayhew GF, Gregor J, Davis NW, Kirkpatrick HA, Goeden MA, Rose DJ, Mau B, Shao Y. 1997. The complete genome sequence of *Escherichia coli* K-12. *Science* 277:1453–1462. <http://dx.doi.org/10.1126/science.277.5331.1453>.
 14. Putze J, Hennequin C, Nougayrède JP, Zhang W, Homburg S, Karch H, Bringer MA, Fayolle C, Carniel E, Rabsch W, Oelschlaeger TA, Oswald E, Forestier C, Hacker J, Dobrindt U. 2009. Genetic structure and distribution of the colibactin genomic island among members of the family *Enterobacteriaceae*. *Infect Immun* 77:4696–4703. <http://dx.doi.org/10.1128/IAI.00522-09>.
 15. Nougayrède J-P, Homburg S, Taieb F, Boury M, Brzuszkiewicz E, Gottschalk G, Buchrieser C, Hacker J, Dobrindt U, Oswald E. 2006. *Escherichia coli* induces DNA double-strand breaks in eukaryotic cells. *Science* 313:848–851. <http://dx.doi.org/10.1126/science.1127059>.
 16. Cuevas-Ramos G, Petit CR, Marcq I, Boury M, Oswald E, Nougayrède J-P. 2010. *Escherichia coli* induces DNA damage *in vivo* and triggers genomic instability in mammalian cells. *Proc Natl Acad Sci U S A* 107:11537–11542. <http://dx.doi.org/10.1073/pnas.1001261107>.
 17. Payros D, Secher T, Boury M, Brehin C, Ménard S, Salvador-Cartier C, Cuevas-Ramos G, Watrin C, Marcq I, Nougayrède J-P, Dubois D, Bedu A, Garnier F, Clermont O, Denamur E, Plaisancié P, Theodorou V, Fioramonti J, Olier M, Oswald E. 2014. Maternally acquired genotoxic *Escherichia coli* alters offspring's intestinal homeostasis. *Gut Microbes* 5:313–325. <http://dx.doi.org/10.4161/gmic.28932>.
 18. Marcq I, Martin P, Payros D, Cuevas-Ramos G, Boury M, Watrin C, Nougayrède J-P, Olier M, Oswald E. 2014. The genotoxin colibactin exacerbates lymphopenia and decreases survival rate in mice infected with septicemic *Escherichia coli*. *J Infect Dis* 210:285–294. <http://dx.doi.org/10.1093/infdis/jiu071>.
 19. McCarthy AJ, Martin P, Cloup E, Stabler RA, Oswald E, Taylor PW. 2015. The genotoxin colibactin is a determinant of virulence in *Escherichia coli* K1 experimental neonatal systemic infection. *Infect Immun* 83:3704–3711. <http://dx.doi.org/10.1128/IAI.00716-15>.
 20. Martin P, Marcq I, Magistro G, Penary M, Garcia C, Payros D, Boury M, Olier M, Nougayrède J-P, Audebert M, Chalut C, Schubert S, Oswald E. 2013. Interplay between siderophores and colibactin genotoxin biosynthetic pathways in *Escherichia coli*. *PLoS Pathog* 9:e1003437. <http://dx.doi.org/10.1371/journal.ppat.1003437>.
 21. Datsenko KA, Wanner BL. 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* 97:6640–6645. <http://dx.doi.org/10.1073/pnas.120163297>.
 22. Homburg S, Oswald E, Hacker J, Dobrindt U. 2007. Expression analysis of the colibactin gene cluster coding for a novel polyketide in *Escherichia coli*. *FEMS Microbiol Lett* 275:255–262. <http://dx.doi.org/10.1111/j.1574-6968.2007.00889.x>.
 23. Krüger J, Rehmsmeier M. 2006. RNAhybrid: microRNA target prediction easy, fast and flexible. *Nucleic Acids Res* 34:W451–W454. <http://dx.doi.org/10.1093/nar/gkl243>.
 24. Payne SM, Wyckoff EE, Murphy ER, Oglesby AG, Boulette ML, Davies NML. 2006. Iron and pathogenesis of *Shigella*: iron acquisition in the intracellular environment. *Biometals* 19:173–180. <http://dx.doi.org/10.1007/s10534-005-4577-x>.
 25. Tobe T, Yen H, Takahashi H, Kagayama Y, Ogasawara N, Oshima T. 2014. Antisense transcription regulates the expression of the enterohemorrhagic *Escherichia coli* virulence regulatory gene *ler* in response to the intracellular iron concentration. *PLoS One* 9:e101582. <http://dx.doi.org/10.1371/journal.pone.0101582>.
 26. Grünig HM, Rutschi D, Schoch C, Lebek G. 1987. The chromosomal *fur* gene regulates the extracellular haemolytic activity encoded by certain *hly* plasmids. *Zentralbl Bakteriol Mikrobiol Hyg A* 266:231–238.
 27. Massé E, Vanderpool CK, Gottesman S. 2005. Effect of RyhB small RNA on global iron use in *Escherichia coli*. *J Bacteriol* 187:6962–6971. <http://dx.doi.org/10.1128/JB.187.20.6962-6971.2005>.
 28. Soberero P, Valverde C. 2012. The bacterial protein Hfq: much more than a mere RNA-binding factor. *Crit Rev Microbiol* 38:276–299. <http://dx.doi.org/10.3109/1040841X.2012.664540>.
 29. Porcheron G, Habib R, Houle S, Caza M, Lépine F, Daigle F, Massé E, Dozois CM. 2014. The small RNA RyhB contributes to siderophore production and virulence of uropathogenic *Escherichia coli*. *Infect Immun* 82:5056–5068. <http://dx.doi.org/10.1128/IAI.02287-14>.
 30. Murphy ER, Payne SM. 2007. RyhB, an iron-responsive small RNA molecule, regulates *Shigella dysenteriae* virulence. *Infect Immun* 75:3470–3477. <http://dx.doi.org/10.1128/IAI.00112-07>.
 31. Arthur JC, Perez-Chanona E, Mühlbauer M, Tomkovich S, Uronis JM, Fan T-J, Campbell BJ, Abujamel T, Dogan B, Rogers AB, Rhodes JM, Stintzi A, Simpson KW, Hansen JJ, Keku TO, Fodor AA, Jobin C. 2012. Intestinal inflammation targets cancer-inducing activity of the microbiota. *Science* 338:120–123. <http://dx.doi.org/10.1126/science.1224820>.
 32. Dalmasso G, Cougnoux A, Delmas J, Darfeuille-Michaud A, Bonnet R. 2014. The bacterial genotoxin colibactin promotes colon tumor growth by modifying the tumor microenvironment. *Gut Microbes* 5:675–680. <http://dx.doi.org/10.4161/19490976.2014.969989>.
 33. Cougnoux A, Dalmasso G, Martinez R, Buc E, Delmas J, Gibold L, Sauvanet P, Darcha C, Déchelotte P, Bonnet M, Pezet D, Wodrich H, Darfeuille-Michaud A, Bonnet R. 2014. Bacterial genotoxin colibactin promotes colon tumour growth by inducing a senescence-associated secretory phenotype. *Gut* 63:1932–1942. <http://dx.doi.org/10.1136/gutjnl-2013-305257>.
 34. Massot M, Daubié A-S, Clermont O, Jauréguy F, Couffignal C, Dahbi G, Mora A, Blanco J, Branger C, Mentré F, Eddi A, Picard B, Denamur E, the COLIVILLE Group. 2016. Phylogenetic, virulence and antibiotic resistance characteristics of commensal strain populations of *Escherichia coli* from community subjects in the Paris area in 2010 and evolution over 30 years. *Microbiology* 162:642–650. <http://dx.doi.org/10.1099/mic.0.000242>.
 35. Johnson JR, Oswald E, O'Bryan TT, Kuskowski MA, Spanjaard L. 2002. Phylogenetic distribution of virulence-associated genes among *Escherichia coli* isolates associated with neonatal bacterial meningitis in The Netherlands. *J Infect Dis* 185:774–784. <http://dx.doi.org/10.1086/339343>.

ABSTRACT

Atypical modulation of the biosynthesis of colibactin, a genotoxin from *Escherichia coli*, or how a genomic island is symbiotic with the bacterial chromosome

The *pks* genomic island codes a complex biosynthetic assembly line that synthesizes the colibactin, a genotoxin produced by some strains of *Escherichia coli*. This genotoxin generates DNA double-strand breaks in eukaryotic cells both *in vitro* and *in vivo*. Colibactin is not a protein, but a secondary metabolite belonging to the chemical family of hybrid polyketide/nonribosomal peptide compounds.

Preliminary results from our research team suggested that certain genes of the *E. coli* core genome (*i.e.* genes present in all strains of the species) could also be involved in the colibactin production. The main goal of this thesis was to identify non-essential *E. coli* genes located outside the *pks* island that are required for colibactin biosynthesis, with the screening of a transposon mutant library. This revealed 29 potential candidate genes, but the project focused specifically on two groups of genes: three genes encoding chaperone proteins, and three genes encoding enzymes involved in polyamine metabolism.

The first project highlighted the role of the molecular chaperone HtpG (or Hsp90_{Ec}), the bacterial homolog of eukaryotic heat shock protein 90, in the production of colibactin, but also yersiniabactin, a siderophore (*i.e.* a bacterial iron uptake system) that belongs to the same chemical family as colibactin. Furthermore, the ClpQ protease was involved in colibactin and yersiniabactin production in combination with Hsp90_{Ec}. These results confirmed the interplay between the biosynthesis of two *E. coli* virulence factors, colibactin and yersiniabactin. Finally, analysis of the effects of *htpG* disruption during systemic infection in animals, using rodent models of sepsis and neonatal meningitis, demonstrated the role of the stress-responsive molecular chaperone Hsp90_{Ec} in *E. coli* virulence.

The second project revealed the involvement of polyamines in the biosynthesis of colibactin. A molecular microbiology approach demonstrated that spermidine was the polyamine required for colibactin production. Preliminary results suggested that spermidine could regulate the expression of some *pks* island genes, and therefore could modulate colibactin production. Further experiments are in progress to elucidate the molecular mechanisms involved in this regulation.

Together, the results of this thesis perfectly illustrate the symbiotic integration of a mobile genetic element acquired during evolution into the bacterial chromosome, through several crosstalks allowing the production of virulence factors in *E. coli*.

AUTEUR :

Christophe GARCIE

Directeur/trice(s) de Thèse :

Dr. Patricia MARTIN et Pr. Éric OSWALD

Lieu et Date de soutenance :

Toulouse, le mercredi 14 décembre 2016

TITRE :

Modulation atypique de la biosynthèse de la colibactine, une génotoxine de *Escherichia coli*, ou comment un îlot génomique est en symbiose avec le chromosome bactérien

RÉSUMÉ en français :

L'îlot génomique *pks* code une machinerie de biosynthèse complexe synthétisant la colibactine, une génotoxine produite par certaines souches de *Escherichia coli*. Cette génotoxine induit des cassures double-brin de l'ADN sur les cellules eucaryotes *in vitro* et *in vivo*. La colibactine n'est pas une protéine, mais un métabolite secondaire de type polycétide/peptide non-ribosomal (PK/NRP).

Des résultats préliminaires de l'équipe semblaient indiquer que certains gènes du *core genome* de *E. coli* seraient également impliqués dans la production de la colibactine. L'objectif de cette thèse était d'identifier les gènes non-essentiels de *E. coli* situés hors de l'îlot génomique *pks* impliqués dans la synthèse de colibactine, en construisant une banque de mutants par insertion de transposons. Ce criblage a permis d'identifier 29 gènes candidats, mais deux groupes de gènes ont été particulièrement étudiés dans la suite du projet : trois gènes codants des protéines chaperons, et trois gènes codant des enzymes impliquées dans le métabolisme des polyamines.

Le premier projet a permis de montrer que la protéine chaperon HtpG (ou Hsp90_{Ec}), homologue bactérien de la protéine de choc thermique eucaryote Hsp90, est requise pour la production de colibactine, mais aussi de yersiniabactine, un sidérophore (ou système bactérien de captation du fer) appartenant à la même famille chimique que la colibactine. De plus, la protéase ClpQ intervient de concert avec Hsp90_{Ec} dans la production de colibactine et de yersiniabactine. Ces résultats confirment ainsi l'interconnexion entre la synthèse des deux facteurs de virulence de *E. coli*, la colibactine et la yersiniabactine. Enfin, l'analyse des effets de la mutation du gène *htpG* au cours d'une infection systémique chez l'animal, dans des modèles de sepsis et de méningite néonatale chez les rongeurs, démontre le rôle de la protéine de réponse au stress Hsp90_{Ec} dans la virulence de *E. coli*.

Le second projet a révélé que les polyamines sont impliquées dans la production de colibactine. L'étude du métabolisme des polyamines par une approche de microbiologie moléculaire a démontré que la spermidine est la polyamine nécessaire à la production de colibactine. Les résultats préliminaires de ce projet indiquent que la spermidine participerait à la régulation de l'expression de certains gènes de l'îlot génomique *pks*, et de fait modulerait la biosynthèse de colibactine. Des études complémentaires sont en cours pour élucider les mécanismes impliqués.

Les résultats de cette thèse sont une illustration parfaite de l'intégration symbiotique d'un élément génétique mobile acquis au cours de l'évolution au sein du chromosome bactérien, grâce à plusieurs connexions bilatérales permettant la production de facteurs de virulence par *E. coli*.

TITRE et RÉSUMÉ en anglais :

au recto de la dernière page de la Thèse

MOTS-CLÉS :

colibactine ; Hsp90 ; spermidine ; sidérophores ; virulence ; *Escherichia coli*

DISCIPLINE administrative :

Microbiologie

INTITULÉ et ADRESSE du LABORATOIRE :

Institut de Recherche en Santé Digestive (INSERM UMR1220 – INRA UMR1416 – ENVT – UT3)
CHU Toulouse Purpan – INSERM Bâtiment B – Place du Docteur Joseph Baylac – CS 60039
31024 Toulouse Cedex 3, France
