# Impact des facteurs micro-environnementaux de l'hôte sur la colonisation instestinale des Escherichia Coli adhérents et invasifs.

Lucie Gibold-Lyonne

▶ **To cite this version:**

Lucie Gibold-Lyonne. Impact des facteurs micro-environnementaux de l'hôte sur la colonisation instestinale des Escherichia Coli adhérents et invasifs.. Médecine humaine et pathologie. Université d'Auvergne - Clermont-Ferrand I, 2016. Français. NNT : 2016CLF1MM22 . tel-01654425

**HAL Id: tel-01654425**
**https://theses.hal.science/tel-01654425**

Submitted on 3 Dec 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ BLAISE PASCAL
Année 2016

UNIVERSITÉ D'AUVERGNE
N° d'ordre

***ECOLE DOCTORALE***
***DES SCIENCES DE LA VIE ET DE LA SANTE***
N° d'ordre :

# *Thèse*

Présentée à l'Université d'Auvergne

pour l'obtention du grade de DOCTEUR
(Décret du 5 juillet 1984)

Spécialité **Microbiologie**

soutenue le 28 septembre 2016

**GIBOLD - LYONNE Lucie**

---

# Impact des facteurs micro-environnementaux de l'hôte sur la colonisation intestinale des *Escherichia coli* adhérents et invasifs

---

**Rapporteurs** : Benoît FOLIGNÉ, Professeur, Faculté de Pharmacie, Lille
Philippe SEKSIK, Professeur, APHP, Paris

**Examinateurs** : Vincent CATTOIR, Professeur, UFR de Médecine, Caen
Mickaël DESVAUX, Directeur de recherche, INRA, Clermont-Fd/Theix
Patrick PLESIAT, Professeur, UFR de Médecine, Besançon

**Directeur de thèse** : Julien DELMAS, Maître de conférences, UFR de Médecine, Clermont-Fd

M2iSH Microbes, Intestin, Inflammation et Susceptibilité de l'Hôte
UMR Inserm 1071 / USC INRA 2018
Université d'Auvergne

*en souvenir de Gégé*
*et de tous ces petits*
*moments de bonheur…*

# *Remerciements*

Je remercie très sincèrement Monsieur le Professeur Philippe Seksik et Monsieur le Professeur Benoît Foligné d'avoir accepté d'être rapporteurs de cette thèse, d'avoir consacré du temps à sa lecture, son analyse et son jugement.

Je tiens également à remercier Monsieur le Professeur Patrick Plésiat, Monsieur le Professeur Vincent Cattoir, ainsi que Monsieur Mickaël Desvaux, pour avoir accepté d'être examinateurs de cette thèse.

Je souhaite particulièrement remercier Vincent de m'avoir accueillie à Caen et de m'avoir formée sur l'extraction d'ARN qui n'a plus de secrets pour moi, et Mickaël d'avoir suivi ce travail durant toutes ces années par le biais de mes comités de thèse, et de m'avoir donné plein de bons conseils.

Enfin, je souhaite tout spécialement remercier mon directeur de thèse Julien Delmas pour m'avoir confié ce sujet très intéressant et m'avoir encadrée tout le long de ce travail.
Julien, il n'y a pas assez d'espace dans ce mémoire pour te remercier, mais le cœur y est !

Il me semble que c'était hier encore que tu me donnais des conseils pour apprendre la bactériologie, alors que j'étais encore une jeune interne, puis pour purifier des cristaux de béta-lactamases pendant mon master 2.
Et alors que je devais travailler sur *E. coli* et le cancer colorectal en poursuivant les travaux d'Antony sur des aspects plutôt cellulaires, tu as trouvé un nouveau sujet intéressant plus centré sur la bactériologie que tu m'as aussitôt proposée.

Tu m'as soutenue tout au long de ce projet, et tu m'as permis de découvrir plein d'aspects différents de la recherche, et tu m'as toujours aidée à mener à bien toutes ces nouvelles expériences. Je n'ai autour de moi jamais vu un directeur de thèse prenant autant part aux manips de son étudiant, à tel point qu'en 5 ans, je n'ai trouvé qu'une seule expérience où tu as été bien moins à l'aise que moi !

Je ne suis pas prête d'oublier la première fois que nous avons manipulé des souris : 120 anses ligaturées en 15 jours sur la musique de magic system et pulp fiction! Et je n'oublierai pas non plus ces expériences parfois un peu loufoques qui nous ont conduit à enfermer un

biophysicien dans une pièce chauffée à 37°C pour chronométrer une bille, ou à récupérer des colorants alimentaires pour faire des tubes multicolores…

Alors merci encore pour tout, merci pour ces moments de franche rigolade, ces petits restos à chaque grève du self, ce petit séjour bostonien bien sympathique, ces séances de ping-pong annulées pour pouvoir finir à temps, et surtout pour ta bonne humeur toujours contagieuse et en toutes circonstances, car je ne pense pas qu'il y ait beaucoup d'encadrants qui prennent aussi bien le congé mater de son étudiante et de sa technicienne et qui se retrouve tout seul à travailler. Et en plus, on t'a fait le coup deux fois !

Bref, tu n'es pas un directeur ordinaire, alors surtout ne change pas car c'est un réel plaisir de travailler avec toi !

Je tiens également à remercier chaleureusement le Professeur Richard Bonnet qui m'a accueillie dans son laboratoire quand j'étais encore une jeune interne, et qui m'a toujours fait confiance, ce qui m'a permis d'arriver là où je suis aujourd'hui.

Richard, les moments où tu prends le temps de discuter avec nous de nos avancées en recherche sont toujours trop rares, mais tellement enrichissants. Tes explications sont toujours incroyablement claires, que ce soit pour nous expliquer comment fonctionne la résistance des entérobactéries aux antibiotiques, comment cristalliser une protéine ou pour expliquer le fonctionnement de linux, le langage python et toutes les lignes de commande réalisées pour l'analyse du RNA-seq.

Merci infiniment pour tout ça, et plus généralement pour m'avoir fait confiance plusieurs fois de suite, entre le master 2, la thèse d'exercice et la thèse d'université sur des sujets bien différents. Les quelques manipulations que j'ai réalisées avec toi m'ont marquée et je suis devenue une pro des colonnes échangeuses d'ions, de la chambre à agitation et de la manipulation des moustaches de chat.

Tu m'as également fait confiance côté hospitalier pour la gestion et l'accréditation d'un secteur, puis le changement informatique, au point même maintenant de me proposer un poste pérenne. Je ne te remercierai jamais assez pour tout ça, et ça me fait vraiment plaisir de pouvoir continuer à travailler dans ce service que l'on m'avait décrit comme « le pays des Bisounours » …

Je tiens aussi à remercier Sylvine, ma technicienne préférée, et ma copine de congé mater…
J'ai adoré travailler avec toi. Tu as toujours été enthousiaste, et tu n'as jamais rechigné à travailler, même pour manipuler des souris. Mon passage de l'autre côté du couloir fait que je ne peux plus autant t'embêter, mais j'espère avoir de nouvelles occasions de travailler avec toi.
Merci également à Sébastien, qui a ouvert ce sujet avec Julien, et grâce à qui j'ai travaillé sur ce sujet, et Camille, Charline, Elise et Cécile, qui m'ont accompagnée pour une partie de ce travail.

Merci également à toute l'équipe du R3 et du CNR : Antony, toujours joyeux, toujours positif, toujours éparpillé aussi, et toujours en tongs, même pour aller remplir des cuves d'azote ; Guillaume, toujours adorable, toujours volontaire, toujours plein de bonnes idées, et toujours meilleur que Julien au ping-pong ; Laëtitia et Anne-Sophie, à la fois râleuses et rigolotes, et vraiment adorables… et aussi Stéphane, Gauthier, Laurent, Alexis, Martine, Marlène, Virginie, Racha….

Merci aussi à tous les membres de l'équipe M2iSH, et plus particulièrement Hang pour avoir passé du temps à nous apprendre à parler anglais courtoisement à des reviewers pas sympathiques, Jérémy pour ta gentillesse, tes conseils et tes anticorps, Benoit et Lucie pour ce cours si particulier de réalisation d'anses ligaturées, Nicolas et Emilie pour vos précieux conseils, Estelle pour tous ces western-blot réalisés en urgence, Allison pour tes pots de high-fat qui ont rendu mes souris vraiment graisseuses, Manu pour tes fils de suture, mais également Jessica, Charlotte, Rafaella, Mathilde, Elisabeth, Jennifer, Sébastien, Pedro, Adeline, Amélie, Marie-Agnès, Caroline, Valérie et tous les autres que j'ai croisé au hasard des couloirs…

Merci également aux voisins du CBRV, et notamment Christiane et Damien pour leurs remarques judicieuses et leurs conseils. Merci à Caroline V., pour tous ces moments passés dans le noir à regarder des bactéries fluorescentes se promener dans du mucus artificiel, à David pour tous ces moments « chauds » à mesurer la vitesse des billes, et Lorraine pour toutes ces coupes réalisées toujours avec le sourire. Enfin, un grand merci à Jean-Louis Couderc pour sa gentillesse et ses conseils.

J'ai également une pensée émue pour Arlette, une meneuse d'équipe pas ordinaire, avec une force et une passion incroyable qui lui ont permis de continuer jusqu'au bout….

Je souhaite par ailleurs remercier toutes les personnes qui m'ont côtoyée et supportée pendant cette grande aventure, et notamment toute l'équipe du service de bactériologie, en particulier mes collègues : Régine, Christophe, Fred, Claire, Titi bien sûr, Mathilde, et Pierre le petit dernier, qui m'ont bien remplacée côté hospitalier pour que je puisse mener à bien ce travail.

Et merci aux internes Benoit, Arnaud, et tous les autres, qui m'ont parfois aidée, et parfois permis de trouver toutes ces longues journées un peu moins longues…

Je souhaite également adresser un grand merci à toute l'équipe du CACN pour tous ces moments de détente tellement sympas entre Aurillac, Lyon et Paris, en passant par chez nous, et tellement appréciables pour faire autre chose que travailler, alors merci à Titi, Gigi, Célinou, Bibi, Alexis, Guéna, Roland…

Merci aussi aux amis de toujours, sur qui on peut toujours compter, même après de nombreuses années, et qui sont toujours là quand on a besoin d'un moment de détente ou d'aide pour construire sa maison pendant qu'on écrit sa thèse, alors merci Pierre-Luc (et encore Titi), Alexandre, Caroline, Vincent…

Je remercie aussi toute ma famille et ma belle-famille pour m'avoir supportée, et pour m'avoir aidée à gérer tout en même temps : le travail, les enfants et la maison…, et en particulier ma maman, qui me soutient depuis toujours, même si elle ne comprend pas toujours pourquoi je fais tout cela…

Enfin, je tiens à remercier la seule personne sans qui rien n'aurait été possible : tu m'as toujours soutenue depuis que je te connais, dans les bons comme les mauvais moments, même quand je suis partie en Afrique, même quand je t'ai annoncé que j'allais peut-être travailler à Strasbourg, et même quand j'ai continué à jouer les étudiantes jusqu'à 30 ans bien sonnés, et même au-delà. Sans toi pour assurer l'intérim, la logistique des enfants et pour me réconforter dans les moments difficiles, cette thèse n'aurait jamais été terminée et pour ça, je te remercie de tout cœur. Merci aussi pour tous ces WE imposés de corvées pour me laisser travailler, merci pour toutes ces nuits de plâtre et d'enduit pour que la maison avance quand même et de gestion des enfants pour que je puisse me reposer…

Et grâce à toi, j'ai pu voir arriver dans cette période de ma vie les deux plus beaux cadeaux du monde…

# SOMMAIRE

*RESUME* ..... *6*

*ABSTRACT* ..... *7*

*LISTE DES ABRÉVIATIONS* ..... *8*

*INDEX DES ILLUSTRATIONS* ..... *12*

**Liste des Figures** ..... **12**

**Liste des tableaux** ..... **14**

*INTRODUCTION* ..... *15*

*ETUDE BIBLIOGRAPHIQUE* ..... *18*

**I. Le système digestif** ..... **20**

1. Physiologie du système digestif ..... 20
2. L'épithélium intestinal ..... 23
   1) Les entérocytes ..... 23
   2) Les cellules caliciformes ..... 24
   3) Les cellules entéro-endocrines ..... 24
   4) Les cellules de Paneth ..... 25
   5) Les cellules M ..... 25
3. Renouvellement cellulaire de l'épithélium intestinal ..... 25
4. Fonctions de la barrière épithéliale intestinale ..... 27
   1) Barrière physique ..... 29
   2) Barrière chimique ..... 29
      a. Le mucus intestinal ..... 29
      b. Les peptides antimicrobiens ..... 40
   3) Barrière immunologique ..... 41
      a. Défenses non spécifiques : immunité innée ..... 42
      b. Acteurs de l'immunité adaptative ..... 44

**II. Le microbiote intestinal** ..... **50**

1. Mise en place du microbiote intestinal ..... 50
2. Composition du microbiote intestinal ..... 52

1) Les bactéries ... 52
a. Axe longitudinal ... 54
b. Axe radial ... 56
2) Les *Archaea* ... 57
3) Les autres micro-organismes ... 57
3. Fonctions du microbiote ... 57
1) Fonctions métaboliques ... 59
a. Métabolisme des polysaccharides ... 59
b. Métabolisme des acides biliaires ... 60
c. Autres métabolismes ... 62
2) Fonctions trophiques ... 62
3) Fonctions de défense ... 64
4) Fonctions immunitaires ... 64
5) Fonctions émergentes ... 66
4. Dysbioses ... 68
**III. La maladie de Crohn ... 71**
1. Epidémiologie ... 71
2. Signes cliniques ... 73
3. Diagnostic et suivi ... 75
1) Diagnostics différentiels ... 75
2) Suivi de la maladie ... 77
3) Marqueurs biologiques ... 77
4. Complications ... 80
5. Traitement ... 82
1) Les médicaments anti-inflammatoires ... 82
a. Les dérivés salicylés ... 82
b. Les corticoïdes ... 83
2) Les médicaments immuno-modulateurs ... 83
a. Azathioprine et 6-mercaptopurine ... 84
b. Méthotrexate ... 84
3) Les anti-TNF-α ... 85
4) La chirurgie ... 85
5) Les nouvelles biothérapies ... 87
6) Les traitements actifs sur le microbiote intestinal ... 88
a. Les antibiotiques ... 88
b. Les probiotiques ... 89
c. La transplantation de microbiote fécal ... 89
6. Etiologies de la MC ... 90
1) Facteurs environnementaux ... 91
2) Facteurs génétiques ... 93
a. Implication du gène *NOD2* ... 95

b. Implication du gène *SLC15A1* ... 96
c. Implication des gènes codant des protéines de l'autophagie ... 97
d. Implication du gène NLRP3 ... 99
e. IL23-R ... 100
f. Les TLR ... 101
3) Facteurs immunologiques ... 101
4) Autres facteurs de l'hôte ... 102
a. Le stress du réticulum endoplasmique ... 102
b. Le récepteur CEACAM6 ... 103
c. Les microARN régulateurs ... 103
d. La glycosylation des mucines ... 104
e. Les acides biliaires ... 104
5) Rôle du microbiote intestinal ... 105
a. Présence d'une translocation excessive ... 105
b. Présence d'une dysbiose ... 106
c. Présence d'agents pathogènes persistant au niveau intestinal ... 109

**IV. *Escherichia coli* et maladie de Crohn ... 113**

1. Description des différents pathovars d'*E. coli* ... 113
1) Généralités ... 113
2) Les *E. coli* pathogènes intestinaux et extra-intestinaux ... 114
2. Le pathovar *Escherichia coli* adhérent et invasif associé à la maladie de Crohn ... 115
1) Propriétés d'adhésion et d'invasion des AIEC ... 116
2) Survie et multiplication dans les macrophages des AIEC ... 116
3) Autres caractéristiques des AIEC ... 117
3. Facteurs de virulence retrouvés chez les AIEC ... 117
1) Les *fimbriae* ... 118
2) Les flagelles ... 118
3) Les vésicules de membrane externe ... 119
4) Le système d'osmorégulation EnvZ/OmpR ... 120
5) Les facteurs d'invasion et de multiplication cellulaire ... 121
6) Les systèmes de sécrétion ... 123
a. Généralités ... 123
b. Les différents systèmes de sécrétion chez les AIEC ... 124
c. Caractéristiques des autotransporteurs monomériques ... 130
d. Sécrétion et fonction des autotransporteurs ... 133

**V. Etude de l'influence de l'environnement physico-chimique de l'intestin sur la colonisation bactérienne ... 146**

1. Influence du pH ... 146
1) Rôle du pH sur la survie et la croissance bactérienne ... 146
2) Influence du pH sur la virulence des bactéries ... 149
2. Influence des acides biliaires ... 150

1) Rôle des acides biliaires sur la survie et la croissance bactérienne ..... 150
2) Influence des acides biliaires sur la virulence des bactéries ..... 151
3. Influence du mucus ..... 153
1) Rôle du mucus dans la colonisation bactérienne ..... 154
2) Influence du mucus sur la virulence des bactéries ..... 155

*TRAVAIL EXPERIMENTAL ..... 156*

**I. Présentation des résultats ..... 157**

**II. La protéase Vat-AIEC favorise la traversée du mucus par les *E. coli* adhérents et invasifs au niveau intestinal ..... 158**

1. Introduction à la publication n°1 ..... 159
2. Publication n°1 ..... 160
3. Commentaires de la publication n°1 ..... 185
4. Résultats complémentaires ..... 189

**III. Impact des sels biliaires sur le transcriptome des AIEC ..... 193**

1. Introduction à la publication n°2 ..... 195
2. Publication n°2 ..... 197
3. Commentaires de la publication n°2 ..... 253

**IV. Discussion – Perspectives ..... 257**

*REFERENCES BIBLIOGRAPHIQUES ..... 261*

*ANNEXES ..... 303*

**Annexe n°1 ..... 303**

**Annexe n°2 ..... 312**

**Annexe n°3 ..... 324**

**Annexe n°4 ..... 337**

**Annexe n°5 ..... 363**

# Résumé

La maladie de Crohn (MC) est une affection inflammatoire chronique du tube digestif dont l'étiologie est multifactorielle. Les lésions intestinales des patients atteints de MC sont anormalement colonisées par des souches pathogènes d'*Escherichia coli* appartenant au pathovar AIEC pour « Adherent-Invasive *Escherichia coli* ». Ces souches sont capables d'adhérer et d'envahir les cellules épithéliales intestinales, et ont la capacité de survivre et de se multiplier en macrophages en induisant une synthèse intense de cytokines pro-inflammatoires. Les AIEC pourraient ainsi être impliquées dans l'induction et/ou l'entretien de l'état inflammatoire de la muqueuse intestinale.

L'objectif de ce travail est d'identifier les déterminants bactériens des AIEC qui vont intervenir dans les étapes précoces de l'implantation des AIEC au niveau intestinal et de définir quel est le rôle des facteurs micro-environnementaux de l'hôte dans cette implantation.

Nous montrons que l'AIEC LF82 possède une activité mucinolytique qui est portée par le gène *vat-AIEC* et qui favorise la traversée du mucus et la colonisation intestinale. Nous avons retrouvé ce gène chez 42% des souches AIEC isolées de patients atteints de MC, et chez 97% des souches AIEC appartenant au phylogroupe B2. Par ailleurs, nous avons montré que les sels biliaires augmentent l'expression de cette mucinase.

Nous avons ensuite étudié l'influence des sels biliaires sur l'expression globale des gènes de la souche LF82. Les sels biliaires modifient profondément le métabolisme de la souche, induisant une diminution globale des voies de biosynthèse (protéines, lipides) et une augmentation des voies de dégradation (alcools, acides carboxyliques, polyamines, …). L'étude du catabolisme de l'éthanolamine et du propanediol indique que les AIEC pourraient se servir de ces substrats pour s'implanter au sein de la flore iléale. De plus, les analyses transcriptomiques révèlent que les sels biliaires augmentent l'expression de gènes codant des facteurs de virulence comme l'invasine IbeA, les systèmes de sécrétion de type VI et la yersiniabactine. Nous montrons également qu'ils favorisent la formation de biofilm chez les souches AIEC.

Ces données indiquent que les sels biliaires constituent un signal permettant à la souche AIEC LF82 de mettre en place différentes voies métaboliques et déterminants bactériens nécessaires à son implantation au niveau intestinal.

**Mots-clé** : *Escherichia coli*, maladie de Crohn, mucines, serine protéase, mucinase, AIEC, sels biliaires, transcriptome

# Abstract

The etiology of Crohn's disease (CD) involves disorders in host genetic factors and intestinal microbiota. Ileal mucosa of CD patients is often abnormally colonized by adherent-invasive *Escherichia coli* (AIEC). These strains isolated from the intestinal mucosa of CD patients are able to adhere to intestinal epithelial cells (IECs). This adhesion to IECs promotes the invasion of cells by the bacteria. Furthermore, the invasive ability of AIEC strains allows bacteria to translocate across the human intestinal epithelium, move into the deep tissues and activate immune cells continuously upon arrival. Thus AIEC could be involved in the inflammatory state of the intestinal mucous membrane.

The aim of this study was to identify components of AIEC virulence, which might favor their implantation in the gut of CD patients and to define the role of several chemical factors from the ileal environment. Here, we reported a protease called Vat-AIEC from AIEC which favors the penetration of AIEC through the mucus layer and enhances gut colonization. The screening of *E. coli* strains isolated from CD patients revealed a preferential *vat-AIEC* association with AIEC strains belonging to the B2 phylogroup. Besides, *Vat-AIEC* transcription was increased with bile salts from the ileum environment.

Then a global RNA sequencing (RNA-seq) of *E. coli* LF82 has been used to observe the impact of bile salts on the expression of bacterial genes. The results demonstrate the explosive effect of bile salts with a dysregulation of about 40% of the genome, with a global upregulation of genes involved in degradation and downregulation of those implicated in several biosynthesis. Our results show that LF82 use ethanolamine as a nitrogen source and propane diol as a carbon source, which can favor their colonization in the gut compared to the other bacteria. We also studied virulence genes expression in the presence of bile salts. They increase the expression of several virulence factors like the IbeA invasion, the type 6 secretion systems and the yersiniabactin. Furthermore, we noticed an increased expression of genes implicated in biofilm formation.

These results improve the understanding of the global regulatory network in the presence of bile salts and thus of AIEC implantation in the human gut of CD patients.

**keywords** : *Escherichia coli*, Crohn’s Disease, mucins, serine protease, mucinase, AIEC, bile salts, transcriptome

# Liste des abréviations

5-ASA : acide 5-aminosalicylique
ADN : acide désoxyribonucléique
AGCC : acide gras à chaînes courtes
AIEC : *E. coli* adhérents et invasifs
AINS : anti-inflammatoires non stéroïdiens
AKT : protéine kinase B
ANCA : antineutrophil cytoplasmic antibodies
APEC : *Escherichia coli* pathogènes aviaires
ARN : acide ribonucléique
ASC : Apotosis-associated speck-like protein containing a caspase recrutment domain
ASCA : anti-*saccharomyces cerevisiae* antibodies
ATG16L1 : autophagy-related 16-like 1
ATP : adénosine triphosphate
Bam : β-barrel assembly machinery
BDNF : facteur neurotrophique dérivé du cerveau
BMP : Bone Morphogenic Proteins
Boa : *bongori* autotransporter
CARD : caspase-recruitment domains
CDAI : index d'activité de la maladie de Crohn
CEACAM6 : Carcinoembryonic Antigen–related Cell Adhesion Molecule 6
CgA : chromogranine A
CLR : *C-lectin-like* récepteurs
CPA : cellules présentatrices d'antigènes
CRC : cancer colorectal
DAEC : *Escherichia coli* à agrégation diffuse
DEG : differentially expressed gene
DSS : Dextran Sodium Sulfate
EAEC : *Escherichia coli* entéro-agrégatifs
EatA : ETEC autotransporter A
EGF : epidermal growth factor
EHEC : *Escherichia coli* entéro-hémorragiques
EIEC : *Escherichia coli* entéro-invasifs
EpeA : EHEC plasmid-encoded autotransporter
EPEC : *Escherichia coli* entéro-pathogènes

EspC : EPEC secreted protein C
EspI : *E. coli* secreted protease, island-encoded
EspP : extracellular serine protease plasmid (pO157)-encoded
ETEC : *Escherichia coli* entéro-toxinogènes
ExPEC : extra-intestinal pathogenic *Escherichia coli*
FODMAP : Fermentable Oligo-, Di-, Mono-saccharides And Polyols
FXR : Farnesoid X receptor
GABA : γ-aminobutyrate
GalNAc : N-acétyl-D- galactosamine
GALT : Gut Associated Lymphoid Tissue
GlcNAc : N-acétyl-D-glucosamine
GSP : General secretion pathway
GWA : Genome-Wide Association
HBI : index de Harvey-Bradshaw
Hbp : hemoglobin protease
Hcp : Haemolysin co-regulated protein
HMP : Human microbiome project
Ig : immunoglobuline
IFN : interféron
IHH : Indian Hedgehog
IL : interleukine
IRAK : kinase associée au récepteur de l'interleukine-1
IRF3 : facteur de transcription 3 régulant l'interféron
IRGM : immunity-related GTPase family M
IRM : imagerie par résonance magnétique
iTreg : lymphocytes T régulateurs induits
JNK : c-jun N-terminal kinase
*LEE* : locus of enterocyte effacement
lpf : long polar *fimbriae*
LPS : lipopolysaccharide
LRR : leucine-rich repeat
MALT : tissu lymphoïde associé aux muqueuses
MAP kinase : mitogen activated protein kinase
MC : maladie de Crohn
MDP : muramyl di-peptide
miARN : microARN
MICI : maladie inflammatoire chronique de l'intestin

MNEC : *Escherichia coli* associés aux méningites
MyD88 : facteur de différenciation myéloïde 88
NBD : nucleotide-binding domain
NF-κB : nuclear factor-kappa B
NGS : next-generation sequencing
NK : natural killer
NOD2 : nucleotide-binding oligomerization domain 2
NLRP3 : NOD like receptor family, pyrin domain contining 3
Omp : outer membrane protein
OMV : outer membrane vesicles
OR : Odd Ratio
PAMP : pathogen associated molecular patterns
PDGF : platelet-derived growth factor
PepT1 : peptide transporter 1
Pet : plasmid-encoded toxin
PI3K : phosphoinositide 3-kinase
Pic : protease involved in intestinal colonization
PRR : pattern recognition receptor
PT : domaine riche en proline, thréonine et/ou sérine
RCH : rectocolite hémorragique
RE : réticulum endoplasmique
RLR : *RIG-I-like* récepteurs
RPKM : reads per kilobase of transcript per million reads mapped
SAAT : Self-Associating AutoTransporter
Sat : secreted autotransporter toxin
SepA : *Shigella* extracellular protein A
SHH : Sonic Hedgehog
SHU : syndrome hémolytique et urémique
SNP : single nucleotide polymorphism
SPATE : serine protease autotransporters of *Enterobacteriaceae*
Spdef : SAM pointed domain-containing Ets transcription factor
SSTn : système de sécrétion de type n
Tafi : thin aggregative *fimbriae*
TAM : translocation and assembly module
Tat : Twin arginine translocation
TFF : trefoil factor
TGF-β : transforming growth factor beta

TMF : transplantation de microbiote fécal
TNF-α : tumor necrosis factor alpha
TLR : *Toll-like* recepteurs
TPH-1 : tryptophane hydroxylase-1
Tr1 : lymphocytes T régulateurs 1
TRAF : facteur associé au récepteur du TNF
TRIF : TIR-domain-containing adapter-inducing interferon-β
Tsh : temperature-sensitive hemagglutinin
UPEC : *Escherichia coli* uropathogènes
Vat : vacuolating autotransporter toxin
VgrG : Valine-Glycine repeat protein G
XBP1 : X Box binding protein-1
ZO : *zonula occludens*

# *Index des Illustrations*

## Liste des Figures

**Figure 1** : Les différentes couches de la paroi du tube digestif

**Figure 2** : L'épithélium intestinal

**Figure 3** : Les différents types de cellules intestinales

**Figure 4** : La barrière épithéliale intestinale

**Figure 5** : Les différents complexes jonctionnels cellulaires au niveau intestinal

**Figure 6** : Structure des mucines

**Figure 7** : Répartition de la couche de mucus le long du tractus gastro-intestinal

**Figure 8** : Voies de signalisation activées par les PRR suite à une stimulation bactérienne

**Figure 9** : La réponse adaptative des lymphocytes T

**Figure 10** : Arbre phylogénétique des principales espèces bactériennes retrouvées dans le microbiote intestinal

**Figure 11** : Facteurs environnementaux influençant le microbiote intestinal

**Figure 12** : Composition et concentration des espèces bactériennes dominantes de la flore commensale le long du tractus digestif

**Figure 13** : Rôles du microbiote intestinal dans les fonctions physiologiques de l'intestin

**Figure 14** : Modulation de l'immunité intestinale par le microbiote intestinal

**Figure 15** : Représentation des interactions entre le microbiote intestinal et l'hôte chez des sujets sains et dans le cadre d'une dysbiose

**Figure 16** : Différentes localisations possibles de la maladie de Crohn

**Figure 17** : Incidence de la maladie de Crohn entre 1960 et 1980 et après 1980

**Figure 18 :** Ulcérations retrouvées dans la maladie de Crohn

**Figure 19** : Stratégies de traitement de la maladie de Crohn

**Figure 20** : Nouvelles thérapies en cours d'essai clinique ou venant d'être commercialisées, en fonction de leur mécanisme d'action et de leur voie d'administration

**Figure 21 :** Facteurs génétiques dans l'étiologie des maladies inflammatoires chroniques de l'intestin

**Figure 22** : Structure 3D du récepteur NOD2

**Figure 23** : Représentation schématique de la voie de l'autophagie

**Figure 24** : Dysbiose retrouvée dans la muqueuse des patients atteints de MC, comparativement aux individus sains

**Figure 25** : Sites colonisés par des souches d'*E. coli* pathogènes

**Figure 26** : Les différents systèmes de sécrétion chez les Entérobactéries

**Figure 27** : Le système de sécrétion de type V

**Figure 28** : Structure primaire des autotransporteurs monomériques

**Figure 29** : Sécrétion des autotransporteurs monomériques

**Figure 30** : Deux modèles suggérés pour la sécrétion du domaine passager à travers la membrane externe

**Figure 31** : Système d'auto-clivage des autotransporteurs de type SPATE

**Figure 32** : Analyse phylogénétique de la séquence en acides aminés du domaine passager des différentes protéines SPATEs avec visualisation des deux groupes

**Figure 33** : Le domaine passager des SPATEs

**Figure 34** : Rôle du pH, des sels biliaires et du mucus sur les différentes fonctions bactériennes

**Figure 35** : Etude *in silico* du promoteur du gène *vat-AIEC*

**Figure 36** : Coupes histologiques d'intestin de souris (anses ligaturées) après 12h de contact avec la bactérie

**Figure 37** : Différentes étapes de l'analyse RNA-seq

**Figure 38** : Principe de la comparaison des données issues du séquençage des ARN en Fonction des conditions de culture

# Liste des tableaux

**Tableau 1** : Caractéristiques des gènes codant les différentes mucines

**Tableau 2** : Graduation de l'activité de la MC d'après le consensus ECCO, Index d'activité de la maladie de Crohn et Index de Harvey-Bradshaw

**Tableau 3** : Apport des marqueurs biologiques dans le diagnostic de MC ou RCH pour des colites indéterminées

**Tableau 4** : Classification des autotransporteurs monomériques en fonction de l'activité du domaine passager

**Tableau 5** : Principaux autotransporteurs de la famille des SPATEs

# *Introduction*

La maladie de Crohn (MC) est une maladie inflammatoire chronique de l'intestin caractérisée par un état d'hyperactivation du système immunitaire intestinal. L'étiologie de cette maladie qui touche les individus jeunes est encore mal connue, mais elle semble être due à une réponse immunitaire inadaptée au microbiote intestinal chez des individus génétiquement prédisposés. Parmi les agents infectieux incriminés dans cette pathologie, plusieurs arguments épidémiologiques, cliniques et expérimentaux sont en faveur d'un rôle de l'espèce *Escherichia coli*. L'espèce *E. coli* comprend des souches commensales et également des bactéries pathogènes qui peuvent être à l'origine d'infections intestinales ou extra-intestinales. Les *E. coli* pathogènes intestinaux sont définis par des facteurs de virulence spécifiques en fonction des différents pathovars. A l'inverse, les *E. coli* pathogènes extra-intestinaux sont des pathogènes facultatifs qui appartiennent à la flore commensale, et qui lorsqu'ils colonisent d'autres niches écologiques peuvent provoquer des infections plus ou moins graves. La pathogénie des *E. coli* extra-intestinaux n'est pas lié un facteur de virulence spécifique mais à une association de différents facteurs qui peuvent varier selon les souches.

Les souches d'*E. coli* retrouvées chez les patients atteints de MC correspondent à un pathovar particulier : les *E. coli* adhérents et invasifs (AIEC). Les AIEC adhèrent fortement à l'épithélium intestinal, envahissent les cellules et se multiplient à l'intérieur de ces cellules ainsi que des macrophages. L'étude des mécanismes impliqués dans la pathogénie des AIEC a montré qu'ils sont capables de transloquer à travers l'épithélium intestinal pour atteindre la *lamina propria*, où ils survivent et se multiplient à l'intérieur des macrophages sans induire d'apoptose cellulaire mais conduisant à la libération de cytokines pro-inflammatoires comme TNF-α, IL-1β et IL-6 et à la formation éventuelle de granulomes. Les AIEC ne se caractérisent pas par des marqueurs génétiques particuliers mais possèdent différents facteurs de virulence observés également chez les *E. coli* responsables d'infections extra-intestinales (Vejborg *et al.*, 2011).

Les conditions physicochimiques retrouvées au niveau digestif, comme le pH et les sécrétions, peuvent jouer un rôle sur la colonisation et la virulence des bactéries (Gupta and Chowdhury, 1997; Harjai *et al.*, 2005; Hamner *et al.*, 2013; Sengupta *et al.*, 2014; Hicks *et al.*, 2015). Par exemple, la bile, qui joue un rôle majeur dans la solubilisation des lipides (Hofmann, 1999), possède également un rôle antimicrobien important (Begley et al., 2005). Mais elle pourrait aussi favoriser la formation de biofilm bactérien au niveau intestinal.

L'objectif de ce travail est d'identifier les déterminants bactériens qui vont intervenir dans les étapes précoces de l'implantation des AIEC au niveau intestinal et de définir quel est le rôle des facteurs micro-environnementaux de l'hôte dans cette implantation. La première partie de ce travail comprend une synthèse bibliographique s'articulant autour de l'épithélium digestif, la maladie de Crohn et les AIEC.

Les résultats de ce travail sont regroupés dans une deuxième partie. Le premier article identifie une enzyme au sein de la souche de référence des AIEC LF82 qui possède une activité mucinolytique. Nous montrons que cette mucinase dénommée Vat-AIEC permet aux AIEC de traverser la couche protectrice de mucus pour atteindre les cellules épithéliales. Le gène codant cette protéine est retrouvé plus fréquemment chez les souches d'AIEC que les non AIEC isolées de patients atteints de MC, et il est associé aux phylogroupes B2 et D des *E. coli*. L'expression de cette mucinase est augmentée par les facteurs micro-environnementaux rencontrés au niveau iléal. Le deuxième article montre l'impact des sels biliaires sur le transcriptome de l'AIEC LF82. En effet, les sels biliaires modifient profondément le métabolisme de la souche, induisant une diminution globale des voies de biosynthèse (protéines, lipides) et une augmentation des voies de dégradation (alcools, acides carboxyliques, polyamines, …). L'étude du catabolisme de l'éthanolamine et du propanediol indique que les AIEC pourraient se servir de ces substrats pour s'implanter au sein de la flore iléale. De plus, les analyses transcriptomiques révèlent que les sels biliaires augmentent l'expression de gènes codant des facteurs de virulence habituellement retrouvés chez les AIEC et qu'ils stimulent la formation du biofilm.

Les sels biliaires jouent donc un rôle clé dans l'implantation des AIEC au niveau iléal.

# *Etude bibliographique*

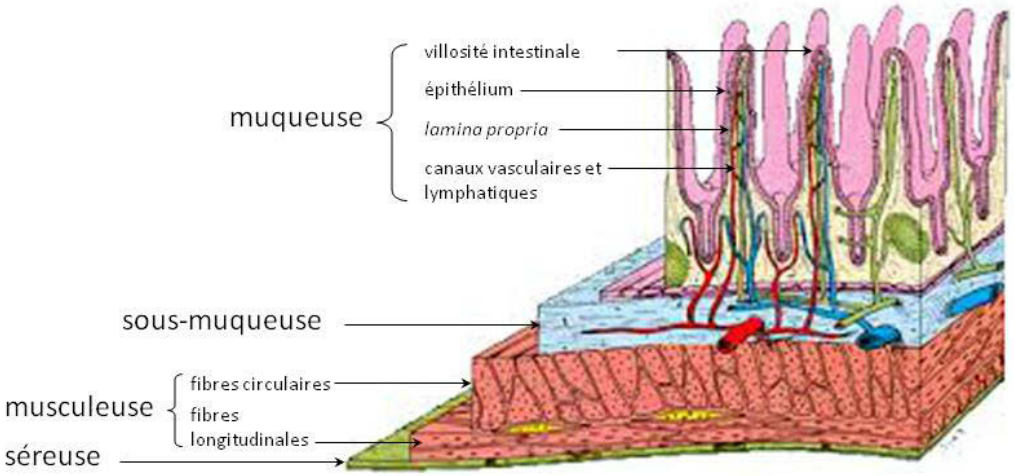



**Figure 1** : Les différentes couches de la paroi du tube digestif
(image modifiée d'après le site internet http://www.corpshumain.ca)

# I. Le système digestif

Le tube digestif, d'une longueur moyenne de 4,5 mètres chez l'Homme, est constitué successivement de la bouche, du pharynx, de l'œsophage, de l'estomac, de l'intestin grêle (duodénum, jéjunum et iléon), du gros intestin (caecum, colon ascendant, colon transverse, colon descendant, colon sigmoïde et rectum) et de l'anus. Les organes digestifs accessoires tels que les glandes salivaires, le pancréas, le foie et la vésicule biliaire, constituent avec le tube digestif proprement dit le système digestif. Ce système joue un rôle majeur dans l'homéostasie en permettant le transfert des nutriments, de l'eau et des électrolytes de l'environnement extérieur vers le milieu intérieur (Sherwood, 2006).

La structure générale de la paroi du tube digestif est globalement la même sur toute la longueur, avec toutefois quelques spécificités régionales. Elle comporte 4 couches qui sont, de la lumière intestinale aux couches les plus profondes, (i) la muqueuse, comprenant un chorion de tissu conjonctif et un épithélium bordant la lumière, (ii) la sous-muqueuse, constituée de tissu conjonctif, (iii) la musculeuse composée de cellules musculaires lisses, et (iv) la séreuse, composée également de tissu conjonctif (**Figure 1**). La muqueuse est la couche la plus complexe : elle est composée d'un épithélium digestif et repose sur une lame basale appelée *lamina propria* riche en vaisseaux sanguins et lymphatiques et sur une fine couche musculaire de soutien appelée sous-muqueuse.

## 1. Physiologie du système digestif

La fonction physiologique principale du système digestif est la digestion et l'assimilation des aliments tout en protégeant notre organisme contre des agents nocifs ou des micro-organismes pathogènes. La digestion permet la transformation des aliments (protéines, glucides et lipides) en molécules qui vont pouvoir passer de la lumière du tube digestif dans le milieu intérieur. Cette digestion commence au niveau de la bouche avec la mastication des aliments, et également la dégradation des glucides par une amylase de la salive. Le bol alimentaire parcourt ensuite l'œsophage pour atteindre l'estomac. A ce niveau, les aliments sont stockés, mélangés au suc gastrique pour devenir le chyme. La sécrétion d'acide

chlorhydrique par l'estomac rend le pH très acide, jusqu'à 2, ce qui permet à la fois d'activer certaines enzymes, de fragmenter les tissus conjonctifs et musculaires pour réduire la taille des aliments, de dénaturer les protéines pour activer leur digestion, et enfin de détruire la plupart des micro-organismes ingérés avec les aliments. Au niveau du duodénum, le chyme est ensuite mélangé avec les sécrétions biliaire et pancréatique contenant les enzymes de la digestion. Le pH acide est alors neutralisé par la bile, de pH basique, ainsi que par des ions bicarbonates provenant du pancréas, afin de ne pas irriter la muqueuse et inactiver les enzymes de la digestion. Les sels biliaires permettent à ce niveau la formation de petites gouttelettes de graisse appelées micelles, ce qui permet aux lipides d'être digérés par la lipase pancréatique puis absorbés au niveau de l'intestin grêle. La bile assure également des fonctions de détoxication, avec la dégradation de certains médicaments et de l'alcool, l'excrétion de déchets comme la bilirubine qui sont toxiques pour l'organisme, l'élimination de l'excès de cholestérol et enfin la limitation de la prolifération bactérienne intestinale dans la partie haute de l'appareil digestif. La bile a ainsi un rôle primordial dans la digestion et l'absorption des nutriments par l'épithélium intestinal. La digestion se termine au niveau de l'intestin grêle (principalement duodénum et jéjunum) par l'absorption des produits issus de la digestion des protéines, lipides et glucides, mais également des vitamines, des électrolytes et de l'eau au niveau de la bordure en brosse de l'épithélium. Seule la partie terminale de l'iléon permet l'absorption de la vitamine B12 et la réabsorption des sels biliaires. Le gros intestin quant à lui ne sécrète pas d'enzyme de la digestion mais seulement du mucus pour faciliter le passage du bol alimentaire. Seules les bactéries présentes dans le colon digèrent un peu du contenu du bol alimentaire pour leur propre métabolisme. Après passage à travers le gros intestin, les résidus non digérés des aliments (fèces) sont finalement éliminés par le rectum (Sherwood, 2006).

Le système digestif peut être considéré comme un écosystème complexe composé d'une alliance stable entre l'épithélium intestinal, le système immunitaire associé et la flore microbienne résidente. Des altérations génétiques ou fonctionnelles de l'un de ces 3 composants de l'écosystème peuvent perturber l'homéostasie intestinale et par conséquent favorisent l'apparition de pathologies (McCracken and Lorenz, 2001).

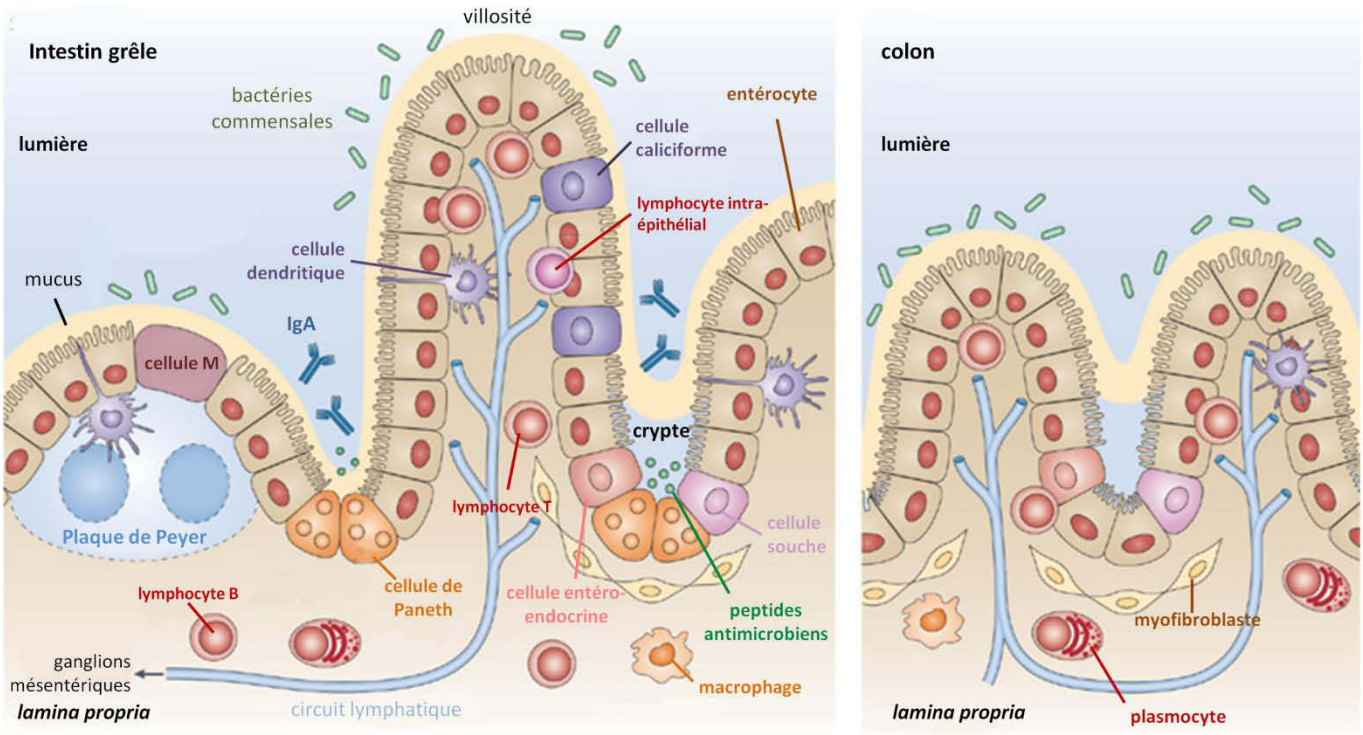



**Figure 2** : L'épithélium intestinal (image modifiée d'après Abreu, 2010)

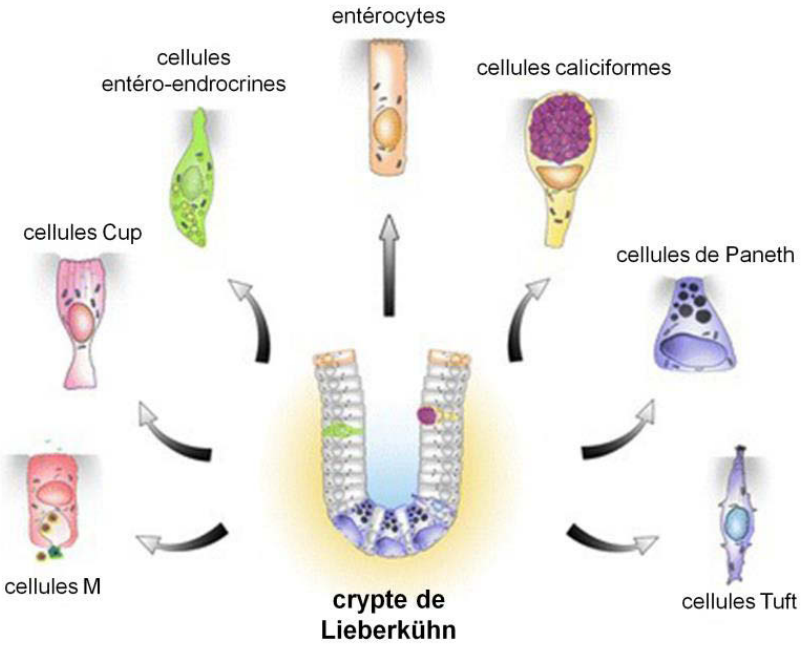



**Figure 3** : Les différents types de cellules intestinales
(image modifiée d'après Gerbe *et al*, 2012)

## 2. L'épithélium intestinal

L'épithélium intestinal est un tissu d'environ 300 $m^2$ de surface, soit l'équivalent d'un terrain de tennis, comprenant une seule couche de cellules. Ce tissu possède le renouvellement le plus actif de tout l'organisme. En effet, il est remplacé tous les 4 à 5 jours, ce qui correspond à une production d'environ 10 milliards de nouvelles cellules chaque jour et autant sont éliminées par desquamation. Cet épithélium joue donc un rôle crucial dans le maintien de l'homéostasie intestinale.

Il possède une structure organisée différemment entre l'intestin grêle et le colon, composée de villosités et d'invaginations appelées cryptes de Lieberkühn (**Figure 2**). L'épithélium au niveau de l'intestin grêle possède également des microvillosités qui vont permettre d'accroitre la surface d'échange et ainsi favoriser l'absorption intestinale. Les cryptes quant à elles jouent un rôle essentiel dans l'homéostasie cellulaire car leur fond est tapissé de cellules souches permettant le renouvellement permanent de l'épithélium intestinal (Clevers, 2013).

L'épithélium intestinal comprend sept types de cellules différentes, dérivant toutes des cellules souches pluripotentes situées au fond des cryptes (Madara, 1982; Snoeck *et al.*, 2005; Roda *et al.*, 2010; Gerbe *et al.*, 2012). Parmi elles, sont retrouvés les entérocytes, des cellules sécrétrices (cellules caliciformes, cellules entéro-endocrines et cellules de Paneth), des cellules impliquées dans les défenses immunitaires (cellules microfold ou M), et des cellules dont la fonction reste encore inconnue à ce jour (cellules Tuft et cellules Cup) (**Figure 3**).

### 1) Les entérocytes

Les entérocytes constituent 90 à 95 % de l'épithélium intestinal et sont caractérisés par la présence de microvillosités appelées bordure en brosse permettant d'augmenter leur capacité d'absorption. Ce sont des cellules responsables de l'absorption des nutriments (au niveau de l'intestin grêle) et de l'eau (au niveau du colon). Elles réabsorbent également les acides biliaires et permettent les échanges d'ions.

Ces cellules peuvent également synthétiser des cytokines pro-inflammatoires comme l'interleukine 6 (IL-6), l'interleukine 8 (IL-8) ou le tumor necrosis factor alpha (TNF-α), ainsi que des peptides antimicrobiens de type β-défensines (Jung *et al.*, 1995; Huttner and Bevins, 1999). Elles participent ainsi aux défenses de l'organisme face aux agressions microbiennes potentielles.

### 2) Les cellules caliciformes

Les cellules caliciformes sécrètent le mucus, qui forme un film protecteur qui recouvre la surface des cellules, ainsi que des peptides en trèfle (TFF, trefoil factor), qui interagissent avec les mucines pour former le mucus. Grâce à leur capacité d'induire la motilité cellulaire, ces peptides prennent également part aux processus de régénération de l'épithélium gastro-intestinal lorsque celui-ci est lésé (Ogata and Podolsky, 1997).

### 3) Les cellules entéro-endocrines

Les cellules entéro-endocrines présentent comme les entérocytes des microvillosités à leur pôle apical et des granules sécrétoires caractéristiques au niveau de leur pôle basal. Elles sécrètent diverses hormones peptidiques telles que la sécrétine, la cholécystokinine et la somatostatine, qui vont intervenir par voie endocrine, paracrine et autocrine dans la régulation de la motricité et des sécrétions digestives. Ces cellules synthétisent également de la sérotonine, des catécholamines, de l'adrénaline et de la noradrénaline. Les catécholamines vont permettre une meilleure perméabilité vasculaire et donc un passage plus facile dans la circulation sanguine des polypeptides. Enfin, ces cellules interviennent également dans la prolifération cellulaire et l'angiogenèse. Parmi les cellules entéro-endocrines, on peut citer les cellules entérochromaffines qui utilisent la tryptophane hydroxylase-1 (TPH-1) pour synthétiser de la sérotonine et ainsi activer certains signaux vers le cerveau (comme par exemple le réflexe nauséeux).

### 4) Les cellules de Paneth

Les cellules de Paneth sécrètent quant à elles de grandes quantités de peptides antimicrobiens, ce qui leur permet de contrôler le microbiote présent dans la lumière intestinale. Ces cellules ne sont présentes qu'au niveau de l'épithélium de l'intestin grêle, afin de ne pas interférer avec la flore bactérienne qui se situe en majorité au niveau du colon. Contrairement aux autres cellules dérivant des cellules souches intestinales, ces cellules restent dans le fond des cryptes et ne seront éliminées par phagocytose qu'au bout d'environ 23 jours. La sécrétion de facteurs de croissance et d'autres molécules régulatrices par ces mêmes cellules suggère qu'elles pourraient contribuer au maintien de la niche de cellules souches *via* une sécrétion paracrine et/ou à la prolifération et différenciation d'autres types cellulaires (Bry *et al.*, 1994; Cadwell *et al.*, 2008; Liu *et al.*, 2013).

### 5) Les cellules M

Les cellules M font partie de l'épithélium associé aux follicules ; elles sont situées dans des zones particulières de l'intestin grêle appelées plaques de Peyer. Elles servent de senseurs de composants présents dans la lumière intestinale en collectant par transcytose des antigènes présents dans la lumière intestinale et en les présentant aux cellules sous-jacentes du système immunitaire (Abreu, 2010).

## 3. Renouvellement cellulaire de l'épithélium intestinal

Toutes les cellules intestinales proviennent de cellules souches pluripotentes localisées au fond des cryptes de Lieberkühn. Ces cellules, ancrées à proximité de la base des cryptes, subissent une division asymétrique qui, d'une part, maintient le stock des cellules souches et, d'autre part, génère une population cellulaire présentant transitoirement une forte activité de prolifération dans les cryptes. Les cellules progénitrices générées se divisent ensuite continuellement pour donner tous les types cellulaires. À l'exception des cellules de Paneth qui restent confinées à la base des cryptes de l'intestin grêle, les cellules épithéliales s'arrêtent de proliférer et se différencient en atteignant le haut des cryptes, puis elles migrent le long des

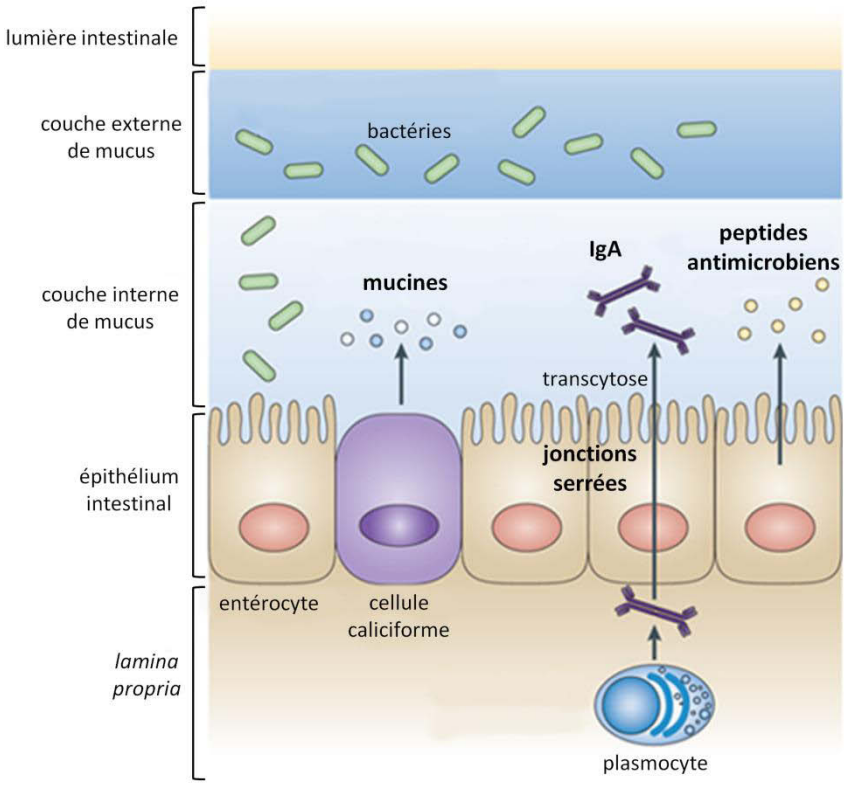



**Figure 4** : La barrière épithéliale intestinale (image modifiée d'après Hooper *et al*, 2010)

villosités de l'intestin grêle ou à la surface des coiffes épithéliales du côlon. Au sommet des villosités, elles sont exfoliées dans la lumière intestinale. Le renouvellement de l'épithélium est ainsi assuré en 4 à 5 jours.

Ces phénomènes de prolifération cellulaire, de différenciation, de migration cellulaire et d'apoptose sont finement contrôlés par l'action coordonnée de différentes voies de régulation cellaire : les voies Wnt et Notch, la voie Hedgehog et la voie de signalisation de PDGF (platelet-derived growth factor) (Crosnier *et al.*, 2006; Noah *et al.*, 2011).

## 4. Fonctions de la barrière épithéliale intestinale

La barrière épithéliale intestinale permet l'absorption de nutriments, d'électrolytes et d'eau grâce à une perméabilité sélective transcellulaire ou paracellulaire, mais elle offre également une protection de l'individu contre d'éventuelles toxines situées au niveau luminal, des antigènes ou des micro-organismes pathogènes (Groschwitz and Hogan, 2009).

Cette barrière est à la fois physique, chimique et immunologique. La barrière physique correspond à la cohésion des cellules entre elles grâce à des jonctions serrées et à leur interaction avec le tissu conjonctif sous-jacent. La composante chimique correspond à 2 éléments principaux : (i) la sécrétion d'une couche de mucus protectrice par les cellules caliciformes et (ii) la sécrétion de peptides antimicrobiens comme les défensines, qui sont synthétisés en grande partie par les cellules de Paneth et qui détruisent ou inhibent la croissance des micro-organismes. Enfin, la barrière immunologique provient de la participation des cellules épithéliales à l'immunité innée et à leurs interactions avec le système lymphoïde associé au tube digestif (**Figure 4**).

Cette barrière épithéliale intestinale est régulée par de multiples facteurs incluant des facteurs exogènes comme l'alcool, la prise d'anti-inflammatoires non stéroïdiens ou encore la présence de micro-organismes pathogènes, mais également l'apoptose des cellules épithéliales, la synthèse de cytokines et le système immunitaire associé au tube digestif (Groschwitz and Hogan, 2009).

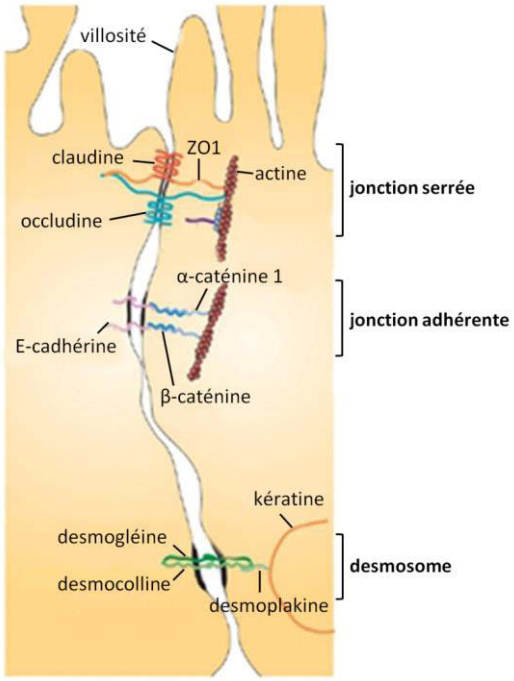



**Figure 5** : Les différents complexes jonctionnels cellulaires au niveau intestinal (image modifiée d'après Turner *et al*, 2009)

### 1) Barrière physique

L'épithélium intestinal est composé d'une monocouche de cellules étroitement reliées entre elles pour former une barrière protectrice vis à vis du contenu de la lumière intestinale grâce à 3 types de structures : (i) les jonctions serrées, (ii) les jonctions adhérentes, et (iii) les desmosomes (Tsukita *et al.*, 2001; Groschwitz and Hogan, 2009; Turner, 2009) (**Figure 5**).

Les jonctions serrées se situent au pôle apical des cellules épithéliales et sont composées de complexes multiprotéiques incluant des protéines transmembranaires de la famille des claudines comme l'occludine, des protéines membranaires d'échafaudage comme les protéines *zonula occludens* 1 et 2 (ZO1, ZO2) et de protéines du cytosquelette comme l'actine. Ces jonctions jouent un rôle majeur dans la perméabilité membranaire par le contrôle des flux paracellulaires.

Sous les jonctions serrées, il existe des jonctions adhérentes qui comme leur nom l'indique permettent une forte adhésion entre les cellules ainsi qu'une communication intercellulaire. Ces jonctions sont constituées par des protéines transmembranaires de la famille des cadhérines, qui sont en contact direct avec les protéines cytoplasmiques caténine δ1 et β-caténine, elle-même liée à l'α-caténine 1 qui permet une liaison de ce complexe aux protéines cytosquelettiques. Outre leur fonction dans la cohésion cellulaire, ces complexes cadhérine-caténine sont également impliqués dans le maintien de la polarisation et la différenciation cellulaire, et sont nécessaires pour l'assemblage des jonctions serrées et des desmosomes.

Enfin, les desmosomes sont des complexes moléculaires formés par des protéines transmembranaires de la famille des cadhérines (desmogléine et desmocolline), la desmoplakine et les filaments de kératine du cytosquelette.

### 2) Barrière chimique

#### a. Le mucus intestinal

L'ensemble de la muqueuse du tractus gastro-intestinal est recouvert d'une couche de mucus. Celui-ci constitue ainsi une barrière protectrice pour le tissu sous-jacent contre

d'éventuels micro-organismes, toxines ou toute autre substance potentiellement nuisible présente dans la lumière intestinale.

### *Constitution du mucus*

Le gel de mucus est un gel viscoélastique composé de plus de 90 % d'eau et de 5 % de polymères de glycoprotéines appelées mucines. Les autres constituants du mucus sont des électrolytes ($Na^+$, $Ca^{2+}$, $K^+$...), des composés organiques tels que protéines et débris cellulaires, des immunoglobulines et des acides nucléiques.

Les mucines sont donc le principal composant structural du mucus, responsable de ses propriétés visco-élastiques et gélifiantes. Les différentes mucines rencontrées chez les mammifères ont des caractéristiques communes : ce sont des glycoprotéines de haut poids moléculaire (entre 250 et 30 000 kDa) formées d'un squelette peptidique (l'apomucine) et de chaînes glucidiques (environ 60 à 80% du poids de la mucine) (Derrien *et al.*, 2010).

L'apomucine peut comprendre de 400 jusqu'à 7000 acides aminés (la protéine MUC2 par exemple en possède 5200). Elle contient dans une région centrale, qui représente à elle seule 50% du corps protéique, des séquences peptidiques répétées en tandem (appelées TR pour Tandem Repeat), riches en sérine, thréonine et proline, dont la séquence initiale est, chez l'homme, composée de 8 à 169 acides aminés selon les mucines. Ces domaines, variables en longueur et en séquences, sont propres à chaque mucine (8 acides aminés pour la mucine MUC5AC et 20 pour MUC1) et à chaque espèce. La région centrale est appelée Variable Number Tandem Repeat (VNTR) du fait du nombre variable de répétitions en tandem selon les individus. Elle est également appelée région PTS à cause de sa richesse en proline, thréonine et sérine.

De nombreux sucres sont liés à ces domaines PTS par des liaisons majoritairement *O*-glycosidiques *via* une liaison à un résidu *N*-acétylgalactosamine, ce qui donne à la protéine une apparence de goupillon. Les chaînes sucrées sont composées de monosaccharides dont le L-fucose, D-galactose, N-acetyl-D-galactosamine (GalNAc), N-acetyl-D-glucosamine (GlcNAc) et l'acide N-acetylneuraminique ou acide sialique (Li *et al.*, 1998). La longueur de ces chaînes varie de 1 à 20 oses, sauf pour les mucines coliques, dont les chaînes vont de 2 à 12 sucres. L'intense *O*-glycosylation des mucines leur confère une forte densité et une grande viscosité et leur permet également d'être protégées des enzymes protéolytiques.

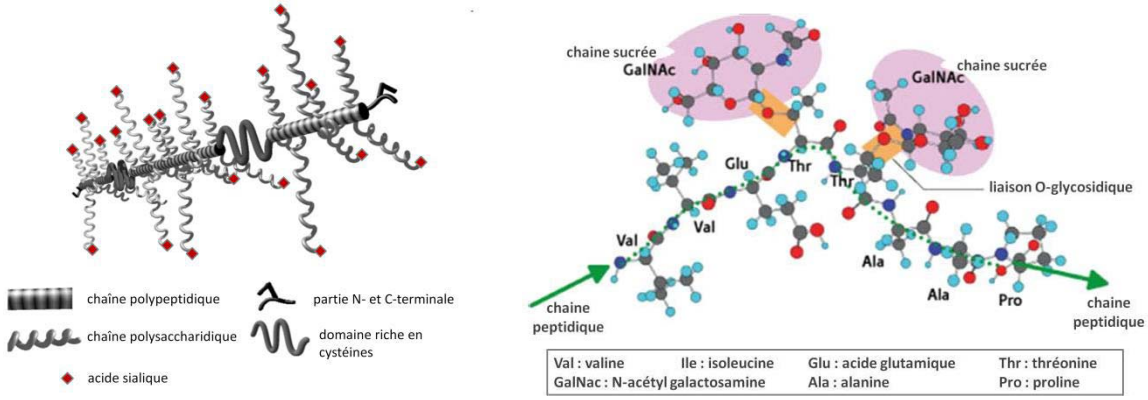



**Figure 6** : Structure des mucines

(image modifiée d'après le site http://lewislabofmicrobialglycobiology.yolasite.com/)

Aux extrémités N- et C-terminales de l'apomucine se situent des séquences uniques, riches en cystéines dans le cas des mucines sécrétées. A ce niveau, les chaînes oligosaccharidiques sont rares et sont fixées à l'axe peptidique par des liaisons N-glycosidiques (**Figure 6**).

Ces chaînes N-glycosidiques sont composées de trois sucres : la GlcNAc, le galactose et le mannose (qu'on ne retrouve pas dans les chaînes O-glycosylées) (Aksoy and Akinci, 2004). La rareté des chaînes glucidiques dans ces régions (par contraste avec les régions hautement glycosylées) rend les mucines sensibles aux protéases bactériennes et endogènes (Dharmani *et al.*, 2009).

Ces extrémités contiennent les caractéristiques permettant de distinguer deux classes de mucines :

- Les mucines membranaires qui possèdent au niveau C-terminal une zone transmembranaire, hydrophobe, et responsable de leur ancrage dans la double couche lipidique de la membrane plasmique. Les mucines membranaires incluent au niveau du tube digestif MUC1, MUC3A/B, MUC4, MUC12, MUC13, MUC15, MUC17, MUC19, MUC20 et MUC21. Elles possèdent donc toutes : (i) un domaine extracellulaire pouvant aller de 1 à 10 MDa environ (excepté pour MUC13 qui est plus petite), avec des séquences répétées en tandem encadrées par des séquences uniques, (ii) un domaine transmembranaire qui les enchâssent dans la membrane cellulaire et (iii) un petit domaine cytoplasmique associé aux protéines du cytosquelette. De par leur localisation à la surface apicale des cellules épithéliales, elles participent au glycocalyx et sont impliquées dans la protection de l'épithélium mais également dans des processus de transduction de signaux cellulaires. Les gènes codant ces mucines sont localisés sur les chromosomes 1q, 3q, 6p, 7q, 11p et 19q (Moniaux *et al.*, 2001; Andrianifahanana *et al.*, 2006; Johansson, Ambort, *et al.*, 2011).

- Les mucines sécrétées qui possèdent des domaines semblables à ceux du facteur de coagulation Von Willebrand (VWF) et appelés domaines D. Ces domaines, riches en cystéines, jouent un rôle dans l'oligomérisation des monomères de mucines et dans l'empaquetage dans des vésicules de sécrétion (Baldus and Hanisch, 2000). Les mucines sécrétées au niveau du système digestif pour former le gel de mucus sont MUC2, MUC5AC, MUC5B, MUC6 et MUC19. La mucine MUC5AC est essentiellement sécrétée au niveau de l'estomac, alors qu'au niveau intestinal, la mucine sécrétée prédominante est MUC2 (McGuckin *et al.*, 2011).

**Tableau 1** : Caractéristiques des gènes codant les différentes mucines (Cys : domaine riche en Cys ; D : domaine D ; PTS : domaine riche en Pro/Thr/Ser)

| gène *MUC* | chromosome | nature de la protéine | domaines | expression tissulaire |
|---|---|---|---|---|
| *MUC1* | 1q21 | membranaire | PTS | poumon, cornée, glande salivaire, œsophage, estomac, pancréas, **colon**, sein, prostate, ovaire, rein, utérus |
| *MUC2* | 11p15.5 | sécrétée, constituant du gel de mucus | Cys, D, PTS | poumon, conjonctive, oreille, estomac, **intestin**, nasopharynx, prostate |
| *MUC3A* | 7q22 | membranaire | PTS | thymus, **intestin**, rein |
| *MUC3B* | 7q22 | membranaire | PTS | **intestin** |
| *MUC4* | 3q29 | membranaire | Cys, PTS | poumon, cornée, glande salivaire, œsophage, **intestin** grêle, rein |
| *MUC5AC* | 11p15.5 | sécrétée, constituant du gel de mucus | Cys, D, PTS | poumon, conjonctive, oreille, **estomac**, vésicule biliaire, nasopharynx |
| *MUC5B* | 11p15.5 | sécrétée, constituant du gel de mucus | Cys, D, PTS | poumon, oreille, glande sublinguale, larynx, glande sous-muqueuse, oesophage, **estomac**, **duodénum**, vésicule biliaire, nasopharynx |
| *MUC6* | 11p15.5 | sécrétée, constituant du gel de mucus | PTS | **estomac, duodénum**, vésicule biliaire, pancréas, rein |
| *MUC7* | 4q13.3 | sécrétée | PTS | poumon, glande lacrymale, glande salivaire, nez |
| *MUC8* | 12q24.3 | sécrétée | - | oviducte |
| *MUC9* | 1p13 | sécrétée | PTS | glande sous-mandibulaire |
| *MUC11/12* | 7q22 | membranaire | PTS | poumon, oreille, thymus, **intestin**, pancréas, foie, rein, utérus, prostate |
| *MUC13* | 3q21.2 | membranaire | - | poumon, conjonctive, **estomac, intestin**, rein |
| *MUC15* | 11p14.3 | membranaire | PTS | conjonctive, amygdale, thymus, nodule lymphoïde, sein, **intestin**, foie, rate, prostate, ovaire, leucocyte, moëlle osseuse |
| *MUC16* | 19p13.2 | membranaire | - | conjonctive, ovaire |
| *MUC17* | 7q22.1 | membranaire | PTS | **intestin**, conjonctive |
| *MUC19* | 12q12 | sécrétée, constituant du gel de mucus | D | poumon, glande salivaire, rein, foie, **colon**, placenta, prostate |
| *MUC20* | 3q29 | membranaire | - | poumon, foie, rein, **colon**, placenta, prostate |
| *MUC21* | 6p21.32 | membranaire | - | poumon, **colon**, thymus, testicule |

Ces mucines forment d'immenses polymères, avec par exemple plus de 100 MDa pour MUC2, possèdant de nombreux domaines riches en cystéines participant à leur polymérisation, et très similaires à certaines régions du facteur Willebrand de la coagulation (Gum *et al.*, 1994). Parmi les 5 gènes existants codant les mucines sécrétées, tous sont localisés sur le chromosome 11p15 dans une zone de 400 kb organisée en cluster (Vinall *et al.*, 1998), à l'exception du gène codant MUC19 situé sur le chromosome 12q12 (Chen *et al.*, 2004; Culp *et al.*, 2004).

Il existe également des mucines sécrétées qui ne rentrent pas dans la composition du gel de mucus, et dont la fonction reste moins connue à ce jour. Ce sont les mucines MUC7, MUC8 et MUC9. Les gènes codant ces mucines sont localisés respectivement sur les chromosomes 4q, 12q et 1p (Moniaux *et al.*, 2001; Andrianifahanana *et al.*, 2006) (**Tableau 1**).

### *Synthèse du mucus*

L'épaisseur du mucus résulte d'un équilibre entre sa synthèse et sa dégradation d'origine mécanique ou protéolytique. A sa surface, il subit l'action d'enzymes dont la pepsine, qui rend ce mucus plus hydrosoluble avec les unités de polymères dépolymérisées. Cette érosion de la couche de mucus doit absolument être compensée par des sécrétions en provenance des cellules caliciformes afin de garantir l'intégrité de la barrière épithéliale intestinale, première ligne de défense contre les agressions extérieures.

La sécrétion des mucines est un phénomène qui peut être régulé au niveau transcriptionnel par de nombreux facteurs. Elle peut être divisée en une sécrétion constitutive basale non régulée, et une sécrétion induite par de nombreux stimuli.

Le mucus est synthétisé de façon continue par les cellules caliciformes à la surface de l'ensemble de l'épithélium digestif, afin de maintenir l'épaisseur de la couche de mucus, sans cesse exposée à des acides et des irritants et régulièrement éliminée par les mouvements péristaltiques de l'intestin (Neutra and Leblond, 1966). Cependant, la libération par les cellules caliciformes de granules contenant du mucus peut être accélérée sous l'effet d'irritations locales, d'agents cholinergiques, d'hormones, de facteurs de croissance, de prostaglandines, de médiateurs de l'inflammation (cytokines et lipides) ou de constituants bactériens (Andrianifahanana *et al.*, 2006; Gaudier and Hoebler, 2006). La libération de

mucus peut ainsi intervenir entre 1 et 3 minutes seulement après la stimulation des cellules caliciformes par des récepteurs de prostaglandines (Belley and Chadee, 1999)

La synthèse des mucines est différente, selon qu'elles soient membranaires ou sécrétées. La synthèse de MUC1, principale mucine membranaire connue, démarre par l'assemblage d'une chaine polypeptidique qui est ensuite clivée en 2 sous-unités avant d'être partiellement glycosylée. Après adressage à la membrane cellulaire, cette protéine est ré-internalisée vers l'appareil de Golgi pour compléter sa *O*-glycosylation avec des résidus d'acide sialique avant de repartir vers la membrane cellulaire sous forme mature (Andrianifahanana *et al.*, 2006).

Concernant les mucines sécrétées, après la transcription du gène *MUC*, les peptides incorporés au niveau du réticulum endoplasmique pendant la phase de traduction sont rapidement *N*-glycosylés sur la séquence consensus suivante : Asn-X-Ser/Thr (X correspondant à n'importe quel acide aminé hormis la proline) (Desseyn *et al.*, 1997). Le processus de polymérisation qui démarre au niveau du réticulum endoplasmique par la formation de dimères grâce à des ponts disulfure est facilité par ces *N*-glycosylations, qui sont également importantes pour le repliement de la protéine et son acheminement vers l'appareil de Golgi. Ces mucines sont alors fortement *O*-glycosylées et polymérisées. D'autres constituants peuvent également établir des liaisons avec la partie protéique centrale, comme les phospholipides ou les acides gras, créant ainsi un centre hydrophobe au polymère, le rendant plus stable et résistant face aux agresseurs hydrophiles (Slomiany *et al.*, 1987). Enfin, ces mucines seront stockées sous forme mature dans des granules de sécrétion afin d'être externalisées au moment voulu (Andrianifahanana *et al.*, 2006; Johansson, Ambort, *et al.*, 2011; Hansson, 2012).

### *Répartition du mucus dans le tractus digestif*

De l'estomac jusqu'au rectum, la muqueuse est constituée d'une monocouche de cellules épithéliales recouvertes d'une couche de mucus, dont l'épaisseur varie tout au long du tractus gastro-intestinal, allant de 150-200 µm dans l'intestin grêle jusqu'à 700 µm dans l'estomac et le colon (Atuma *et al.*, 2001). Au niveau de l'estomac, la principale mucine sécrétée est

MUC5AC, alors qu'au niveau de l'intestin grêle et du colon, la mucine largement prédominante est MUC2 (**Tableau 1**).

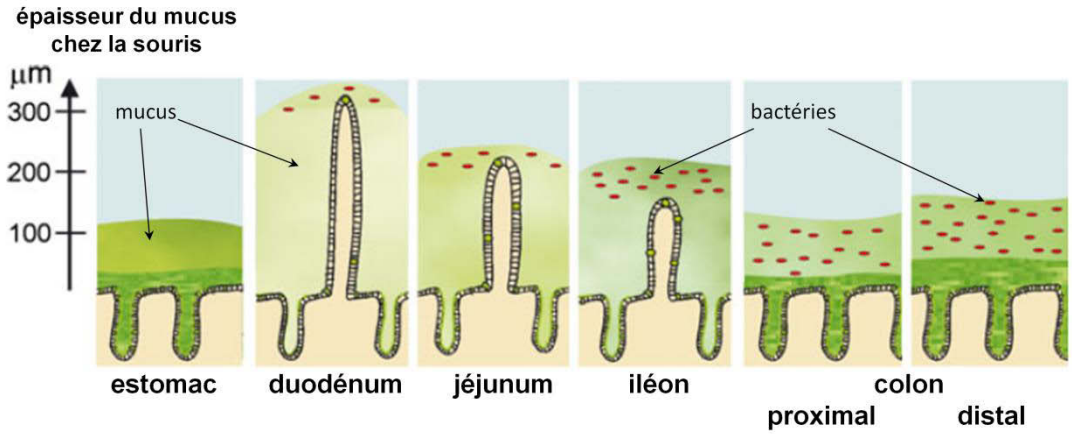



**Figure 7** : Répartition de la couche de mucus le long du tractus gastro-intestinal (image modifiée d'après Pelaseyed *et al.*, 2014)

La couche de mucus est organisée différemment selon sa localisation dans le tractus-gastro-intestinal (Johansson *et al.*, 2008; McGuckin *et al.*, 2011; Johansson and Hansson, 2011; Pelaseyed *et al.*, 2014) (**Figure 7**). Dans l'estomac et le colon, le mucus est divisé en deux couches. La couche interne est continue, fine et très dense. Elle est continuellement renouvelée à partir des cellules caliciformes situées à la surface de l'épithélium. Ce renouvellement a été estimé à environ 1 à 2 heures au niveau du colon de souris (Johansson, 2012). Ce renouvellement est important pour garder la couche interne de mucus exempte de micro-organismes et avec une densité suffisante pour se comporter comme un filtre empêchant la pénétration des bactéries ou des particules de plus de 0.5 µm (Johansson *et al.*, 2014). Dans des conditions physiologiques, cette couche interne reste donc totalement impénétrable aux bactéries. La couche externe de mucus est plus lâche, plus volumineuse et colonisée par le microbiote intestinal. Il n'existe pas de différence de composition entre les deux couches de mucus qui pourrait expliquer les propriétés physico-chimiques divergentes. La seule différence entre ces deux couches relève du niveau de condensation et d'hydratation du mucus.

Au niveau de l'intestin grêle, il n'existe qu'une seule couche de mucus qui remplit la lumière entre les villosités et les recouvrent. Cette couche d'environ 100 à 150 µm est assez lâche et non ancrée à la surface de l'épithélium. Elle est ainsi éliminée progressivement avec les mouvements péristaltiques de l'intestin. La synthèse continue du mucus par les cellules caliciformes assure le maintien de son intégrité (Johansson, Ambort, *et al.*, 2011). Les cellules de Paneth localisées dans les cryptes, mais aussi les entérocytes, produisent des peptides antimicrobiens qui génèrent un gradient antibactérien qui garde les bactéries à distance de la surface de l'épithélium. Le mucus de l'intestin grêle n'est donc pas imperméable aux bactéries mais il constitue tout de même une barrière que les bactéries qui veulent atteindre l'épithélium doivent traverser (Pelaseyed *et al.*, 2014). Cette notion pourrait expliquer que la mobilité flagellaire soit un facteur de virulence majeur chez les bactéries entéropathogènes. La couche de mucus est moins épaisse au niveau des cellules M surplombant les plaques de Peyer et est produite par les cellules caliciformes se trouvant dans les villosités voisines (Ermund *et al.*, 2013).

### *Fonctions du mucus*

Les fonctions du mucus dans le tractus gastro-intestinal sont multiples. Tout d'abord, il permet de protéger l'épithélium digestif contre l'effet abrasif du bol alimentaire, les chocs thermiques ou les forces mécaniques du péristaltisme intestinal, l'acidité gastrique de l'estomac et les enzymes de la digestion. Une fois dépolymérisé, il possède un rôle lubrifiant qui permet de minimiser les différentes forces exercées par le bol alimentaire, ce qui facilite sa progression (Gaudier and Hoebler, 2006).

Le mucus forme également une barrière chimique contre certaines toxines bactériennes, des radicaux libres et d'autres molécules potentiellement nocives (Cross *et al.*, 1984). En effet, les différents constituants du gel de mucus font de lui un véritable réseau de mailles plus ou moins imperméable. De plus, l'intense glycosylation des mucines rend l'action d'enzymes protéolytiques bien plus difficile. Ainsi, il ne permet qu'une diffusion lente et sélective des nutriments destinés à être absorbés par les entérocytes.

Le gel de mucus participe également à l'interaction entre l'hôte et le microbiote intestinal tant pour la mise en place de la flore endogène que dans la lutte contre les micro-organismes pathogènes. En effet, la grande hétérogénéité des glycanes portés par MUC2 servent tout d'abord de source de nutriments pour les bactéries commensales telles que *Akkermansia muciniphila*, *Bacteroides thetaiotaomicron, Bifidobacterium spp* ou *Ruminococcus spp* qui possèdent des glycosidases (Corfield *et al.*, 1992; Sonnenburg *et al.*, 2005; Katayama *et al.*, 2005; Ouwerkerk *et al.*, 2013). Mais le mucus offre également une matrice idéale pour l'attachement et le développement du microbiote commensal, ce qui permet de contribuer à la sélection d'espèces bactériennes spécifiques du colon (Johansson, Larsson, *et al.*, 2011). Par ailleurs, afin de lutter contre d'éventuels micro-organismes pathogènes, le mucus contient du lysozyme et diverses molécules antimicrobiennes (Gaudier and Hoebler, 2006; McGuckin *et al.*, 2011) soit en provenance de la circulation sanguine (immunoglobulines) soit directement sécrétées par les cellules épithéliales ou les cellules de Paneth (défensines). Ces molécules peuvent se fixer aux mucines *via* des résidus sucrés pour rester au niveau de la couche de mucus où elles vont avoir une efficacité maximale (Iontcheva *et al.*, 2000; Royle *et al.*, 2003). Le mucus permet également le masquage de certaines adhésines présentes à la surface de bactéries pathogènes, ce qui inhibe leur adhésion et les élimine du tractus gastro-intestinal (Paerregaard *et al.*, 1991). En outre, le mucus possède certains motifs glycaniques portés par

les mucines, comme par exemple l'α-1-4-*N*-acétyl-glucosamine, capables d'inhiber la croissance de pathogènes (Kawakubo *et al.*, 2004).

Les cellules caliciformes sécrètent également des peptides dits en feuille de trèfle (*trefoil factor*, TFF) codés par des gènes situés sur le chromosome 21 (Rio, 2002). Les TFF interagissent avec les mucines et participent à la formation du mucus, pour renforcer sa viscosité et ses propriétés de barrière (Thim *et al.*, 2002). Ces petites protéines agissent en conditions physiologiques comme des gardiens de l'intégrité de la muqueuse digestive. Lors des agressions bactériennes, virales ou médicamenteuses, ainsi que dans certaines pathologies inflammatoires ou ulcéreuses, elles jouent un rôle clé dans la restitution et la régénération des muqueuses. Ces peptides TFF voient donc leur niveau d'expression varier de manière considérable en réponse à des agressions ou à des modifications de l'environnement (Ogata and Podolsky, 1997; Vandenbroucke *et al.*, 2004).

Enfin, le mucus possède également un rôle dans la réparation, car il participe à la formation d'un épais manteau gélatineux protégeant l'épithélium lésé du contenu agressif endo-luminal et permettant ainsi la reconstitution de l'intégrité épithéliale à la suite d'une agression mécanique ou chimique (Allen *et al.*, 1988).

De par ses propriétés physico-chimiques, le mucus joue donc un rôle de barrière à la fois physique, chimique et antimicrobienne. Cependant, les mucines qui le constituent peuvent également avoir d'autres fonctions plus spécifiques que la seule formation du mucus, comme l'échange d'informations pour les cellules épithéliales ou le renouvellement des épithéliums. MUC2 est par exemple impliquée dans la différenciation et la croissance cellulaire. La suppression du gène *Muc2* chez des souris entraine en effet une morphologie aberrante des cryptes intestinales ainsi qu'une différenciation et une migration cellulaire altérées, conduisant au cancer (Velcich *et al.*, 2002). MUC5AC et MUC6, quant à elles, semblent impliquées dans la réparation épithéliale car leur synthèse est augmentée dans les cellules situées à la marge des ulcères chez les patients atteints de maladie de Crohn (MC) (Buisine *et al.*, 2001). Enfin, MUC1, qui est la mucine la plus étudiée, possède de nombreuses propriétés propres. Elle peut induire tout d'abord une immunomodulation par induction d'apoptose ou inhibition de la prolifération des lymphocytes T (Agrawal *et al.*, 1998). Elle est également capable de s'associer au récepteur du facteur EGF (epidermal growth factor), ce qui induit sa

phosphorylation conduisant à sa liaison intracellulaire avec des protéines comme la β-caténine, ou ZO1. Elle permet ainsi la transduction de signaux et joue un rôle dans l'interaction entre les cellules (Gaudier and Hoebler, 2006). Enfin, sa surexpression au niveau de cellules cancéreuses va déstabiliser les interactions cellule-cellule, déstabiliser la cohésion du tissu et induire des métastases par transition epithélio-mésenchymateuse des cellules (Hollingsworth and Swanson, 2004; Kufe, 2009).

### b. Les peptides antimicrobiens

Les peptides antimicrobiens constituent un élément important des mécanismes de défense intestinaux, et sont largement retrouvés chez les animaux, les insectes et les plantes (Bulet *et al.*, 2004; Diamond *et al.*, 2009). Le premier peptide antimicrobien a été découvert en même temps que les antibiotiques au début du XIXème siècle. Il s'agissait du lysozyme, qui est un peptide possédant une activité enzymatique permettant la destruction du peptidoglycane des bactéries à Gram positif. Depuis lors, plusieurs molécules antimicrobiennes ont été caractérisées chez l'Homme, dont les principales sont les défensines, les cathélicidines et les lectines de type C. Ces molécules une fois sécrétées par les cellules se concentrent au niveau de la couche de mucus, formant ainsi une barrière protectrice bactéricide (Huttner and Bevins, 1999; Bulet *et al.*, 2004).

La plupart de ces peptides sont petits, chargés positivement et possèdent une structure amphipathique, qui permet l'interaction avec les membranes microbiennes, indispensable à leur efficacité. Leur activité est le plus souvent basée sur une interaction électrostatique entre leur partie cationique avec les phospholipides chargés négativement des membranes bactériennes. L'accumulation de ces peptides conduisent à des modifications de la structure de la membrane bactérienne avec la création de pores conduisant à la mort (Wimley *et al.*, 1994). Les cellules eucaryotes sont constituées de phospholipides chargés négativement que dans la partie interne de leur membrane, c'est pourquoi ces peptides n'agissent que sur les micro-organismes de type bactéries, champignons, parasites et/ou virus. En plus d'attaquer les membranes, ces peptides antimicrobiens sont également capables d'inhiber des activités enzymatiques et la synthèse de composants de la paroi cellulaire, des acides nucléiques ou des protéines.

A côté de leur activité antimicrobienne, certains peptides sont également capables de recruter des cellules immunitaires, de moduler l'expression de cytokines et/ou chémokines et de provoquer un relarguage d'histamine. A titre d'exemple, les défensines, qui sont de petits peptides cationiques de 3 à 5 kDa, ne possèdent pas seulement un rôle antibiotique mais jouent un rôle central dans la connexion entre l'immunité innée et adaptative. Certaines d'entre elles sont en effet capables de recruter les monocytes, les cellules dendritiques et les lymphocytes T (Diamond *et al.*, 2009).

Au niveau de l'intestin, les peptides antimicrobiens peuvent être sécrétés par tous les types cellulaires de l'épithélium, incluant les entérocytes, les cellules caliciformes et les cellules de Paneth ; ces dernières sont des cellules sécrétrices spécialisées dans la production des susbstances antimicrobiennes et localisées au niveau de l'intestin grêle. Les peptides antimicrobiens sont indispensables au maintien de l'intégrité de la barrière intestinale. Ils semblent également impliqués dans la régulation du microbiote intestinal et de la relative stérilité observée à l'intérieur des cryptes. La régulation de la sécrétion de ces peptides est complexe et mal connue. Certains comme l'α-défensine semblent être sécrétés de manière constitutive, alors que d'autres comme RegIIIγ, une lectine de type C, sont sécrétés de manière induite suite à l'activation de récepteurs de reconnaissance de motifs (PRR, pattern recognition receptor). La maladie de Crohn, et en particulier la forme iléale, a été associée à une diminution de la concentration des α-défensines (hBD-2 et hBD-3) par rapport à celles observées chez les patients sains. De même, l'expression déficiente des α-défensines HD-5 et HD-6 dans les cellules de Paneth a été reliée à cette maladie (Wehkamp *et al.*, 2004; Wehkamp *et al.*, 2005; Guaní-Guerra *et al.*, 2010).

### 3) Barrière immunologique

Le système immunitaire intestinal peut se diviser en une composante innée à laquelle participent de nombreuses cellules de la barrière épithéliale intestinale, et une composante adaptative constituée des cellules présentatrices d'antigènes et des lymphocytes.

### a. Défenses non spécifiques : immunité innée

Le système immunitaire inné, non spécifique, est primordial pour l'Homme dans le laps de temps entre le premier contact avec un micro-organisme et les mécanismes de défense spécifiques qui ont besoin de temps pour être efficaces. Ces défenses incluent notamment le mouvement péristaltique, la sécrétion d'acides biliaires, la production de peptides antimicrobiens et de mucus, et le recrutement et/ou l'activation de cellules natural killer (NK) ou de macrophages, qui assurent la destruction des bactéries commensales réussissant à franchir l'épithélium lorsqu'une perte de son intégrité survient (Kelsall, 2008; Hooper and Macpherson, 2010).

La muqueuse intestinale étant au contact d'un grand nombre de molécules microbiennes et de micro-organismes commensaux et/ou pathogènes, les cellules intestinales doivent rapidement détecter la présence de micro-organismes pathogènes afin de mettre en place une réponse immunitaire adaptée, tout en maintenant une "tolérance" vis à vis des micro-organismes commensaux faisant partie du microbiote intestinal. Pour cela, de nombreux types cellulaires au niveau intestinal peuvent également intervenir dans l'immunité innée, comme les entérocytes ou les monocytes, grâce à la présence à leur surface de PRR. Ces PRR ont la faculté de reconnaitre des motifs associés aux pathogènes (PAMP, pathogen associated molecular patterns) qui sont des motifs moléculaires propres aux micro-organismes et conservés au sein d'une même classe. Le lipopolysaccharide (LPS) des bactéries à Gram négatif et la flagelline en font partie, de même que les acides nucléiques bactériens ou viraux comme par exemple l'ARN double brin. Les PRR permettent de détecter les micro-organismes et constituent ainsi des récepteurs de l'immunité innée. Une partie d'entre eux sont exprimés de manière constitutive mais dans certaines circonstances ils peuvent également être induits.

Il existe différents types de PRR. Les plus connus sont les *Toll-like* recepteurs (TLR). Ce sont des récepteurs transmembranaires présents à la surface des cellules ou des endosomes. Ils sont au nombre de dix chez l'Homme et sont capables de reconnaitre une grande diversité de PAMPs bactériens, fongiques, parasitaires et viraux. Cependant, lorsque l'épithélium intestinal est intact, l'expression limitée des TLRs à la surface des cellules permet de contrôler la réponse immunitaire face à de fortes concentrations de PAMPs dans la lumière intestinale (Abreu, 2010; McGuckin *et al.*, 2011).

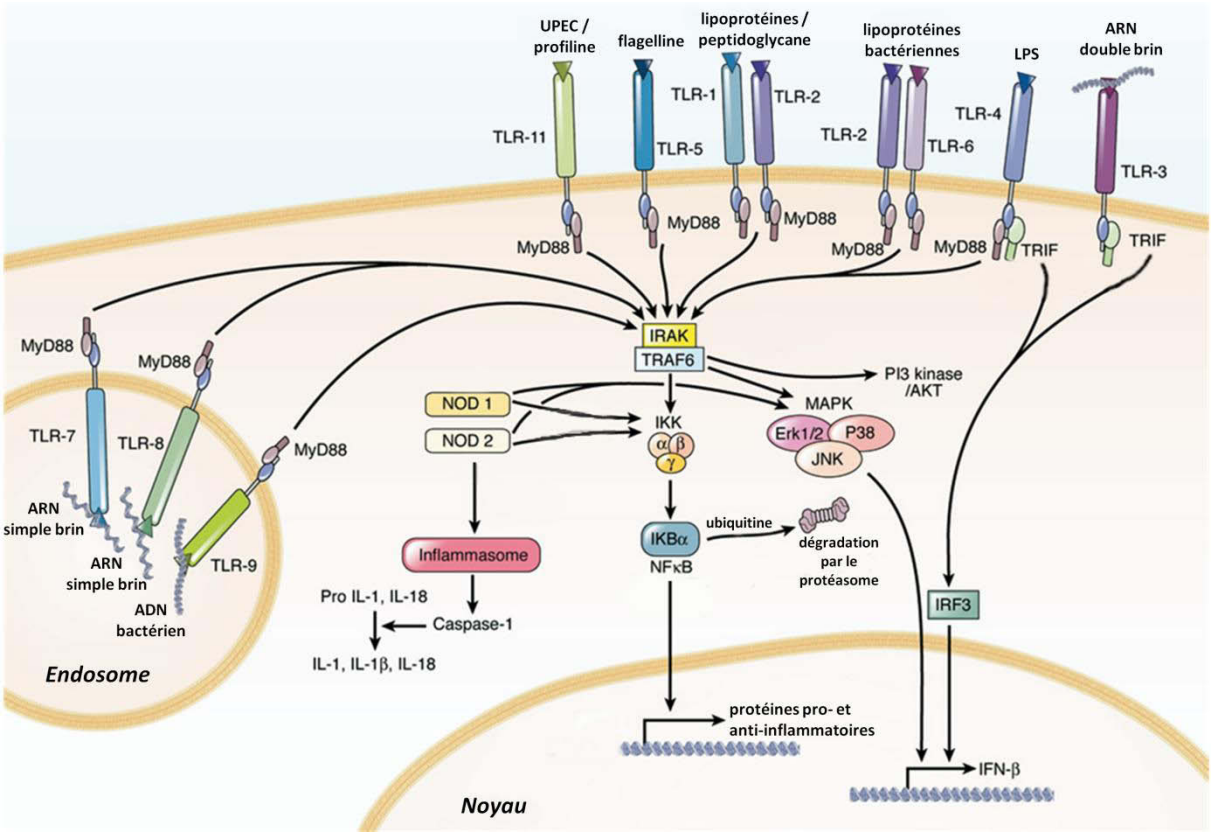



**Figure 8** : Voies de signalisation activées par les PRR suite à une stimulation bactérienne (image modifiée d'après Sartor, 2008)

Les *NOD-like* récepteurs (NLR, NOD = nucleotide-binding oligomerization domain-containing molecules) sont quant à eux des récepteurs intracellulaires. Il en existe plus de 20 différents chez l'Homme. Leur fonction principale est la détection de PAMPs intracellulaires et de signaux de danger. Parmi eux, le récepteur NOD2 (également appelé CARD15 pour Caspase-Activating Recruitment Domain 15) qui reconnait un muramyl dipeptide bactérien, est principalement exprimé dans les macrophages et les cellules de Paneth où il est impliqué dans la sécrétion de défensines antimicrobiennes (Claes *et al.*, 2015). De nombreux polymorphismes de ce gène ont été mis en évidence chez les patients atteints de la maladie de Crohn.

Il existe également les *RIG-I-like* récepteurs (RLR) qui sont des récepteurs cytoplasmiques aux ARN viraux. Enfin, les *C-lectin-like* récepteurs (CLR) sont une grande famille de récepteurs membranaires détectant des motifs hydrocarbonés contenus principalement dans les parois fongiques.

La fixation des PAMPs sur les PRR induit le recrutement d'adaptateurs cytoplasmiques : le facteur de différenciation myéloïde 88 (MyD88) et l'adaptateur TRIF (TIR-domain-containing adapter-inducing interferon-β). Ceci engendre une cascade de phosphorylations impliquant le facteur associé au récepteur du TNF (TRAF) et une kinase associée au récepteur de l'interleukine-1 (IRAK), qui conduit à l'activation des voies de signalisation des MAP kinases (mitogen activated protein kinase), PI3K/AKT (phosphoinositide 3-kinase et phosphokinase B) et de la voie NF-κB (Nuclear Factor kappa B), provoquant la dégradation de l'inhibiteur IκB par le protéasome. L'activation de ces voies conduit à la translocation dans le noyau des facteurs de transcription tels que NF-κB ou le facteur de transcription 3 régulant l'interféron (IRF3), ce qui entraine une réponse associant la synthèse de cytokines et chémokines pro-inflammatoires comme IL6, IL8 ou TNF-α, la production de mucus et de peptides antimicrobiens, ainsi que le recrutement de polynucléaires neutrophiles et de macrophages (**Figure 8**).

### b. Acteurs de l'immunité adaptative

Le système immunitaire adaptatif peut être plus précisément divisé en sites inducteurs, impliqués dans la capture et la présentation des antigènes, et en sites effecteurs de la réponse immunitaire adaptative.

### *Cellules impliquées dans la capture et la présentation des antigènes*

Les antigènes de la lumière intestinale peuvent être capturés de 3 manières différentes : (i), directement par les cellules épithéliales, (ii) par des cellules présentatrices d'antigènes (CPA) comme les cellules dendritiques, et (iii) par les plaques de Peyer ou les nodules lymphoïdes isolés, qui représentent les sites inducteurs majeurs de l'immunité adaptative intestinale.

Les cellules présentatrices d'antigène (CPA) jouent un rôle important dans l'immunité. Dans l'intestin, elles sont présentes au niveau de la *lamina propria*. Classiquement, elles sont distinguées en cellules dendritiques et macrophages, mais de nombreuses CPA au niveau intestinal possèdent des caractéristiques de ces deux types cellulaires. La fonction principale des cellules dendritiques intestinales est l'échantillonnage d'antigènes présents dans la lumière intestinale *via* des dendrites étendues entre deux cellules épithéliales. Ces cellules vont ensuite migrer au niveau des ganglions mésentériques de proximité afin d'induire une réponse immunitaire adaptative.

Les plaques de Peyer sont de volumineux agrégats de follicules lymphoïdes primaires et secondaires siégeant dans le chorion de la muqueuse de la partie terminale de l'iléon. Apparentés aux plaques de Peyer, les follicules lymphoïdes constituent des structures plus petites mais très nombreuses et réparties dans toute la longueur du tube digestif, avec néanmoins une prédominance dans l'iléon. Ces sites inducteurs de la réponse immunitaire possèdent un épithélium particulier qui comporte des cellules épithéliales dédifférenciées appelées cellules M. Ces cellules présentent de nombreuses microvésicules et une forme particulière qui leur permet un contact étroit avec les cellules dendritiques, macrophages ou lymphocytes sous-jacents. Elles captent de manière sélective les microparticules antigéniques qui arrivent à leur contact. Ces particules traversent leur cytoplasme sous forme de vésicules puis sont libérées dans le micro-environnement sur lequel ces cellules reposent.

### *Cellules effectrices impliquées dans la réponse immunitaire adaptative*

Le système immunitaire adaptatif intestinal comprend comme effecteurs essentiellement des lymphocytes et des plasmocytes répartis dans l'épithélium intestinal (lymphocytes T intraépithéliaux) et dans le tissu conjonctif du chorion de la muqueuse et de la sous-muqueuse (follicules lymphoïdes et cellules lymphoïdes dispersées, où prédominent largement les lymphocytes B et les plasmocytes sécréteurs d'IgA). Le tissu lymphoïde associé au tube digestif (Gut Associated Lymphoid Tissue ou GALT) comporte en plus des cellules lymphoïdes dispersées au niveau de l'épithélium, de la *lamina propria* et des follicules lymphoïdes, les amygdales, l'appendice iléocæcale et les plaques de Peyer. Les cellules du GALT comprennent des lymphocytes B mémoires et plasmocytes sécréteurs d'IgA pour 70 à 90% d'entre eux, ainsi que des lymphocytes T CD4+ (helpers) et CD8+ (cytotoxiques). Les CD8+ prédominent dans l'épithélium alors que les CD4+ sont retrouvés principalement au niveau de la *lamina propria*.

### *La réponse adaptative des lymphocytes B*

Les cellules lymphoïdes naïves B sont informées et sélectionnées via les cellules M notamment, mais également par les CPA et/ou les entérocytes eux-mêmes, qui peuvent capter certains antigènes et les présenter directement aux lymphocytes sous-jacents grâce à une molécule du complexe majeur d'histocompatibilité de type II. Les cellules B prolifèrent et constituent des centres germinatifs au niveau des nodules ou des plaques de Peyer. Les ganglions mésentériques de voisinage peuvent également participer à cette réponse immunitaire.

Les lymphocytes B provenant d'un nodule lymphoïde, d'une plaque de Peyer ou d'un ganglion mésentérique vont emprunter le système lymphatique pour gagner la circulation systémique. Ils vont ensuite coloniser tous les territoires muqueux par voie sanguine et grâce à leurs propriétés d'adhésion spécifique, ils vont finir par rejoindre la *lamina propria*. Ces lymphocytes B activés par un antigène finissent alors leur processus de différenciation en plasmocytes et sécrètent des IgA spécifiques de l'antigène rencontré. Les IgA sécrétoires présentent la particularité de se regrouper en dimères et de s'associer à une pièce sécrétoire synthétisée par les cellules épithéliales lors d'un phénomène de transcytose au niveau de ces cellules. Les IgA sont ensuite libérées dans la lumière intestinale au niveau du pôle apical des

cellules épithéliales. En tapissant la surface des muqueuses, elles vont ainsi capturer les antigènes et les empêcher de pénétrer dans les tissus.

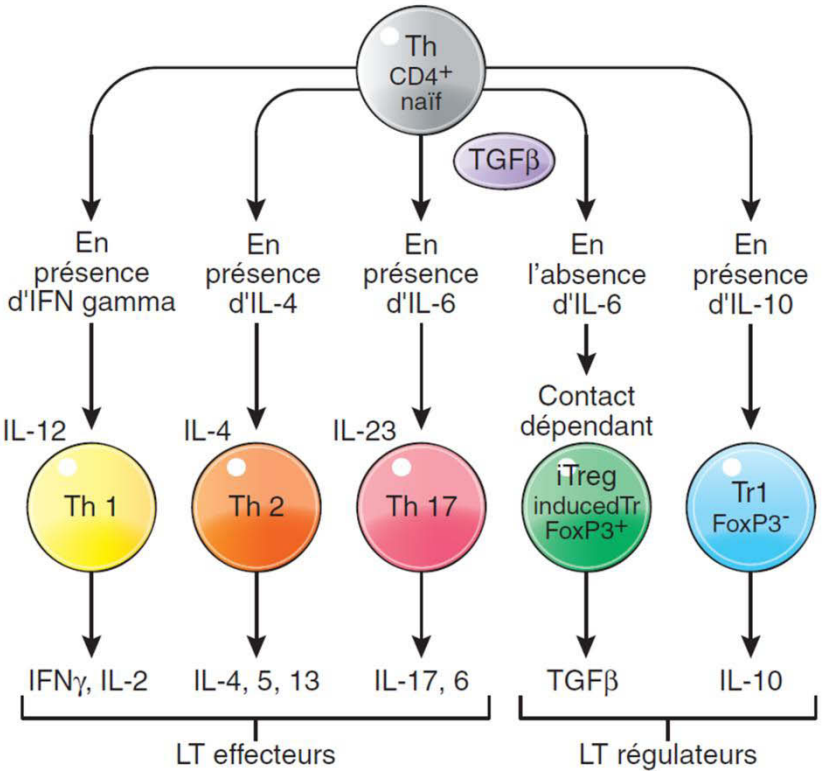



**Figure 9** : La réponse adaptative des lymphocytes T
(image issue du livre : Les fondamentaux de la pathologie digestive,
CDU-HGE/ed Elsevier Masson - octobre 2014)

### *La réponse adaptative des lymphocytes T*

La présentation des antigènes par les CPA aux lymphocytes T résidents de la *lamina propria* active ces derniers. Selon l'environnement inflammatoire, ces lymphocytes prendront un phénotype pro-inflammatoire (effecteurs) ou anti-inflammatoire (régulateurs). Pour leur permettre de choisir entre ces deux phénotypes, les cellules dendritiques vont jouer un rôle majeur car elles peuvent intégrer tous les paramètres environnementaux et génétiques qui vont conduire à cette réponse. Par exemple, des signaux d'induction de tolérance peuvent provenir de cellules épithéliales ou mésenchymateuses par la sécrétion de TGF-β ou d'IL-10, alors que des signaux de danger proviendront de n'importe quel type cellulaire présent, *via* la sécrétion de TNF-α, IL-6, IL-8 ou IL-12.

Schématiquement, 2 grands types de lymphocytes T régulateurs sont distingués : iTreg et Tr1, et 3 grands types de lymphocytes T effecteurs : Th1, Th2 et Th17 (**Figure 9**).

Au niveau intestinal, il existe une activation basale des lymphocytes T, même en l'absence d'infection. Cette activation permet d'établir et de maintenir une tolérance vis à vis de la flore commensale. Il existe notamment un équilibre entre les populations de lymphocytes Th17 et T régulateurs. Parmi ces lymphocytes T régulateurs, il existe tout d'abord les lymphocytes T régulateurs induits (iTreg) de la *lamina propria* qui dépendent de la présence de TGF-β et du facteur de transcription FoxP3 combinés à l'absence d'IL-6, et qui synthétisent du TGF-β. Les autres lymphocytes T régulateurs sont les T régulateurs 1 (Tr1) localisés en position intra-épithéliale et dans la *lamina propria* de l'intestin grêle qui dépendent de la présence d' IL-10 et qui synthétisent également de l'IL-10. Ces lymphocytes T régulateurs sont très importants pour limiter l'inflammation en réponse à la flore commensale. (Barnes and Powrie, 2009; Hooper and Macpherson, 2010).

Parmi les lymphocytes T effecteurs, les lymphocytes Th1 dépendent de la présence d'IFN-γ et d'IL-12 ainsi que des facteurs de transcription Stat1, Stat4 et Tbet. Ils synthétisent de l'IFN-γ et de l'IL-2 et sont impliqués dans la réponse aux infections bactériennes intracellulaires. Les lymphocytes Th2 dépendent de la présence d'IL-4 et des facteurs de transcription Stat6 et Gata3. Ils synthétisent de l'IL-4, IL-5 et IL-13 et sont impliqués dans la réponse aux infections

parasitaires. Enfin, les lymphocytes Th17 dépendent de la présence d'IL-6, d'IL-23 et de TGF-β, ainsi que des facteurs Stat3 et ROR-γt. Ils synthétisent quant à eux de l'IL-17 et de l'IL-6 et sont impliqués dans la réponse aux infections bactériennes extracellulaires et aux infections fongiques (Hooper and Macpherson, 2010). Tous ces lymphocytes effecteurs, par l'induction de cytokines pro-inflammatoires induisent également une différenciation de cellules caliciformes, conduisant à une augmentation de production de mucus, en réponse à l'inflammation (McGuckin *et al.*, 2011).

Enfin, il existe également parmi les cellules du GALT des lymphocytes T intraépithéliaux au niveau des muqueuses. Ces lymphocytes sont au contact direct des cellules épithéliales et réparties tout au long de l'intestin grêle. Contrairement aux lymphocytes T de la *lamina propria*, ces cellules expriment le marqueur CD8+. Elles ont donc une activité cytotoxique et sont capables de sécréter des cytokines pro-inflammatoires comme l'IFN-γ.

# II. Le microbiote intestinal

Les progrès spectaculaires du séquençage des génomes microbiens, les techniques de biologie moléculaire de dernière génération et les outils de bio-informatique ont conduit à l'identification de métagénomes, c'est-à-dire de séquences exhaustives des génomes d'une flore complexe. Ces techniques ont enfin permis de s'affranchir des techniques de culture longues, fastidieuses et incomplètes. Les flores commensales ont ainsi été rebaptisées en « microbiote ».

Le microbiote correspond à l'ensemble des communautés microbiennes (bactéries, archéobactéries, protozoaires, champignons et virus) qui colonisent les différents écosystèmes de l'homme dans des conditions normales et/ou pathologiques. La grande majorité des micro-organismes présents dans le microbiote intestinal correspond à des bactéries. Le microbiote est en effet composé d'environ 100 000 milliards de bactéries, soit dix fois plus que de cellules humaines. Les virus représentent quant à eux 5,8 % des communautés présentes, les archées seulement 0,8 % et les eucaryotes 0,5 % (Arumugam *et al.*, 2011). Les micro-organismes qui constituent ce microbiote se sont adaptés à l'environnement intestinal et vivent en synergie avec l'Homme avec lequel ils ont co-évolué depuis plusieurs milliers d'années. Ainsi, le microbiote, de par le dialogue qu'il établit avec la muqueuse intestinale, influence plusieurs aspects de la physiologie humaine et en particulier la digestion mais aussi le métabolisme, le comportement et l'homéostasie tissulaire (Sommer and Bäckhed, 2013).

## 1. Mise en place du microbiote intestinal

La mise en place du microbiote intestinal débute dès la naissance. En l'absence d'un système immunitaire mature, le tube digestif du nouveau-né constitue un système très permissif, avec une augmentation rapide des populations bactériennes qui vont rapidement atteindre $10^{11}$ bactéries par gramme de selles. La colonisation bactérienne dépend tout d'abord de l'exposition du nouveau-né à la flore maternelle fécale, vaginale et/ou cutanée ainsi qu'à la flore environnementale. L'alimentation ou une antibiothérapie précoce ont également des effets majeurs sur la mise en place du microbiote. La colonisation va également dépendre des sécrétions du tube digestif du nouveau-né ainsi que des premiers micro-organismes colonisateurs, qui vont conditionner la physicochimie du biotope mis en place. Lors d'une

naissance par voie basse, les bactéries anaérobies sont parmi les premières bactéries à s'installer

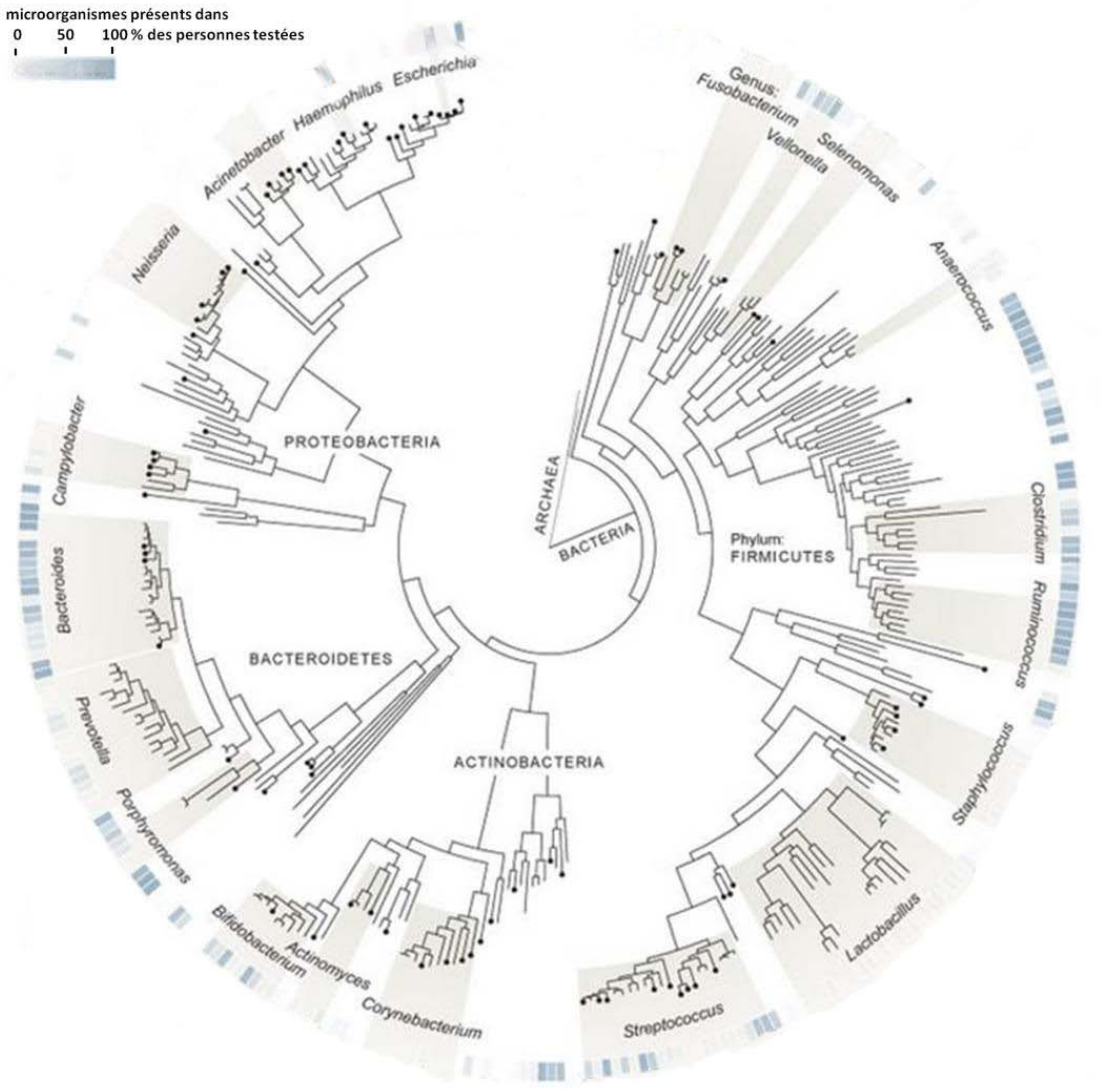



**Figure 10** : Arbre phylogénétique des principales espèces bactériennes retrouvées dans le microbiote intestinal.

(image modifiée depuis le site internet : http://www.drschweig.com/american-food-project/)

dans l'intestin mais leur expansion ne se produit que lorsque l'oxygène présent a été consommé par les bactéries anaérobies facultatives. La mise en place successive des différents micro-organismes, notamment lors du passage à la nourriture solide, conduit finalement à un microbiote stable et fonctionnel vers l'âge de 2 à 4 ans (Mackie *et al.*, 1999; Koenig *et al.*, 2011).

## 2. Composition du microbiote intestinal

### 1) Les bactéries

L'étude du microbiome humain, c'est-à-dire de l'ensemble des génomes ou des gènes constituant le microbiote, montre que la diversité intra-individuelle et inter-individuelle varie en fonction des microbiotes. Le « Human microbiome project » (HMP) a identifié au niveau de 18 sites prélevés, 30 phylums, 51 classes, 125 ordres, 493 familles et 929 genres bactériens. Ainsi, il a été observé, dans le microbiote digestif la dominance de 4 phyla majeurs : les *Firmicutes*, les *Bacteroïdetes*, les *Actinobacteria* et les γ-*Proteobacteria* (Eckburg *et al.*, 2005; Manichanh *et al.*, 2006; Morgan *et al.*, 2013) (**Figure 10**).

Le *phylum* des *Firmicutes* (bactéries à Gram positif) est toujours fortement représenté, avec le groupe des *Clostridium coccoïdes* (14 à 31 % des bactéries selon les études) comprenant des espèces bactériennes des genres *Eubacterium*, *Clostridium*, *Ruminococcus* et *Butyrovibrio*, et le groupe des *Clostridium leptum* avec notamment *Faecalibacterium prausnitzii*, *Ruminococcus albus* et *Ruminococcus flavefaciens* (16 à 22 % des bactéries en moyenne). Ce phylum comprend également les lactobacilles et les streptocoques.

Le *phylum* des *Bacteroïdetes* comprend principalement les genres *Bacteroïdes*, *Prevotella* et *Porphyromonas*, toujours présents dans le microbiote intestinal, représentant 9 à 42 % des bactéries totales du microbiote.

Le phylum des *Actinobacteria* est moins représenté et comprend essentiellement les bifidobactéries (0,7 à 10 % en moyenne).

Les γ-*Proteobacteria* qui comprennent principalement les entérobactéries représentent environ 1% des bactéries du microbiote (CDU-HGE, 2014).

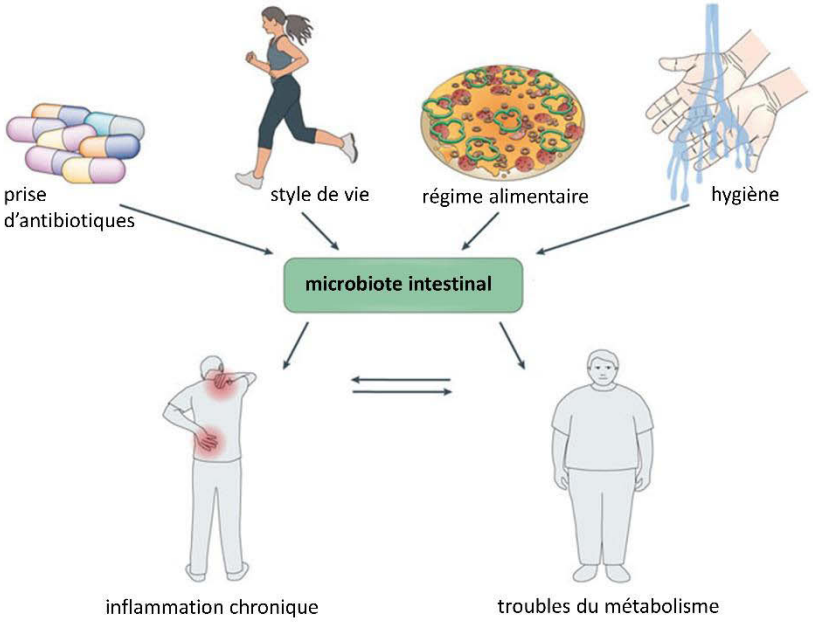



**Figure 11** : Facteurs environnementaux influençant le microbiote intestinal (image modifiée d'après Sommer and Bäckhed, 2013)

Toutefois, chaque individu semble unique en terme de diversité d'espèces bactériennes et 3 entérotypes, c'est-à-dire des grands types de communautés microbiennes, ont pu être définis selon le taxon prédominant, *Bacteroidetes, Prevotella* ou *Ruminococcus* (Arumugam *et al.*, 2011). Ainsi, toutes les communautés sont dominées par un à plusieurs taxons conservés dans l'espèce humaine, suivis par une longue liste de micro-organismes moins abondants et à forte variabilité interindividuelle. Cette variabilité est à mettre en relation avec le style de vie, le régime alimentaire, l'environnement (conditions plus strictes d'hygiène dans les pays industrialisés), la prise d'antibiotiques, etc... (**Figure 11**).

Malgré cette biodiversité et cette variabilité inter-individuelle, le métabolome associé au microbiote apparaît conservé entre les individus et quel que soit le site du microbiote (Turnbaugh *et al.*, 2009; Walter and Ley, 2011). La notion de répertoire génique commun à un grand nombre d'individus (core genome) est donc en passe d'être remplacée par celle d'un répertoire de fonctions bactériennes communes à une grande majorité d'individus (core fonctionnel métabolique) (Lozupone *et al.*, 2012). De fait, le microbiote humain est aujourd'hui considéré comme un organe à part entière formant avec les autres organes humains un hétéro-organisme ou super-organisme complexe (O'Hara and Shanahan, 2006; Evans *et al.*, 2013).

Dans le tractus intestinal humain, la distribution spatiale du microbiote s'étend selon deux axes, l'un longitudinal (de la cavité orale jusqu'au rectum) et l'autre radial (de la lumière intestinale jusqu'à la couche du mucus en contact avec les cellules épithéliales).

### a. Axe longitudinal

La répartition des bactéries varie quantitativement le long du tube digestif. L'estomac, le duodénum et le jéjunum sont peu colonisés par le microbiote intestinal, avec environ 10 à $10^2$ bactéries par gramme de contenu luminal. La prolifération bactérienne au niveau de l'intestin grêle est normalement limitée par différents facteurs : le pH acide de l'estomac, le flux luminal qui est rapide à ce niveau, la présence du gel de mucus contenant des IgA et de peptides antimicrobiens, ainsi que la présence de sels biliaires sécrétés au niveau du

duodénum et qui possèdent une activité anti-bactérienne (Inagaki *et al.*, 2006; Merritt and Donaldson, 2009).

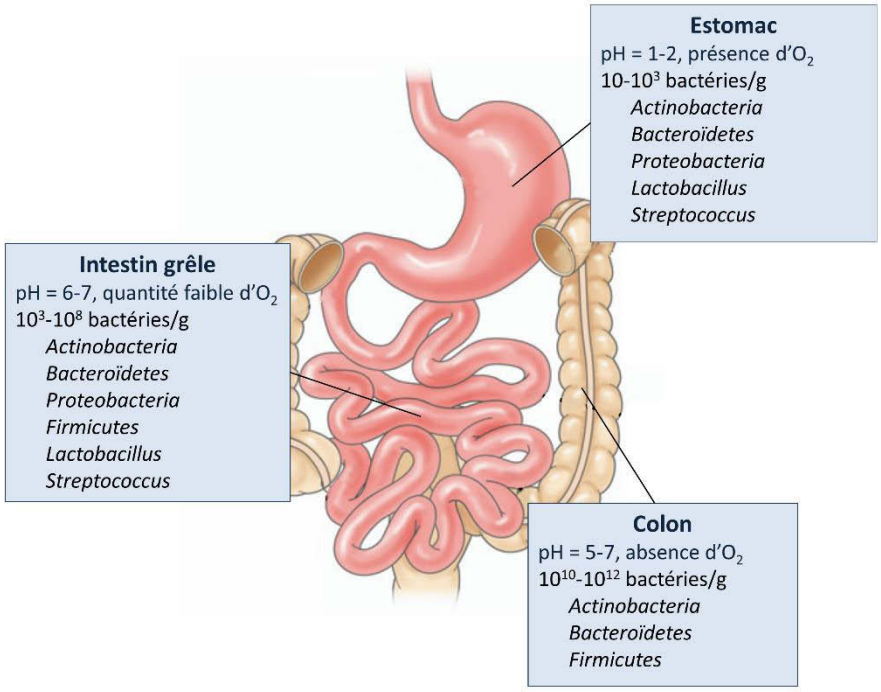



**Figure 12** : Composition et concentration des espèces bactériennes dominantes de la flore commensale le long du tractus digestif (image modifiée d'après Sartor, 2008)

Puis plus l'on s'éloigne de l'estomac, plus le pH augmente, plus le dioxygène se raréfie et plus le nombre de bactéries augmente. L'iléon comprend environ $10^7$ à $10^8$ bactéries par gramme de matière. Au niveau de l'intestin proximal (jéjunum, duodénum et iléon**)**, les trois phylums principaux de la flore du colon (*Firmicutes*, *Bacteroidetes* et *Actinobacteria*) sont présents, accompagnés de bactéries buccales ainsi que d'entérobactéries.

La diversité bactérienne ainsi que la densité atteignent leur maximum au niveau du colon avec $10^{10}$ à $10^{12}$ bactéries par gramme de contenu (Sartor, 2008) (**Figure 12**).

Le colon est plus favorable à la prolifération bactérienne avec un pH proche de la neutralité, un temps de rétention plus important lié à un péristaltisme plus lent, une couche de mucus externe épaisse colonisée par les bactéries, peu de sels biliaires, ces derniers ayant été réabsorbés au niveau de l'iléon terminal et une absence des plaques de Peyer. La flore du colon est pratiquement identique à celle des selles avec la présence majoritaire des 3 phylums (*Firmicutes*, *Bacteroidetes* et *Actinobacteria*) et en petite proportion des centaines d'espèces différentes (Savage, 1977; Walter and Ley, 2011; Marchesi, 2011).

### b. Axe radial

Au niveau de l'intestin grêle, les *Bacteroïdetes* et certains membres des *Firmicutes* sont largement retrouvés au niveau de la flore intestinale associée à la muqueuse, alors que les autres *phyla* dont les γ-*Proteobacteria*, qui incluent les *Enterobacteriaceae,* sont plutôt retrouvés au niveau de la lumière intestinale. Au niveau du colon, la flore intestinale associée à la muqueuse comprend également majoritairement les *Firmicutes* (principalement *Lachnospiraceae* et *Ruminococcaceae*) puis les *Bacteroïdetes,* mais également d'autres *phyla* comme les *Actinobacteria* et quelques γ-*Proteobacteria.* Par contre, dans la lumière intestinale, les *Lactobacillaceae* sont largement majoritaires et les *Firmicutes* sont beaucoup moins présents (les *Lachnospiraceae* et les *Ruminococcaceae* ont par exemple totalement disparus) (Nava and Stappenbeck, 2011; Walter and Ley, 2011).

### 2) Les *Archaea*

Les archées ou *Archaea* ont été découverts vers la fin des années 70 dans des environnements extrêmes. Ils sont assez similaires en taille et en forme aux bactéries et comme elles, ils appartiennent aux procaryotes. Cependant, ils possèdent des propriétés métaboliques et génétiques proches des eucaryotes. De plus, les membranes des archées sont composées de lipides d'éther de glycérol alors que les bactéries et les eucaryotes sont composés principalement de lipides d'ester de glycérol. Cette différence confère à la membrane une plus grande stabilité et donc de pouvoir survivre à des conditions extrêmes.

Dans une étude de 2011, l'extraction et la détection de l'ADN des archées a été réalisée sur 650 selles humaines. Il a été détecté une grande prévalence de *Methanobrevibacter smithii* (95,5% des cas) et *Methanosphaera stadtmanae* (29,4%) dans l'intestin humain. Ces procaryotes sont capables de produire du méthane en condition d'anaérobiose. Cette production est importante pour prévenir l'accumulation d'acides et de produits de fin de réaction dans l'intestin (Dridi *et al.*, 2011).

### 3) Les autres micro-organismes

A côté des bactéries et des archées, d'autres types de micro-organismes composent le microbiote, comme les protozoaires et les champignons, dont les fonctions restent à ce jour moins connues (Scanlan and Marchesi, 2008), ainsi que des virus, avec majoritairement des prophages et des phages dont les hôtes sont les bactéries de la flore. La flore fongique est dominée par des Ascomycètes et des Basidiomycètes, avec principalement *Saccharomyces*, *Debaryomyces*, *Penicillium*, *Kluyveromyces* et *Candida.* La plupart de ces champignons proviennent de notre nourriture (Sokol *et al.*, 2016). Les virus sont quant à eux fortement retrouvés au niveau du colon, et jouent donc probablement un rôle important dans cet écosystème intestinal (Reyes *et al.*, 2010).

## 3. Fonctions du microbiote

Lors de notre évolution, les interactions entre l'hôte et le microbiote intestinal ont fini par aboutir à une relation équilibrée, bénéfique pour les deux parties. En effet, les bactéries

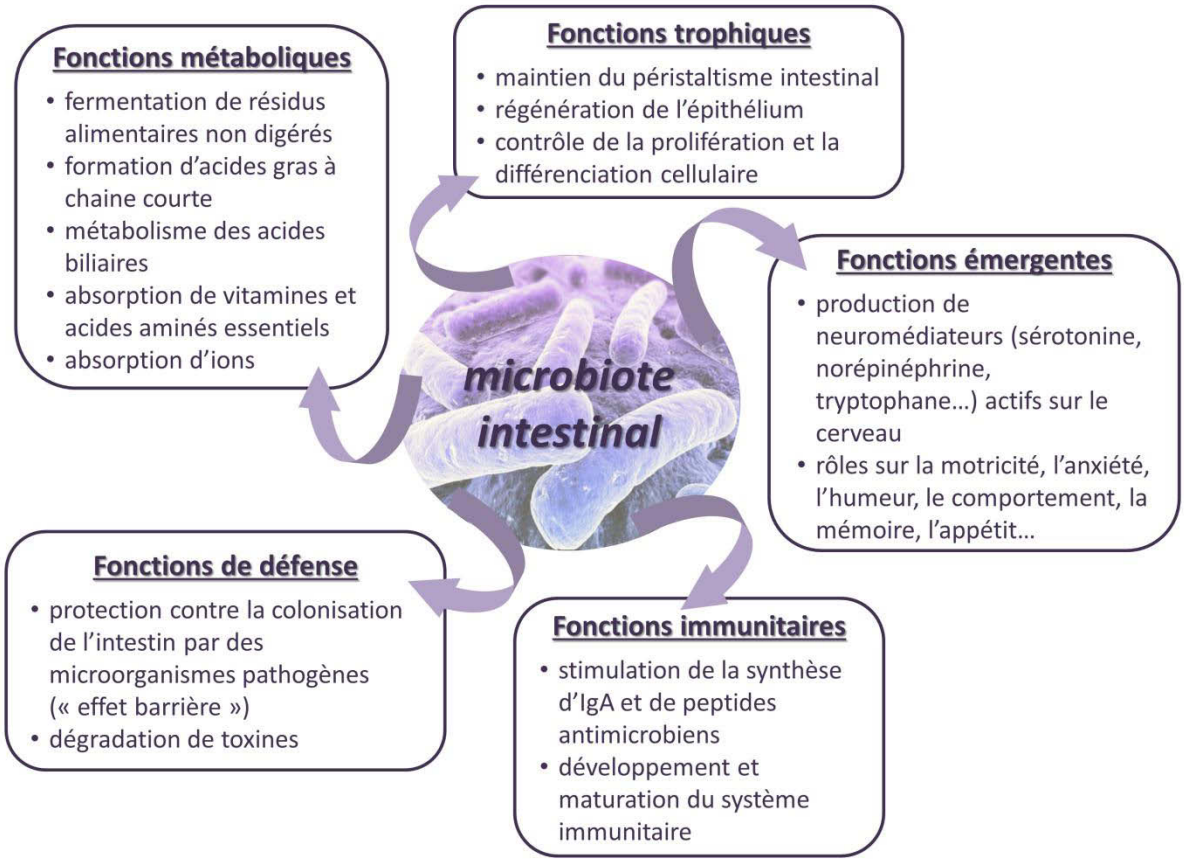



**Figure 13** : Rôles du microbiote intestinal
dans les fonctions physiologiques de l'intestin

intestinales sont dans un habitat tempéré avec une source continuelle de nutriments. En même temps, l'hôte bénéficie de la présence du microbiote intestinal qui possède un rôle important dans de nombreuses fonctions physiologiques, notamment sur le plan métabolique, trophique, immunitaire et protecteur (Guarner and Malagelada, 2003) (**Figure 13**).

La centaine de gènes présents chez les bactéries et absents chez l'Homme a fait émerger l'idée que nous formons avec notre flore endogène un "super-organisme", avec des échanges réciproques de nutriments et d'énergie (Cerf-Bensussan and Gaboriau-Routhiau, 2010).

## 1) Fonctions métaboliques

Le microbiote intestinal dégrade et fermente les aliments qui sont non digérables par l'hôte. Ces substrats alimentaires sont essentiellement composés d'amidon résistant, de polysaccharides végétaux (paroi cellulaire) et de certains oligosaccharides et sucres comme l'inuline, les gommes, les mucilages ou les fructo-oligosaccharides. Le microbiote participe également au métabolisme des substrats produits par l'hôte (mucus, cellules épithéliales desquamées) (Papillon *et al.*, 1999).

### a. Métabolisme des polysaccharides

Le métabolisme des polysaccharides comprend deux phénomènes anaérobies, la dégradation par hydrolyse et la fermentation par glycolyse.

La dégradation des polysaccharides fait intervenir une série d'enzymes hydrolytiques qui permet la formation de composés intermédiaires (type lactate, formate) et des fragments de polysaccharides. Les principales espèces bactériennes pour lesquelles une activité hydrolytique a été démontrée appartiennent aux genres *Bacteroïdes*, *Bifidobacterium*, *Ruminococcus* et *Roseburia* ainsi qu'aux *Clostridium*, *Eubacterium* et *Enterococcus*. Une même fonction hydrolytique peut être retrouvée dans différentes espèces bactériennes, par l'exemple l'amidon peut être hydrolysé par des bactéries du genre *Bacteroïdes* mais aussi par des *Bifidobacterium*, *Ruminococcus* et *Roseburia*. Ces composés intermédiaires et ces fragments de polysaccharides sont alors transformés en acide gras à chaînes courtes (AGCC), les principaux étant l'acétate, le butyrate et le propionate, et en gaz ($CO_2$, $H_2$ et $SO_4$) par

d'autres communautés bactériennes spécialisées. La dégradation d'une structure aussi complexe exige ainsi la contribution de plusieurs espèces bactériennes dotées d'activités complémentaires (Bernalier-Donadille, 2010).

La majorité des espèces bactériennes utilisent la glycolyse pour convertir les hydrates de carbone en pyruvate. Comme pour la voie de dégradation par hydrolyse, les principaux métabolites de la fermentation sont les AGCC (l'acétate, le butyrate et le propionate).

Dans des conditions normales, la concentration en AGCC dans le colon est relativement constante, avec respectivement 50 à 70 mmol/kg de contenu luminal pour l'acétate, et 15 à 25 mmol/kg pour le propionate et le butyrate (Cummings, 1981; Cummings *et al.*, 1987). Les quantités d'AGCC synthétisés varient selon la nature de la flore bactérienne colique, le temps de transit intestinal et les substrats disponibles. La majorité des bactéries intestinales produisent de l'acétate, alors que le propionate est produit par les *Bacteroïdes* et le butyrate principalement par les groupes *Roseburia* et *Faecalibacterium*. Les AGCC générés par le microbiote intestinal sont ensuite absorbés par les cellules épithéliales du colon. Environ 70% de l'énergie utilisée par les colonocytes proviennent de la transformation du butyrate en corps cétoniques. Les AGCC non utilisés par les colonocytes se retrouvent ensuite dans la circulation sanguine et sont métabolisés dans différents organes. Le propionate est essentiellement utilisé pour la néoglucogenèse hépatique alors que l'acétate est utilisé comme source d'énergie par de nombreux tissus (reins, cœur, muscles…) (Papillon *et al.*, 1999).

### b. Métabolisme des acides biliaires

Les acides biliaires sont des acides de structure stéroïdienne qui proviennent de la dégradation du cholestérol dans le foie. En milieu neutre, les acides biliaires se trouvent sous forme de sels biliaires qui sont les composants principaux de la bile avec le cholestérol et les phospholipides. Chez l'homme, les principaux acides biliaires primaires sont l'acide cholique et chénodésoxycholique. Ces acides biliaires primaires sont conjugués à la glycine ou à la taurine avant d'être excrétés dans les voies biliaires.

Dans l'iléon, sous l'action des bactéries intestinales, les acides biliaires vont être déconjugués par des enzymes appartenant à la classe des hydrolases, enzymes présentes chez de nombreuses bactéries comme *Lactobacillus, Bifidobacterium, Bacteroïdes* ou *Roseburia*

(Ridlon *et al.*, 2006; Labbé *et al.*, 2014). Les autres transformations des acides biliaires par le microbiote intestinal incluent des oxydations et des épimérisations.

D'un point de vue physiologique, la biotransformation la plus significative chez l'homme est la 7-déshydroxylation des acides primaires cholique et chénodésoxycholique en acides secondaires déoxycholique et lithocholique, respectivement (Begley *et al.*, 2005; Ridlon *et al.*, 2006). Cette transformation, qui permet la réabsorption des acides biliaires, est effectuée grâce aux enzymes bactériennes (hydroxystéroïdes deshydrogénases). La communauté responsable de la 7-déshydroxylation des acides biliaires n'est détectée qu'à un faible niveau dans les selles humaines ($10^3$ à $10^5$ par g de selles) et est principalement composée d'espèces de *Clostridium* et *Eubacterium* (White *et al.*, 1980; Baron and Hylemon, 1995). Quatre vingt dix pour cent des acides biliaires sont réabsorbés tout au long de l'intestin grêle, mais surtout au niveau de l'iléon terminal et sont transportés au foie par la veine porte : c'est le cycle entéro-hépatique. Six à dix cycles entérohépatiques ont lieu chaque jour correspondant à 40 à 60 mmol d'acides biliaires sécrétés. Le taux de synthèse hépatique des acides biliaires est d'environ 1,2 mmol/24h et correspond à la perte fécale (CDU-HGE, 2014; Dictionnaire médical de l'Académie de Médecine).

Les acides biliaires secondaires réabsorbés, comme l'acide lithocholique, peuvent être sulfatés. Ainsi, l'acide lithocholique n'est pas conjugué à un acide aminé comme les autres acides biliaires, mais subit l'action d'une sulfotransférase à PAPS (phosphoadénosine phosphosulfate) qui estérifie la fonction alcool en position C-3 par un ion sulfate. Le sulfolithocholate est ensuite réexcrété dans la bile avec les autres sels biliaires conjugués, mais il ne sera pas réabsorbé par le cycle entéro-hépatique au niveau de l'iléon et pourra être définitivement excrété dans les matières fécales. Une désulfatation des acides biliaires par l'intermédiaire de sulfatases bactériennes spécifiques a également été décrite et concerne certaines bactéries comme *Clostridium, Faecalibacterium* et *Peptococcus.* Ainsi, plus de 20 acides biliaires secondaires ont été mis en évidence dans les selles, témoignant de la grande diversité des voies de transformation des sels biliaires primaires par le microbiote intestinal (Baron and Hylemon, 1995; Humbert *et al.*, 2012).

Les acides biliaires sont des composés amphipathiques qui rendent hydrosolubles les lipides alimentaires, jouant ainsi un rôle important dans l'absorption des lipides et des vitamines liposolubles (A, D, E, K) en formant des micelles. D'autres rôles impliquent la fixation des acides biliaires sur des récepteurs spécifiques. Il est désormais connu que les acides biliaires agissent sur le métabolisme des lipides, sur le métabolisme du glucose, sur

l'homéostasie ainsi que sur l'inflammation *via* leur liaison à des récepteurs tels que FXR (Farnesoid X receptor) et TGR5 (Swann *et al.*, 2011; Zhang *et al.*, 2012; de Aguiar Vallim *et al.*, 2013).

### c. Autres métabolismes

Certaines bactéries anaérobies facultatives (*E. coli*, *E. aerogenes*) sont capables de synthétiser *in vitro* un large éventail de vitamines (biotine, riboflavine, acide pantothénique, pyridoxine et vitamine K). Des bactéries anaérobies strictes (*C. butyricum*, *Veillonella sp*.) sont capables de synthétiser la vitamine B12, d'une grande utilité pour la croissance locale bactérienne.

Les bactéries intestinales jouent également un rôle dans l'absorption du magnésium, du calcium et du fer par l'intermédiaire de la fermentation de certains glucides complexes comme l'amidon (Berg, 1996; Younes *et al.*, 2001; Guarner and Malagelada, 2003).

## 2) Fonctions trophiques

Des modifications anatomiques et histologiques liées à la présence de la microflore sont mises en évidence en comparant ce qui se passe chez l'animal conventionnel par rapport à l'animal axénique (dépourvu de bactéries). C'est ainsi que l'on constate que l'absence de flore entraîne un ralentissement du transit intestinal et une dilatation du caecum, avec comme conséquence un effet sur la motricité. Le microbiote joue également un rôle important dans la maturation du tube digestif, en particulier sur l'épaisseur de la muqueuse intestinale, la taille des villosités, la production de mucus, la vascularisation épithéliale et l'activité enzymatique de la muqueuse. L'effet trophique sur l'épithélium intestinal pourrait faire intervenir la synthèse d'AGCC, et notamment le butyrate (Papillon *et al.*, 1999; Guarner and Malagelada, 2003). De plus, les bactéries stimulent le récepteur NOD2 présent dans les cellules souches par leur motif commun muramyl di-peptide (MDP) et conduisent à la régénération cellulaire (Berg, 1996; Sherwood, 2006; Nigro *et al.*, 2014).

Le microbiote intestinal possède également un rôle important dans la formation du réseau microvasculaire intestinal (Sekirov *et al.*, 2010). Plusieurs espèces, comme par exemple

*Bacteroïdes tetaiotaomicron*, sont capables d'induire la transcription de l'angiogénine-3, une protéine à activité angiogénique (Stappenbeck *et al.*, 2002).

### 3) Fonctions de défense

La flore résidente s'oppose à l'implantation et à la multiplication de bactéries exogènes, potentiellement pathogènes, par deux mécanismes.

D'une part, la flore endogène rentre en compétition pour l'utilisation de sources de carbone et d'énergie, ainsi que pour les sites d'adhésion à la muqueuse intestinale (Nyangale *et al.*, 2012). De plus, certaines bactéries commensales sont également capables de synthétiser elles-mêmes des peptides antimicrobiens (bactériocines) qui vont également empêcher l'implantation d'éventuels pathogènes (Konisky, 1982; Cerf-Bensussan and Gaboriau-Routhiau, 2010; Walter and Ley, 2011).

D'autre part, certains organismes de la flore résidente protègent l'hôte des pathogènes en stimulant le système immunitaire (Hooper and Macpherson, 2010). Par exemple, l'invasion et la dissémination de *Salmonella enterica* Typhimurium reste limitée grâce à la stimulation des récepteurs TLRs par des bactéries commensales (Vaishnava *et al.*, 2008). Les micro-organismes commensaux protègent également l'hôte contre le protozoaire *Toxoplasma gondii* en activant la production des cytokines par les cellules dendritiques (Benson *et al.*, 2009).

### 4) Fonctions immunitaires

Le microbiote intestinal est primordial dans la maturation et l'éducation du système immunitaire intestinal et donc dans la protection de l'hôte contre des micro-organismes pathogènes comme le confirme les études chez les animaux axéniques (dépourvus de flore commensale). Ces animaux ont une plus faible synthèse de peptides antimicrobiens et moins de lymphocytes intra-épithéliaux. Leurs plaques de Peyer sont également moins actives et plus petites, et le nombre de plasmocytes sécréteurs d'IgA est fortement diminué. Le niveau des lymphocytes Th1 et Th17 est également diminué au niveau local et systémique. La réduction du taux basal de lymphocytes Th1 contribue à l'augmentation du risque d'infections par des bactéries pathogènes telles que *Shigella flexneri* ou *Listeria monocytogenes* (Round and Mazmanian, 2009). De même la diminution de lymphocytes Th17 retrouvée chez des souris axéniques, par défaut de stimulation du système immunitaire par le microbiote intestinal, entraine un risque accru d'infections à *Citrobacter rodentium* (Ivanov *et al.*, 2009).

L'éducation du système immunitaire se fait par l'échantillonnage de la flore résidente par le

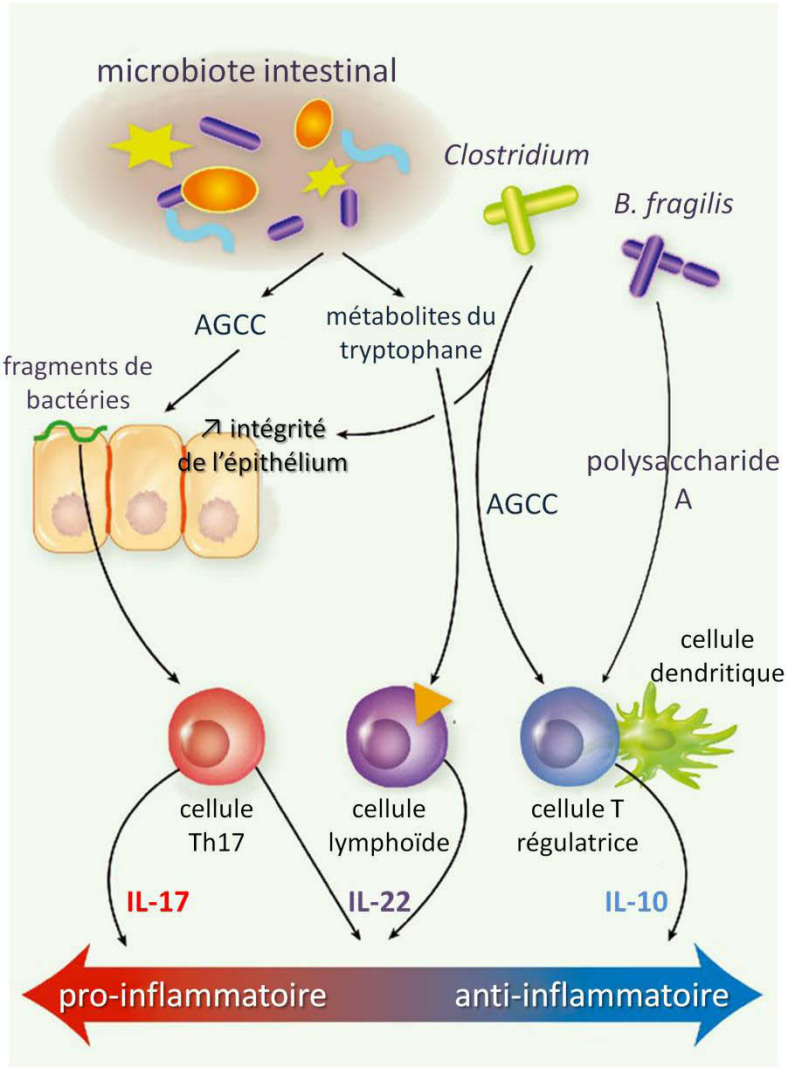



**Figure 14** : Modulation de l'immunité intestinale par le microbiote intestinal (image modifiée d'après McDermott and Huffnagle, 2014)

système immunitaire, notamment *via* les PRR, ce qui permet la différenciation et l'activation de certains lymphocytes comme les lymphocytes T régulateurs et les plasmocytes, ainsi que la stimulation de la synthèse d'IgA et de peptides antimicrobiens (Sartor, 2008; Hooper *et al.*, 2012).

De plus, certains micro-organismes constituant notre microbiote intestinal, comme par exemple *Bacteroïdes spp* ou *Clostridium spp*, sont capables de générer par fermentation des AGCC, comme le butyrate et le propionate, qui possèdent une activité anti-inflammatoire. Ils induisent une diminution de dégradation de la protéine IκB, ce qui empêche la translocation nucléaire du facteur de transcription pro-inflammatoire NF-κB (Yin *et al.*, 2001; Tedelind *et al.*, 2007; Puertollano *et al.*, 2014).

L'espèce commensale *Bacteroïdes fragilis* génère également du polysaccharide A qui permet la stimulation d'une réponse T régulatrice avec production d'IL-10 anti-inflammatoire, et de l'α-galactosylcéramide qui stimule les cellules NK. Le microbiote génère également des métabolites du tryptophane, qui peuvent induire une réponse lymphocytaire (**Figure 14**) (McDermott and Huffnagle, 2014).

Des altérations du microbiote intestinal pourraient donc affecter l'immunité intestinale, mais également la régulation du système immunitaire en général et pourraient donc conduire à des modifications fonctionnelles profondes rompant l'homéostasie intestinale (McDermott and Huffnagle, 2014).

### 5) Fonctions émergentes

Plus récemment, les chercheurs se sont intéressés au rôle du microbiote sur le développement du système nerveux central et périphérique. Ces interactions constituent l'axe « intestin-cerveau » bidirectionnel, avec en cas de dysfonction des implications potentielles sur le système nerveux entérique, pouvant jouer un rôle clé dans des pathologies gastro-intestinales aiguës ou chroniques, comme par exemple le syndrome de l'intestin irritable ou des maladies inflammatoires chroniques de l'intestin (Rhee *et al.*, 2009). La régulation de la communication entre l'intestin et le cerveau s'effectue par les cellules entérochromaffines, qui sont des cellules entéro-endocrines du tube digestif. Cet axe « intestin-cerveau » joue un rôle important dans la douleur, la modulation de la réaction immunitaire, le contrôle d'émotions, et d'autres fonctions homéostatiques.

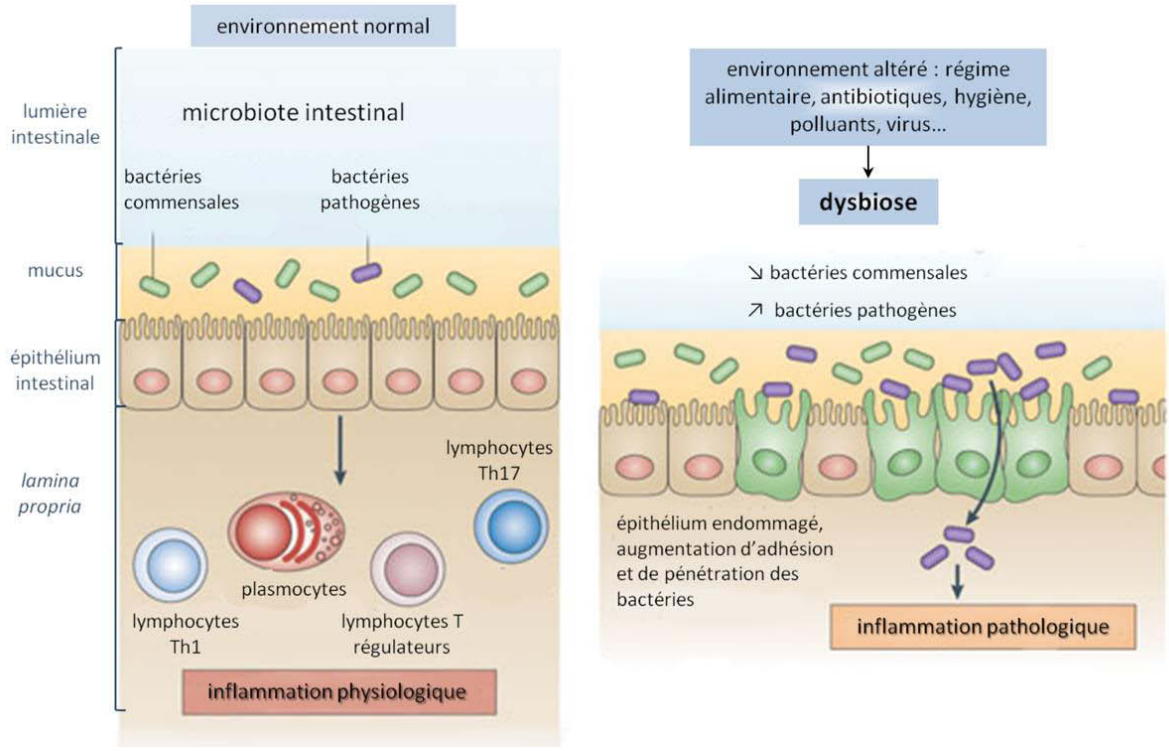



**Figure 15** : Représentation des interactions entre le microbiote intestinal et l'hôte chez des sujets sains et dans le cadre d'une dysbiose

(image modifiée d'après Cerf-Bensussan and Gaboriau-Routhiau, 2010)

Une étude chez des souris axéniques montre une augmentation de l'activité motrice associée à une diminution de l'anxiété comparativement aux souris conventionnelles (Heijtz *et al.*, 2011). Par ailleurs, l'étude de Collins et coll. en 2012 révèle le rôle du microbiote intestinal sur le développement de l'axe hypothalamo-hypophyso-surrénalien et sur le contrôle du niveau de certaines molécules comme le facteur neurotrophique dérivé du cerveau (BDNF), impliqué dans la croissance et la survie des neurones, la norépinéphrine et le tryptophane (Collins *et al.*, 2012). Le microbiote est donc impliqué dans la réponse au stress, la mémoire, l'humeur et le comportement.

## 4. Dysbioses

Le microbiote intestinal est un écosystème en équilibre qui s'autorégule en permanence. Cependant, cet équilibre peut être rompu avec des conséquences néfastes pour l'hôte : c'est ce qui s'appelle une dysbiose. De nombreux facteurs peuvent être à l'origine de cette dysbiose. C'est le cas par exemple d'une alimentation riche en graisses, sucres et viande rouge, d'une consommation excessive d'alcool, ou d'agressions extérieures comme l'administration d'antibiotiques (**Figure 15**).

La présence d'une dysbiose intestinale a été associée à des pathologies comme la colite à *Clostridium difficile,* le syndrome de l'intestin irritable, les maladies inflammatoires chroniques de l'intestin et même le cancer colorectal (Nyangale *et al.*, 2012). Concernant plus particulièrement la maladie de Crohn, différents scénarios ont été proposés : (i) une augmentation de bactéries à potentiel pro-inflammatoire, comme les *E. coli* adhérents et invasifs, ou (ii) une diminution de bactéries à potentiel anti-inflammatoire, comme *Faecalibacterium prausnitzii*. D'autres études ont mis en évidence une véritable dysbiose du microbiote intestinal, avec une augmentation de bactéries associées à la muqueuse, une plus faible diversité bactérienne, et une diminution globale du nombre de *Firmicutes* et de *Bacteroïdetes* au profit des γ-*Proteobacteria*. Une telle dysbiose a été observée chez des souris après avoir induit une inflammation avec un pathogène ou en injectant des cellules T transgéniques attaquant l'épithélium intestinal (Cerf-Bensussan and Gaboriau-Routhiau, 2010).

Sans microbiote intestinal efficient, le système immunitaire est moins actif, et peut ainsi conduire à des pathologies inflammatoires ou allergiques (Sartor, 2008; Cerf-Bensussan and Gaboriau-Routhiau, 2010; Nyangale *et al.*, 2012). En effet, des études épidémiologiques ont montré que le microbiote intestinal des enfants asthmatiques différait de celui des autres enfants, et que la prise d'antibiotiques aux premiers âges de la vie conduisait à un risque accru de survenue d'un asthme (Kalliomäki *et al.*, 2001; Björkstén *et al.*, 2001). Il existe donc une période critique, au début de la vie, durant laquelle une modification du microbiote intestinal peut perturber le développement du système immunitaire (Verhulst *et al.*, 2008). D'autre part, l'altération du microbiote intestinal provoquée par des changements nutritionnels et/ou des traitements antibiotiques répétés conduit à une perturbation des évènements conduisant normalement au développement de la « tolérance » immunitaire. Cette « hypothèse de la microflore » serait responsable de l'augmentation de l'incidence des hypersensibilités allergiques (Noverr and Huffnagle, 2005).

Des études utilisant des animaux axéniques ont également montré la contribution du microbiote intestinal dans le développement de l'obésité et de pathologies associées (Musso *et al.*, 2011). Plus récemment, des études effectuées chez l'homme ont confirmé que les personnes obèses hébergent un microbiote intestinal déséquilibré.

Le microbiote intestinal pourrait également avoir un rôle sur la physiopathologie de certains désordres nutritionnels, ou sur des maladies comme l'autisme, la maladie de Parkinson ou la maladie d'Alzheimer (Sekirov *et al.*, 2010; Ghaisas *et al.*, 2016)

Le microbiote intestinal est donc un élément essentiel qui participe à la santé de l'Homme, mais également à la survenue et au maintien des maladies quand il est dérégulé. Toutefois, il reste à déterminer si ces dysbioses participent à la survenue de la pathologie ou si elles réflètent l'adaptation des bactéries à un nouvel environnement.

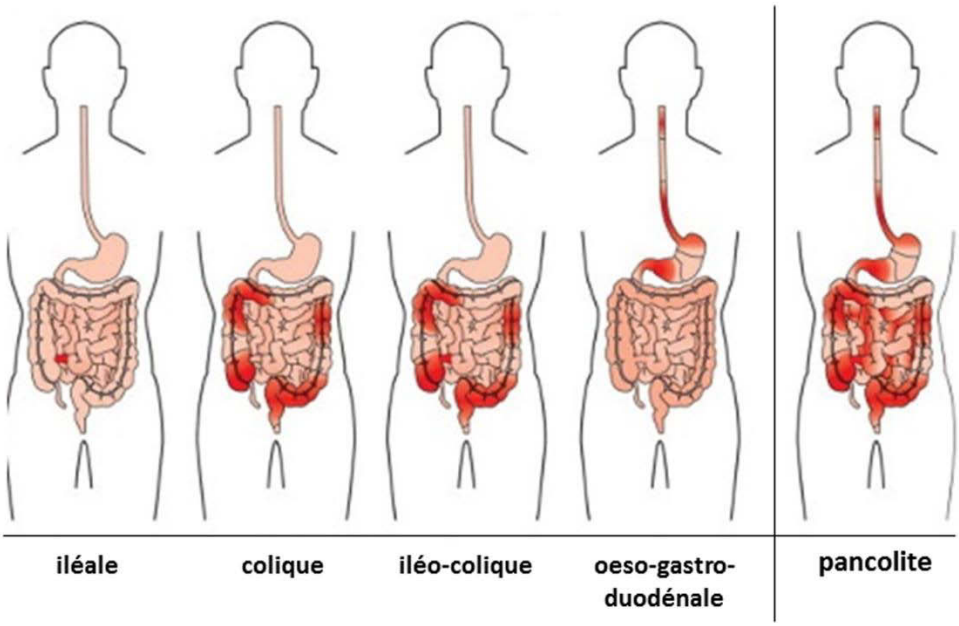



**Figure 16** : Différentes localisations possibles de la maladie de Crohn (image modifiée d'après Baumgart and Sandborn, 2012)

# III.La maladie de Crohn

La maladie de Crohn (MC) est une maladie inflammatoire chronique de l'intestin (MICI) qui a été décrite pour la première fois par Burrus B. Crohn en 1932. Il décrivit 14 cas d'une affection appelée iléite terminale ou régionale localisée au niveau de l'iléon terminal (Crohn *et al.*, 2000). Par la suite, les lésions tissulaires identiques ont pu être observées dans d'autres régions du tube digestif.

La MC est caractérisée par une inflammation chronique du tube digestif qui est entrecoupée par des poussées symptomatiques de gravité et de durée variable. Les lésions intestinales sont la conséquence d'une activation non régulée du système immunitaire lié à la muqueuse, et peuvent atteindre la totalité du tube digestif, de la bouche à l'anus. Ces lésions sont habituellement segmentaires, asymétriques et séparées par des intervalles de muqueuse saine. Au moment du diagnostic, l'iléon terminal est la région la plus fréquemment touchée (45 % des cas), puis le côlon (32 % des cas) et l'iléo-côlon (19 % des cas). A l'inverse, les atteintes de la partie supérieure du tractus gastro-intestinal ne concernent que très peu de patients (4 % des cas) (**Figure 16**) (Baumgart and Sandborn, 2012).

L'atteinte iléale, qui touche environ deux tiers des patients, est également l'atteinte provoquant le plus de complications (sténoses, fistules, abcès…). De plus, l'inflammation chronique de la paroi intestinale finit par conduire à une perte de la fonction physiologique de l'intestin.

Cette maladie constitue un problème majeur en gastro-entérologie car elle touche des sujets jeunes et retentit à la fois sur leur vie professionnelle et personnelle, en raison de la fréquence des poussées et des complications engendrées par cette maladie nécessitant souvent un recours à la chirurgie.

## 1. Epidémiologie

La MC peut survenir à tout âge, y compris chez les enfants, mais elle est le plus souvent diagnostiquée chez les jeunes adultes entre 20 et 30 ans. Un deuxième pic moins important survient entre 50 et 60 ans et il existe également des cas de MC pédiatrique. Les femmes sont plus fréquemment touchées (59 %) (Cosnes *et al.*, 2011).

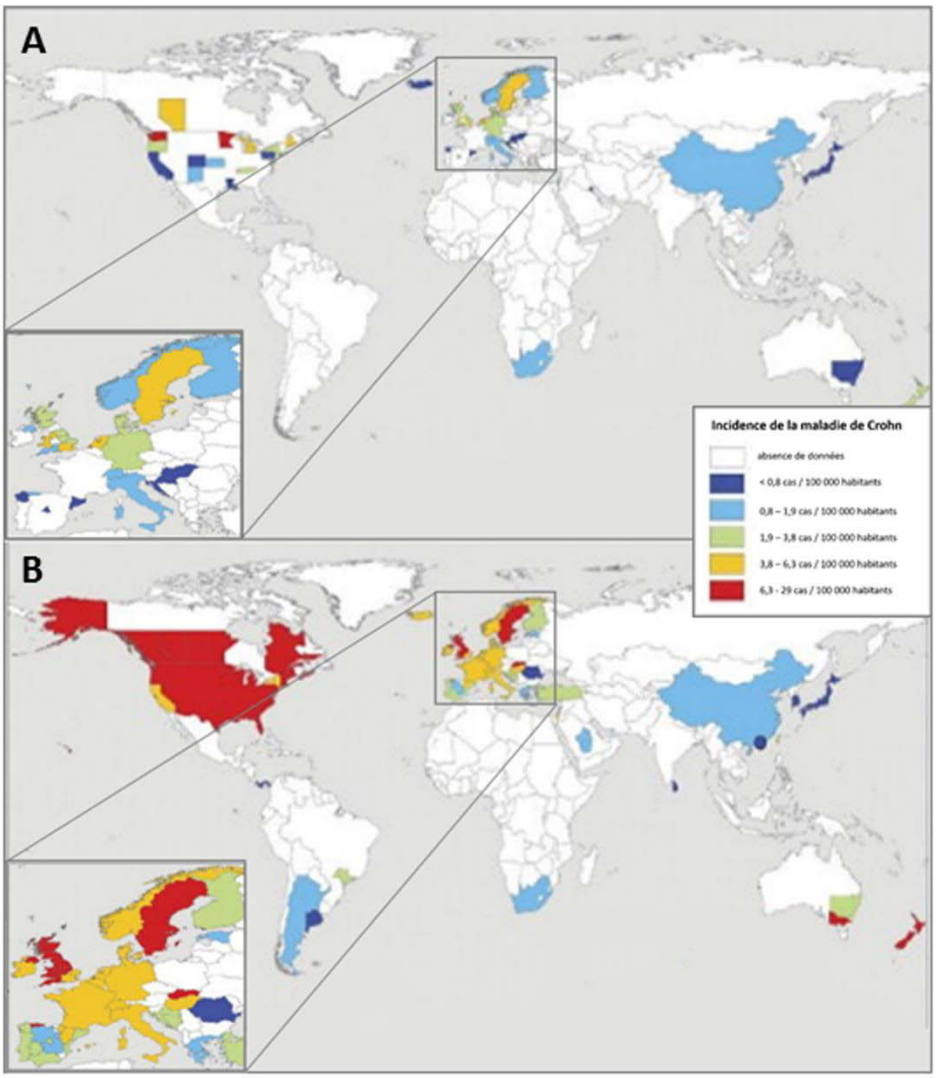



**Figure 18** : Incidence de la maladie de Crohn entre 1960 et 1980 (**A**) et après 1980 (**B**).
(image modifiée d'après Molodecky et al, 2012)

Au total, environ 120 000 personnes sont touchées en France, ce qui représente à peu près une personne sur 1000. Son incidence est de l'ordre de 5 pour 100 000 habitants par an, avec une progression globale des MICI d'environ 50 % entre 2000 et 2013 (données au 31 décembre 2013 sur *www.ameli.fr*). Le gradient Nord-Sud qui se dessinait clairement il y a quelques années tend aujourd'hui à disparaître.

Au niveau mondial, l'incidence de la MC ne fait qu'augmenter, avec de fortes disparités entre les pays (entre 0,3 et 20,2 pour 100 000 habitants, selon les régions du monde). Les populations les plus touchées par cette maladie sont les populations d'Europe (environ 5 cas / 100 000 habitants, et jusqu'à 10,6 cas / 100 000 habitants au Royaume Uni) et d'Amérique du Nord (environ 7 cas / 100 000 habitants, mais jusqu'à 20,2 cas / 100 000 habitants au Canada), mais également les populations blanches d'Australie, de Nouvelle-Zélande et d'Afrique du Sud (**Figure 18**) (Economou and Pappas, 2008; Molodecky *et al.*, 2012).

## 2. Signes cliniques

La MC provoque des ulcérations du tube digestif, qui peuvent parfois constituer des fistules qui traversent la paroi intestinale et font anormalement communiquer des organes entre eux. Cette maladie se manifeste dans la majorité des cas par une triade de symptômes non spécifiques : douleurs abdominales, diarrhées persistantes plus ou moins glaireuses et sanglantes, altération de l'état général (asthénie, perte de poids, dénutrition…) entrainant chez les enfants un retard de croissance (Laass *et al.*, 2014). D'autres signes peuvent également être retrouvés comme des nausées, des vomissements, de la fièvre et de l'essoufflement. En cicatrisant, ces lésions peuvent également aboutir à des sténoses, qui conduisent à un rétrécissement plus ou moins important du diamètre du tube digestif.

Chez 30 à 40 % des patients atteints, les symptômes de la MC sont accompagnés de manifestations extra intestinales de type articulaire (rhumatisme articulaire, polyarthrite, spondylarthrite ankylosante), cutané (aphte, érythème noueux), oculaire (uvéite), hépatique (stéatose, hépatite, cirrhose), biliaire ou urinaire (Ott and Schölmerich, 2013).

Les lésions retrouvées dans la MC sont plurisegmentaires et discontinues, avec des intervalles de muqueuse saine, et touchent toute l'épaisseur de la paroi intestinale. La lésion la

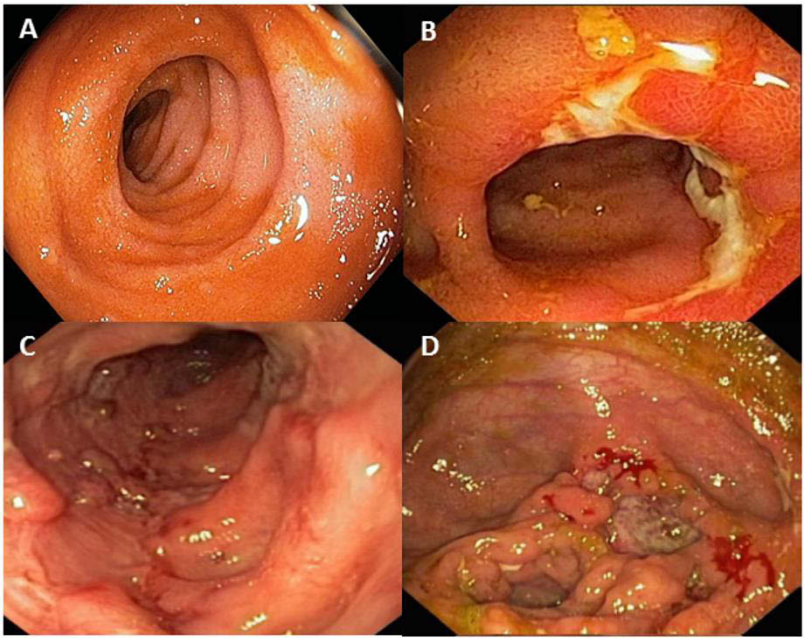



**Figure 18 :** Ulcérations retrouvées dans la maladie de Crohn.
(A) muqueuse iléale normale, (B) ulcérations aphtoïdes de l'iléon, (C) ulcérations profondes au niveau du cecum, (D) ulcérations étendues du colon
(image d'après http://www.gastrolab.net et Benitez et al., 2013).

plus précoce est l'ulcération aphtoïde, issue des cellules M au niveau des plaques de Peyer ou des follicules lymphoïdes (Fujimura et al., 1996; Gullberg and Söderholm, 2006). Puis apparaissent des lésions plus ou moins profondes et étendues, pouvant atteindre la sous-muqueuse et la séreuse, et conduisant à la formation de fistules et/ou d'abcès. Au niveau microscopique, les lésions sont des granulomes épithéloïdes et gigantocellulaires, présents dans 30 à 60 % des cas, formés par l'agglomération d'histiocytes associés à des lymphocytes et quelques cellules géantes (Sura et al., 2011) (**Figure 18**).

## 3. Diagnostic et suivi

La MC est principalement diagnostiquée lors d'une poussée par l'association d'un examen clinique, un interrogatoire, un bilan biologique, une coloscopie avec des biopsies pour différencier la MC des autres MICI, une endoscopie oeso-gastro-duodénale pour rechercher une localisation haute de la maladie (voire dans certains cas une vidéocapsule), et d'examens radiologiques (échographie abdominale, scanner abdominal et IRM) (Van Assche *et al.*, 2010; Baumgart and Sandborn, 2012; Laass *et al.*, 2014). Cependant, l'absence de marqueur spécifique et la forte variabilité inter-individuelle de la maladie nécessitent d'exclure toutes les autres pathologies digestives (virales, bactériennes, parasitaires, maladies de système, cancer, médicaments…) capables d'induire le même tableau clinique avant de pouvoir évoquer le diagnostic de MC.

### 1) Diagnostics différentiels

Les principaux diagnostics différentiels de la MC sont divisés en causes infectieuses et non infectieuses. Parmi les causes infectieuses à l'origine des même symptômes, nous pouvons retrouver des colites infectieuses d'origine bactérienne (*Salmonella*, *Shigella*, *Campylobacter*, *Yersinia*, *Clostridium*, *Escherichia coli* entérohémorragiques et/ou entéroaggrégatifs…), la maladie de Whipple (*Tropheryma whippelii),* des infections sexuellement transmissibles (*Chlamydia trachomatis*, *Neisseria gonorrhoeae…),* la tuberculose intestinale (*Mycobacterium tuberculosis*…), mais également des infections parasitaires (*Entamoeba histolytica*, *Giardia lamblia*, *Cryptospora*, *Isospora*, *Trichuris*…), des infections virales (*Enterovirus*, *Norovirus*,

**Tableau 2 :** Index d'activité de la maladie de Crohn, Index de Harvey-Bradshaw et graduation de l'activité de la MC d'après le consensus ECCO (European Crohn's and Colitis Organisation)

| CDAI | J1 à J7 : somme | coefficient multiplicateur | total |
|---|---|---|---|
| nombre de selles liquides ou molles | | 2 | |
| **douleurs abdominales** :<br>absentes = 0<br>légères = 1<br>moyennes = 2<br>intenses = 3 | | 2 | |
| **bien-être général** :<br>bon = 0<br>moyen= 1<br>médiocre = 2<br>mauvais = 3<br>très mauvais = 4 | | 2 | |
| **autres manifestations** : | | | |
| arthrites ou arthralgies | | 20 | |
| iritis ou uvéite | | 20 | |
| érythème noueux, pyoderma, aphtes buccaux | | 20 | |
| fistules, fissures, abcès anal ou périrectal | | 20 | |
| autre fistule intestinale | | 20 | |
| fièvre (>38 °C dans la semaine) | | 20 | |
| **traitement antidiarrhéïque** (lopéramine, opiacés) :<br>non = 0<br>oui = 1 | | 30 | |
| **masse abdominale** :<br>absente = 0<br>douteuse = 1<br>certaine = 5 | | 10 | |
| **hématocrite** :<br>homme = 47 – hématocrite<br>femme = 42 – hématocrite | | 6 | |
| **poids** :<br>100x (1-poids actuel/poids théorique) | | | |
| | | Total | |

| HBI | Valeur |
|---|---|
| **bien-être général** :<br>bon = 0<br>moyen= 1<br>médiocre = 2<br>mauvais = 3<br>très mauvais = 4 | |
| **douleurs abdominales** :<br>absentes = 0<br>faibles = 1<br>moyennes = 2<br>intenses = 3 | |
| selles liquides : nombre / jour | |
| **masse abdominale** :<br>absente = 0<br>douteuse = 1<br>certaine = 2<br>certaine avec défense | |
| signes extra-digestifs, fistule, fissure anale :<br>1 point par item présent | |
| **score (= somme)** | |

| | |
|---|---|
| ***Activité minime*** | CDAI entre 150 et 220 ; HBI entre 4 et 8 |
| patient en ambulatoire, sans gène alimentaire, avec une perte de poids < 10 %<br>absence de signe d'occlusion, d'hyperthermie, de déshydratation, de masse abdominale ou de sensibilité à la palpation<br>CRP en général augmentée au-dessus de limite supérieure à la normale | |
| ***Activité modérée*** | CDAI entre 220 et 450 ; HBI entre 8 et 12 |
| vomissements intermittents ou perte de poids > 10 %<br>traitement des poussées minimes inefficaces ou masse sensible<br>absence de signe évident d'occlusion<br>CRP augmentée au-dessus de limite supérieure à la normale | |
| ***Activité sévère*** | CDAI > 450 ; HBI > 12 |
| cachexie (IMC < 18 kg/m²) ou occlusion évidente ou abcès<br>symptômes persistants malgré un traitement intensif<br>CRP augmentée | |

*Astrovirus*, *Adenovirus*, *Rotavirus*, *Cytomegalovirus*, *Herpes simplex virus*…) ou des infections fongiques (*Candida*, *Aspergillus*…).

Les diagnostics différentiels de la MC d'origine non infectieuse incluent le syndrome de l'intestin irritable, la maladie coeliaque, le syndrome de Behcet, les colites non infectieuses, les gastroentérites à éosinophiles, la réaction du greffon contre l'hôte chez les patients allogreffés, la sarcoïdose, la diverticulite, la granulomatose familiale chronique, la prise de médicaments (AINS, laxatifs à forte dose, chimiothérapie anticancéreuse…) et les tumeurs (cancer colorectal, lymphome du MALT, syndrome néoplasique…) (Nikolaus and Schreiber, 2007).

Une fois les autres causes écartées, il reste à différencier la MC d'une autre MICI, la rectocolite hémorragique (RCH). La RCH, contrairement à la MC, induit une inflammation continue du colon en débutant par le bas du rectum, mais ne touche ni l'intestin grêle ni l'anus. Le diagnostic différentiel entre ces deux maladies nécessite donc des examens endoscopiques et anatomo-pathologiques (Xavier and Podolsky, 2007). Dans la majorité des cas, la distinction est simple, surtout lorsqu'il existe des atteintes ano-périnéales typiques de la MC. Cependant, lors d'une atteinte isolée du rectum et/ou du colon, la différence peut être plus difficile et le diagnostic rester indéterminé pendant plusieurs années. C'est le cas d'environ 10 à 12 % des colites dues aux MICI.

### 2) Suivi de la maladie

Lors des poussées de la MC, l'activité clinique de la maladie est mesurée, afin d'aider les cliniciens à évaluer la sévérité de la poussée (**Tableau 2**). Pour cela, deux scores peuvent être utilisés : l'index d'activité de la maladie de Crohn (CDAI) et l'index de Harvey-Bradshaw (HBI), étroitement corrélé au CDAI mais dont l'utilisation s'avère plus simple en pratique quotidienne (**Tableau 2**) (Best *et al.*, 1976; Best, 2006).

### 3) Marqueurs biologiques

La recherche de marqueurs sériques spécifiques de la MC présente un grand intérêt car ils pourraient apporter un bénéfice considérable dans le diagnostic non-invasif de la MC, avec une détection plus précoce de la maladie et une meilleure prise en charge liée au suivi de la réponse au traitement mis en place. Dans la mesure où cette maladie est due à une

dérégulation du système immunitaire, des marqueurs de type auto-anticorps ont été recherchés en premier lieu.

**Tableau 3** : Apport des marqueurs biologiques dans le diagnostic de MC ou RCH pour des colites indéterminées (d'après Beaven and Abreu, 2004)

| Diagnostic | marqueurs standards | valeur prédictive positive | nouveaux marqueurs | valeur prédictive positive |
|---|---|---|---|---|
| maladie de Crohn | ASCA +<br>pANCA - | 67 % | OmpC | 75 % |
| Rectocolite hémorragique | ASCA -<br>pANCA + | 70 % | ASCA -<br>pANCA +<br>OmpC - I2 - | 83 % |

De nombreux marqueurs ont été étudiés dans le but de définir un marqueur « idéal » permettant le diagnostic et le suivi de l'évolution de la maladie au fil du temps (Meuwis *et al.*, 2007). Cependant, aucun marqueur n'est aujourd'hui assez sensible et spécifique pour remplacer les outils actuels de diagnostic.

Outre les marqueurs classiques de l'inflammation (protéine C réactive, haptoglobine, orosomucoïde, TNF-α et IL-6), plusieurs marqueurs sérologiques ont été mis en évidence et permettent d'améliorer la fiabilité du diagnostic et la différenciation entre la MC et la RCH (**Tableau 3**).

Les ASCA (anti-*saccharomyces cerevisiae* antibodies) sont des anticorps dirigés contre des phosphopeptidomannanes de la membrane cellulaire des levures, mais également des mycobactéries et autres micro-organismes. La présence de ces anticorps chez les patients reflète une perte de tolérance immunologique vis-à-vis des antigènes retrouvés dans la lumière intestinale. Ces anticorps sont présents dans 50 à 70% des patients atteints de MC, et sont plutôt associés à une atteinte iléale de la maladie. Les pANCA (perinuclear antineutrophil cytoplasmic antibodies) quant à eux sont des anticorps dirigés contre les antigènes cytoplasmiques des polynucléaires neutrophiles et retrouvés en zone périnucléaire (à la différence des cANCA cytoplasmiques retrouvés dans la granulomatose de Wegener). Ces anticorps sont principalement retrouvés dans la RCH (Beaven and Abreu, 2004). L'association de ces deux types d'anticorps permet d'affiner le diagnostic, avec une valeur prédictive positive d'environ 70 à 80% pour l'association ASCA+, pANCA- dans la MC, et d'environ 70% pour l'association ASCA-/pANCA+ dans la RCH (Joossens *et al.*, 2002; Beaven and Abreu, 2004). D'autres marqueurs comme la calprotectine et la lactoferrine fécales, communes aux deux maladies aident au diagnostic de MICI, et servent également à mesurer l'activité de la maladie, l'état de la muqueuse intestinale et le risque de poussée évolutive (Benitez *et al.*, 2013; Wright *et al.*, 2014; Mosli *et al.*, 2015). Des anticorps dirigés contre des antigènes des sécrétions pancréatiques ont également été reportés chez des patients atteints de MC, de même que des anticorps anti-OmpC dirigés contre la porine membranaire *d'Escherichia coli*, des anti-I2 ciblant une séquence de *Pseudomonas fluorescens*, ainsi que des anticorps anti-flagelline CBir1. La présence de ces anticorps est plutôt associée à la sévérité de la maladie (Nakamura *et al.*, 2003; Bossuyt, 2006).

Plus récemment, de nouveaux marqueurs biologiques ont été reportés comme étant augmentés dans les poussées actives de MICI, sans toutefois pouvoir différencier la MC de la RCH. Ce sont les neuropeptides, qui constituent le lien entre le système nerveux entérique et le système nerveux central (axe « intestin-cerveau »). Le plus étudié est la chromogranine A (CgA), qui peut être dosée dans le sang et dans les selles, et qui augmente dans les poussées évolutives de la maladie en parallèle du TNF-α (Massironi *et al.*, 2016).

## 4. Complications

L'espérance de vie des patients atteints de MC est légèrement diminuée par rapport à la population générale, avec seulement un quart des décès imputable à la pathologie. Bien pris en charge, les patients atteints de MC peuvent donc mener une vie presque normale. Cependant, la moitié des patients atteints de MC vont souffrir de complications au bout de 10 ans (Louis *et al.*, 2001). Ainsi, la forte baisse de la qualité de vie des patients ainsi que les surcoûts engendrés par cette pathologie en font un véritable problème de santé publique (Floyd *et al.*, 2015).

L'évolution de la MC diffère d'une personne à l'autre. Les complications les plus fréquemment rencontrées sont des fistules (ano-périnéales, entéro-vésicales…), des fissures anales ou des abcès, pouvant aller jusqu'à l'occlusion intestinale, la péritonite, l'appendicite, la perforation intestinale ou l'hémorragie digestive nécessitant une prise en charge chirurgicale en urgence. Au fil du temps, la maladie devient également réfractaire aux traitements médicaux et la chirurgie reste alors le dernier recours possible. Le risque cumulé d'exérèse intestinale est ainsi de 80 % à 20 ans et les patients sont fréquemment réopérés suite à des récidives de la maladie.

Par ailleurs, l'inflammation chronique ainsi que des résections étendues de l'intestin grêle peuvent conduire à des problèmes de dénutrition et/ou d'anémie. Chez les enfants, cette maladie entraine souvent des retards de croissance et des difficultés scolaires. Chez les adultes, des répercutions psychologiques de même qu'un retentissement socioprofessionnel sont également à craindre.

D'autres affections sont plus rarement retrouvées, comme les pyodermites, les uvéites, les thromboses ou certaines maladies associées (polyarthrite…).

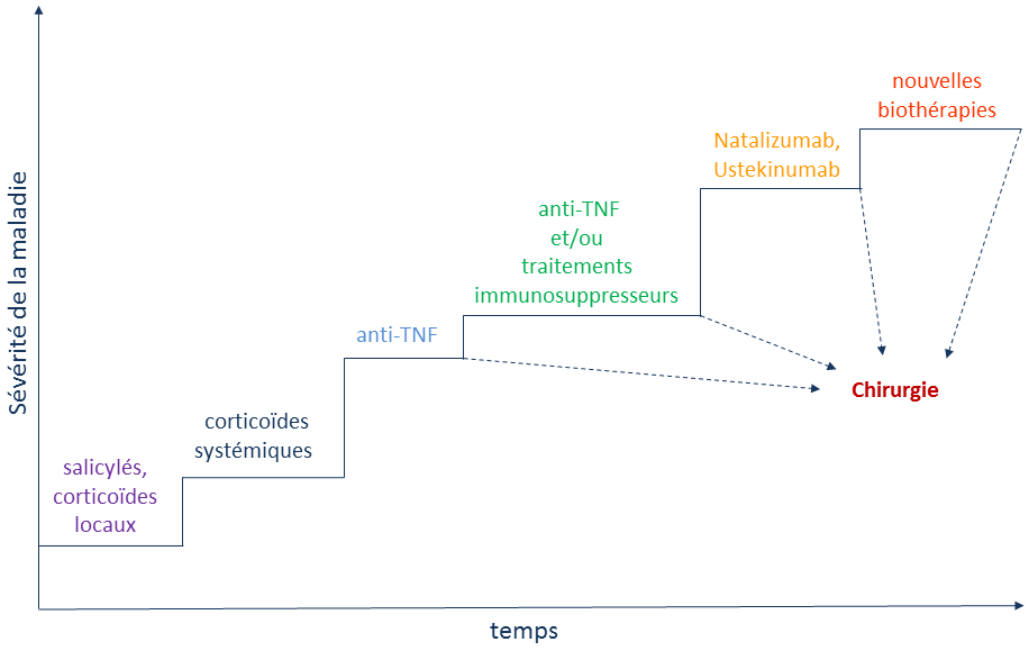



**Figure 19** : Stratégies de traitement de la maladie de Crohn
(image d'après Hanauer *et al.*, 2009 ; Simon *et al.*, 2016)

Enfin, sur le long terme, les patients atteints de MC présentent un risque plus élevé de développer un cancer colorectal après plusieurs années, vraisemblablement lié à la chronicité de l'inflammation intestinale (Ullman and Itzkowitz, 2011; Averboukh *et al.*, 2011).

## 5. Traitement

La MC est jusqu'à aujourd'hui une maladie incurable. Le traitement vise donc essentiellement à calmer les signes cliniques associés aux poussées de la maladie, prévenir les rechutes et participer au confort de vie des patients. Le traitement est assuré par une équipe multidisciplinaire menée par le médecin traitant et le gastroentérologue qui ont établi le diagnostic. Il varie en fonction de la nature, la localisation, la sévérité et l'évolution de la maladie (**Figure 19**).

Les principaux médicaments utilisés dans la MC sont les anti-inflammatoires et les immuno-modulateurs. Lors des poussées inflammatoires, des traitements symptomatiques sont également instaurés : anti-spasmodiques, anti-diarrhéiques, pansements intestinaux et alimentation sans résidu. Lorsque la maladie évolue vers une forme très inflammatoire conduisant à une forte dénutrition, une alimentation entérale ou parentérale peut être également efficace pour couvrir les besoins nutritionnels.

### 1) Les médicaments anti-inflammatoires

#### a. Les dérivés salicylés

Les dérivés salicylés font partie de la famille des anti-inflammatoires non stéroïdiens (AINS). Le principe actif est toujours le même : l'acide 5-aminosalicylique (5-ASA). Il agit au niveau de la muqueuse intestinale lésée par inhibition des prostaglandines, des leucotriènes et du facteur de transcription NF-κB, qui joue un rôle clé dans la production des facteurs de l'inflammation. En fonction du site de libération du 5-ASA, il existe différents dérivés salicylés.

La sulfasalazine, qui associe le 5-ASA et un sulfamide, est libérée au niveau du colon sous l'effet d'une azoréductase produite par les bactéries de la flore commensale. Elle est

donc essentiellement utilisée dans les formes coliques peu sévères de la MC car son efficacité reste limitée (Burger and Travis, 2011). D'autres dérivés salicylés sont plus utilisés : la mesalazine et l'olsalazine qui ne comportent que du 5-ASA et dont le site de libération varie dans l'intestin. Un traitement local (lavement ou suppositoire) de 5-ASA peut suffire en cas d'atteinte basse isolée.

Les principaux effets indésirables de ces médicaments sont des réactions immuno-allergiques (rash, syndrome de Stevens-Johnson, érythème multiforme) ainsi que des leucopénies, thrombopénies, une toxicité hépatique, pancréatique et surtout rénale, limitant souvent leur utilisation (Curkovic *et al.*, 2013).

### b. Les corticoïdes

Le traitement par corticoïdes (prednisone et prednisolone) est utilisé sur des formes moyennes à sévères de la MC. C'est le traitement le plus rapidement efficace, quelle que soit la localisation des lésions. La rémission de la MC est obtenue en 3 à 4 semaines dans plus de 90 % des cas. Par contre, les corticoïdes s'avèrent inefficaces pour maintenir les patients en rémission (Steinhart *et al.*, 2003).

Les effets indésirables des corticoïdes associent œdème, surcharge pondérale, hypertension artérielle, hyperglycémie… mais sont généralement réversibles à l'arrêt du traitement. Pour éviter ces effets indésirables systémique, une forme locale (budésonide) a été développée pour être libérée au niveau de l'iléon terminal et du colon droit afin d'obtenir la même efficacité avec des posologies plus faibles.

## 2) Les médicaments immuno-modulateurs

Dans les formes réfractaires de la MC ou dans les cas de cortico-dépendance, un traitement par immuno-modulateurs (azathioprine, 6-mercaptopurine et méthotrexate) peut être mis en place en traitement d'entretien pour limiter les poussées. Leur activité immuno-modulatrice repose sur l'inhibition de la prolifération et de l'activation des lymphocytes T (Triantafillidis *et al.*, 2011). Ces traitements restent réservés aux patients stéroïdes résistants ou stéroïdes dépendants et nécessitent une surveillance biologique régulière, de par leurs effets

indésirables fréquents et graves (toxicité, induction de maladies opportunistes, augmentation du risque de lymphome).

### a. Azathioprine et 6-mercaptopurine

L'azathioprine et son métabolite actif, la 6-mercaptopurine, sont très efficaces dans le maintien de la rémission de la maladie, ainsi que pour le sevrage des MC cortico-dépendantes (Brooke *et al.*, 1969; Louis *et al.*, 2014). L'efficacité de ce traitement à long terme est bonne avec un taux de rechute d'environ 5 à 10 % par an, hormis dans les formes fistulisantes. Mais les effets indésirables de ces médicaments sont malheureusement fréquents et graves. Ils peuvent être immuno-allergiques (fièvre, hépatite, pancréatite…), ou plus rarement hématologiques (leucopénie, thrombopénie), nécessitant dans ce cas l'arrêt du traitement. De plus, l'inhibition des voies de dégradation du médicament (pour 10% de la population déficitaire en thiopurine méthyl transférase) ainsi que la prise simultanée d'allopurinol accentuent la toxicité médullaire du médicament et imposent un arrêt du traitement.

### b. Méthotrexate

Le méthotrexate est un analogue de l'acide folique. Son administration se fait par voie intramusculaire ou sous-cutanée car son absorption intestinale est trop variable d'un individu à l'autre. Le méthotrexate conduit à une amélioration de 80 % des formes cortico-dépendantes de MC, avec une rémission dans 40 à 50 % des cas (Feagan *et al.*, 1995; Fraser, 2003). Ce traitement est utilisé chez les patients qui ne tolèrent pas l'azathioprine, mais il possède également de nombreux effets indésirables nécessitant une surveillance de la numération formule sanguine et du bilan hépatique toutes les 6 à 8 semaines, ainsi qu'une contraception rigoureuse. Ses effets indésirables peuvent être réduits par la prise concomitante d'acide folique.

### 3) Les anti-TNF-α

Des thérapies plus ciblées basées sur des anticorps monoclonaux anti-TNF-α s'avèrent efficaces chez les patients avec une atteinte sévère de la MC ne répondant pas aux traitements conventionnels (anti-inflammatoires ou immuno-modulateurs). Ces anticorps comprennent l'infliximab, l'adalimumab et le certolizumab et sont indiqués lors des poussées et/ou en traitement d'entretien (Sandborn and Hanauer, 2002; Sandborn, 2007).

Le TNF-α est une cytokine pro-inflammatoire qui exerce un rôle majeur dans l'inflammation en général, et dont la concentration est augmentée dans la muqueuse intestinale et dans les selles des patients atteints de MC (Gross *et al.*, 1991). Le TNF-α est associé à la présence de granulomes et joue un rôle central dans l'expansion des lymphocytes T (Breese *et al.*, 1994; Glasser *et al.*, 2001).

Les anti-TNF-α peuvent donc être utilisés sur des patients atteints de MC et réfractaires et/ou intolérants aux autres traitements (corticoïdes, immuno-modulateurs). Une réponse clinique est atteinte en environ 10 semaines pour 65 % des patients. Ils sont également très efficaces dans les formes fistulisantes de la MC (D'Haens *et al.*, 1999; Present, 1999; Rutgeerts *et al.*, 1999; Seminerio *et al.*, 2012).

Les principaux effets indésirables de ces médicaments sont liés à l'immunogénicité des anticorps avec une réaction immédiate et/ou retardée à la perfusion (bouffées de chaleur, rash, prurit, fièvre, maux de tête…) ainsi que l'apparition d'anticorps anti-médicaments nécessitant parfois l'arrêt du traitement (Baert *et al.*, 2003). Un traitement préalable ou concomitant par immunosuppresseurs peut engendrer une meilleure réponse au traitement tout en limitant l'apparition de ces anticorps anti-médicament (Colombel *et al.*, 2010). Le suivi d'apparition d'anticorps anti-médicaments peut également être utilisé pour le suivi du traitement ou le switch plus rapide vers une alternative thérapeutique (Lewis, 2011). Un risque plus élevé d'infections respiratoires et/ou urinaires ou la réactivation possible d'une tuberculose latente sont également à prendre en compte chez les patients traités (Hommes *et al.*, 2006).

### 4) La chirurgie

Quand l'utilisation de médicaments s'avère insuffisante pour traiter les symptômes, la chirurgie peut s'avérer nécessaire. La résection de tout ou partie de l'intestin ulcéré s'avère

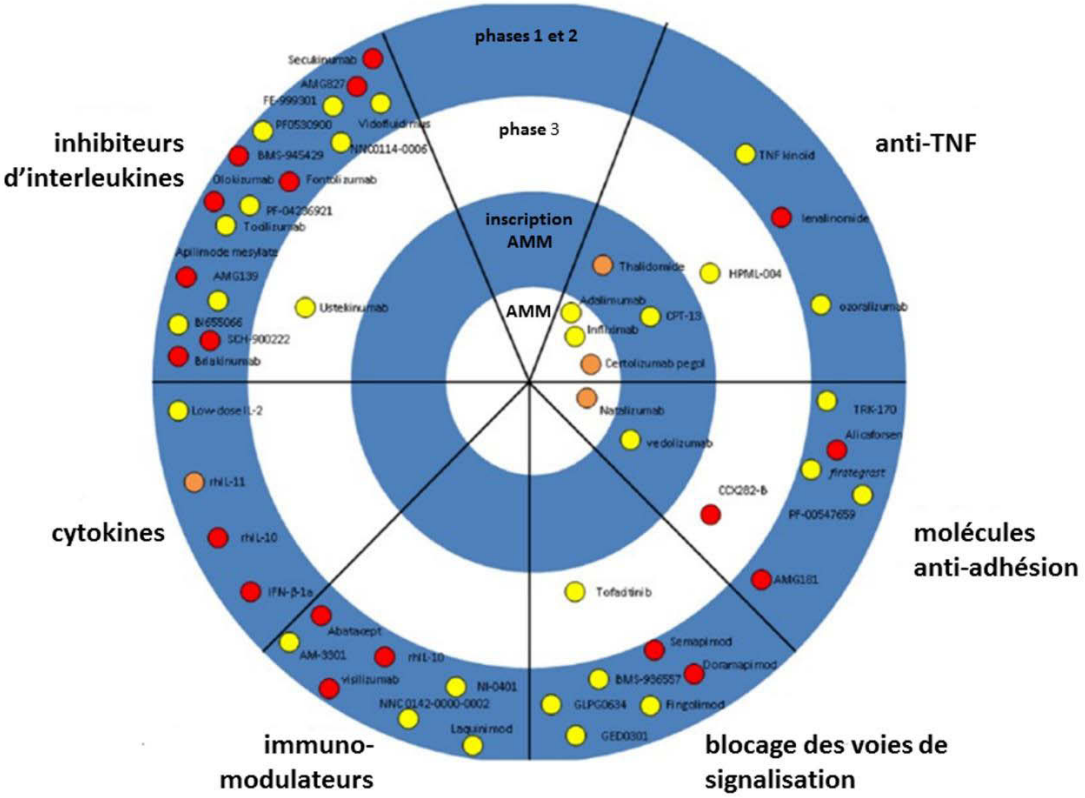



**Figure 20** : Nouvelles thérapies en cours d'essai clinique ou venant d'être commercialisées, en fonction de leur mécanisme d'action et de leur voie d'administration (rh : recombinant human ; IL : interleukine ; AMM : autorisation de mise sur le marché) (image modifiée d'après Amiot and Peyrin-Biroulet, 2015)

inévitable en cas d'évolution défavorable de la maladie. Elle consiste à enlever les parties ulcérées du tube digestif et recoudre les parties saines entre elles. Mais parfois, l'intestin peut être provisoirement abouché à la peau au niveau de l'abdomen. Cette stomie est alors utilisée pour collecter le contenu de l'intestin dans une poche jusqu'à ce que la continuité de l'intestin soit rétablie après cicatrisation des zones opérées.

Les poussées évolutives de la MC peuvent également conduire à des complications (occlusions, perforations, abcès, nécroses, fistules…) qui nécessiteront une intervention chirurgicale en urgence.

Dans les 2 ans suivant le diagnostic de la maladie, environ 45 % des patients vont subir une intervention chirurgicale, et plus de 90 % vont subir une résection partielle de l'intestin au moins une fois au cours de leur vie. Il est à noter que l'exérèse des segments atteints ne met pas à l'abri des récidives sur l'intestin restant (Lee, 1984).

### 5) Les nouvelles biothérapies

Aujourd'hui, les stratégies thérapeutiques pour lutter contre la MC portent sur des thérapies ciblées modulant l'expression des cytokines impliquées dans la réponse inflammatoire de la MC (**Figure 20**) (Bandzar *et al.*, 2013; Amiot and Peyrin-Biroulet, 2015).

Un anticorps monoclonal dirigé contre des protéines de la famille des intégrines, le natalizumab, a montré des résultats prometteurs en empêchant la migration des cellules immunitaires circulantes à travers l'endothélium vasculaire (Ghosh *et al.*, 2003).

Des essais cliniques sont également en cours pour tester certains anticorps dirigés contre d'autres cytokines pro-inflammatoires comme l'Ustekinumab, dirigé contre la sous-unité p40 commune à IL-12 et IL-23. L'expression d'IL-12 dans l'intestin des patients atteints de MC est corrélée à celle de TNF-α et d'interféron- γ (IFN-γ). L'IL-23 est également augmentée dans la MC et induit l'expression d'autres cytokines (IL-6 et IL-17) orientant les lymphocytes T vers un phénotype Th17 (Peluso *et al.*, 2006; Kobayashi *et al.*, 2008; Toussirot, 2012). L'Ustekinumab montre une réponse des patients atteints de la MC dans environ 50 % des cas (Simon *et al.*, 2016).

Par ailleurs, les cellules souches sont également une voie de recherche intéressante pour le traitement de la MC. En effet, leur potentiel pluripotent leur alloue un fort pouvoir de régénération des tissus (cellules épithéliales, cellules immunitaires). Cette capacité permettrait une restauration de l'épithélium intestinal ainsi qu'un rééquilibrage de la balance immunitaire chez les patients. Plusieurs essais cliniques ont déjà montré une action bénéfique des cellules souches mésenchymateuses sur la réparation tissulaire intestinale, ouvrant la voie à une utilisation potentielle pour induire une rémission à long terme de la MC (Ciccocioppo *et al.*, 2011; Liew *et al.*, 2014).

La transplantation de cellules souches hématopoïétiques chez les patients de MC a également été testée, mais elle reste réservée aux seuls patients avec une MC sévère réfractaire aux traitements médicaux et dont la chirurgie n'est pas envisageable de par l'étendue des lésions. Une autogreffe de cellules souches hématopoïétiques peut alors être le dernier recours afin de restaurer le système immunitaire. Cette solution permet une rémission clinique et endoscopique de la maladie avec cependant un effet probablement transitoire (Kreisel *et al.*, 2003; Clerici *et al.*, 2011; Hommes *et al.*, 2011).

## 6) Les traitements actifs sur le microbiote intestinal

Le microbiote intestinal joue un rôle central dans la MC (voir § III.6. Etiologies de la MC). D'autres stratégies alternatives à la chirurgie et aux traitements anti-inflammatoires et immuno-modulateurs ont donc été développées, comme l'utilisation d'antibiotiques et/ou de probiotiques (Prantera and Scribano, 2009) et plus récemment la transplantation de microbiote fécal.

### a. Les antibiotiques

Les antibiotiques sont utiles pour diminuer le nombre de bactéries associées à la muqueuse intestinale et ainsi réduire le taux de translocation bactérienne. C'est pourquoi des antibiotiques à large spectre sont fréquemment utilisés avec succès dans le traitement symptomatique de la MC (K. J. Khan *et al.*, 2011; Scribano and Prantera, 2013).

Le métronidazole, par exemple, est efficace contre certaines atteintes coliques ou fistules péri-anales, mais avec des doses très importantes et donc potentiellement neurotoxiques (Ursing and Kamme, 1975; Blichfeldt *et al.*, 1978; Bernstein *et al.*, 1980; Colombel *et al.*, 2001).

La ciprofloxacine et la clarithromycine peuvent s'avérer des alternatives efficaces au métronidazole en cas d'effets indésirables de ce dernier, en association avec des traitements standards chez les patients résistants (Colombel *et al.*, 1999; Arnold *et al.*, 2002).

La combinaison de la ciprofloxacine avec la métronidazole peut être un traitement alternatif aux stéroïdes dans les phases aiguës de MC (Prantera *et al.*, 1996). Cette combinaison est aussi utilisée pour le traitement des manifestations péri-anales de la MC (Schwartz and Herdman, 2004). Il faut noter que la ciprofloxacine est un antibiotique à tropisme intracellulaire et permet donc de cibler les bactéries invasives (Subramanian *et al.*, 2008).

Des études cliniques ayant pour objectif d'étudier l'impact éventuel de la rifaximine ont montré que cet antibiotique induisait des rémissions cliniques chez 52% des patients atteints de maladie de Crohn, contre seulement 33% chez le groupe placebo (Prantera *et al.*, 2006).

### b. Les probiotiques

Plusieurs micro-organismes tels que *Escherichia coli* Nissle 1917, *Lactobacillus* GG et *Saccharomyces boulardii* ont été étudiés dans le traitement de la MC, mais leur utilisation reste à ce jour très discutée (Prantera, 2006; Prantera and Scribano, 2009; Reid *et al.*, 2010).

### c. La transplantation de microbiote fécal

Une meilleure connaissance du microbiote intestinal ainsi que des moyens efficaces pour le rééquilibrer afin d'influer sur des pathologies associées à une dysbiose comme la MC est à ce jour un enjeu important.

L'efficacité des antibiotiques dans la MC semble liée à un traitement prolongé, mais qui va souvent de pair avec de lourds effets indésirables, comme par exemple les colites à

*Clostridium difficile* causées notamment par les fluoroquinolones comme la ciprofloxacine (Scribano and Prantera, 2013).

La transplantation de microbiote fécal (TMF) consiste en l'introduction des selles d'un donneur sain dans le tube digestif d'un patient receveur afin de rééquilibrer la flore intestinale altérée de l'hôte. Elle est aujourd'hui très utilisée pour rééquilibrer le microbiote fécal après des colites récidivantes à *Clostridium difficile* (Borody and Khoruts, 2012). Une étude multicentrique récente a montré que la diarrhée se résorbe dans les 90 jours après la transplantation fécale chez 91% des patients ayant une infection à *Clostridium difficile*. L'utilisation de cette thérapeutique dans la MC semble également prometteuse (Cui *et al.*, 2015). Cependant, cette approche est pour l'instant réservée aux situations graves, lors d'un échec des traitements conventionnels et en l'absence d'autre alternative thérapeutique car l'effet bénéfique de la TMF sur des patients ayant reçu des thérapies immuno-suppressives semble plus limité (Khoruts *et al.*, 2016) et ses effets indésirables sont pour l'instant inconnus.

## 6. Etiologies de la MC

Bien que de nombreux progrès aient été faits, en particulier dans le domaine de la génétique et dans la compréhension de la physiologie, les causes de la MC ne sont à ce jour pas totalement connues.

La prévalence de la MC est plus importante en Europe et en Amérique du Nord, ce qui suggère un rôle du mode de vie dans les pays industrialisés. Le tabagisme est également un facteur favorisant de la maladie, ainsi que certaines prédispositions génétiques et/ou immunologiques. La MC a donc une origine multifactorielle, faisant intervenir à la fois des facteurs environnementaux, génétiques et immunologiques.

Le microbiote intestinal semble également jouer un rôle clé dans cette maladie. En effet, de nombreuses études suggèrent que la MC provient d'un déséquilibre entre l'hôte et son microbiote intestinal (Hooper and Macpherson, 2010; Chassaing and Darfeuille–Michaud, 2011; Gevers *et al.*, 2014). Chez les patients atteints de MC, le nombre de bactéries associées à la muqueuse est augmenté, ce qui sous-entend la présence d'une faille dans les mécanismes de défense de l'hôte qui permettent en temps normal de limiter les contacts entre l'épithélium intestinal et les bactéries du microbiote.

### 1) Facteurs environnementaux

L'influence de facteurs environnementaux dans la MC a été soulevée avec l'augmentation importante de la prévalence de la maladie en Europe et en Amérique du Nord, comparativement à l'Asie ou l'Afrique. Des études portant sur les populations migrantes ont montré que cette différence de prévalence de la MC était plutôt liée au style de vie des individus qu'à une différence raciale d'ordre génétique (Loftus Jr, 2004; Burisch *et al.*, 2014). En effet, la MC était rarement retrouvée dans les pays asiatiques, mais elle était bien retrouvée dans les mêmes populations émigrées en Europe (Yang *et al.*, 2001). Plusieurs facteurs environnementaux ont été associés à la MC, comme par exemple la cigarette ou le régime alimentaire.

Plusieurs études ont montré l'influence de la cigarette sur la survenue de la MC. Les fumeurs auraient au moins deux fois plus de risque de développer cette maladie que les non-fumeurs (Calkins, 1989). Un autre facteur de risque pour l'apparition de la MC est une appendicectomie préalable (Andersson *et al.*, 2003). L'utilisation de contraceptifs oraux semble également augmenter le risque de contracter une MC. En effet, un Odd Ratio (OR) de 1,44 pour la MC a été retrouvé chez les personnes utilisant ce moyen de contraception (Godet *et al.*, 1995).

Par ailleurs, l'apparition de la MC semble être influencée par des évènements survenus dans la petite enfance, comme par exemple le type d'allaitement (maternel ou artificiel), l'hygiène domestique et l'existence d'infections périnatales. En effet, l'allaitement maternel semble être un élément protecteur contre l'apparition de la MC (Rigas *et al.*, 1993). D'autre part, le risque d'apparition de la maladie parait plus élevé (OR = 5) chez les personnes ayant bénéficié de meilleures conditions d'hygiène dans l'enfance (installation d'eau chaude, salle de bain privative…) (Gent *et al.*, 1994). Ceci peut expliquer en partie l'augmentation de la MC dans les pays développés ces cinquante dernières années. Enfin, la survenue de complications en période périnatale (infections périnatales ou autres problèmes de santé, conditions socio-économiques défavorables) est associée à un risque plus important (3 à 4 fois plus élevé) de développer une MC (Ekbom *et al.*, 1990).

Le régime alimentaire pourrait être un des principaux facteurs environnementaux impliqués dans la survenue de la MC. En effet, le type de régime alimentaire peut expliquer la différence significative d'incidence de cette maladie entre les différentes régions du monde, de même que l'augmentation de cette incidence dans les populations migrantes. Le premier lien entre l'alimentation et la MC a été découvert il y a plus de 35 ans avec la consommation de sucres

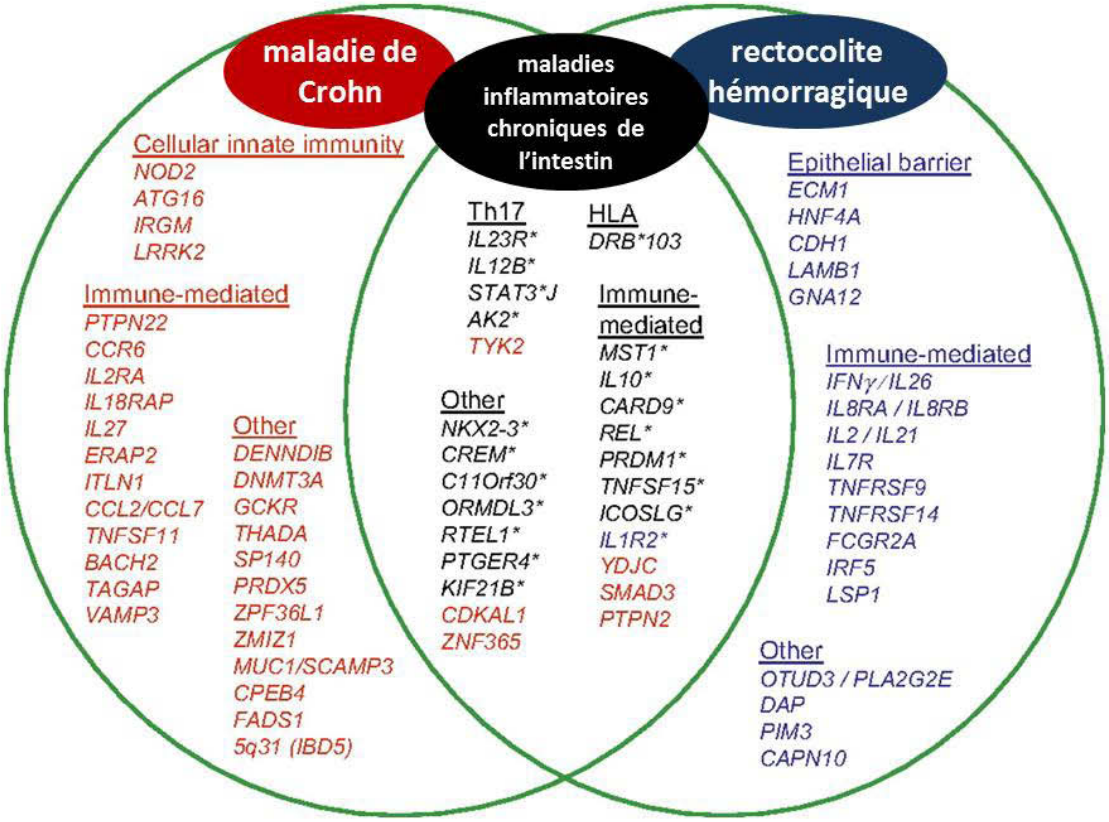



**Figure 21 :** Facteurs génétiques dans l'étiologie des maladies inflammatoires chroniques de l'intestin. Les gènes de susceptibilités significativement associés à la maladie de Crohn sont représentés en rouge, et ceux associés à la rectocolite hémorragique en bleu (image modifiée d'après Lees *et al.*, 2010).

(bonbons, gâteaux, alcool…) (Martini and Brandes, 1976). Plus récemment, il a été suggéré un rôle des FODMAP (Fermentable Oligo-, Di-, Mono-saccharides And Polyols) dans la survenue de la MC. Ces hydrates de carbone sont des dérivés sucrés peu absorbés par l'intestin grêle, et donc fermentés par le microbiote intestinal au niveau du colon. Ils sont abondants dans l'alimentation occidentale moderne (Gibson and Shepherd, 2005) et sont présents par exemple dans les céréales comme le blé ou le seigle, le lait et les fruits (pommes et poires entre autres). D'autres études épidémiologiques ont également suspecté le rôle de certains acides gras dans la survenue de la maladie (Shoda *et al.*, 1996). A l'inverse, la consommation de fibres (fruits et légumes) semble avoir un rôle protecteur (Reif *et al.*, 1997; Ananthakrishnan *et al.*, 2013).

La MC semble donc liée entre autres choses à un haut niveau d'hygiène, de nouvelles habitudes alimentaires et une industrialisation de l'alimentaire.

### 2) Facteurs génétiques

Des études épidémiologiques indiquent que le risque de développer une MC lorsqu'un parent du premier degré est atteint est 10 à 21 fois plus élevé que dans la population générale, et la fréquence de formes familiales de MC peut atteindre 40% (Küster *et al.*, 1989; Monsén *et al.*, 1991; Orholm *et al.*, 1991; Hugot *et al.*, 1994; Binder and Orholm, 1996; Halme *et al.*, 2006). De même, la concordance de la MC est d'environ 40 % chez les jumeaux monozygotes alors qu'elle est pratiquement nulle chez les jumeaux dizygotes (Tysk *et al.*, 1988; M. Orholm, 2000; Halfvarson *et al.*, 2003).

La génétique inverse utilisant des marqueurs génétiques de polymorphisme dont la localisation est connue sur le génome et plus récemment les études d'association à l'échelle du génome (GWA, *Genome-Wide Association*) ont permis l'exploration de la composante génétique de la MC avec l'identification de plus de 160 loci (**Figure 21**) (Abraham and Cho, 2009; Franke *et al.*, 2010; Jostins *et al.*, 2012; Liu *et al.*, 2015).

Les gènes de susceptibilité identifiés ont permis de mettre en lumière des processus cellulaires qui semblent jouer un rôle clé dans l'étiologie de la MC : (i) l'immunité innée avec la prise en charge de micro-organismes (récepteurs NOD et TLR, autophagie, stress du réticulum endoplasmique), la barrière mucosale et la synthèse de cytokines pro-

inflammatoires et (ii) l'immunité adaptative avec l'établissement de la réponse lymphocytaire T (Voie Th17).

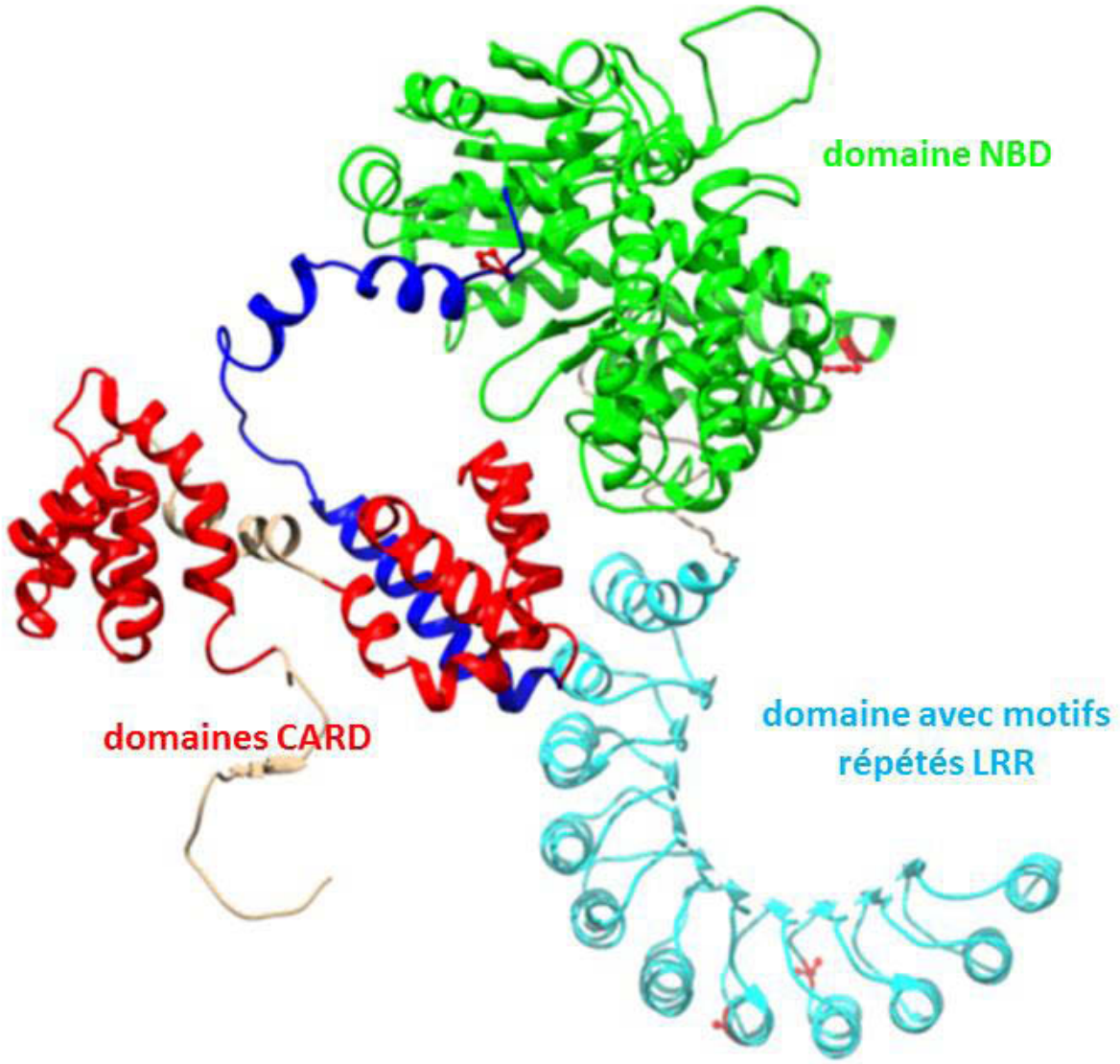



**Figure 22** : Structure 3D du récepteur NOD2
(image modifiée d'après Nakagome *et al.*, 2012)

Les gènes les plus étudiés à ce jour sont ceux codant les protéines NOD2, TLR, ATG16L1, IRGM et IL-23R.

### a. Implication du gène *NOD2*

En 2001, deux études ont permis d'établir le premier lien entre la MC et la présence de polymorphismes au niveau du gène *NOD2* (Hugot *et al.*, 2001; Ogura *et al.*, 2001). Comme détaillé au paragraphe I.4.3.a., NOD2 est un récepteur cytoplasmique de reconnaissance de motifs exprimé préférentiellement par les cellules de Paneth et par des cellules immunitaires comme les macrophages, et qui intervient dans la réponse immunitaire aux micro-organismes.

Chez les patients atteints de MC, de nombreux polymorphismes dans le gène *NOD2* ont été mis en évidence dont trois principaux qui représentent plus de 80% des variants rencontrés dans la population caucasienne associée à la MC. Ces mutations conduisent à la production de protéines NOD2 variantes par mutations faux-sens (Arg702Trp, Gly908Arg) ou à des protéines tronquées (Leu1007fsinsC) par insertion d'un codon stop (Hugot *et al.*, 2001; Ogura *et al.*, 2001; Begue *et al.*, 2006; Borzutzky *et al.*, 2010). Il a été estimé que le risque de développer une MC est 1,5 à 3 fois supérieur chez les sujets hétérozygotes, et environ 40 fois supérieur chez les sujets homozygotes pour un allèle à risque ou hétérozygotes composites (deux allèles à risque différents) par rapport à des sujets n'ayant pas ces polymorphismes, ce qui représente un des facteurs de risque les plus important de la MC connus à ce jour (Bouma and Strober, 2003). Des mutations de *NOD2* ont par la suite été associées à un phénotype particulier de la MC : la maladie iléale, ainsi qu'à une maladie plus précoce et d'évolution sténosante.

Le récepteur NOD2 est constitué de deux domaines CARD (caspase-recruitment domains) à la partie N-terminale, un domaine NBD (nucleotide-binding domain) central, ainsi qu'un domaine composé de motifs répétés LRR (leucine-rich repeat) à la partie C-terminale (**Figure 22**) (Nakagome *et al.*, 2012).

Le domaine LRR est impliqué dans la reconnaissance du MDP bactérien (Girardin *et al.*, 2003). La fixation de ce ligand induit un changement de conformation du récepteur, lui permettant de recruter la kinase RIP2, ce qui conduit à l'activation des voies de signalisation NF-κB et des MAP-kinases régulant la transcription de gènes cibles qui codent entre autres de

nombreuses cytokines pro-inflammatoires et des peptides antimicrobiens (Girardin *et al.*, 2003; Philpott and Viala, 2004).

Les trois principaux polymorphismes du gène *NOD2* associé à la MC sont situés au niveau du domaine de fixation du ligand LRR, et des études *in vitro* ont montré une diminution de l'activation de la voie de signalisation NF-κB par ces trois protéines NOD2 variantes comparativement à la protéine sauvage suite à une stimulation par du MDP (Girardin *et al.*, 2003). Des souris invalidées pour le gène *NOD2* présentent un développement anormal des plaques de Peyer, porte d'entrée pour les bactéries entéro-invasives, ainsi qu'une augmentation de la perméabilité intestinale et une réponse immunitaire mucosale exagérée, parfaitement en concordance avec la physiopathologie de la MC (Barreau *et al.*, 2007).

Enfin, chez les patients atteints d'une MC iléale, et plus particulièrement chez ceux porteurs d'un variant du gène *NOD2*, il a été retrouvé une diminution de la synthèse des α-défensines par les cellules de Paneth, sans diminution des autres peptides antimicrobiens (Wehkamp *et al.*, 2004).

### b. Implication du gène *SLC15A1*

Le transporteur PepT1 (peptide transporter 1), codé par le gène *SLC15A1*, est impliqué dans le transport de produits bactériens tels que le MDP à l'intérieur des cellules épithéliales intestinales et des cellules immunitaires. Ce transport induit l'activation de la voie de signalisation NF-κB passant par le récepteur NOD2 et la mise en place d'une inflammation (Charrier and Merlin, 2006; Dalmasso *et al.*, 2010). Des patients atteints de la MC présentent un polymorphisme du gène *SLC15A1* qui induit une augmentation de l'activité du transporteur PepT1, à laquelle peut se surajouter une augmentation de l'expression de ce transporteur (Merlin *et al.*, 2001; Zucchelli *et al.*, 2009). Il a également été démontré que la surexpression de PepT1 exacerbe les colites induites en modèle murin d'une manière NOD2 dépendante (Dalmasso *et al.*, 2011). L'ensemble de ces données suggère que le transporteur intestinal PepT1 pourrait jouer un rôle important dans la survenue de l'inflammation intestinale chez les patients atteints de MC *via* un mécanisme passant par la voie de signalisation NOD2.

### c. Implication des gènes codant des protéines de l'autophagie

L'autophagie est un processus ubiquitaire des cellules eucaryotes qui correspond à la dégradation d'éléments cytoplasmiques divers (mitochondries, peroxysomes, agrégats protéiques, protéines mal repliées…) par les lysosomes. Ce processus se déroule en plusieurs grandes étapes : (i) l'initiation de l'autophagie, (ii) la nucléation, c'est-à-dire la mise en place d'une membrane d'isolation des éléments à détruire (phagosome), (iii) l'élongation de la membrane, (iv) la fusion du phagosome avec un lysosome pour former un autophagolysosome, et enfin (v) la dégradation du contenu intraluminal et de la membrane interne suite à l'acidification de la vacuole et l'action des enzymes lysosomiales telles que les cathepsines (**Figure 23**) (Eskelinen, 2005; Mizushima, 2007; Marzban *et al.*, 2015).

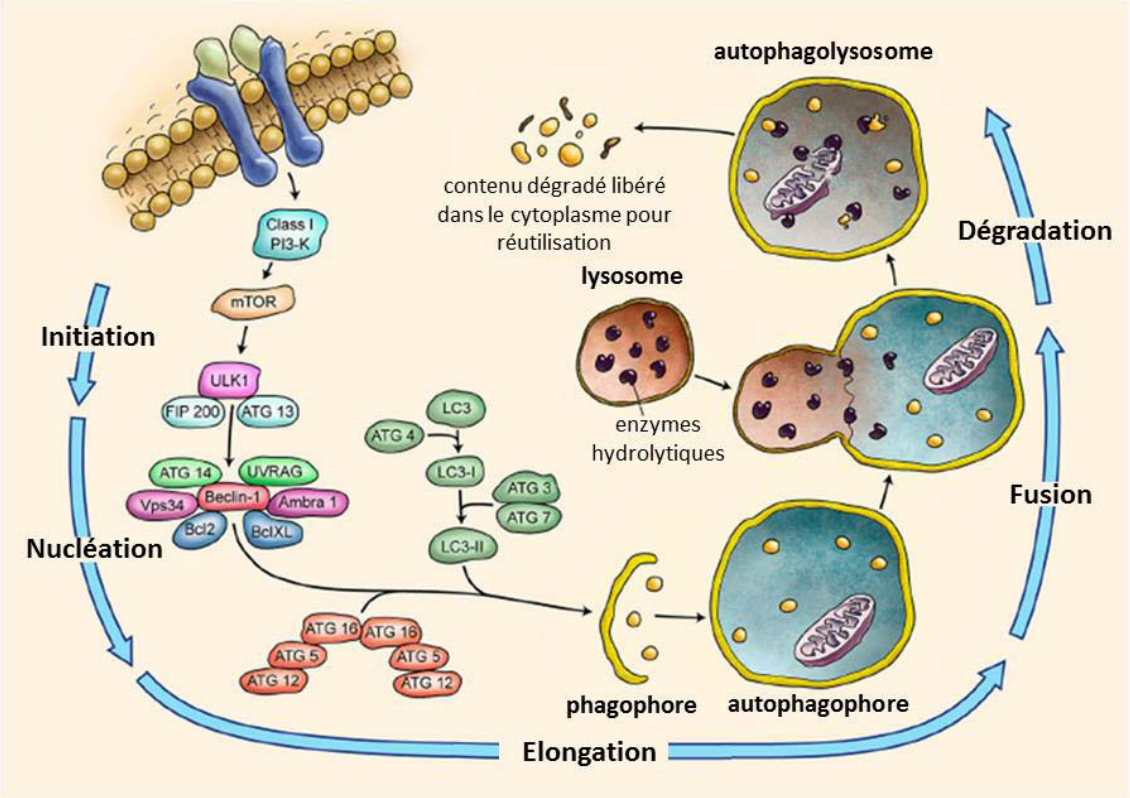



**Figure 23** : Représentation schématique de la voie de l'autophagie
(image modifiée d'après Marzban *et al.*, 2015)

L'autophagie est un processus de recyclage nécessaire au maintien de l'homéostasie cellulaire et au renouvellement des organites cellulaires. De plus, elle joue un rôle dans la réponse immunitaire de l'hôte vis-à-vis des micro-organismes pathogènes, et pourrait ainsi jouer également un rôle important dans la MC (Nakagawa *et al.*, 2004; Gutierrez *et al.*, 2004; Amer *et al.*, 2005; Deretic, 2005; Singh *et al.*, 2006; Levine and Deretic, 2007).

Les récentes études d'analyse du génome ont permis d'associer deux gènes codant des protéines directement impliquées dans l'autophagie à la MC : la protéine ATG16L1 (autophagy-related 16-like 1) et la protéine IRGM (immunity-related GTPase family M) (Nguyen *et al.*, 2013).

### *Le gène ATG16L1*

ATG16L1 joue un rôle fondamental dans l'autophagie car sa présence est indispensable pour l'élongation de la membrane du phagosome. Il existe un polymorphisme du gène *ATG16L1* (Thr300Ala) dont l'association avec la MC a été mise en évidence en 2007 (Hampe *et al.*, 2007) et a été confirmée depuis par au moins 28 études d'association sur des populations de types caucasiennes (Brest *et al.*, 2010). Ce gène est plus particulièrement associé à la forme iléale de la MC (Rioux *et al.*, 2007; Wellcome Trust Case Control Consortium, 2007; Parkes *et al.*, 2007; Barrett *et al.*, 2008). Cette association suggère un rôle de l'autophagie dans la forme la plus fréquente de la MC, qui est également associée à un polymorphisme de *NOD2* et la présence anormale d'*Escherichia coli* adhérents et invasifs (AIEC) (cf § IV). Il a également été démontré l'induction du processus autophagique en réponse à l'infection par des AIEC associés à la MC, ainsi qu'une survie accrue des AIEC dans les cellules lors d'un défaut d'expression du gène *ATG16L1* (Lapaquette *et al.*, 2012). De plus, la protéine variante d'ATG16L1 retrouvée associée à la MC perd sa capacité à être recrutée à la membrane par la protéine NOD2 lors d'une infection bactérienne (Cooney *et al.*, 2010; Travassos *et al.*, 2010; Homer *et al.*, 2010). Un allèle à risque du gène *NOD2* ou *ATG16L1* conduit à l'altération de la même voie cellulaire : la prise en charge des bactéries intracellulaires par l'autophagie. L'altération de la protéine ATG16L1 conduit également à une altération de sécrétion de peptides antimicrobiens par les cellules de Paneth (Cadwell *et al.*, 2008).

### *Le gène IRGM*

Le second gène associé à la MC et impliqué dans le processus autophagique code une guanosine triphosphatase appelée IRGM. Il a également été identifié en 2007, avec une association préférentielle à la MC iléale (Wellcome Trust Case Control Consortium, 2007; Parkes *et al.*, 2007).

Le rôle de la protéine IRGM est encore flou. Une altération d'IRGM diminue la proportion de bactéries prises en charge par l'autophagie suggérant un rôle dans le ciblage des bactéries intracellulaires (Singh *et al.*, 2006; McCarroll *et al.*, 2008).

Initialement, la MC est associée à un polymorphisme nucléotidique (SNP) silencieux dans la région codante du gène *IRGM* (C313T), donc n'entrainant pas de changement d'acide aminé au niveau de la protéine. Cependant ce SNP a ensuite été associé dans tous les cas à une délétion de 20kb en amont du gène *IRGM* qui pourrait interférer avec la fixation d'éventuels facteurs de transcription et ainsi altérer l'expression du gène (McCarroll *et al.*, 2008; Prescott *et al.*, 2010). Plus récemment, il a été également découvert l'existence de microARN (miR-196A et B) surexprimés chez les patients atteints de MC, ayant pour cible l'ARNm d'*IRGM*. Chez les patients porteurs de l'allèle à risque 313T d'*IRGM*, le miR-196 ne peut plus se fixer, et cela conduit à la perte de la régulation de l'expression de la protéine IRGM, et donc de l'autophagie. Cette étude a également montré que l'allèle à risque 313T d'*IRGM* conduit à une persistance accrue des bactéries AIEC intracellulaires (Brest *et al.*, 2011).

## d. Implication du gène NLRP3

La protéine NLRP3 (NOD like receptor family, pyrin domain contining 3) est une protéine intra-cytoplasmique appartenant à la famille des NLR. Elle se compose d'un domaine central de liaison aux nucléotides et un domaine C-terminal constitué de motifs répétés riches en leucine (LRR). Cette protéine participe à la formation d'un complexe oligomérique « l'inflammasome ». Ce dernier est constitué de multiples copies de NLRP3 liées grâce à son domaine pyrine à une protéine adaptatrice ASC (Apotosis-associated speck-like protein containing a caspase recrutment domain) et à la caspase-1. L'activation de l'inflammasome-NLRP3 conduit au clivage protéolytique de la pro-caspase-1 en caspase-1 qui induit à son

tour la maturation des pro-interleukines IL-1β et IL-18 en cytokines pro-inflammatoires actives (Zhong *et al.*, 2013).

L'inflammasome-NLRP3 joue un rôle important dans le maintien de l'intégrité de l'épithélium digestif et la défense contre les agents pathogènes. Il est d'ailleurs exprimé par les monocytes, les macrophages, les neutrophiles, les cellules dendritiques ainsi que certaines cellules épithéliales. Une étude a mis en évidence des variants génétiques de *NLRP3* qui seraient associés à la MC et qui conduiraient à une diminution de l'expression de NLRP3 (Villani *et al.*, 2009).

Le rôle exact de l'inflammasome-NLRP3 dans la pathogénèse de la MC reste à ce jour inconnu. Cependant, il semblerait que les mutations affectant la région non codante du gène *NLRP3* induisent une diminution de l'activité antimicrobienne et une augmentation de la perméabilité de la muqueuse digestive. Cela induirait également un défaut de régénération des tissus (Zaki *et al.*, 2011).

### e. IL23-R

L'inflammation de la MC implique la cytokine pro-inflammatoire IL-12, cytokine hétérodimérique comprenant les sous-unités IL-12p40 et IL-12p35 (Bouma and Strober, 2003). La sous-unité IL-12p40 peut également se combiner avec la sous-unité IL-23p19 pour former l'IL-23 (Oppmann *et al.*, 2000). Des études cliniques ont mis en évidence plusieurs polymorphismes dans le gène *IL23R*, qui code une sous-unité du récepteur à l'IL-23 (Duerr *et al.*, 2006; Wellcome Trust Case Control Consortium, 2007; Parkes *et al.*, 2007). La sous-unité p40 et le facteur de transcription STAT3 (impliqué dans la transduction des signaux en aval d'IL-23R) présentent également des polymorphismes associés à la MC (Barrett *et al.*, 2008). De plus, il semble exister un risque encore supplémentaire de développer la MC chez les personnes cumulant plusieurs allèles à risque dans les gènes de cette voie, indiquant des interactions possibles entre ces gènes de susceptibilité (McGovern *et al.*, 2009). Les activités pro-inflammatoires de l'IL-23 sont supportées par la sous-population de lymphocytes T CD4+ Th17, produisant de l'IL-17, de l'IL-6 et du TNF-α (Bettelli and Kuchroo, 2005). L'IL-17 induit à son tour la production de peptides antimicrobiens et le maintien de l'intégrité de la barrière intestinale. L'association de plusieurs gènes de la voie IL-23/IL-17 avec la MC est un argument fort en faveur d'un rôle clé de ces médiateurs de la réponse immunitaire dans la génération de l'inflammation exacerbée observée dans la maladie, ce qui a conduit au

développement d'un anticorps monoclonal dirigé contre la sous unité p40 commune à IL-12 et IL-23 (Ustekinumab) dans le traitement de la MC (cf §III.5.5.).

### f. Les TLR

Les récepteurs TLR, principalement localisés à la surface des monocytes et des macrophages, jouent un rôle majeur dans la détection des bactéries et dans l'initiation des réponses de l'immunité innée (cf §I.4.3.a) (Cario, 2005; Kawai and Akira, 2006). Leur activation par des bactéries ou des produits bactériens entraîne des cascades de voies de signalisation conduisant à une augmentation de l'expression de gènes codant des cytokines pro-inflammatoires.

Une augmentation de l'expression à la surface des cellules épithéliales intestinales de TLR4, récepteur activé par le LPS bactérien, a été observée lors des MICI (Cario and Podolsky, 2000). Par ailleurs, deux polymorphismes de TLR4 ont été mis en évidence, constituant un facteur de risque mineur pour la MC (Franchimont *et al.*, 2004; Brand *et al.*, 2005). Il a également été observé la présence d'un polymorphisme du promoteur du récepteur TLR9, activé par des motifs CpG de l'ADN bactérien (Török *et al.*, 2004).

Ces mutations ne semblent pas avoir d'impact majeur, mais l'interaction de plusieurs susceptibilités génétiques peut conduire à une augmentation de susceptibilité à la MC.

## 3) Facteurs immunologiques

La MC est caractérisée par un état d'hyperactivation du système immunitaire qui pourrait être la conséquence d'une dérégulation de la réponse immunitaire au niveau de la muqueuse intestinale de l'hôte. En effet, des défauts dans la barrière intestinale et les interactions entre microbiote et cellules de l'hôte qui en découlent sont probablement à l'origine de la maladie.

L'inflammation intestinale se traduit par un recrutement et une activation des lymphocytes et des macrophages. La libération consécutive de cytokines solubles et d'autres médiateurs de l'inflammation conduit à l'amplification de la réponse immunitaire locale, responsable des lésions tissulaires. Les cytokines pro-inflammatoires telles que l'IL-1, l'IL-6, l'IL-8 et le TNF-α sont libérées en grande quantité dans la muqueuse intestinale des patients atteints de

MC, consécutivement à l'activation du facteur de transcription NF-κB (Niessner and Volk, 1995; Desreumaux *et al.*, 1997; Schreiber *et al.*, 1998; Ellis *et al.*, 2014). Ces cytokines jouent également un rôle clé dans la mise en place et le maintien de l'inflammation intestinale car elles participent au recrutement des lymphocytes T CD4+ naïfs et à leur différenciation en lymphocytes T helpers (Th1, Th2 ou Th17). La forte concentration des cytokines TNF-α, IFN-γ, IL-1β, IL-6, IL-23 durant la phase chronique de la pathologie laissent supposer que la MC est plutôt une pathologie de type Th1/Th17 (Sartor, 2006). Cependant, une réponse Th2 dans la phase précoce de la maladie a également été mise en évidence du fait de la détection de taux élevés d'IL4 (Desreumaux *et al.*, 1997).

### 4) Autres facteurs de l'hôte

#### a. Le stress du réticulum endoplasmique

Dans certaines conditions cellulaires, la quantité de protéines de conformation anormale dans le réticulum endoplasmique (RE) augmente si fortement que cela sature l'équipement cellulaire en protéines chaperonnes et en protéines de dégradation ; les protéines incorrectes ont alors tendance à s'accumuler, plaçant les cellules dans une situation de « stress du réticulum endoplasmique » (Garlatti and Barouki, 2002).

Une réponse se met alors en place avec l'activation de trois protéines (IRE1, PERK et ATF6) qui vont inhiber la machinerie traductionnelle et induire la translocation de facteurs de transcription comme XBP1 (X Box binding protein-1), induisant l'expression de gènes codant des enzymes de dégradation (Wu and Kaufman, 2006). L'absence de XBP1 empêche donc une réponse correcte au stress du RE et mène à une forte activation de la kinase JNK (c-jun N-terminal kinase) conduisant à la transcription de gènes pro-inflammatoires, un dysfonctionnement des cellules de Paneth associé à une perte de l'activité bactéricide au fond des cryptes, de même qu'une autophagie accrue (Kaser *et al.*, 2008; Hetz *et al.*, 2009). Tous ces éléments conduisent à des colites inflammatoires spontanées. Des polymorphismes du gène *XBP1*, responsables d'une activité réduite de la protéine, semblent avoir été retrouvés associés à la MC (Kaser *et al.*, 2008).

Par ailleurs, une autre étude a montré l'expression anormale d'une protéine de stress du RE, Gp96, au niveau de l'épithélium iléal de patients atteints de MC (Rolhion *et al.*, 2010). Gp96 est une glycoprotéine très abondante du RE, responsable de la maturation des TLRs et de certaines intégrines (Randow and Seed, 2001). Elle semble également être un facteur d'internalisation important des *E. coli* AIEC associés à la MC (Rolhion *et al.*, 2010).

Une altération du stress du RE peut ainsi conduire à une susceptibilité accrue des patients atteints de MC à être colonisés par des bactéries.

### b. Le récepteur CEACAM6

Le récepteur CEACAM6 (carcinoembryonic antigen-related cell adhesion molecule 6) est anormalement exprimé dans la muqueuse iléale de patients atteints de MC, au niveau de la bordure en brosse des entérocytes (Barnich *et al.*, 2007). Cette expression anormale est également retrouvée au niveau des zones saines de la muqueuse iléale des patients. Au niveau colique, l'expression de *CEACAM6* n'est pas différente entre les patients atteints de MC et les personnes saines, ce qui suggère une association de CEACAM6 avec la forme iléale de la maladie. Ce récepteur CEACAM6 reconnait *via* ses résidus mannose les pili de type 1 des *E. coli* AIEC retrouvés chez les patients atteints de la MC (Barnich *et al.*, 2007). Chez des souris CEABAC10 porteuses du gène *CEACAM6* humain, les AIEC sont capables de persister au niveau intestinal et provoquent une colite sévère (Carvalho *et al.*, 2009). La cause de l'expression anormale de ce récepteur n'est pas véritablement connue mais il a été montré une surexpression de ce récepteur suite à une stimulation par des cytokines pro-inflammatoires ainsi qu'une infection par des bactéries AIEC. Ainsi, les AIEC seraient capables de promouvoir leur propre colonisation chez les patients atteints de MC *via* l'expression de CEACAM6 au niveau iléal (Barnich *et al.*, 2007).

### c. Les microARN régulateurs

Les microARN (miARN) sont des acides ribonucléiques (ARN) non codants d'environ 20 nucléotides, dont la fonction est de réguler l'expression de gènes de manière post-transcriptionnelle. Ces miARN reconnaissent la partie 3'UTR des ARN messagers (ARNm) cibles, induisant leur répression traductionnelle ou leur destruction. De nombreuses études ont

montré le rôle régulateur majeur de ces miARN dans de nombreux processus physiologiques (croissance et différenciation cellulaire, métabolisme, apoptose…). De plus, l'expression anormale de certains miARN a été associée à des pathologies comme le cancer (Roush and Slack, 2008; Palanichamy and Rao, 2014; Rupaimoole *et al.*, 2016).

Plusieurs études ont montré une expression modifiée de certains miARN chez des patients atteints de MC (Kalla *et al.*, 2015). Certains de ces miARN régulent des protéines dont le rôle a déjà été démontré dans la MC, comme par exemple NOD2, ATG16L1, IRGM et IL-23…(Brest *et al.*, 2011; Brain *et al.*, 2013; Chuang *et al.*, 2014; Nguyen *et al.*, 2014). Ainsi, l'étude de l'expression anormale des miARN et de leurs impacts fonctionnels dans la physiopathologie de la MC ouvre de nouvelles voies dans la compréhension de cette maladie.

### d. La glycosylation des mucines

Certaines variations dans la quantité et/ou la qualité de la couche de mucus pourraient également compromettre sa fonction protectrice et ainsi jouer un rôle important dans la MC. En effet, il n'a pas été retrouvé de différence d'expression de gènes codant les mucines au niveau de la muqueuse non enflammée des patients atteints de MC (Buisine *et al.*, 2001), mais il y a un certain nombre d'anomalies de glycosylation sur ces mucines, avec notamment une diminution de O-glycosylations (Bodger *et al.*, 2006; Rhodes *et al.*, 2008). Des études génétiques ont en effet montré des *loci* de susceptibilité sur des gènes impliqués dans les voies de glycosylation chez les patients atteints de MC (Theodoratou *et al.*, 2014), comme par exemple un SNP (single nucleotide polymorphism) sur le gène *FUT2* codant une α(1,2)fucosyltransférase nécessaire à l'ajout d'un résidu α(1,2)fucose sur des mucines (McGovern *et al.*, 2010).

### e. Les acides biliaires

Des modifications de la composition des acides biliaires ont été reportées dans les maladies inflammatoires chroniques avec une diminution des acides biliaires secondaires présents dans les selles associée à une augmentation des acides biliaires conjugués et sulfatés (Duboc *et al.*,

2013; Dior *et al.*, 2016). Plus spécifiquement, l'analyse de la bile issue du duodénum de patients atteints de MC révèle une augmentation de la concentration des acides biliaires primaires (Marks et al., 1977; Lapidus and Einarsson, 1998; Pereira et al., 2003; Lapidus et al., 2006) et une diminution des acides biliaires secondaires comme l'acide désoxycholique (Rutgeerts *et al.*, 1986). Compte tenu de la place centrale du microbiote dans la biotransformation des acides biliaires, le dysmétabolisme des acides biliaires observé chez ces patients pourrait être relié à des modications du microbiote.

### 5) Rôle du microbiote intestinal

Comme nous venons de le voir, différents facteurs impliqués dans la survenue de la MC peuvent être reliés au microbiote intestinal qu'ils soient génétiques, immunologiques ou environnementaux. De plus, plusieurs éléments suggèrent que des micro-organismes pourraient jouer un rôle clé dans l'initiation et/ou le maintien de l'inflammation : (i) les lésions engendrées présentent des caractéristiques anatomo-pathologiques similaires à celles qui sont associées à des infections bactériennes intestinales (ulcérations, abcès, fissures, fistules, granulomes) (Sartor, 1995) ; (ii) l'utilisation d'antibiotiques s'avère parfois efficace dans le traitement de la MC (K. J. Khan *et al.*, 2011) ; (iii) des taux élevés d'anticorps dirigés contre au moins un épitope microbien (ASCA, OmpC, I2 ou flagelline) sont mis en évidence chez 80% des patients atteints de MC (Quinton *et al.*, 1998; Sutton *et al.*, 2000; Landers *et al.*, 2002; Mow *et al.*, 2004; Targan *et al.*, 2005).

Différentes théories, qui ne s'excluent pas mutuellement, ont été élaborées sur le rôle des micro-organismes dans la MC : (i) une translocation excessive de micro-organismes, (ii) une dysbiose et/ou (iii) un agent pathogène persistant au niveau intestinal.

#### a. Présence d'une translocation excessive

La translocation bactérienne se caractérise par le passage de bactéries de la lumière intestinale vers les nodules mésentériques et les organes internes. Elle se retrouve augmentée chez les patients atteints de MC. Ainsi certaines bactéries ou antigènes bactériens présents

dans la lumière intestinale, tels que *E. coli*, *Enterococcus spp*, *Clostridium perfringens*, *Proteus spp* et *Bacteroïdes fragilis*, sont retrouvés au niveau des ganglions mésentériques indiquant un franchissement de la barrière épithéliale intestinale (Laffineur *et al.*, 1992; Takesue *et al.*, 2002). Une augmentation du passage d'*E. coli* à travers la muqueuse iléale au niveau des follicules lymphoïdes et des plaques de Peyer a également été rapportée (Keita *et al.*, 2008; Salim *et al.*, 2009).

Les lésions précoces de la MC sont plutôt retrouvées au niveau des follicules lymphoïdes et des plaques de Peyer chez les sujets jeunes et des follicules du colon à tout âge (Morson, 1972; Fujimura *et al.*, 1996; Gullberg and Söderholm, 2006). Il est également à noter que des souches d'*E. coli* sont plus fréquemment isolées à partir de séreuses intestinales et de ganglions mésentériques. Ces différents éléments supposent l'implication d'une translocation bactérienne excessive dans la formation des fistules et des abcès dans la MC (Laffineur *et al.*, 1992).

### b. Présence d'une dysbiose

Il a été montré qu'au sein du microbiote des patients atteints de MC, il existe une véritable dysbiose, à la fois localisée et généralisée, avec une diminution de la diversité bactérienne globale, une diminution du nombre total de *Firmicutes* et de *Bacteroïdetes* au profit des γ-*Proteobacteria,* ainsi qu'une augmentation du nombre de bactéries associées à la muqueuse. L'analyse taxonomique des populations bactériennes retrouvées dans les selles montre une différence de populations entre les patients atteints de MC et les individus sains (Qin *et al.*, 2010).

Dans la flore fécale des patients atteints de MC, la quantité de *Bacteroïdes vulgatus* et d'entérobactéries ainsi que des bactéries mucinolytiques telles que *Ruminococcus gnavus* et *Ruminococcus torques*, est plus importante que celle des individus sains (Keighley *et al.*, 1978; Seksik *et al.*, 2003; Willing *et al.*, 2009 ; Png *et al.*, 2010; Willing *et al.*, 2010; Joossens *et al.*, 2011). Il a également été observé une diminution des *Firmicutes*, et notamment de *Clostridium leptum* (Sokol *et al.*, 2006), et une diminution des *Lactobacillus* et *Bifidobacterium* (Favier *et al.*, 1997). D'autre part, environ 30% de la flore retrouvée chez les patients appartient à des groupes phylogénétiques non présents chez des individus sains et la

composition de cette flore varie au cours du temps, contrairement aux individus sains (Seksik *et al.*, 2003).

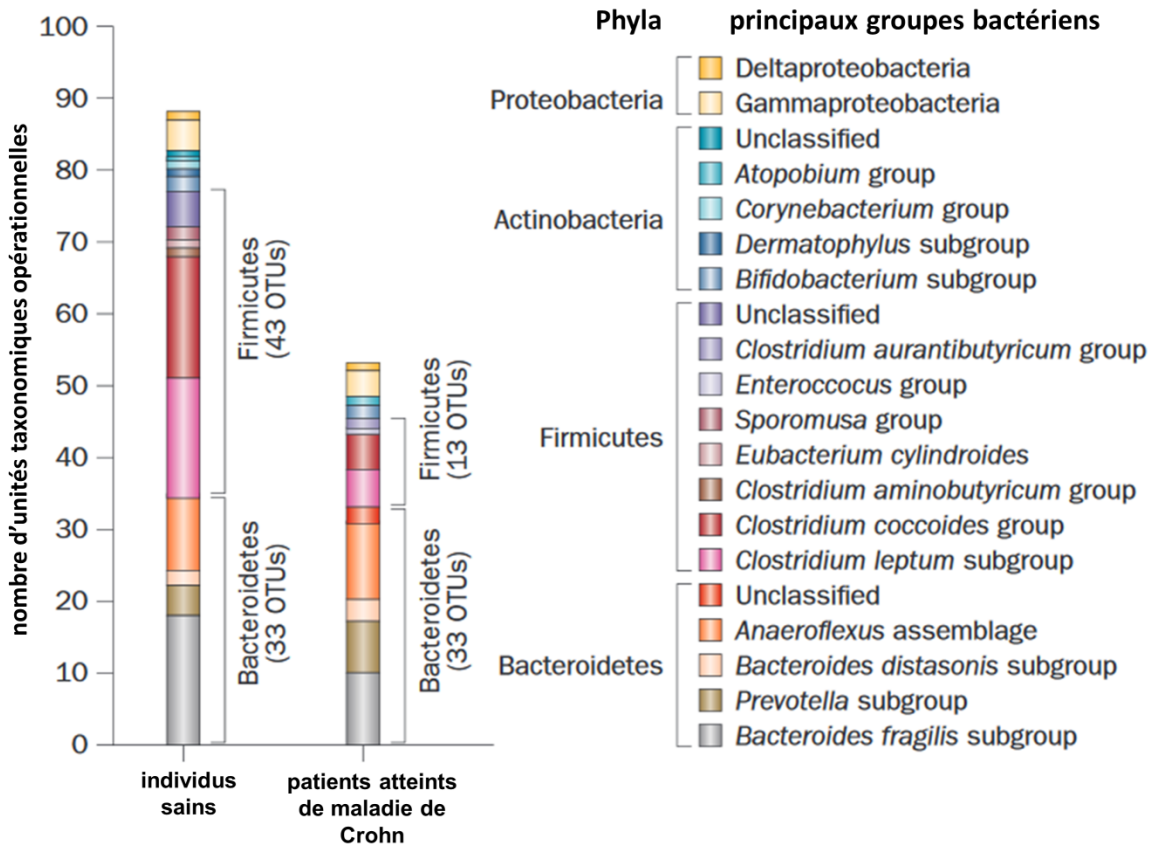



**Figure 24** : Dysbiose retrouvée dans la muqueuse des patients atteints de MC, comparativement aux individus sains (image modifiée d'après Manicahn *et al.*, 2012)

La flore associée à la muqueuse est généralement différente de la flore fécale (Lepage *et al.*, 2005). Plusieurs études ont montré que la population bactérienne totale est plus importante chez les patients, que ce soit au niveau du mucus ou au niveau des cellules épithéliales (Schultsz *et al.*, 1997; Kleessen *et al.*, 2002; Swidsinski *et al.*, 2005; Swidsinski *et al.*, 2009). Il a en outre été retrouvé une diminution de bactéries bénéfiques telles que *Bifidobacterium* ou *Lactobacillus,* associée à une augmentation de bactéries potentiellement délétères comme *Bacteroïdes* ou *E. coli* associées à la muqueuse intestinale des patients atteints de MC (**Figure 24**) (Keighley *et al.*, 1978; Giaffer *et al.*, 1992; Swidsinski *et al.*, 2002; Neut *et al.*, 2002; Martin *et al.*, 2004; Darfeuille-Michaud *et al.*, 2004; Kotlowski *et al.*, 2007; Baumgart *et al.*, 2007; Willing *et al.*, 2009; Manichanh *et al.*, 2012; Gevers *et al.*, 2014).

Contrairement aux patients contrôle, les patients atteints de MC possèdent un nombre plus important de bactéries invasives, ainsi qu'un biofilm à la surface de l'épithélium intestinal avec principalement l'espèce *Bacteroïdes fragilis* et des *Enterobacteriacaeae*. Chez les patients, ces bactéries sont capables d'adhérer aux cellules épithéliales, d'envahir les cryptes, et ont été retrouvées de temps en temps à l'intérieur des cellules (Swidsinski *et al.*, 2009).

En parallèle, des analyses de flore iléale ont montré une diminution significative de *Faecalibacterium prausnitzii*, bactérie appartenant au phylum des *Firmicutes*, productrice de butyrate et qui aurait un potentiel immuno-modulateur et anti-inflammatoire (Sokol *et al.*, 2008; Sokol *et al.*, 2009; Sokol and Seksik, 2010; Miquel *et al.*, 2013).

Comme évoqué précédemment, la présence d'une dysbiose pourrait induire un dysmétabolisme des acides biliaires et affecter l'homéostasie intestinale. Les *Firmicutes*, les *Bifidobacterium* et les *Lactobacillus* sont particulièrement impliqués dans l'activité de déconjugaison des acides primaires. De plus, les *Firmicutes* (*Clostridium, Eubacterium*) sont impliqués dans la transformation (7-déshydroxylation) des acides primaires en secondaires. Leur diminution au niveau de la muqueuse intestinale des patients atteints de MC pourrait donc être reliée à la diminution des acides biliaires secondaires observée chez ces patients. Les *Firmicutes* (*Clostridium, Faecalibacterium*) ont également une activité de désulfatation. Or, une augmentation d'acides aminés sulfatés est observée chez les patients atteints d'une MICI (Duboc *et al.*, 2013). Ces modifications pourraient avoir des répercutions importantes sur l'hôte puisque les acides biliaires secondaires sont les principaux ligands du récepteur TGR5 dont l'activation inhibe la sécrétion de TNF-α, IL-1β et d'IL-6 par les macrophages. De

plus, les acides biliaires secondaires exercent des effets anti-inflammatoires au niveau des cellules épithéliales du côlon et cet effet est perdu après sulfatation.

La dysbiose présente dans les MICI a été associée au dysmétabolisme des acides biliaires et à une aggravation de l'inflammation (Duboc *et al.*, 2013), suggérant que la dysbiose *via* le métabolisme des acides biliaires pourrait influencer l'évolution des MICI par son action sur l'inflammation intestinale. Récemment, l'abondance des gènes impliqués dans les transformations des acides biliaires a été évaluée par des analyses métagénomiques qui montrent que chez les patients atteints de MC, il existe une diminution de l'abondance des bile salt hydrolases des *Firmicutes*, enzymes responsables de la déconjugaison (Labbé *et al.*, 2014). Ces résultats sont à mettre en parallèle avec ceux de Duboc et coll. (2013) qui montrent un dysfonctionnement de cette déconjugaison en mesurant l'activité métabolique des échantillons fécaux des patients atteints de MICI. Suite à la dysbiose, le défaut de déconjugaison pourrait représenter une étape limitante de la voie de biotransformation des acides biliaires chez les patients atteints de MC.

La dysbiose constatée dans la MC ne concerne pas seulement des espèces bactériennes, mais également des espèces fongiques. Ainsi chez les patients atteints de MC, il a été observé une diminution significative de *Saccharomyces cerevisiae*, ainsi qu'une augmentation de *Candida albicans* (Sokol *et al.*, 2016).

L'ensemble de ces résultats conforte l'hypothèse d'une perturbation de la balance entre les micro-organismes bénéfiques et les micro-organismes délétères, qui pourraient contribuer à l'initiation et/ou à la chronicité de la maladie (Tamboli *et al.*, 2004). Cette hypothèse est renforcée par la présence d'immunoglobulines dirigées contre *E. coli*, *Bacteroïdes spp.* et *Clostridium spp*. en quantités significativement supérieures dans la muqueuse intestinale de patients atteints de MC comparativement à la muqueuse de sujets contrôles (Macpherson *et al.*, 1996).

### c. Présence d'agents pathogènes persistant au niveau intestinal

Les similitudes entre la MC et certaines formes de colites infectieuses (avec sécrétion de cytokines, augmentation du taux des bactéries dans la muqueuse, altération de la phagocytose…) ont suggéré que certains agents infectieux à tropisme intestinal pourraient participer à l'étiologie de la MC (Sartor, 2008). De plus, la maladie survient fréquemment

suite à des épisodes de gastro-entérites ou de diarrhées post-antibiotiques. Plusieurs micro-organismes tels que des virus, des levures et des bactéries non commensales du tube digestif ont ainsi été suspectés (Kallinowski *et al.*, 1998; Chiba *et al.*, 1998). Mais, leur rôle dans le processus pathologique n'a jamais été clairement prouvé.

Pendant des décennies, l'apparition de la MC a été associée à des infections virales aiguës, et de nombreuses études épidémiologiques ont associé la pathologie à une consommation d'eau de source, suggérant un virus aquatique. L'hypothèse de survenue de la maladie correspondrait à une vasculite au niveau de la muqueuse intestinale consécutive à une infection virale persistante de l'endothélium mésentérique. Plusieurs virus ont été suspectés, comme les *Herpesviridae* (cytomégalovirus, Epstein-Barr virus…), le virus de la rougeole ou encore le Parvovirus B19 (Wakefield *et al.*, 1992; Wakefield *et al.*, 1993; Miyamoto *et al.*, 1995; Daszak *et al.*, 1997; Pironi *et al.*, 2009; Weinberg *et al.*, 2009), mais les résultats des différentes études n'ont pas permis de déterminer le rôle exact de ces virus dans la survenue de la pathologie. L'implication de bactériophages a également été évoquée, car leur nombre, jusqu'à 10 fois plus élevé que les bactéries dans certains écosystèmes, pourrait influencer la diversité et la composition de la population bactérienne et pourrait ainsi être à l'origine de la dysbiose retrouvée dans la MC (Riley, 2004; Lepage *et al.*, 2008; Norman *et al.*, 2015; Brooks and Watson, 2015).

Le rôle des levures dans la survenue de la MC a également été envisagé puisque le phosphopeptidomannane de la paroi des levures a été identifié comme antigène entrainant la production des anticorps ASCA (Joossens *et al.*, 2002). Ces derniers sont retrouvés au niveau sanguin chez 20% des patients, avant même les premiers symptômes de la MC. *Candida albican*s a été proposé comme pouvant induire la production d'ASCA (Standaert–Vitse *et al.*, 2006; Schaffer *et al.*, 2007; Li *et al.*, 2014). Ce pathogène fongique peut coloniser la surface mucosale de l'épithélium digestif d'une grande partie de la population sans symptômes associés, et il a été démontré qu'il était capable d'aggraver une colite induite au DSS (Dextran Sodium Sulfate) par la production de TNF-α (Jawhara and Poulain, 2007).

Parmi les agents bactériens suspects d'être impliqués dans la MC, *Mycobacterium avium spp. Paratuberculosis* (MAP) a été le premier organisme proposé comme agent pathogène pouvant être responsable de la maladie. Cette bactérie intracellulaire provoque une entérite granulomateuse avec des caractéristiques similaires à la MC. Malgré le fait que cette bactérie ait été isolée au niveau de biopsies intestinales et de granulomes de patients, son rôle dans le déclenchement de la maladie n'a pas été établi (Naser *et al.*, 2004; Sartor, 2005; Feller *et al.*,

2007). Cette bactérie, responsable d'une maladie touchant le bétail, se trouve dans notre alimentation (viande, produits laitiers…) car elle résiste à des conditions extrêmes comme la pasteurisation. L'autophagie est un processus nécessaire afin d'inhiber cette bactérie présente dans les macrophages infectés, mais l'altération de l'autophagie présente chez les patients atteints de MC soutient l'idée du lien entre cette bactérie et la MC (Hansen *et al.*, 2010).

D'autres bactéries ont également été incriminées dans la survenue de la MC, et notamment *Helicobacter spp, Listeria monocytogenes*, *Yersinia enterocolitica*, *Campylobacter spp, Salmonella spp* et *Pseudomonas spp* (Hugot *et al.*, 2003; Saebo *et al.*, 2005; Gradel *et al.*, 2009; Madi *et al.*, 2010; Hansen *et al.*, 2010; Jess *et al.*, 2011).

Enfin, il a été montré une colonisation anormale de la muqueuse intestinale des patients atteints par des souches d'*E. coli*. La caractérisation d'un phénotype d'adhésion et d'invasion de ces souches isolées des muqueuses et l'absence de déterminants génétiques ont conduit à définir un nouveau pathovar d'*E. coli* adhérents et invasifs ou AIEC (Darfeuille-Michaud *et al.*, 1998; Boudeau *et al.*, 1999).

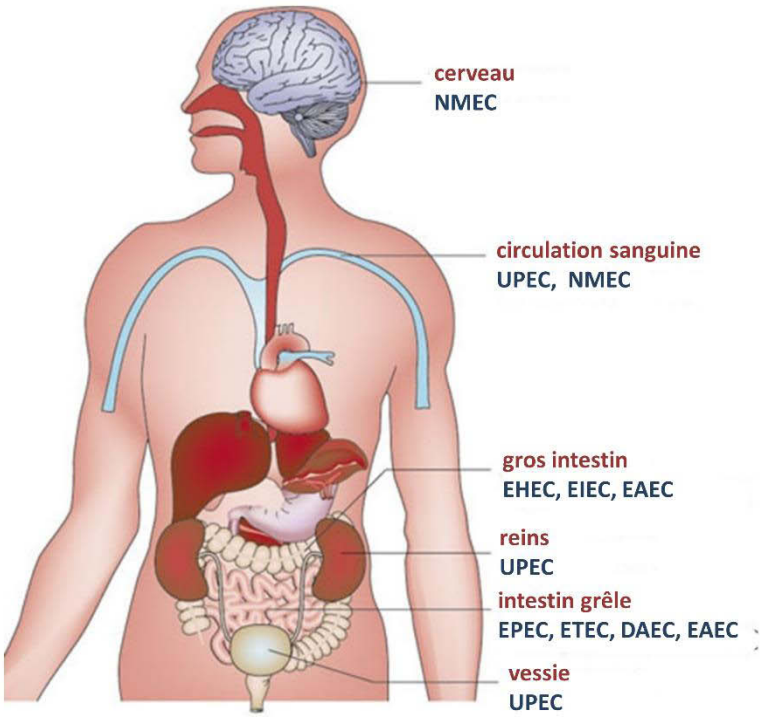



**Figure 25** : Sites colonisés par des souches d'*E. coli* pathogènes (image modifiée d'après Croxen and Finlay, 2010)

# IV. *Escherichia coli* et maladie de Crohn

## 1. Description des différents pathovars d'*E. coli*

### 1) Généralités

Découvert par Théodore Escherich dans des selles de nourrisson en 1885, *Escherichia coli* est un bacille à Gram négatif non sporulé, aérobie-anaérobie facultatif (Escherich, 1989). Bien que cette espèce représente moins de 1% du microbiote intestinal, qui comprend plus de 500 espèces bactériennes, c'est l'espèce aérobie la plus représentée au niveau du tube digestif de l'Homme et des animaux à sang chaud (Tenaillon *et al.*, 2010). Cette bactérie reste une des premières à coloniser l'intestin des nouveau-nés dès l'accouchement par contact avec la flore périnéale de la mère, ou dans le cas des césariennes par contact avec la mère et le personnel (Wall *et al.*, 2009). Elle perdure ensuite au niveau intestinal, principalement dans le caecum et le colon, durant toute la vie (Tenaillon *et al.*, 2010). L'espèce *E. coli* comprend des souches commensales mais également des bactéries pathogènes qui peuvent être à l'origine d'infections intestinales ou extra-intestinales (**Figure 25**) (Kaper *et al.*, 2004; Tenaillon *et al.*, 2010).

Cette espèce bactérienne est particulièrement intéressante du fait des nombreuses stratégies qu'elle a développé pour s'adapter à différents environnements. Ces évolutions se retrouvent en analysant les génomes des souches de l'espèce. Le génome d'*E. coli* présente un core-genome d'environ 2000 gènes essentiels aux fonctions cellulaires de base et un ensemble génétique flexible pouvant être composé de 200 à plus de 3000 gènes par souche, issus d'un pool d'environ 15800 gènes codant des fonctions optionnelles souvent impliquées dans la pathogénèse (Hacker and Carniel, 2001). Des analyses phylogénétiques ont montré que l'espèce *E. coli* peut être divisée en 7 grands phylogroupes : A, B1, B2, C, D, E et F (Clermont *et al.*, 2013). D'après Lecointre *et al*, les phylogroupes A et B1 sont proches, contrairement au phylogroupe B2 qui correspond à une branche plus ancestrale des *E. coli* (Lecointre *et al.*, 1998). Ces phylogroupes de *E. coli* diffèrent entre eux de par leur niche écologique, l'histoire de leur vie et certaines caractéristiques comme leur capacité à exploiter certaines sources de carbone, leur profil de résistance aux antibiotiques et leur taux de

croissance (Gordon and Cowling, 2003; Carlos *et al.*, 2010). Par exemple, la majorité des *E. coli* qui persistent dans l'environnement appartiennent au phylogroupe B1 (Walk *et al.*, 2007). La taille du génome est également variable selon le phylogroupe : les *E. coli* appartenant aux phylogroupes A et B1 possèdent un génome plus petit que ceux des phylogroupes B2 et D (Bergthorsson and Ochman, 1998). Par ailleurs, les souches les plus virulentes, dont les *E. coli* responsables d'infections extra-intestinales ainsi que les *E. coli* retrouvés dans les maladies inflammatoires chroniques de l'intestin, appartiennent majoritairement aux phylogroupes B2 et D, contrairement aux souches commensales et aux souches pathogènes intestinales qui sont retrouvées essentiellement au sein des phylogroupes A et B1 (Bingen *et al.*, 1998; Picard *et al.*, 1999; Johnson *et al.*, 2001; Kotlowski *et al.*, 2007; Desilets *et al.*, 2016).

## 2) Les *E. coli* pathogènes intestinaux et extra-intestinaux

Les *E. coli* responsables de pathologies extra-intestinales sont des pathogènes facultatifs faisant partie de la flore commensale du tube digestif d'une partie de la population. Ces souches ont la capacité d'atteindre de nouvelles niches, de les coloniser et d'induire des maladies (pathologie urinaire, méningite, sepsis…) (Köhler and Dobrindt, 2011). Trois pathovars d'ExPEC sont souvent individualisés : les *E. coli* uropathogènes UPEC, les *E. coli* associés aux méningites néonatales NMEC et les *E. coli* responsables d'infections systémiques en élevage aviaire APEC (Johnson and Russo, 2002).

Dans le cadre des pathologies intestinales, il existe 6 principaux pathovars d'*E. coli* qui ne sont pas retrouvés dans les souches commensales à la différence des *E. coli* responsables d'infections extra-intestinales. Ces pathovars sont capables d'utiliser un large éventail de facteurs de virulence (Kaper *et al.*, 2004; Croxen and Finlay, 2010; Croxen *et al.*, 2013). Les *E. coli* entéro-pathogènes (EPEC) sont responsables de gastro-entérites graves chez l'enfant, notamment dans les pays en voie de développement. Les *E. coli* entéro-hémorragiques (EHEC) sont responsables d'épidémies de gastroentérites sévères dans les pays développés. Le réservoir de ces bactéries est le bétail, et la transmission à l'Homme s'effectue par la nourriture contaminée ou l'eau. Ces infections peuvent se compliquer, plus particulièrement chez les jeunes enfants et les personnes âgées, avec un syndrome hémolytique et urémique (SHU) caractérisé par une anémie hémolytique, une thrombopénie et une insuffisance rénale aigue. Les *E. coli* entéro-toxinogènes (ETEC) sont responsables de la diarrhée du voyageur

(« turista »), ainsi que de la plupart des diarrhées infantiles dans les pays en voie de développement. Les *E. coli* entéro-invasifs (EIEC) sont responsables d'infections aussi bien chez l'adulte que chez l'enfant. Enfin, les *E. coli* entéro-agrégatifs (EAEC) et les *E. coli* à agrégation diffuse (DAEC) sont principalement responsables de diarrhées du nourrisson dans les pays en voie de développement. Ces souches pathogènes de *E. coli* codent des facteurs d'adhésion spécifiques leur permettant de coloniser le tractus intestinal ainsi que des toxines ou des protéines sécrétées capables d'affecter un certain nombre de fonctions eucaryotes.

A côté de ces souches commensales et pathogènes, certaines souches d'*E. coli* qui ne sont pas pathogènes pour un sujet sain, sont capables de contribuer au développement et à l'exacerbation de maladies inflammatoires chroniques intestinales (Carvalho *et al.*, 2008; Agus *et al.*, 2014; Dalmasso *et al.*, 2014; Chassaing *et al.*, 2014; Yang and Jobin, 2014; Delmas *et al.*, 2015).

## 2. Le pathovar *Escherichia coli* adhérent et invasif associé à la maladie de Crohn

De nombreux arguments sont en faveur de l'implication d'*E. coli* dans la MC. Des études ont ainsi rapporté que le titre d'anticorps dirigé contre cette bactérie était plus élevé chez les patients atteints de MICI (Tabaqchali *et al.*, 1978). De même, la présence d'antigènes spécifiques d'*E. coli* a pu être mise en évidence dans des biopsies intestinales de sujets ayant la MC notamment au niveau des macrophages et des cellules géantes à la base des ulcères, le long des fissures, dans les granulomes et la *lamina propria* (Liu *et al.*, 1995). Il a d'ailleurs été démontré par la suite une augmentation du nombre d'*E. coli* associés à la muqueuse intestinale dans la MC : ils colonisent anormalement les lésions iléales aiguës et chroniques, pouvant représenter jusqu'à 100% de la flore totale aéro-anaérobie comparativement aux sujets sains (Darfeuille-Michaud *et al.*, 1998).

Les souches d'*E. coli* retrouvées dans l'épithélium des patients atteins de MC sont particulièrement adhérentes aux cellules épithéliales intestinales, sont capables d'envahir les cellules et peuvent également survivre et se multiplier dans les cellules épithéliales ainsi que les macrophages. Ces propriétés ont permis de définir un nouveau pathovar d'*E. coli* : les *E. coli* adhérents et invasifs (AIEC), retrouvés dans 36 % des MC iléales (Darfeuille-Michaud *et al.*, 2004; Barnich and Darfeuille-Michaud, 2007; Agus *et al.*, 2014). La souche LF82

(sérogroupe O83:H1) isolée de la muqueuse iléale d'un patient atteint de la MC constitue une souche de référence des AIEC.

### 1) Propriétés d'adhésion et d'invasion des AIEC

Les AIEC adhèrent *via* leurs pili de type I au récepteur humain CEACAM6 qui a été montré comme étant surexprimé au niveau iléal chez les patients prédisposés à la MC (Barnich *et al.*, 2007; Carvalho *et al.*, 2009). De plus, l'adhésion des AIEC augmente l'expression de ce récepteur en stimulant la sécrétion de TNF-α et d'IFN-γ ; ils favorisent ainsi leur propre colonisation.

Plusieurs études ont montré la capacité des *E. coli* associés à la muqueuse d'envahir les cellules épithéliales intestinales (Martin *et al.*, 2004; Darfeuille-Michaud *et al.*, 2004; Conte *et al.*, 2006; Kotlowski *et al.*, 2007; Sasaki *et al.*, 2007). Ainsi, les AIEC sont capables de transloquer au niveau de l'épithélium intestinal vers la *lamina propria*, où ils survivent et peuvent se multiplier dans les macrophages. Ils exacerbent l'inflammation par activation de la voie NF-κB qui conduit à la sécrétion de cytokines pro-inflammatoires comme le TNF-α, l'IL-1β ou l'IL-6.

### 2) Survie et multiplication dans les macrophages des AIEC

Comme d'autres souches invasives de type EIEC, *Shigella* ou *Salmonella*, les AIEC sont capable de survivre, mais également de se multiplier à l'intérieur des cellules épithéliales intestinales. Il a en effet été montré que la souche LF82 était capable, comme *Shigella*, de lyser la membrane de la vacuole d'endocytose et de se retrouver libre dans le cytoplasme de la cellule hôte, où elle peut ensuite se multiplier (Boudeau *et al.*, 1999). Mais contrairement à ces dernières, la souche LF82 est également capable de se multiplier dans les macrophages sans induire de mort cellulaire par apoptose ou nécrose (Glasser *et al.*, 2001). Les macrophages infectés par la LF82 sécrètent des taux élevés de TNF-α. Cette sécrétion favorise la multiplication de la bactérie à l'intérieur des macrophages (Bringer *et al.*, 2012).

### 3) Autres caractéristiques des AIEC

L'étude des souches AIEC et notamment de la souche répresentative de ce pathovar LF82 a montré que ces souches étaient également caractérisées par leur interaction avec l'hôte.

Les AIEC peuvent également former des granulomes (Meconi *et al.*, 2007; Agus *et al.*, 2014). Elles sont capables de former d'épais biofilms (Martinez-Medina, Naves, *et al.*, 2009). Elles peuvent moduler l'autophagie *via* des microARN régulateurs (miR-30C et miR-130A) qui contrôlent l'expression des protéines de l'autophagie ATG5 et ATG16L1 (Lapaquette *et al.*, 2010; Nguyen *et al.*, 2014), conduisant à une augmentation du nombre de bactéries internalisées ainsi qu'une inflammation associée (Nguyen *et al.*, 2014). Parmi les gènes qui ont été associés à la MC, nous retrouvons les gènes *ATG16L1* et *IRGM*, impliqués dans l'autophagie (Nguyen *et al.*, 2013). Ce lien entre bactérie invasive et susceptibilité génétique renforce l'hypothèse d'une impossibilité pour les patients atteints de MC de contrôler l'infection par les AIEC aboutissant à une persistance de ces souches dans l'intestin.

## 3. Facteurs de virulence retrouvés chez les AIEC

Le pathovar AIEC n'est pas défini sur la possession d'un facteur de virulence spécifique comme les *E. coli* entéropathogènes. Par ailleurs, les souches AIEC peuvent posséder des facteurs de virulence retrouvés chez les UPEC (Vejborg *et al.*, 2011; Dreux *et al.*, 2013). Toutefois, le pathovar AIEC n'est pas associé à des facteurs de virulence spécifiques des UPEC tels que les pili P (*pap*), S (*sfa*), et F1C (*foc*), qui permettant aux UPEC de coloniser le tractus urinaire, et les toxines telles que l'hémolysine (*hly*) et le facteur de nécrose (CNF1), induisant des dommages tissulaires (Wiles *et al.*, 2008). A l'inverse des AIEC, les ExPEC ne sont pas des souches adhérentes aux cellules épithéliales intestinales (Martinez-Medina, Mora, *et al.*, 2009). L'analyse des génomes des souches AIEC LF82 et NRG857C a ainsi révélé que le pathovar AIEC possède (i) des gènes flexibles, c'est-à-dire des gènes ne faisant pas partie du core-génome de la bactérie et qui sont spécifiques donc non retrouvés chez les autres *E. coli* (ii) des gènes de virulence qui pourraient expliquer en partie les mécanismes de pathogénie des souches AIEC et (iii) des mutations pathoadaptatives qui pourraient jouer un rôle majeur dans l'amélioration des capacités de colonisation et

d'inflammation du tractus intestinal (Miquel, Peyretaillade, *et al.*, 2010; Clements *et al.*, 2012).

### 1) Les *fimbriae*

Les pili de type I, synthétisés par la majorité des *Enterobacteriaceae*, sont impliqués dans la pathogénicité de par leur rôle dans l'adhésion bactérienne à différents types de cellules eucaryotes. Ces pili reconnaissent des résidus mannose (Ofek et al., 1977; Giampapa et al., 1988). Ils sont également impliqués dans l'invasion de certains *E. coli* uropathogènes ou de *Salmonella enterica* serovar Typhimurium (Horiuchi et al., 1992; Bower et al., 2005).

Les AIEC possèdent des pili de type I qui leur permettent d'adhérer au récepteur humain CEACAM6 au niveau des cellules épithéliales intestinales, première étape indispensable à l'invasion des cellules (Barnich *et al.*, 2007). Ils jouent également un rôle clé dans l'invasion des cellules épithéliales par les AIEC en induisant la formation d'élongations membranaires à la surface des cellules cibles, contribuant à leur internalisation (Boudeau *et al.*, 2001).

De plus, il a été rapporté un variant de l'adhésine FimH, située à l'extrémité des pili de type I, qui confère aux AIEC une liaison de meilleure affinité avec les résidus mannose du récepteur CEACAM6 (Dreux *et al.*, 2013).

A côté des cellules épithéliales, les AIEC sont également capables de traverser l'épithélium intestinal au niveau des plaques de Peyer en passant par les cellules M, afin de rejoindre la *lamina propria.* L'adhésion des AIEC aux cellules M implique d'autres pili particuliers : les long polar *fimbriae* (lpf) (Chassaing *et al.*, 2011). L'operon *lpf* est également présent chez *Salmonella enterica* serovar Typhimurium, *Shigella sp*, ainsi que chez des EHEC (Bäumler *et al.*, 1996; Torres *et al.*, 2002). Ces lpf sont impliqués dans l'adhésion des bactéries aux plaques de Peyer, et contribuent ainsi à la colonisation intestinale des AIEC.

### 2) Les flagelles

Les flagelles, synthétisés par de nombreuses espèces bactériennes, sont responsables de la mobilité des bactéries. La mobilité conférée par les flagelles est indispensable au pouvoir invasif de certaines bactéries comme les *E. coli* pathogènes, *Salmonella enterica*, ou

*Campylobacter jejuni* (Yao *et al.*, 1994; Dibb-Fuller *et al.*, 1999; Schmitt *et al.*, 2001; Parker and Guard-Petter, 2001; Girón *et al.*, 2002; Kakkanat *et al.*, 2015). Certains flagelles peuvent également jouer le rôle de système de sécrétion de type 3 pour l'injection de facteurs de virulence comme la phospholipase YlpA de *Yersinia enterocolitica* ou certaines protéines de *Campylobacter jejuni* et *Campylobacter coli* (Young *et al.*, 1999; Guerry, 2007).

Les flagelles bactériens des AIEC agissent de concert avec les pili de type I pour permettre l'adhésion des AIEC aux cellules épithéliales intestinales (Barnich *et al.*, 2003). En effet, la souche LF82 délétée du gène *fliC* codant la flagelline FliC n'est plus capable de synthétiser de pili de type I, et d'adhérer et d'envahir les cellules épithéliales intestinales.

### 3) Les vésicules de membrane externe

Les vésicules de membrane externe (OMV, outer membrane vesicles) sont produites par de très nombreuses bactéries à Gram négatif, pathogènes ou non, dont *E. coli*, *Shigella sp*, *Salmonella* Typhi, *Pseudomonas aeruginosa*, *Bacteroïdes spp* ou encore *Helicobacter pylori* (Hoekstra *et al.*, 1976; Grenier and Mayrand, 1987; Kadurugamuwa and Beveridge, 1999; Fiocca *et al.*, 1999; Wai *et al.*, 2003). La formation d'OMV est habituelle chez ces bactéries, et s'observe dans de nombreuses conditions, y compris lorsque ces dernières sont organisées en biofilms (Schooling and Beveridge, 2006).

Ces vésicules résultent d'un bourgeonnement de la membrane externe à la surface des bactéries et sont donc naturellement constituées de LPS, de protéines de la membrane externe ainsi que de phospholipides. Des protéines périplasmiques peuvent être emprisonnées dans ces vésicules lors de leur formation (Wai *et al.*, 2003). Les vésicules constituent ainsi un moyen pour les bactéries d'interagir avec des cellules eucaryotes et/ou d'autres cellules procaryotes. Elles permettent de véhiculer des adhésines, des toxines ou des protéines immunomodulatrices impliquées dans la virulence bactérienne vis-à-vis des cellules de l'hôte (Kuehn and Kesty, 2005).

L'étude de la production d'OMV chez les AIEC a montré pour la première fois l'implication de ces vésicules dans l'invasion de cellules-hôtes par des bactéries pathogènes (Rolhion *et al.*, 2005). La délétion du gène *yfgL* codant une protéine de la membrane externe, entraine une incapacité de la bactérie LF82 à produire des OMV. En effet, la protéine YfgL régule positivement des transglycosidases lytiques, impliquées dans le renouvellement du

peptidoglycane. Ainsi, l'absence de YfgL entraine une diminution de l'accumulation de produits de dégradation du peptidoglycane au niveau du périplasme, et ainsi une diminution de pression exercée sur la membrane externe, conduisant à une libération diminuée d'OMV. Mais l'absence de cette protéine provoque aussi une incapacité de la souche mutante à envahir les cellules eucaryotes. En effet, les OMV sécrétées par la souche LF82 interagissent avec les cellules-hôte *via* la protéine bactérienne de membrane externe OmpA sécrétée dans les OMV qui reconnait le récepteur Gp96, glycoprotéine du stress du réticulum endoplasmique. Ce dernier est anormalement surexprimé au niveau de la muqueuse iléale des patients atteints de MC. Cette surexpression facilite donc l'invasion des cellules par les AIEC (Rolhion *et al.*, 2010).

Ainsi, les OMV des AIEC représentent un système de transport permettant de délivrer aux cellules-hôte des effecteurs bactériens impliqués dans le réarrangement du cytosquelette des cellules eucaryotes, et donc dans l'internalisation des bactéries (Rolhion *et al.*, 2005)

### 4) Le système d'osmorégulation EnvZ/OmpR

En fonction des conditions environnementales, comme par exemple les changements d'osmolarité rencontrés le long du tractus intestinal, la pression osmotique cytoplasmique des bactéries est variable. Pour résister à une éventuelle entrée d'eau, la bactérie doit mettre en place des mécanismes de transport pour maintenir la différence d'osmolarité entre le cytoplasme bactérien et le milieu extérieur. Les systèmes à deux composantes sont des mécanismes de régulation transcriptionnelle pour l'adaptation rapide des bactéries à leur environnement. Plus de 250 systèmes ont été identifiés dans le monde bactérien, dont une trentaine chez *E. coli* (Mizuno, 1997). Ces systèmes répondent à un stimulus spécifique et font intervenir deux protéines : un senseur présent dans la membrane externe des bactéries à Gram négatif et un régulateur de la transcription localisé dans le cytoplasme.

Chez *E. coli*, la protéine qui joue le rôle de senseur, EnvZ, est une histidine kinase située au niveau du périplasme de la bactérie. Elle comprend deux domaines transmembranaires qui séparent une région N-terminale qui joue le rôle de senseur, et une région C-terminale cytoplasmique catalytique qui permet de transmettre un signal. EnvZ surveille d'éventuels stimuli extérieurs et transfère ainsi ces informations au régulateur OmpR par un signal de

phosphorylation. Ce dernier possède deux domaines : un domaine N-terminal de régulation, et un domaine C-terminal effecteur qui joue le rôle de facteur de transcription grâce à une région de liaison à l'ADN. OmpR permet ainsi d'activer ou de réprimer la transcription de gènes spécifiques comme *ompC* et *ompF* qui codent pour des porines de la membrane externe (Ann M. Stock *et al.*, 2000). Cette régulation mise en place permet ainsi de modifier la perméabilité membranaire de la bactérie en fonction de l'osmolarité du milieu mais également en fonction du pH et de la disponibilité en nutriments (Heyde and Portalier, 1987; Liu and Ferenci, 2001).

Chez les AIEC, lorsque la forte osmolarité présente dans la lumière intestinale favorise l'expression des porines OmpC et OmpF, leur accumulation périplasmique ainsi que leur mauvais repliement éventuel activent la voie σE. Cette voie permet l'activation de la synthèse des pili de type I, des flagelles et d'autres facteurs de virulence impliqués dans l'interaction entre la bactérie et la cellule hôte. Cette voie active également la formation de biofilm et semble indispensable pour la colonisation de la muqueuse intestinale par les AIEC (Rolhion *et al.*, 2007; Chassaing and Darfeuille-Michaud, 2013).

### 5) Les facteurs d'invasion et de multiplication cellulaire

D'autres gènes codant des protéines impliquées dans l'invasion des cellules eucaryotes et dans la multiplication intracellulaire ont également été mis en évidence chez les AIEC. Parmi eux, le gène *ibeA* code une invasine, pour laquelle plusieurs récepteurs eucaryotes ont été identifiés, notamment sur les cellules endothéliales microvasculaires situées au niveau du cerveau. En effet, cette invasine a été impliquée dans l'invasion des souches *E. coli* K1 responsables de méningites néonatales (Zou *et al.*, 2007). Plus récemment, il a été montré que l'invasine IbeA joue également un rôle majeur dans l'invasion des cellules épithéliales par les AIEC et dans leur survie en macrophage sans que le mécanisme d'action ne soit élucidé (Cieza *et al.*, 2015). D'autre part, la souche LF82 possède dans son génome le gène *gipA* qui code une transposase-like et dont la délétion diminue la capacité des bactéries à multiplier dans les macrophages, (Klumpp and Fuchs, 2007; Vazeille *et al.*, 2016).

Il a également été observé la présence du cluster de gènes *pdu* chez la souche LF82, impliqués dans le catabolisme du 1,2-propanediol dépendant du coenzyme B12, qui est une source de carbone nécessaire à la multiplication de *Salmonella* à l'intérieur des macrophages (Heithoff *et al.*, 1999; Klumpp and Fuchs, 2007). Ce cluster *pdu* pourrait donc permettre aux

souches AIEC de mieux coloniser la muqueuse intestinale et de se multiplier à l'intérieur des macrophages (Dogan *et al.*, 2014).

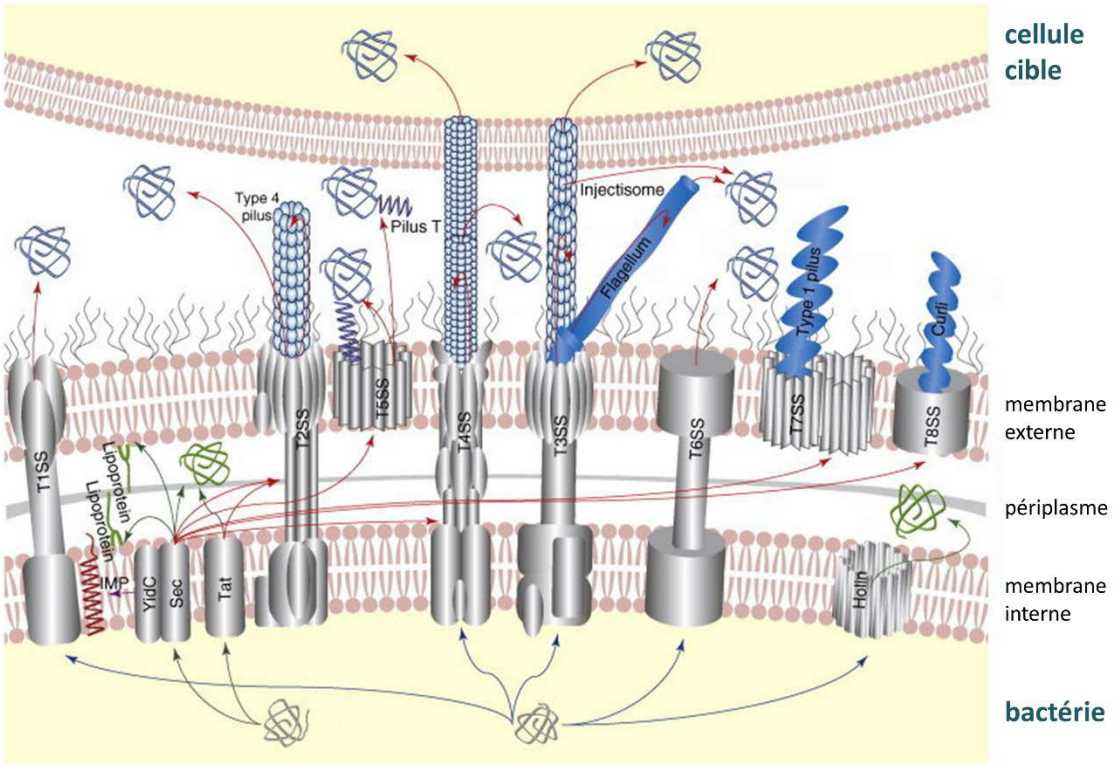



**Figure 26** : Les différents systèmes de sécrétion chez les Entérobactéries (image modifiée d'après Desvaux *et al.*, 2009)

## 6) Les systèmes de sécrétion

### a. Généralités

Les protéines impliquées dans la colonisation, l'invasion et/ou la survie des bactéries dans un organisme hôte peuvent être sécrétées dans le milieu extracellulaire ou directement dans le cytoplasme des cellules cibles, procaryotes ou eucaryotes, par des systèmes de sécrétion. Il existe huit grands systèmes de sécrétion chez les entérobactéries, dont chacun permet la sécrétion de molécules spécifiques (**Figure 26**) (Tseng *et al.*, 2009; Desvaux *et al.*, 2009; Fronzes *et al.*, 2009; Waksman, 2012).

Ces systèmes de sécrétion sont de véritables machines moléculaires qui permettent le transport de protéines effectrices à travers la membrane interne et la membrane externe de la bactérie, et parfois même l'injection des effecteurs directement dans une cellule cible. Les effecteurs ainsi sécrétés jouent un rôle important dans la perturbation de processus cellulaires de l'hôte et dans l'induction d'une réponse immunitaire (Anderson and Frank, 2012). Ils sont donc considérés comme des facteurs de virulence car ils jouent un rôle important dans l'infection des cellules hôte.

La plupart des effecteurs sont sécrétés et injectés dans les cellules hôtes grâce aux systèmes de sécrétion de type 3, 4 et 6, qui permettent le franchissement des deux membranes bactériennes en une seule étape, alors que les toxines sécrétées dans le milieu extracellulaire passent par les systèmes de sécrétion de type 1, 2 et 5 (Cambronne and Roy, 2006; Jayamani and Mylonakis, 2014). Les systèmes de sécrétion de type 7 et 8 ont été rapportés plus récemment (Desvaux *et al.*, 2009). Les systèmes de sécrétion de type 2, 5, 7 et 8 ne permettent le franchissement que de la membrane externe. Ces derniers nécessitent donc une première étape de transfert des protéines à travers la membrane interne vers le périplasme grâce au système Sec ou GSP (General secretion pathway) ou au système Tat (Twin arginine translocation).

Le système Sec est le principal système utilisé pour la translocation des protéines à travers la membrane cytoplasmique. Il comprend un translocon SecYEG formant un pore dans la membrane pour permettre le transport les polypeptides. Ce système fait également intervenir une ATPase cytoplasmique SecA permettant de générer l'énergie nécessaire à la translocation

et entrainer les polypeptides destinés à être exportés grâce à son interaction avec la séquence signal du peptide. Lors de la translocation, le peptide signal est clivé par une peptidase (Oliver and Beckwith, 1982; Gouridis *et al.*, 2009).

Le système Tat est quant à lui composé d'un complexe membranaire TatBC qui reconnaît une séquence signal possédant un motif « double arginine » des peptides, puis se lie à TatA pour former un pore dans la membrane (Gohlke *et al.*, 2005). Ce système se distingue du système Sec par sa capacité à transporter des protéines repliées, ce qui représente un défi plus important pour les bactéries car il est capable de sécréter des protéines allant jusqu'à 70 Å chez *E. coli* (Berks *et al.*, 2000; Voulhoux *et al.*, 2001). Ce système doit donc être bien régulé afin de conserver une perméabilité membranaire pour la bactérie.

## b. Les différents systèmes de sécrétion chez les AIEC

L'étude du génome de la souche de référence LF82 a montré l'existence potentielle de plusieurs systèmes de sécrétion au sein du pathovar AIEC (Miquel, Peyretaillade, *et al.*, 2010). Son génome possède entre autres deux systèmes de sécrétion de type VI (SST6), correspondant aux îlots de pathogénicité I et III. Il a également été mis en évidence la présence des gènes appartenant à l'opéron *gsp* codant un système de sécrétion de type II (SST2), ainsi que l'opéron *csg* codant un système de sécrétion de type VIII (SST8). Enfin, la souche LF82 possède également un gène, annoté *LF82_p043*, qui code une protéine semblant appartenir à la famille des autotransporteurs, soit le système de sécrétion de type V (SST5).

### *Le système de sécrétion de type VI*

Le SST6 a été découvert plus récemment chez *Vibrio cholerae* et *Pseudomonas aeruginosa* (Pukatzki *et al.*, 2006; Mougous *et al.*, 2006). Ce système permet la translocation d'effecteurs directement dans la cellule cible, mais la cellule cible peut être une cellule eucaryote tout comme une autre bactérie (Jani and Cotter, 2010; Silverman *et al.*, 2012). C'est un système à composants multiples de 12 à 25 sous-unités avec une ATPase hexamérique (ClpV) et deux protéines extracellulaires Hcp (Haemolysin co-regulated protein) et VgrG (Valine-Glycine repeat protein G), qui ressemble à un bactériophage inversé (Kanamaru, 2009). Hcp forme un cylindre à l'extérieur de la bactérie, alors que VgrG s'organise en complexe trimérique

rappelant la queue d'un bactériophage pour traverser la membrane bactérienne (Pukatzki *et al.*, 2007). Ce système forme une structure tubulaire permettant le transport de protéines vers le milieu extracellulaire. Cet export ne nécessite pas le clivage d'un peptide signal. Ce système est présent chez de nombreuses bactéries pathogènes comme par exemple *Vibrio cholerae*, *Pseudomonas aeruginosa*, *Salmonella enterica* ou *Yersinia pestis* (Filloux *et al.*, 2008; Pukatzki *et al.*, 2009). Chez *Pseudomonas aeruginosa*, le SST6 servirait à éliminer les autres bactéries afin d'être la seule à persister dans le tractus respiratoire des patients atteints de mucoviscidose (Russell *et al.*, 2011). Chez les AIEC, deux SST6 ont été décrits au niveau des îlots de pathogénicité I et III. On peut supposer que les AIEC possèderaient ce système de sécrétion pour éliminer les bactéries commensales du tube digestif et favoriser leur colonisation.

### *Le système de sécrétion de type II*

Le SST2 est un système complexe permettant le transport de protéines du périplasme vers le milieu extracellulaire. Les protéines sont donc sécrétées en deux étapes distinctes, et possèdent une séquence signal en N-terminal ce qui, selon la nature de ce peptide, permet aux précurseurs protéiques d'être pris en charge par le système Sec ou Tat au niveau de la membrane interne (Voulhoux *et al.*, 2001; Johnson *et al.*, 2006). Le peptide signal est ensuite clivé par une peptidase au niveau du périplasme et la protéine mature est transportée à travers la membrane externe grâce à un pore formé majoritairement par un multimère de la sécrétine GspD. L'énergie nécessaire à la translocation est fournie par l'ATPase cytoplasmique GspE, qui est reliée à GspD grâce à un pseudopilus formé par le complexe GspGHIJK. Ce pseudopilus agirait comme un piston pour pousser la protéine destinée à être sécrétée du périplasme vers le pore membranaire. Ce système de sécrétion (ou secreton), codé par 12 à 16 gènes, permet de sécréter des protéines repliées allant jusqu'à 55Å. Il est spécifique de certaines enzymes et toxines, comme la pullulanase PulA de *Klebsiella oxytoca* impliquée dans le catabolisme de l'amidon ou la toxine cholérique de *Vibrio cholerae* (Hirst *et al.*, 1984; Sandkvist, 2001). Le génome de la souche LF82 contient l'opéron *gsp*, mais le fait que cet opéron code un système fonctionnel n'a aujourd'hui pas été démontré.

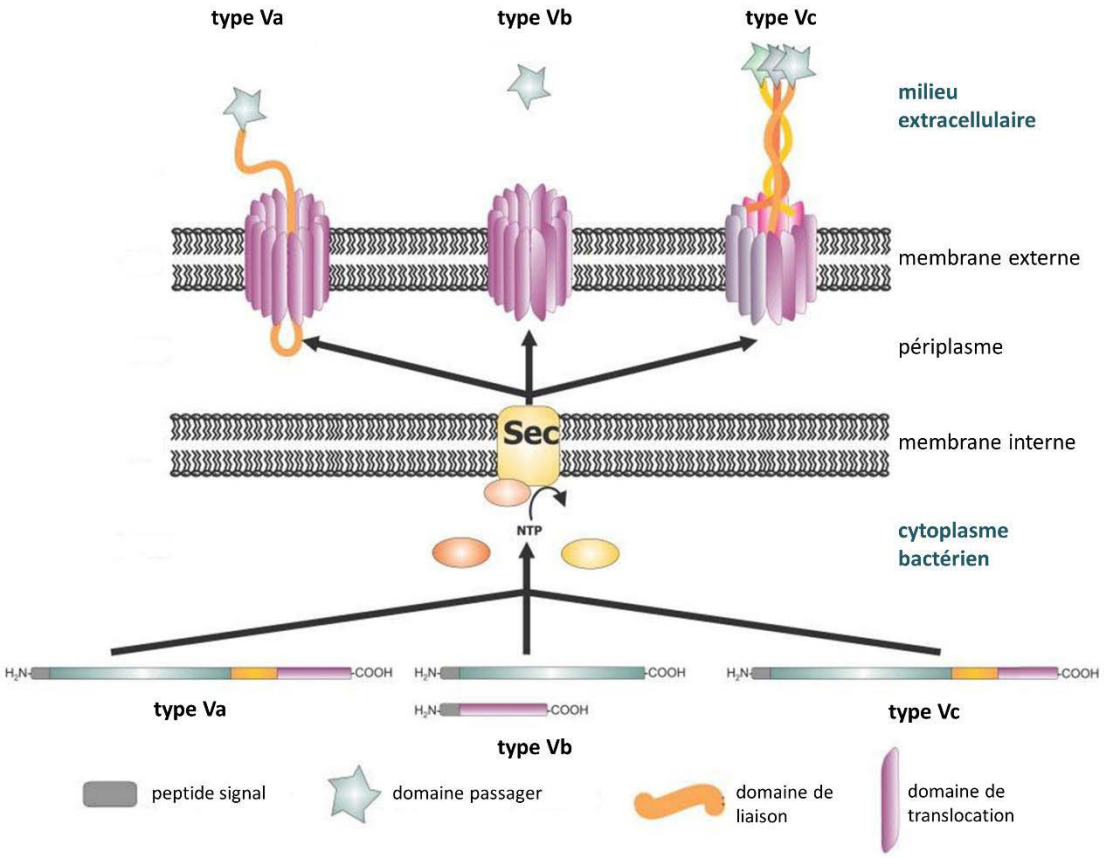



**Figure 27** : Le système de sécrétion de type V
(image modifiée d'après Henderson et al., 2004)

### *Le système de sécrétion de type VIII*

Le SST8 a été récemment suggéré pour la voie ENP (extracellular nucleation-precipitation pathway) qui permet à des bactéries pathogènes comme *E. coli* d'interagir spécifiquement avec son hôte et former un biofilm (Desvaux *et al.*, 2009), grâce à la sécrétion et l'assemblage de *fimbriae* aggrégatifs. A la différence des autres *fimbriae*, le composé protéique Curli, appelé également Tafi (thin aggregative *fimbriae*) s'assemble dans le milieu extracellulaire pour former le biofilm des Protéobactéries, notamment *E. coli* et *Salmonella spp*. Il est constitué de deux sous-unités CsgA et CsgB, qui forment des fibres amyloïdes et est sécrété en deux étapes successives par une machinerie spécifique, à la différence des fibres amyloïdes d'origine eucaryote. Le SST8 comprend ainsi une lipoprotéine membranaire CsgG formant un tonneau composé de 36 brins β, et deux protéines accessoires CsgE et CsgF. Ces 3 protéines sont adressées au périplasme *via* le système Sec, de même que les sous-unités CsgA et B qui seront sécrétées dans le milieu extracellulaire (Barnhart and Chapman, 2006; Goyal *et al.*, 2014). Bien que l'opéron *csg* soit présent dans le génome de la souche LF82, la présence d'un SST8 fonctionnel chez les souches AIEC n'a pas non plus été démontrée à l'heure actuelle.

### *Le système de sécrétion de type V*

Le SST5 permet la libération de protéines dans le milieu extracellulaire en deux étapes successives, dont la première étape consiste au passage à travers la membrane interne et fait intervenir le système Sec. Avec plus de 700 membres identifiés, les protéines sécrétées par le SST5 font partie de la plus grande famille de facteurs de virulence bactériens. Il existe trois grands types de SST5 : le système d'autotransporteur monomérique Va, le système à deux partenaires Vb et le système d'autotransporteur trimérique Vc (**Figure 27**) (Desvaux, Parham, *et al.*, 2004; Leo *et al.*, 2012).

Le système Va est le système de sécrétion le plus simple décrit dans la littérature, car toute l'information nécessaire à la sécrétion est détenue dans un seul peptide. Les autotransporteurs sont des protéines comprenant 4 grands domaines : un peptide signal en N-terminal leur permettant d'être adressés vers le périplasme *via* le système Sec ; un domaine passager qui supporte l'activité catalytique de la protéine et qui sera exposé à la surface de la membrane

externe de la bactérie ou relargué dans le milieu extracellulaire ; un domaine de liaison nécessaire à la translocation du domaine passager à travers le dernier domaine ; et enfin en C-terminal un domaine de translocation composé de feuillets β impliqués dans la formation d'un pore transmembranaire. Le gène annoté *LF82_p043* chez la souche de référence LF82 code une protéine qui fait à priori partie de ces autotransporteurs monomériques.

Le système à deux partenaires diffère des autotransporteurs monomériques par le fait que les domaines passager et de translocation sont traduits en tant que deux protéines séparées. Une protéine translocatrice appelée TpsB permet la sécrétion de la protéine effectrice TpsA à travers la membrane externe. Les gènes codant ces deux protéines sont généralement organisés en opéron. La protéine TpsB diffère également du domaine de translocation par le nombre de feuillets β et par son implication dans la maturation de la protéine TspA (Jacob-Dubuisson *et al.*, 2001; Henderson *et al.*, 2004). Les exemples les plus étudiés de ce système sont l'hémagglutinine filamenteuse FHA de *Bordetella pertussis* et les adhésines HMW1 et HMW2 de *Haemophilus influenzae* (Mazar and Cotter, 2006; St Geme and Yeo, 2009).

Enfin le système Vc comprend les autotransporteurs trimériques, tous impliqués dans l'adhésion. L'organisation est similaire à celle des autotransporteurs monomériques, avec un peptide signal Sec en position N-terminale, un domaine passager composé d'une tête et d'une tige, un domaine de liaison et en C-terminal un domaine de translocation (Henderson *et al.*, 2004; Linke *et al.*, 2006). La tête supporte l'activité de la protéine sécrétée et la tige comprend une structure « coiled-coil » qui éloigne le domaine d'adhésion de la membrane. Les protéines doivent cette fois former un trimère pour être stables et fonctionnelles (Cotter *et al.*, 2006). Chaque monomère contribue pour 4 brins β du tonneau constitué de 12 brins β qui permettra la sécrétion du domaine passager. L'architecture trimérique confère une résistance accrue aux protéases et aux détergents, et permet également à ces adhésines de former des interactions multivalentes avec leur récepteur, conférant ainsi une plus grande affinité. Parmi les protéines utilisant ce système se trouve par exemple l'adhésine YadA de *Yersinia enterocolitica* (Nummelin *et al.*, 2004).

Deux autres systèmes Vd et Ve ont été récemment proposés (Grijpstra *et al.*, 2013). Le système Vd, correspondant à la protéine PlpD de *Pseudomonas aeruginosa*, est un système hybride entre un autotransporteur monomérique et un système à deux partenaires. Le domaine *translocator* ressemble en effet à celui retrouvé dans le système à deux partenaires avec la présence de 16 brins β (Salacha *et al.*, 2010). Le système Ve se rapproche plus d'un autotransporteur monomérique inversé, avec le domaine *translocator* en N-terminal et le

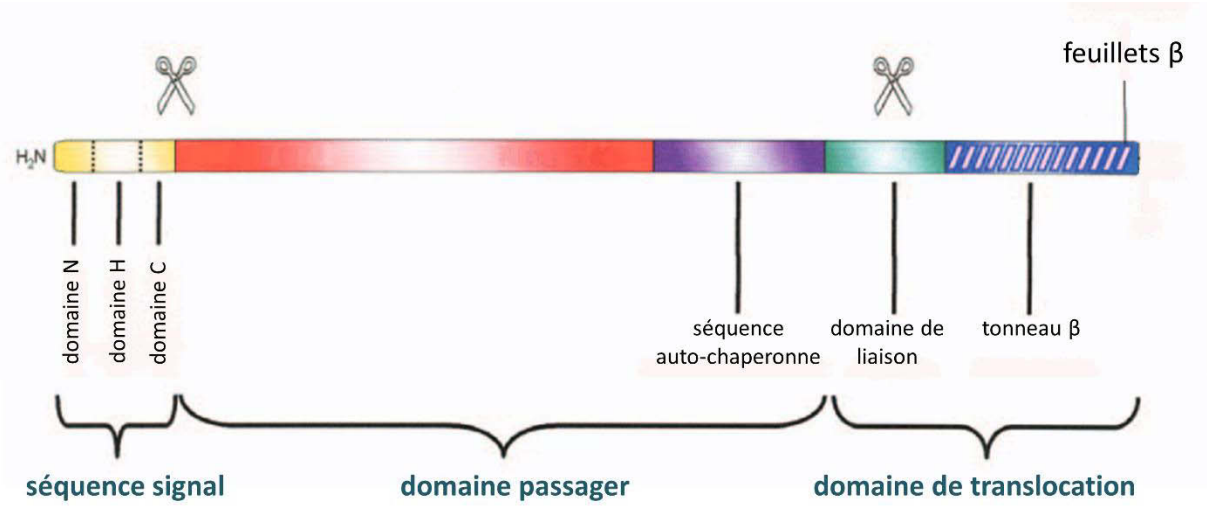



**Figure 28** : Structure primaire des autotransporteurs monomériques
(image modifiée d'après Desvaux *et al.*, 2004)

domaine *passenger* en C-terminal. C'est le cas pour l'intimine de *E. coli* ou l'invasine de *Yersinia pestis* (Tsai *et al.*, 2010).

### c. Caractéristiques des autotransporteurs monomériques

Les autotransporteurs monomériques font partie du SST5 (type Va). Leur nom provient du fait que toute l'information nécessaire à leur sécrétion est présente dans un seul polypeptide. En effet, ils contiennent une séquence signal Sec-dépendante, un domaine passager qui supporte la fonction de la protéine, un domaine de jonction ainsi qu'un domaine de translocation formant un tonneau β permettant la sécrétion de la protéine dans le milieu extracellulaire (**Figure 28**).

Ce système a été décrit pour la première fois avec l'immunoglobuline protéase IgA1 de *Neisseria gonorrhoeae* (Pohlner *et al.*, 1987), et correspond au système le plus simple utilisé par les entérobactéries. Toutefois, ce mécanisme de sécrétion est un peu plus complexe qu'il n'en a l'apparence.

Une fois le précurseur protéique présent dans le périplasme grâce au système Sec, la séquence signal est clivée et l'autotransporteur est libéré dans le périplasme sous forme mature ou immature. Le tonneau β formé par le domaine de translocation s'insère ensuite dans la membrane externe pour former un pore, sans besoin d'énergie. Le domaine passager passe par ce pore où il pourra subir d'autres modifications post-traductionnelles nécessaires à son activité. Ce processus est assisté par une séquence auto-chaperonne interne au domaine passager et par des protéines chaperonnes périplasmiques (Henderson *et al.*, 2004). Le domaine passager pourra être sécrété dans le milieu extracellulaire ou rester fixé au domaine β.

#### *La séquence signal*

La plupart des autotransporteurs présentent une séquence signal typique pour le système Sec, comprise entre 18 et 26 acides aminés. Cette séquence comprend trois parties : une partie chargée positivement riche en résidus lysine et arginine (domaine N) permettant l'adressage de la pré-protéine à la translocase SecA ; une partie hydrophobe (domaine H) dont la longueur

est déterminante pour l'efficacité de la translocation ; et une séquence consensus (domaine C) reconnue par les signal peptidases pour le clivage du peptide signal (Fekkes and Driessen, 1999).

Toutefois, certains autotransporteurs possèdent une séquence signal plus grande, de 50 à 60 acides aminés, composée de cinq régions N1-H1-N2-H2-C (Henderson *et al.*, 1998). La partie N1-H1 correspond à l'extension par rapport à la séquence signal normale et reste très conservée entre les autotransporteurs, suggérant une fonction additionnelle comme le recrutement de protéines accessoires. Par exemple, la séquence signal allongée des autotransporteurs Hbp et EspC permet son interaction avec la protéine YidC, ce qui retarde le relâchement de la protéine dans le périplasme et prévient ainsi un mauvais repliement (Jong *et al.*, 2010).

Enfin, certains autotransporteurs comme NalP chez *Neisseria meningitidis*, SphB1 chez *Bordetella pertussis* ou CapA chez *Campylobacter jejuni* possèdent un motif caractéristique des lipoprotéines. Après passage à travers la membrane interne, ces protéines sont adressées dans le périplasme vers une acétyltransférase qui ajoute une chaine lipidique à la première cystéine de la protéine. Le rôle de la lipidation n'est pas encore défini, mais elle semble indispensable à l'activité de la protéine SphB1 (Coutte *et al.*, 2003; van Ulsen *et al.*, 2003; Ashgar *et al.*, 2007).

### *Le domaine passager*

Le domaine passager est le domaine qui supporte la fonction de la protéine et qui va être sécrété dans le milieu extracellulaire. C'est la partie la plus variable des autotransporteurs car il n'existe pas de séquence, de taille ou de fonction consensus liées au domaine passager. Cependant, la plupart des autotransporteurs possèdent une activité en rapport avec la virulence des bactéries (Henderson and Nataro, 2001). Certains domaines passagers ont en effet une activité enzymatique (protéase, lipase ou estérase), une action sur la mobilité des bactéries dépendante de l'actine, une action toxique, cytotoxique ou immunomodulatrice.

Malgré la diversité de séquence primaire, ce domaine est construit sur une structure centrale en hélice β droite, comme la pertactine, une adhésine de *Bordetella pertussis* impliquée dans l'activation de l'hémagglutinine filamenteuse et ainsi la fixation de la bactérie aux hématies (Junker *et al.*, 2006). D'autres autotransporteurs, tels que Hbp et EspP chez *E.*

*coli* (Otto *et al.*, 2005; Khan *et al.*, 2011) ou Hap chez *Haemophilus influenzae* (Meng *et al.*, 2011), ont montré en plus de l'hélice β centrale plusieurs domaines globulaires associés sortant de la structure hélicoïdale de base. Parmi ces domaines globulaires, nous pouvons citer un domaine α en N-terminal responsable de l'activité sérine protéase de certains autotransporteurs comme Hbp dénommés SPATEs (serine protease autotransporters of *Enterobacteriaceae*).

Une séquence auto-chaperonne associée à ce domaine a été récemment découverte chez la protéine BrkA (Oliver *et al.*, 2003). Cette région est considérée comme faisant partie du domaine passager ou comme un domaine à part (domaine de liaison) selon les études. Elle forme deux brins β parallèles très stables et qui est la première région du domaine passager exposée au milieu extracellulaire, semble être indispensable pour le bon repliement de la protéine à la surface de la bactérie.

### *Le domaine de translocation*

Le domaine de translocation est très conservé entre les différents autotransporteurs (Yen *et al.*, 2002). Il se divise en deux sous-domaines : un domaine de liaison, qui est une petite région structurée en hélice α, et un domaine β qui forme un pore membranaire en s'insérant dans la membrane externe des bactéries.

Le domaine en hélice α de 21 à 39 acides aminés précédent le tonneau β est impliqué dans la formation d'une épingle à cheveux avec le domaine passager, permettant la sécrétion du domaine passager à travers le tonneau β (Desvaux, Nicholas J Parham, *et al.*, 2004). Le tonneau β est composé chez les autotransporteurs de 14 brins β de 9 à 12 résidus, reliés entre eux par des boucles de longueur variable. Une hélice α est également présente au centre de ce tonneau, pour maintenir sa forme et le laisser fermer dans le but de préserver l'intégrité membranaire. Il est constitué d'un motif en C-terminal reconnu par le complexe Bam (β-barrel assembly machinery) pour son insertion dans la membrane.

## d. Sécrétion et fonction des autotransporteurs

### *Sécrétion des autotransporteurs*

Les autotransporteurs sont synthétisés dans le cytoplasme de la bactérie et sont pris en charge par le système Sec pour traverser la membrane interne. Comme toutes les autres protéines de la

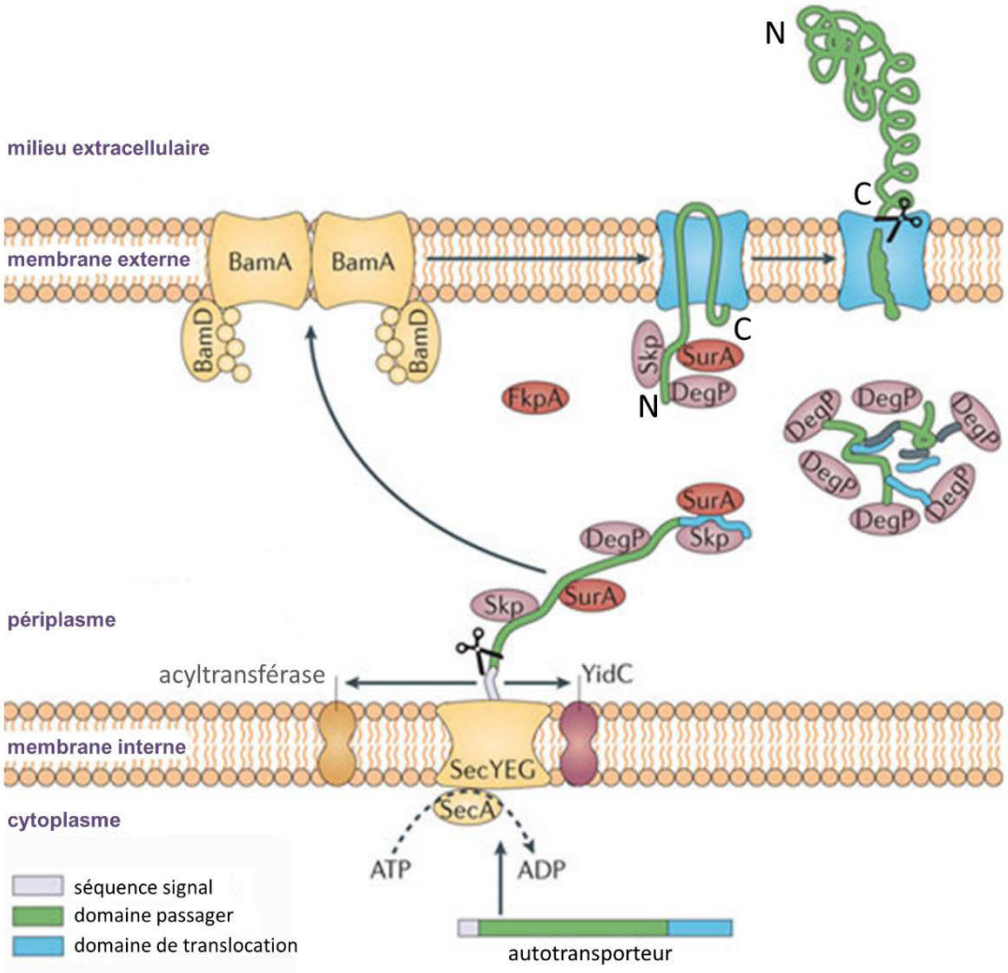



**Figure 29** : Sécrétion des autotransporteurs monomériques
(image modifiée d'après Leyton *et al.*, 2012)

membrane externe, l'interaction avec des protéines chaperonnes périplasmiques appartenant au complexe Bam est nécessaire pour conserver une conformation compatible avec la translocation et intervient dès la sortie du système Sec, avant même le clivage de la séquence signal (**Figure 29**) (Knowles *et al.*, 2009).

Ces protéines, parmi lesquelles sont retrouvées les 3 principales SurA, Skp et DegP, sont nécessaires à la fois pour le domaine passager, mais également pour le domaine de translocation des autotransporteurs (Ieva *et al.*, 2011). Une fois la protéine dans le périplasme, une interaction du domaine de translocation avec le complexe Bam va lui permettre de se replier et de s'insérer dans la membrane externe (Sauri *et al.*, 2009; Ruiz-Perez *et al.*, 2009). Enfin, la dernière étape correspond à la sécrétion du domaine passager dans le milieu extracellulaire. La partie C-terminale du domaine passager comprenant la séquence auto-chaperonne est sécrétée en premier, et la partie N-terminale en dernier. Cette étape semble plus compliquée qu'il n'y parait, au regard de l'absence d'énergie au niveau de la membrane, du petit diamètre du tonneau β et de sa fermeture par une hélice α. Deux principaux modèles ont été mis en avant pour cette étape (**Figure 30**) (Leyton *et al.*, 2012; Gawarzewski *et al.*, 2013).

Le premier modèle (modèle « hairpin ») suppose que la partie terminale du domaine passager traverse le tonneau β et forme une épingle à cheveux à la sortie du tonneau β, et ainsi se retrouve en premier dans le milieu extracellulaire. Chez certains autotransporteurs, la partie C-terminale du domaine passager est bien celle retrouvée en premier à l'extérieur, ce qui concorde avec ce modèle (Junker *et al.*, 2006). Cependant, cela suppose que les 2 brins β traversent en même temps le tonneau β qui est stabilisé en permanence par une hélice α, ce qui ne semble pas possible, au vu de l'espace à l'intérieur du tonneau. Il a donc été proposé un modèle où le tonneau β s'étire légèrement pour agrandir le pore et permettre la sécrétion des autotransporteurs (Pavlova *et al.*, 2013).

Dans le deuxième modèle (modèle « BamA »), le complexe Bam ne sert pas uniquement à l'insertion du domaine de translocation dans la membrane externe mais participe également à la sécrétion du domaine passager. Cependant, le rôle de la protéine BamA reste flou actuellement. Il a été émis l'hypothèse que BamA servait de pore pour le domaine passager, et que le domaine de translocation ne servirait que « d'adressage » des protéines à la membrane externe, mais cette hypothèse est contredite par le fait que la substitution du domaine de translocation par un autre domaine β affecte la sécrétion du domaine passager (Saurí *et al.*, 2012). Une autre hypothèse suggère que le domaine de translocation s'insère entièrement dans

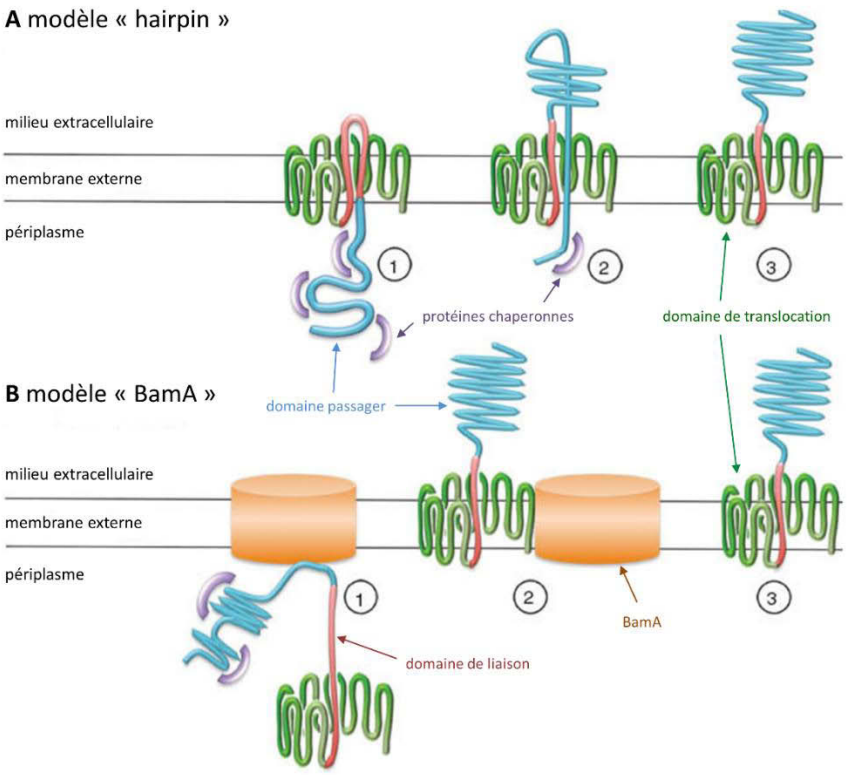



**Figure 30** : Deux modèles suggérés pour la sécrétion du domaine passager à travers la membrane externe (image modifiée d'après Gawarzewski *et al.*, 2013)

le pore formé par BamA, mais cela suppose une taille de pore de BamA suffisante pour accommoder le tonneau β d'un autotransporteur. Une troisième hypothèse a donc été mise en avant. Elle suggère que l'interaction de BamA avec le domaine de translocation de l'autotransporteur permette à ce dernier d'adopter une conformation transitoire compatible avec la sécrétion du domaine passager.

La sécrétion du domaine passager, quel que soit le modèle utilisé, nécessite de l'énergie. Cependant, aucune source d'énergie connue n'est présente au niveau du périplasme ou de la membrane externe. La principale hypothèse avancée fait donc intervenir le repliement de la séquence auto-chaperonne ou domaine de liaison comme source d'énergie pour la sécrétion du domaine passager. En effet, il a déjà été démontré le rôle de ce domaine dans le bon repliement des protéines sécrétées, mais cette séquence pourrait également apporter l'énergie nécessaire à cette sécrétion. Ce domaine, plus stable que le domaine passager, est le premier domaine exposé à la surface des bactéries. Son repliement dans le milieu extracellulaire permettrait d'attirer le domaine passager à travers la membrane externe et ainsi de fournir l'énergie nécessaire à sa complète sécrétion (Junker *et al.*, 2006; Peterson *et al.*, 2010). Les autotransporteurs contiennent donc eux-mêmes l'énergie nécessaire à leur sécrétion.

Cependant, une nouvelle hypothèse concernant la sécrétion de certains autotransporteurs est récemment apparue. Elle correspond à une combinaison entre les deux précédents modèles et implique un complexe TAM (translocation and assembly module), composé d'une protéine TamA présente dans la membrane externe et faisant partie de la superfamille BamA, et la protéine TamB présente dans la membrane interne. L'association de ces deux protéines serait nécessaire à la sécrétion des autotransporteurs, avec un rôle de la protéine TamA proche de celui de la protéine TpsB du système de sécrétion SST5 de type Vb (Selkrig *et al.*, 2012; Gawarzewski *et al.*, 2013).

### *Clivage des autotransporteurs*

Une fois la protéine sécrétée, elle peut rester ancrée à la membrane externe, ou être clivée afin de libérer le domaine passager dans le milieu extracellulaire. Les mécanismes de clivage sont variés.

Certains autotransporteurs sont clivés par d'autres protéines présentes dans la membrane externe. C'est par exemple le cas pour IcsA de *Shigella flexneri* clivée par IcsP (Shere *et al.*,

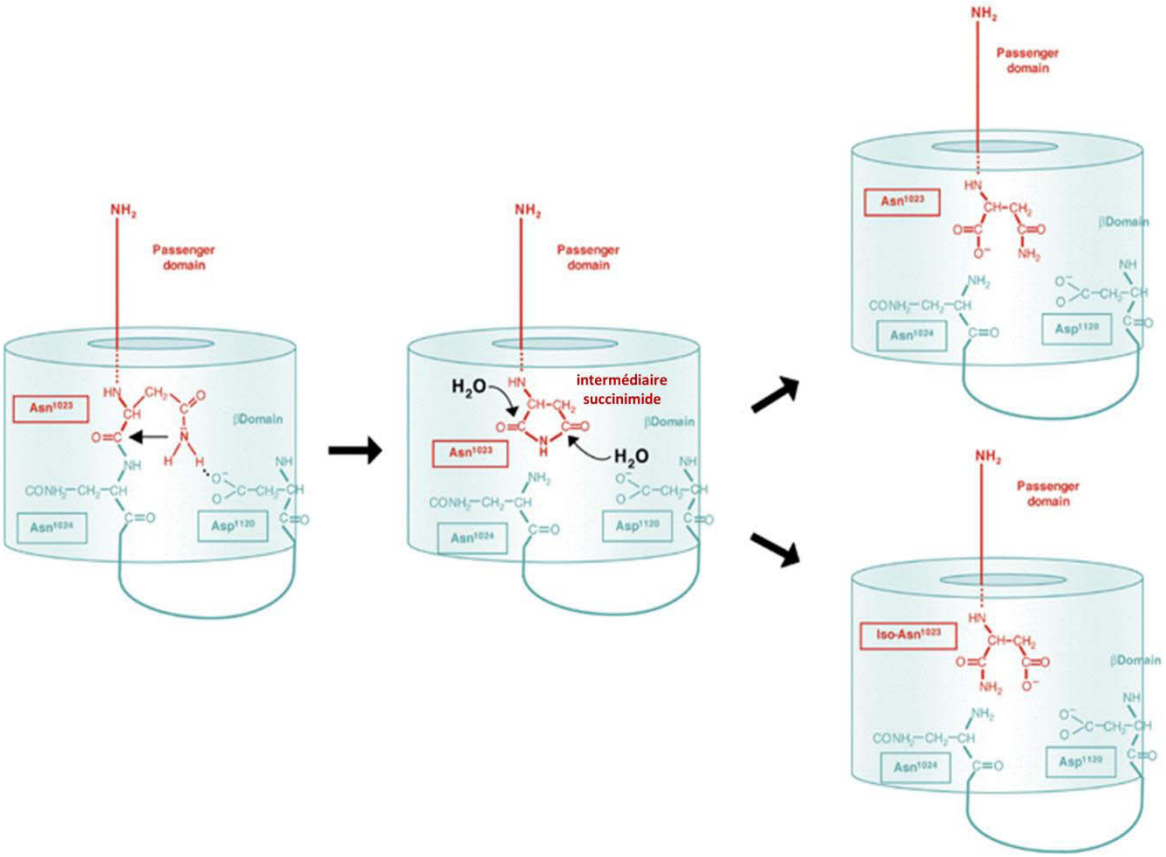



**Figure 31** : Système d'auto-clivage des autotransporteurs de type SPATE (image modifiée d'après Dautin *et al.*, 2007)

1997) ou l'IgA protease de *Neisseria meningitidis* clivée par NalP, un autre autotransporteur présent chez la bactérie (van Ulsen *et al.*, 2003). Mais dans la majorité des cas, l'autotransporteur est clivé par un système autocatalytique. Par exemple pour l'autotransporteur Hap chez *Haemophilus influenzae*, le mécanisme de clivage est intermoléculaire : une molécule de Hap en coupe une autre, grâce à un domaine sérine protéase (Fink *et al.*, 2001).

Pour d'autres autotransporteurs, comme ceux appartenant aux SPATEs, le clivage est intramoléculaire et fait intervenir deux résidus très conservés au sein des autotransporteurs : un résidu aspartate (Asp) présent dans le domaine de translocation et une asparagine (Asn) présente au début du domaine de liaison. La liaison hydrogène présente entre ces deux résidus conduit à une attaque nucléophile du groupement amide de l'Asn sur la liaison peptidique entre les deux asparagines successives. Cette réaction conduit à un auto-clivage faisant suite à la cyclisation de l'asparagine en succinimide. L'hydrolyse du succinimide conduit à un mélange d'asparagine et d'iso-asparagine (**Figure 31**) (Dautin *et al.*, 2007; Barnard *et al.*, 2007; Weiss and Brockmeyer, 2012).

Chez de nombreux autres autotransporteurs, le système de clivage reste à ce jour inconnu.

Certains autotransporteurs possèdent des particularités qui leur sont propres et qui sont souvent nécessaires à la fonction de la protéine. Après la sécrétion de la protéine dans le milieu extracellulaire, elle peut intégrer des groupements lipidiques (NalP, *Neisseria meningitidis*) ou sucrés (AIDA-I, *E. coli*) (Benz and Schmidt, 2001; van Ulsen *et al.*, 2003). D'autres autotransporteurs peuvent avoir une sécrétion polaire, comme par exemple IcsA chez *Shigella flexneri*, impliqué dans les mouvements intracellulaires de la bactérie grâce à son interaction avec les protéines du cytosquelette (Goldberg *et al.*, 1993) .

### *Fonctions des autotransporteurs*

Les autotransporteurs possèdent un domaine de translocation bien conservé entre les différentes familles. Par contre, la diversité du domaine passager permet à ces autotransporteurs d'avoir des activités très différentes. A ce jour, tous les domaines passagers

caractérisés semblent être impliqués dans la virulence des bactéries. Ils ont été classés en 11 groupes, selon

**Tableau 4** : Classification des autotransporteurs monomériques en fonction de l'activité du domaine passager (tableau d'après Henderson *et al.*, 2004)

| Protéines | Fonction(s) | espèces bactériennes |
|---|---|---|
| **cluster 1** | | |
| **Ssp** | protéase | *Serratia marcescens* |
| **Ssp-h1, Ssp-h2** | inconnue | *Serratia marcescens* |
| **PspA, PspB** | inconnue | *Pseudomonas fluorescens* |
| **Ssa1** | inconnue | *Pasterurella haemolytica* |
| **SphB1** | maturation de l'adhésine FHA | *Bordetella pertussis* |
| **AspA/NalP** | maturation d'autotransporteurs | *Neisseria meningitidis* |
| **cluster 2** | | |
| **VacA** | cytotoxine vacuolisante | *Helicobacter pylori* |
| **clusters 3 et 4** | | |
| **AIDA-1** | adhésion/invasion, formation de biofilm | *Escherichia coli* |
| **IcsA** | mobilité intracellulaire | *Shigella flexneri* |
| **MisL** | inconnue | *Salmonella enterica* |
| **TibA, Ag43** | adhésion/invasion, formation de biofilm | *Escherichia coli* |
| **ShdA** | adhésion | *Salmonella enterica* |
| **AutA** | inconnue | *Neisseria meningitidis* |
| **cluster 5** | | |
| **Tsh/Hbp** | hémagglutinine, activité mucinase | *Escherichia coli* |
| **SepA** | invasion, inflammation | *Shigella flexneri* |
| **EspC** | activité cytotoxique | *Escherichia coli* |
| **EspP** | activité cytotoxique, dégradation du complément | *Escherichia coli* |
| **Pet** | inflammation, activité cytotoxique | *Escherichia coli* |
| **Pic** | activité mucinase | *Escherichia coli* |
| **SigA** | activité cytopathique | *Shigella flexneri* |
| **Sat** | cytotoxine vacuolisante, autophagie | *Escherichia coli* |
| **Vat** | cytotoxine vacuolisante | *Escherichia coli* |
| **EatA, EpeA** | activité mucinase | *Escherichia coli* |
| **EspI** | dégradation de protéines du plasma | *Escherichia coli* |
| **EaaA, EaaC** | inconnue | *Escherichia coli* |
| **cluster 6** | | |
| **Pertactine** | adhésion | *Bordetella pertussis* |
| **BrkA** | résistance au sérum | *Bordetella pertussis* |
| **Tef** | colonisation trachéale | *Bordetella pertussis* |
| **Vag8** | inconnue | *Bordetella pertussis* |
| **cluster 7** | | |
| **PmpD** | adhésion | *Chlamydia trachomatis* |
| **Pmp20, Pmp21** | adhésion | *Chlamydophila pneumoniae* |
| **cluster 8** | | |
| **IgA1 protease** | dégradation des IgA1 | *Neisseria meningitidis, Haemophilus influenzae* |
| **Hap** | adhésion | *Haemophilus influenzae* |
| **App** | adhésion | *Neisseria meningitidis* |
| **cluster 9** | | |
| **rOmpA, rOmpB** | adhésion/invasion | *Rickettsia ricketsii* |
| **cluster 10** | | |
| **ApeE** | hydrolyse d'esters | *Salmonella enterica* |
| **EstA** | activité lipolytique | *Pseudomonas aeruginosa* |
| **McaP** | activité lipolytique, adhésion | *Morawella catarrhalis* |
| **cluster 11** | | |
| **BabA, SabA, AlpA** | adhésion | *Helicobacter pylori* |

les relations phylogénétiques entre les domaines passager (**Tableau 4**) (Henderson *et al.*, 2004; Weiss and Brockmeyer, 2012).

Certains autotransporteurs avec une fonction identique ont également été classés par familles. Ainsi, nous pouvons retrouver en sein des clusters 3 et 4 la famille des SAATs (Self-Associating AutoTransporters) comprenant notamment AIDA-I, TibA et Ag43 (Klemm *et al.*, 2006). Cette famille comprend des protéines multifonctionnelles impliquées dans l'auto-agrégation des bactéries, l'adhésion et l'invasion des cellules épithéliales et la formation de biofilm. D'autres autotransporteurs possèdent une activité adhésine, avec notamment la pertactine, présente chez *Bordetella pertussis*. De la même manière, la famille des SPATEs (Serine Protease AutoTransporters of *Enterobacteriaceae*) regroupe plusieurs autotransporteurs du cluster 5 ayant une activité serine protease, mais avec des activités diverses.

Cependant, cette classification par famille est imparfaite, car certains autotransporteurs possèdent plusieurs fonctions. C'est par exemple le cas de la protéine Hap de *Haemophilus influenzae* qui possède un domaine impliqué dans l'agrégation entre les bactéries commun aux SAATs et un domaine sérine protéase commun aux SPATEs (Meng *et al.*, 2011).

### *Les autotransporteurs à activité sérine peptidase des Enterobacteriaceae*

Les sérine peptidases sont parmi les enzymes sécrétées les plus abondantes et les plus diversifiées sur le point fonctionnel (Dautin, 2010; Ruiz-Perez and Nataro, 2014). Elles sont retrouvées aussi bien chez des organismes procaryotes qu'eucaryotes, et clivent des liaisons peptidiques grâce au résidu sérine actif, ainsi que deux autres résidus (souvent l'histidine et l'aspartate) qui forment avec la sérine le site catalytique de l'enzyme.

La famille des SPATEs comprend entre autres les protéines suivantes : EspP (extracellular serine protease plasmid (pO157)-encoded) et EpeA (EHEC plasmid-encoded autotransporter) des EHEC, Pet (plasmid-encoded toxin) des EAEC, Pic (protease involved in intestinal colonization) des EAEC, UPEC et *Shigella*, EspC (EPEC secreted protein C) et Hbp (hemoglobin protease) chez les EPEC, Sat (secreted autotransporter toxin) chez les UPEC, Tsh (temperature-sensitive hemagglutinin) et Vat (vacuolating autotransporter toxin) chez les

**Tableau 5** : Principaux autotransporteurs de la famille des SPATEs ; STEC : *E. coli* producteurs de Shiga toxines (tableau d'après Yen *et al.*, 2008)

| protéine | organisme | maladie |
|---|---|---|
| **AidA** | ETEC | diarrhée |
| **Boa** | *Salmonella bongori* | diarrhée |
| **EaaA** | *E. coli* | ? |
| **EaaC** | *E. coli* | ? |
| **EatA** | ETEC | diarrhée |
| **EpeA** | EHEC | diarrhée sanglante, colite hémorragique |
| **EspC** | EPEC | diarrhée |
| **EspI** | STEC | diarrhée, insuffisance rénale |
| **EspP** | EHEC | diarrhée sanglante |
| **Pet** | EAEC | diarrhée |
| **Pic** | EAEC, *Shigella flexneri* | diarrhée sanglante, shigellose |
| **PicU** | UPEC | infection urinaire |
| **PssA** | STEC | diarrhée, insuffisance rénale |
| **Sat** | UPEC | infection urinaire |
| **SepA** | *Shigella flexneri* | shigellose |
| **SigA** | *Shigella flexneri* | shigellose |
| **Tsh/Hbp** | APEC, UPEC, *E. coli* responsables de sepsis | colibacillose, infection urinaire, sepsis |
| **Vat** | APEC | colibacillose, sepsis |

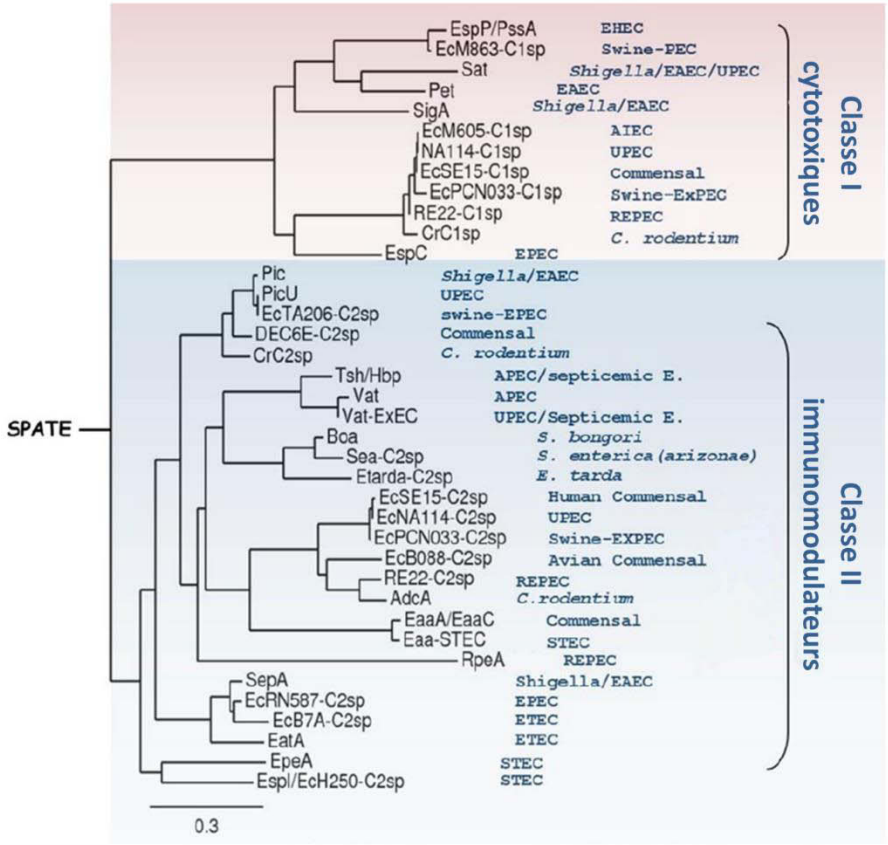



**Figure 32** : Analyse phylogénétique de la séquence en acides aminés du domaine passager des différentes protéines SPATEs avec visualisation des deux groupes (image modifiée d'après Ruiz-Perez et Nataro, 2014)

APEC, EatA (ETEC autotransporter A) chez les ETEC, EspI (*E. coli* secreted protease, island-encoded) chez les EHEC, mais également des autotransporteurs sécrétés par *Shigella flexneri* comme SepA (*Shigella* extracellular protein A) ou SigA, ainsi que l'autotransporteur Boa (*bongori* autotransporter) sécrété par *Salmonella bongori* (**Tableau 5**).

Ces autotransporteurs possèdent une séquence signal anormalement longue avec plus de 49 acides aminés. Ce type de séquence signal est retrouvé dans environ 10% des autotransporteurs, de la même manière que les protéines sécrétées par le système de sécrétion à deux partenaires Vb.

Selon l'homologie structurale et la fonctionnalité du domaine passager, les SPATEs ont ensuite été divisées en 2 groupes : les SPATEs de classe I qui possèdent une activité cytotoxique, et les SPATEs de classe II qui ont des propriétés immunomodulatrices lectine-like (**Figure 32**).

Le domaine passager est donc très variable, avec de 954 à 1050 résidus, et 25 à 50% de similarité. Cependant, l'activité protéase repose toujours sur une sérine active, située au niveau N-terminal du domaine passager, au sein d'un motif conservé GD**S**GSP entre les résidus 250 et 270. Deux autres résidus, une histidine (H) et un aspartate (D), permettent l'activation de cette sérine. De plus, la plupart des domaines passager se replient comme des hélices β, formant une sorte de colonne entourée des domaines fonctionnels (**Figure 33**) (Kajava and Steven, 2006).

Enfin, le domaine de translocation est très conservé. Il est constitué de 277 résidus, avec de 60 à 99% de similarité. Certaines sérine protéases restent intactes au niveau de la membrane externe après leur sécrétion, alors que d'autres sont clivées entre le domaine passager et le domaine de translocation β, ce qui permet de libérer le domaine passager dans le milieu extracellulaire à travers la membrane externe. Le clivage s'effectue au niveau du motif conservé FxxEV**NN**LNK localisé du côté C-terminal du domaine de liaison, entre les deux motifs Asn (cf §IV.3.7.d.). Ce processus est intramoléculaire et autocatalytique, catalysé par le domaine de translocation.

La plupart des gènes codant ces autotransporteurs sont situés sur des éléments génétiques mobiles (plasmides, phages, ilots de pathogénicité) et sont entourés de séquences d'insertion IS. De plus leur pourcentage en bases GC est significativement différent du reste du génome bactérien, ce qui peut supposer que ces gènes ont été acquis lors de transferts horizontaux de

gènes, ou de spéciation et duplication de gènes (Yen *et al.*, 2002; Dautin, 2010). Par ailleurs, la conservation quasi-parfaite du domaine de translocation et la diversité des domaines passager

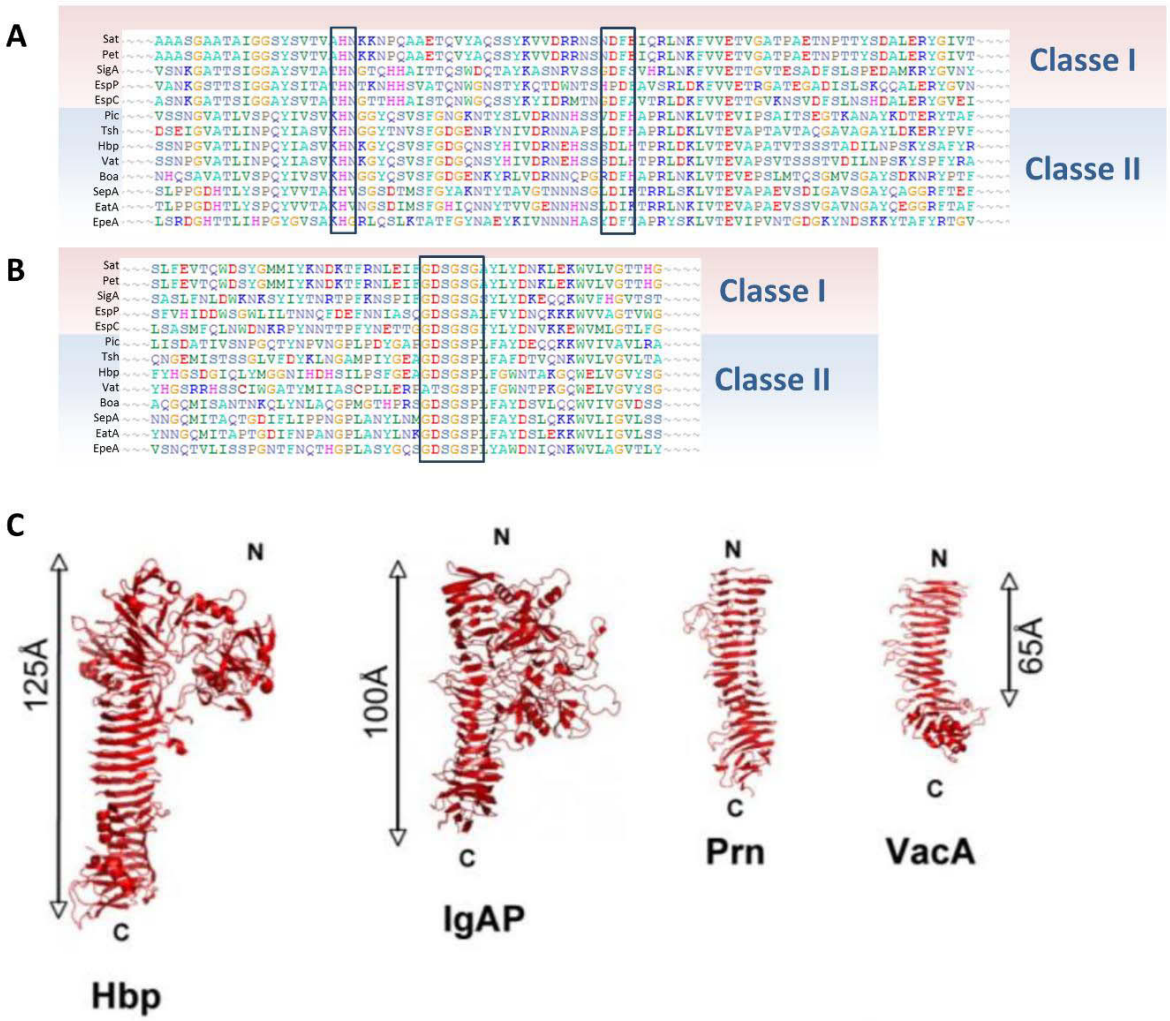



**Figure 33** : Le domaine passager des SPATEs. (**A**) et (**B**) Alignement du domaine passager des SPATEs avec les différents motifs conservés ; (**C**) structure cristallographique du domaine passager des SPATEs (N : N-terminal ; C : C-terminal)
(image d'après Dautin *et al.*, 2010)

des différentes SPATEs supposent que les autotransporteurs sont nés de la fusion entre une protéine de translocation et différentes protéines de virulence bactérienne (Loveless and Saier, 1997).

L'expression des différentes SPATEs semble régulée par des stimuli environnementaux (température, oxygène, pH...) et/ou par quorum sensing. Par exemple, les autotransporteurs SigA, Pic, EpeA, EspP, EspC, Tsh/Hbp et SepA sont régulés par la température, avec une expression optimale à 37°C. De même, l'expression de certaines SPATEs dépend du pH du milieu, avec une expression optimale à pH alcalin : 7 à 9 pour EspP, 8 pour EpeA ou 9 pour Pic (Benjelloun-Touimi *et al.*, 1995; Brunder *et al.*, 1997; Stathopoulos *et al.*, 1999; Henderson *et al.*, 1999; Al-Hasani *et al.*, 2000; Parreira and Gyles, 2003; Leyton *et al.*, 2003). Enfin, il a également été montré un rôle du quorum sensing dans l'expression de la protéine EspC par la voie de régulation Per/Ler (Elliott *et al.*, 2000).

Le role des autotransporteurs et leur regulation par le micro-environnement n'a pas été étudié chez les AIEC.

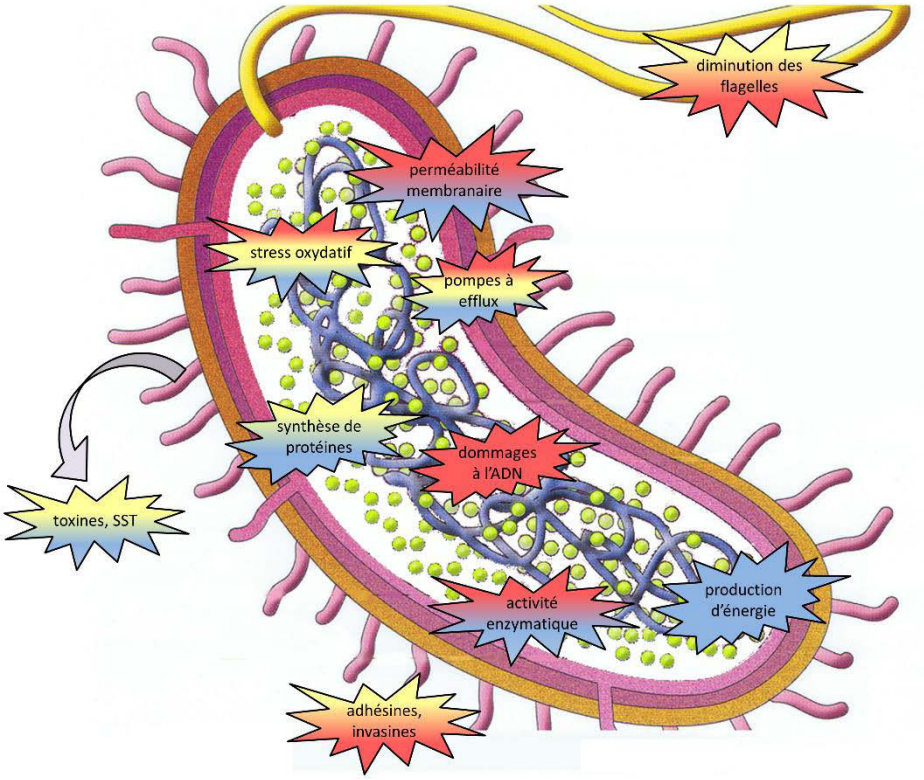



**Figure 34** : Rôle du pH, des sels biliaires et du mucus sur les différentes fonctions bactériennes (l'impact du pH est représenté en bleu, celui des sels biliaires en rouge et celui du mucus en jaune).

# V. Etude de l'influence de l'environnement physico-chimique de l'intestin sur la colonisation bactérienne

Au sein du tractus digestif, les conditions physico-chimiques varient afin de permettre une digestion optimale des aliments tout en protégeant l'épithélium des agressions extérieures. Ainsi, le pH varie tout le long du tractus, de même que l'épaisseur de la couche de mucus recouvrant les cellules épithéliales. Ces deux éléments vont donc avoir un impact sur la colonisation bactérienne. D'autre part, au niveau de l'intestin grêle, la présence de sels biliaires joue également un rôle important sur les bactéries. Les effets de ces trois facteurs environnementaux sur les bactéries sont regroupés **Figure 34** et détaillés par la suite.

## 1. Influence du pH

Le pH peut être mesuré directement *in vivo* à l'aide de capsules radiotélémétriques ou de sondes contenant des électrodes sensibles au pH. Le pH de la lumière intestinale a ainsi été déterminé chez des volontaires sains. Il est très acide au niveau de l'estomac (1,5-3,5), augmente aux alentours de 5,5 dans le duodénum, puis monte régulièrement pour atteindre environ 7,5 au niveau de l'iléon terminal. Le pH diminue à 5,7 au niveau du caecum avant de remonter régulièrement dans le colon, jusqu'à atteindre entre 6,1 et 7,5 au niveau du rectum (Fallingborg, 1999; Nugent *et al.*, 2001). Le pH très acide au niveau de l'estomac possède trois grandes fonctions : (i) l'activation du pepsinogène et la dénaturation des protéines, (ii) l'augmentation de l'absorption du calcium et du fer, et (iii) la neutralisation des micro-organismes (Tennant *et al.*, 2008).

### 1) Rôle du pH sur la survie et la croissance bactérienne

Le pH est un élément clé pour l'homéostasie cellulaire bactérienne. En effet, il joue un rôle important au niveau de la production d'énergie par la chaine respiratoire, de la perméabilité membranaire, ainsi que de l'activité métabolique des enzymes bactériennes. Chez les bactéries à Gram négatif, la membrane externe n'est pas une barrière imperméable aux mouvements des protons car les porines ont un diamètre suffisant pour les laisser passer. C'est

donc au niveau de la membrane interne que les bactéries ont des mécanismes de résistance aux variations extracellulaires du pH, afin de conserver leur intégrité (Lund *et al.*, 2014).

L'effet du pH du milieu extracellulaire sur le métabolisme des bactéries a été observé pour la première fois il y a plus de 70 ans (Gale and Epps, 1942). Il avait été démontré qu'un changement de pH extracellulaire conduisait à une altération de la constitution en enzymes des bactéries, afin de lutter contre ce changement physico-chimique tout en maintenant certaines activités essentielles à un niveau constant.

Les bactéries peuvent être divisées en trois grands groupes de bactéries en fonction de leur résistance au pH du milieu : (i) les bactéries acidophiles, qui préfèrent un milieu acide pour se multiplier, comme par exemple *Lactobacillus* ; (ii) les bactéries alcalophiles, qui préfèrent des milieux à pH alcalin, comme *Pseudomonas* ou *Vibrio* ; et (iii) les bactéries neutrophiles qui se développent dans des milieux dont le pH est compris entre 5,5 et 8,5, avec une préférence pour un pH voisin de 7. La plupart des bactéries pathogènes et/ou liées à l'écosystème humain font partie de cette dernière catégorie (Booth, 1985).

La croissance bactérienne à un pH très différent du pH intracellulaire conduit à une réponse bactérienne qui a pour but, tout en maintenant le pH intracellulaire à sa valeur optimale, de préparer la bactérie à une exposition à des valeurs de pH encore plus délétères (Small *et al.*, 1994; Maurer *et al.*, 2005). Ainsi, pour lutter contre les conditions extérieures, les bactéries sont capables de diminuer leur perméabilité membranaire aux ions H+ pour tamponner le pH du cytoplasme, mais également de produire des déchets métaboliques acides ou alcalins et/ou de développer des systèmes de pompe afin d'améliorer le pH extracellulaire. La réponse des bactéries à un pH délétère implique donc des systèmes de détection et de régulation (Dilworth and Glenn, 1999).

Une augmentation du pH extracellulaire conduit à une entrée de protons, avec une répression de gènes codant des protéines impliquées dans la synthèse des flagelles, et une régulation de nombreuses protéines périplasmiques et membranaires (Stancik *et al.*, 2002; Maurer *et al.*, 2005). Des désaminases sont également induites, comme la tryptophane déaminase et la sérine déaminase qui génèrent des acides pour lutter contre l'augmentation de pH (Yohannes *et al.*, 2004).

Une diminution du pH extracellulaire accélère la consommation d'acides et l'export de protons *via* des pompes à protons. Cela conduit également à l'induction d'un stress oxydatif

avec synthèse de protéines de chocs thermiques (Maurer *et al.*, 2005). Le pH régule également l'expression de certaines enzymes et de transporteurs du catabolisme. Par exemple, l'opéron maltose chez *E. coli* et la porine OmpF sont diminués à pH acide (Chagneau *et al.*, 2001). Une diminution du métabolisme enzymatique est également constatée puisque le pH n'est plus optimal pour le fonctionnement des enzymes et la disponibilité des cofacteurs enzymatiques est moindre. Ainsi, la croissance bactérienne est partiellement ou totalement inhibée (Gale and Epps, 1942; Dilworth and Glenn, 1999; Lund *et al.*, 2014). Si le pH est très acide, comme par exemple dans l'estomac, les bactéries pourront être détruites.

Les bactéries ont trouvé différents moyens pour survivre à un pH très acide. Le système Gad (glutamate décarboxylase), retrouvé chez de nombreuses bactéries, intervient dans le maintien du pH intracellulaire en consommant des protons dans une réaction de décarboxylation produisant du γ-aminobutyrate (GABA) à partir du glutamate. Le GABA est ensuite transformé en succinate qui rentre dans le cycle de Krebs pour former de l'adénosine triphosphate (ATP) (Castanie-Cornet *et al.*, 1999; Tucker *et al.*, 2002; Feehily and Karatzas, 2013).

Ainsi, l'homéostasie bactérienne intestinale est bien sûr liée aux pressions partielles en oxygène de l'intestin, mais également aux différentes zones de pH du tube digestif.

Les bactéries entéropathogènes doivent passer par l'estomac (pH < 3) pendant minimum 2 heures avant de pouvoir coloniser l'intestin (Small *et al.*, 1994). Certaines bactéries retrouvées au niveau intestinal comme *Escherichia coli* ou *Shigella flexneri* peuvent survivre plusieurs heures à un pH compris entre 2 et 3. Il a également été observé une forte résistance à l'acidité gastrique des EHEC O157:H7 (Benjamin and Datta, 1995). Cette résistance à l'acidité est portée par le gène *rpoS* codant le facteur σ38, régulateur transcriptionnel de nombreux gènes exprimés pendant la phase stationnaire de croissance (Small *et al.*, 1994; Cheville *et al.*, 1996). *Yersinia enterocolitica* semble l'espèce la plus résistante à l'acidité, avec une survie possible avec un pH de l'ordre de 2. Cette meilleure résistance de *Yersinia* au pH est due à la présence d'une uréase. A l'inverse, *Salmonella enterica* Typhimurium a une survie de seulement 4 % lorsque le pH est aux alentours de 4 (Tennant *et al.*, 2008), ce qui explique que la dose infectieuse soit plus importante que celle des shigelles pour contaminer l'homme.

## 2) Influence du pH sur la virulence des bactéries

Le pH semble être un élément important dans l'expression des facteurs de virulence des bactéries pathogènes. Chez *Pseudomonas aeruginosa,* une induction de la synthèse de sidérophores, comme la pyoverdine ou la pyochéline, est observée à pH 5 alors que l'augmentation de la synthèse d'alginates et de protéinases se fait à pH 8 (Harjai *et al.*, 2005). De même, l'acidité du pH du milieu induit des variation importantes de la composition physico-chimique de la membrane de *Neisseria meningitidis*, avec notamment une augmentation de la synthèse de polysaccharide capsulaire conduisant à une augmentation de la virulence des souches (Masson and Holbein, 1985).

Chez les bactéries entéropathogènes, les variations de pH peuvent également induire la production de facteurs de virulence. Un pH acide induit l'expression de l'opéron *ToxR-ToxT* chez *Vibrio cholerae*, l'antigène pH6 de *Yersinia pestis* ou encore l'opéron *phoP-phoQ* chez *Salmonella enterica*, (Lindler *et al.*, 1990; Bearson *et al.*, 1998; Payne *et al.*, 1998; Behari *et al.*, 2001). Un pH alcalin favorise la synthèse et la sécrétion de la toxine thermolabile LT des ETEC (Gonzales *et al.*, 2013). Un pH acide ou basique induit une augmentation de l'expression des gènes codant les toxines A et B de *Clostridium difficile* (Stewart and Hegarty, 2013).

Le pH a également été impliqué dans la capacité d'adhésion des AIEC aux cellules épithéliales intestinales. Pour un pH compris entre 5 et 7, soit le pH retrouvé au début de l'intestin grêle et du colon, la force d'adhésion de la protéine FimH aux résidus mannose est maximale, conduisant à une meilleure adhésion des AIEC au récepteur CEACAM6 (Jacquot *et al.*, 2014). Par ailleurs, la synthèse du facteur GipA, impliquée dans le pouvoir invasif des AIEC ainsi que de leur survie en macrophages, est augmentée à pH acide (Vazeille *et al.*, 2016).

Le pH du milieu extracellulaire peut donc induire la production des facteurs de virulence des bactéries.

## 2. Influence des acides biliaires

Les acides biliaires proviennent de la dégradation du cholestérol par le foie. Ils sont synthétisés sous la forme d'acides biliaires primaires, l'acide cholique et chénodésoxycholique. Ils sont ensuite conjugués avec de la glycine ou de la taurine avant d'être excrétés dans l'intestin. Sous l'action du microbiote intestinal, ils sont ensuite déconjugués et transformés en acides biliaires secondaires (acide désoxycholique et acide lithocolique) (Begley *et al.*, 2005; Ridlon *et al.*, 2006). Outre leur rôle physiologique sur la digestion des lipides, les acides biliaires ont également une action majeure sur les bactéries.

### 1) Rôle des acides biliaires sur la survie et la croissance bactérienne

L'effet des acides biliaires sur les membranes est concentration dépendant. En effet, une concentration élevée de glycocholate est par exemple capable de dissoudre rapidement les lipides membranaires et provoque une dislocation des protéines membranaires, qui conduisent directement à la mort cellulaire (Coleman *et al.*, 1980). A plus faible concentration, les acides biliaires peuvent interagir avec les membranes cellulaires des bactéries et induire des modifications de la perméabilité membranaire, une altération de l'activité de certaines enzymes liées à la membrane et une augmentation du passage transmembranaire de cations divalents (Noh and Gilliland, 1993; Fujisawa and Mori, 1996). Ils vont également modifier les propriétés physico-chimiques de la surface cellulaire, avec notamment une modification de l'hydrophobicité et du potentiel zêta (Waar *et al.*, 2002; Gómez Zavaglia *et al.*, 2002). En fonction de la concentration des acides biliaires, la croissance bactérienne pourra être ainsi ralentie ou stoppée.

En plus de leur effet sur les membranes cellulaires, les acides biliaires, qu'ils soient primaires ou secondaires, sont capables d'induire un stress oxydatif conduisant à des dommages de l'ADN bactérien. Ces dommages provoquent une réponse SOS de la part de la bactérie *via* l'expression du gène *sulA* qui inhibe une étape clé de la division cellulaire (Jones and Holland, 1985). Ceci conduit donc à un arrêt de la croissance allant même parfois jusqu'à la mort cellulaire (Merritt and Donaldson, 2009).

Par ailleurs, les acides biliaires sont des molécules « signal » qui peuvent agir *via* des récepteurs membranaires ou nucléaires pour réguler le métabolisme des lipides, du cholestérol, des acides biliaires et la motilité intestinale (Katsuma *et al.*, 2005; Fiorucci *et al.*, 2010). Par ce biais, ils peuvent agir sur l'immunité locale intestinale et hépatique. L'activation du récepteur nucléaire FXR et du récepteur à la vitamine D par l'acide chénodésoxycholique induit la production de peptides antimicrobiens (D'Aldebert *et al.*, 2009; Sipka and Bruckner, 2014). L'activation du récepteur membranaire TGR5 par les acides biliaires secondaires conduit à une diminution de la synthèse de cytokines proinflammatoires, TNF-α, d'IL-1β et d'IL-6, par les macrophages circulants et hépatiques (Duboc *et al.*, 2014).

Ainsi, les acides biliaires possèdent un effet immunomodulateur et bactéricide, qui se révèle protecteur pour notre organisme. En l'absence d'acides biliaires, une véritable prolifération bactérienne aurait lieu, pouvant conduire à des lésions de la muqueuse et à des translocations bactériennes à travers l'épithélium intestinal (Begley *et al.*, 2005; Inagaki *et al.*, 2006).

Les bactéries intestinales ont développé des mécanismes de résistance aux acides biliaires afin de pouvoir survivre et se multiplier dans l'intestin. Un mécanisme de résistance des bactéries à Gram négatif aux acides biliaires provient d'un constituant du lipopolysaccharide de surface (LPS), l'antigène O. En effet, cet antigène à la surface de la paroi bactérienne constitue une barrière contre les composés extérieurs à la bactérie. L'absence de cet antigène O induit une moindre résistance aux acides biliaires (Gunn, 2000). De plus, de nombreuses bactéries intestinales ou entéropathogènes possèdent des pompes à efflux permettant d'expulser les acides biliaires hors de la cellule. C'est par exemple le cas pour *Escherichia coli*, *Vibrio cholerae* ou *Campylobacter jejuni* (Thanassi *et al.*, 1997; Lin *et al.*, 2003; Chatterjee *et al.*, 2004). Certaines bactéries pathogènes sont même capables de résister à de grandes quantités d'acides biliaires et de provoquer des infections de la vésicule biliaire comme par exemple *Listeria monocytogenes*, *Salmonella enterica* ou les EAEC (Hardy *et al.*, 2004; Joo *et al.*, 2007; Kristoffersen *et al.*, 2007; Crawford *et al.*, 2008).

### 2) Influence des acides biliaires sur la virulence des bactéries

Les acides biliaires primaires et secondaires possèdent un pouvoir antibactérien. Cependant, dans certains cas, les acides biliaires favorisent la colonisation d'espèces pathogènes en induisant la synthèse de leurs facteurs de virulence. Ainsi, la présence d'acide cholique et d'acide désoxycholique dans l'environnement intestinal conduit à l'expression des gènes *ospE1*/*ospE2* qui codent des protéines à la surface de la membrane externe de *Shigella flexneri* impliquées dans l'adhésion initiale de la bactérie à l'épithélium intestinal (Faherty *et al.*, 2012). Les acides chénodésoxycholique et désoxycholique favorisent quant à eux la sécrétion des invasines Ipa *via* le SST3 de *Shigella* (Pope *et al.*, 1995). De la même manière, ces deux acides biliaires favorisent la synthèse de *pili* chez *Campylobacter jejuni*, *C. coli* et *C. fetus*, conférant une meilleure adhésion des bactéries à la muqueuse intestinale (Doig *et al.*, 1996).

La présence d'acides biliaires affecte également la régulation de l'opéron *PhoP-PhoQ* chez *Salmonella enterica*. Cet opéron régule des gènes nécessaires à la virulence des bactéries, leur survie à l'intérieur des macrophages ainsi que la résistance aux peptides antimicrobiens (Guo *et al.*, 1997; Crawford *et al.*, 2008).

Chez *Vibrio cholerae*, les acides biliaires induisent une augmentation de la mobilité de la bactérie et à l'inverse une diminution de la synthèse de la toxine cholérique et des *pili* co-régulés avec la toxine (Gupta and Chowdhury, 1997; Wibbenmeyer *et al.*, 2002). Chez *Vibrio parahaemolyticus,* une bactérie responsable de la majorité des empoisonnements alimentaires au Japon et à Taiwan, il a été observé une augmentation de la synthèse du polysaccharide capsulaire et d'adhérence de la bactérie (Hsieh *et al.*, 2003).

Le biofilm est un élément important dans la virulence des bactéries car il favorise l'implantation des bactéries pathogènes et leur procurent une protection contre les défenses immunitaires de l'hôte et d'éventuels traitements antibiotiques (Sanchez *et al.*, 2016). Il a été montré que la présence de bile dans le milieu module la formation du biofilm. Les sels biliaires favorisent sa formation chez *Salmonella enterica* ou *Listeria monocytogenes* (Guo *et al.*, 1997; Begley *et al.*, 2009) et l'entravent ou provoque sa dispersion chez *Vibrio cholerae* ou *Pseudomonas aeruginosa* (Sanchez *et al.*, 2016).

La bile est capable de réprimer certains facteurs de virulence, comme par exemple le régulateur HilD des protéines du SST3 de *Salmonella enterica* Typhimurium. Ainsi, les protéines d'invasion ne peuvent plus être sécrétées par ce système dans les cellules intestinales cibles, ce qui diminue le pouvoir invasif de la bactérie. Cependant, la répression de la virulence de *Salmonella* en présence de bile reste très limitée car après passage à travers le mucus pour arriver au contact des cellules épithéliales, la concentration en acides biliaires diminue fortement ce qui permet à la bactérie de récupérer son pouvoir invasif (Prouty *et al.*, 2004; Wang *et al.*, 2010; Eade *et al.*, 2016).

Chez *E. coli* O157:H7, l'exposition aux acides biliaires conduit à une diminution de l'expression des gènes du locus *LEE* (locus of enterocyte effacement) ainsi que de certains gènes codant le filament et le « moteur » du flagelle (gène *fliC* et opérons *fliDST* et *flgKL* notamment). A l'inverse, la structure (« corps ») du flagelle est induite par les acides biliaires (opérons *fliLMNOPQR*, *flgAMN*, *flgBCDEFGHIJ* et *fliAZY*) (Hamner *et al.*, 2013). Malgré ces modifications d'expression des gènes codant les constituants des flagelles, aucune modification de leur longueur et/ou de leur nombre n'a pu être montrée.

Chez les AIEC, la présence d'acides biliaires tels que l'acide cholique stimule la synthèse de facteurs de virulence impliqués dans l'adhésion et l'invasion des souches aux cellules intestinales, comme les lpf ou le facteur GipA (Chassaing *et al.*, 2013; Vazeille *et al.*, 2016).

Ces nombreux exemples soulignent le rôle majeur des acides biliaires dans la modulation de la virulence chez les bactéries entéropathogènes. La bile semble donc utilisée par les bactéries comme un signal pour s'adapter à leur nouvel environnement.

## 3. Influence du mucus

La surface de l'épithélium du tractus gastro-intestinal de l'homme est revêtue d'une couche de mucus. Ce dernier constitue une barrière protectrice entre le tissu sous-jacent et les micro-organismes, les toxines qu'ils produisent et d'autres substances potentiellement nocives présentes dans la lumière intestinale.

## 1) Rôle du mucus dans la colonisation bactérienne

Bien que le mucus soit produit par l'hôte pour protéger la surface de l'épithélium gastro-intestinal, il peut aussi favoriser la colonisation par des micro-organismes commensaux ou pathogènes en fonctionnant comme source de nutriments et comme matrice dans laquelle ils peuvent proliférer. La microflore abondante dans l'intestin des mammifères, particulièrement au niveau du côlon, semble être principalement concentrée dans la couche de mucus.

Le mucus est principalement constitué de mucines dont la partie glucidique peut être extrêmement variable. Ce sont essentiellement ces sucres qui participent à la colonisation bactérienne :

- Ils constituent en effet des sites potentiels pour l'adhésion de micro-organismes via leurs adhésines. Ils permettent la colonisation de la muqueuse intestinale, mais dans le même temps l'adhésion des bactéries aux mucines leur empêchent d'adhérer aux cellules épithéliales. Par exemple, Yersinia enterocolitica est capable d'adhérer au mucus via son adhésine YadA, mais cette liaison modifie l'hydrophobicité à la surface de la bactérie et diminue ainsi son adhésion aux cellules épithéliales intestinales (Paerregaard et al., 1991).
- Ils servent au métabolisme des bactéries de la flore commensale qui possèdent des glycosidases comme *Akkermansia muciniphila, Bacteroides thetaiotaomicron*, *Bifidobacterium spp* ou *Ruminococcus spp* (Katayama et al., 2005; Ouwerkerk et al., 2013). Lactobacillus rhamnosus est quant à lui capable d'adhérer aux mucines et d'utiliser leurs fractions glucidiques comme unique source de carbone pour la synthèse d'acides gras à chaine courte comme l'acétate avec génération d'ATP (Sanchez et al., 2010).

Pour franchir cette barrière et atteindre les cellules intestinales, les bactéries entéropathogènes doivent être mobiles et altérer les propriétés physico-chimiques du gel de mucus pour le rendre plus lâche. Ainsi, *Helicobacter pylori* produit une uréase, en milieu acide, qui transforme l'urée en ammoniac et en dioxyde de carbone. Il en résulte une augmentation du pH qui diminue la viscoélasticité du mucus et permet le passage de la bactérie (Celli *et al.*, 2009). D'autres bactéries produisent des protéases qui dégradent les mucines comme Hap pour *Vibrio cholerae* (Silva *et al.*, 2003), Pic pour *Shigella flexneri* (Henderson *et al.*, 1999), StcE et EpeA pour les EHEC (Leyton *et al.*, 2003; Grys *et al.*, 2005), Tsh pour les APEC (avian pathogenic *E. coli*) (Dozois *et al.*, 2000). Le clivage de la

partie protéique des mucines dépolymérise le gel de mucus, ce qui diminue sa viscosité et favorise son franchissement.

## 2) Influence du mucus sur la virulence des bactéries

Le mucus peut induire l'expression de facteurs de virulence chez certaines espèces bactériennes. Chez *Pseudomonas aeruginosa,* le mucus présent au niveau respiratoire favorise la synthèse d'un SST6. Ce système de sécrétion délivre divers effecteurs dans des cellules eucaryotes cibles mais également dans d'autres bactéries ce qui a pour but de favoriser la colonisation de *P. aeruginosa*. La présence de mucus induit également la transcription du facteur OxyR, qui régule la synthèse de protéines impliquées dans la protection des bactéries contre le stress oxydatif, ainsi que d'une chitinase ChiC et d'une protéine de liaison à la chitine CbpD, qui sont considérées comme des adhésines bactériennes mais qui ont également un rôle dans l'acquisition de nutriments (Cattoir *et al.*, 2013).

Le mucus est également utilisé comme un signal chimioattractant pour *Campylobacter jejuni* et induit la surexpression de gènes codant plusieurs facteurs de virulence, dont des cytotoxines, l'antigène B d'invasion, la lipoprotéine de surface JlpA, des systèmes d'efflux et des protéases apparentées à des mucinases (Hugdahl *et al.*, 1988; Tu *et al.*, 2008).

Enfin, le mucus peut également jouer un rôle sur la synthèse des protéines constituant les flagelles. Par exemple, il réprime le gène *flaA* codant la flagelline A chez *C. jejuni* (Tu *et al.*, 2008) et les gènes *flg* codant les constituants du flagelle chez *E. coli* O157:H7, avec une réduction associée de la mobilité de cette dernière (Kim *et al.*, 2012). Le mucus constitue donc, comme les acides biliaires, un signal environnemental qui régule la production des flagelles de bactéries entéropathogènes.

# *Travail expérimental*

# I. Présentation des résultats

L'étiologie de la MC implique à la fois des facteurs génétiques et environnementaux, avec également un rôle central du microbiote intestinal dans la maladie, et notamment les *E. coli* adhérents et invasifs particulièrement présents chez les patients par rapport aux sujets contrôle.

L'ensemble des travaux présentés dans ce travail avait pour objectif de comprendre comment les AIEC arrivaient jusqu'à l'épithélium intestinal, première étape indispensable à leur implantation chez les patients atteints de MC, et quel était l'impact du micro-environnement de l'hôte sur les AIEC.

Une première étude présente un nouveau facteur de virulence des AIEC, une mucinase dénommée Vat-AIEC, favorisant le passage à travers le mucus intestinal, et permettant ainsi aux bactéries d'atteindre plus facilement le récepteur humain CEACAM6 nécessaire à leur adhésion aux cellules. Elle montre notamment le rôle de cette protéine dans l'adhésion et l'invasion de la souche de référence LF82 *in vitro* au sein d'une lignée cellulaire productrice de mucus, et la colonisation de cette même souche *in vivo* dans le tractus gastro-intestinal des souris. Enfin, cette étude montre une surexpression du gène codant cette protéine en présence de conditions physico-chimiques retrouvées au niveau de l'iléon terminal.

Une deuxième étude présente l'effet des sels biliaires sur l'expression globale des gènes de la souche de référence LF82 par technique RNA-seq. L'analyse du transcriptome de la souche LF82 montre une augmentation de l'expression des gènes impliqués dans des processus métaboliques comme la dégradation des amines et des carboxylates, ainsi qu'une diminution d'expression des gènes impliqués dans les synthèses bactériennes. Nous avons notamment montré que les AIEC utilisent l'éthanolamine et le 1,2-propanediol comme sources de nutriments, leur conférant un avantage pour coloniser l'intestin. De plus, les sels biliaires ont un impact sur la synthèse des facteurs de virulence des AIEC et sur la formation de biofilm.

# II. La protéase Vat-AIEC favorise la traversée du mucus par les *E. coli* adhérents et invasifs au niveau intestinal

Publication n°1 :

The Vat-AIEC protease promotes crossing of the intestinal mucus layer by Crohn's disease-associated *Escherichia coli*.

Gibold, L., Garenaux, E., Dalmasso, G., Gallucci, C., Cia, D., Mottet-Auselo, B., Faïs, T., Darfeuille-Michaud, A. †, Nguyen, H. T. T., Barnich, N., Bonnet, R.*, et Delmas, J.*

*Cellular Microbiology*, mai 2016, 18: 617–631.

† Ce travail est dédié à Arlette Darfeuille-Michaud

(* : Contribution équivalente des auteurs)

## 1. Introduction à la publication n°1

De nombreuses études ont montré une augmentation de la prévalence des *E. coli* adhérents et invasifs (AIEC) au niveau des lésions intestinales des patients atteints de MC. Cette présence anormale de bactéries AIEC s'observe à la fois pour des atteintes iléales et coliques de la maladie.

Il a également été montré l'implication des pili de type 1 dans le processus d'adhésion des AIEC aux cellules épithéliales intestinales *via* le récepteur humain CEACAM6. Les flagelles et les protéines de la membrane externe agissent de concert avec les pili de type 1 pour promouvoir l'adhésion bactérienne. Cette adhésion permet par la suite une invasion des cellules épithéliales intestinales par les AIEC, permettant à ces derniers de résister au péristaltisme intestinal et de coloniser la muqueuse, puis d'envahir les tissus sous-jacents et d'activer le système immunitaire.

Le but de cette étude était de comprendre comment les AIEC arrivaient à atteindre les cellules épithéliales malgré la couche de mucus protectrice qui les recouvre. Certaines bactéries pathogènes comme *Shigella flexneri* ou certains pathovars d'*E. coli* sécrètent une protéase à activité mucinolytique (Stathopoulos *et al.*, 1999; Henderson *et al.*, 1999). La présence d'une zone de lyse observée sur un gel de mucine prétraitée avec le surnageant de la souche AIEC LF82 suggérait la présence d'une mucinase. Nous n'avons pas trouvé de séquence de mucinases connues dans le génome de la LF82. Toutefois, une protéine de la souche LF82 annotée « hemoglobin protease » présentait 68 % d'homologie avec la protéine Hbp, un autotransporteur présent chez les APEC qui appartient au sous-groupe des SPATE (Kobayashi *et al.*, 2010) et qui possède une activité mucinolytique. L'analyse des domaines de l'homologue de Hbp dans la souche LF82 suggérait qu'il s'agissait également d'une SPATE. Le domaine enzymatique de la protéine issue de la souche LF82 a été modélisé et comparé à la structure cristallographique de Hbp. La grande similarité entre ces deux structures nous a conduits à étudier l'activité mucinolytique de cette protéine jamais caractérisée et qui a été dénommée Vat-AIEC. Nous avons également démontré son implication dans la colonisation de l'intestin par les AIEC associés à la MC, ainsi que l'impact de l'environnement physico-chimique de l'intestin sur sa synthèse.

## 2. Publication n°1

Cellular Microbiology (2015) doi:10.1111/cmi.12539

# The Vat-AIEC protease promotes crossing of the intestinal mucus layer by Crohn's disease-associated *Escherichia coli*

**Lucie Gibold,[1,2] Estelle Garenaux,[2] Guillaume Dalmasso,[2] Camille Gallucci,[1] David Cia,[3] Benoit Mottet-Auselo,[1,2] Tiphanie Faïs,[1,2] Arlette Darfeuille-Michaud,[2,‡] Hang Thi Thu Nguyen,[2] Nicolas Barnich,[2] Richard Bonnet[1,2,†] and Julien Delmas[1,2*,†]**

[1]*Laboratoire de Bactériologie, Centre Hospitalo-Universitaire Clermont-Ferrand, Clermont-Ferrand, France.*

[2]*Microbes, Intestins, Inflammation et Susceptibilité de l'Hôte, Université d'Auvergne, INSERM U1071, INRA USC2018, Clermont-Ferrand, France.*

[3]*Equipe Biophysique Neurosensorielle, Faculté de Pharmacie, Université d'Auvergne, UMR INSERM 1107, Clermont-Ferrand, France.*

## Summary

**The aetiology of Crohn's disease (CD) involves disorders in host genetic factors and intestinal microbiota. Adherent-invasive *Escherichia coli* (AIEC) are receiving increased attention because in studies of mucosa-associated microbiota, they are more prevalent in CD patients than in healthy subjects. AIEC are associated both with ileal and colonic disease phenotypes. In this study, we reported a protease called Vat-AIEC from AIEC that favours the mucosa colonization. The deletion of the Vat-AIEC-encoding gene resulted in an adhesion-impaired phenotype *in vitro* and affected the colonization of bacteria in contact with intestinal epithelial cells in a murine intestinal loop model, and also their gut colonization *in vivo*. Furthermore, unlike LF82Δ*vat-AIEC*, wild-type AIEC reference strain LF82 was able to penetrate a mucus column extensively and promoted the degradation of mucins and a decrease in mucus viscosity. *Vat-AIEC* transcription was stimulated by several chemical conditions found in the ileum environment. Finally, the screening of *E. coli* strains isolated from CD patients revealed a preferential *vat-AIEC* association with AIEC strains belonging to the B2 phylogroup. Overall, this study revealed a new component of AIEC virulence that might favour their implantation in the gut of CD patients.**

Received 15 March, 2015; revised 7 October, 2015; accepted 20 October, 2015. *For correspondence: E-mail jdelmas@chu-clermontferrand.fr; Tel. (+33) 473754920; Fax (+33) 473754922.
†R.B. and J.D. contributed equally to this work.
‡This work is dedicated to Arlette Darfeuille-Michaud.

© 2015 John Wiley & Sons Ltd

## Introduction

Crohn's disease (CD) is a chronic intestinal disease caused by an abnormal inflammatory response to intestinal microbiota in genetically susceptible hosts (Strober *et al.*, 2007; Xavier and Podolsky, 2007; Kaser *et al.*, 2010; Man *et al.*, 2011). Changes in mucosa-associated microbiota have been observed in CD patients, particularly a decrease in *Faecalibacterium prausnitzii*, a bacterium that has anti-inflammatory properties (Sokol *et al.*, 2008; Sokol *et al.*, 2009; Sokol and Seksik, 2010; Miquel *et al.*, 2013), and an increase in *Escherichia coli* (Neut *et al.*, 2002; Darfeuille-Michaud *et al.*, 2004; Martin *et al.*, 2004; Mylonaki *et al.*, 2005; Baumgart *et al.*, 2007; Kotlowski *et al.*, 2007; Martinez-Medina *et al.*, 2009; Willing *et al.*, 2009; Gevers *et al.*, 2014; Martinez-Medina and Garcia-Gil, 2014). The increased prevalence of *E. coli* in CD patients led to the hypothesis that this bacterium is involved in the aetiology of CD.

As a commensal organism, *E. coli* coexists in harmony with its mammalian host and rarely causes disease except when the gastrointestinal barrier is breached. However, some *E. coli* strains have acquired specific virulence factors that increase their ability to adapt to new niches and allow them to cause a broad spectrum of diseases (Delmas *et al.*, 2015). A hallmark of *E. coli* strains isolated from the intestinal mucosa of CD patients is the ability to adhere to intestinal epithelial cells (IECs). Adhesion is in part mediated by the interaction between type 1 pili expressed by these strains and a specific host receptor overexpressed in CD patients, the glycoprotein carcinoembryonic antigen-related cell adhesion molecule 6 (CEACAM6) (Barnich and Darfeuille-Michaud, 2007; Carvalho *et al.*, 2009). In addition, bacterial flagella and outer membrane proteins act in concert with type 1 pili to promote bacterial

adhesion (Boudeau *et al.*, 2001; Barnich *et al.*, 2003; Rolhion and Darfeuille-Michaud, 2007). This adhesion to IECs promotes the invasion of cells by the bacteria, which have been designated adherent-invasive *E. coli* (AIEC). These features enable AIEC to resist mechanical removal from the intestine and to colonize the mucosa. Furthermore, the invasive ability of AIEC strains allows bacteria to translocate across the human intestinal epithelium, move into the deep tissues and activate immune cells continuously upon arrival. To interact with IECs, AIEC should overcome the protective mucus barrier that lines the surface of the intestinal epithelium.

In the small intestine, a thin mucus layer limits the number of bacteria that can reach the epithelium and the Peyer's patches (Ermund *et al.*, 2013; Pelaseyed *et al.*, 2014). In the large intestine, the mucus layer is dense and comprises an inner and outer mucus layer (Johansson *et al.*, 2014). The outer mucus layer is the natural habitat for commensal bacteria whereas the inner mucus layer is impermeable to them and separates the bacteria from host epithelium. The question of how AIEC can cross this mucus layer has not yet been explored.

The mucus layer consists primarily of gel-forming glycoprotein polymers, the mucins, formed by a protein backbone designated as apomucin that harbours a high density of O-linked glycans. These glycoprotein polymers result in very high molecular weight complexes of many millions of Daltons, and it is responsible for the viscoelastic and antibacterial properties of the mucus gel (Derrien *et al.*, 2010). In the small and large intestines, the predominant secretory mucin is MUC2 (McGuckin *et al.*, 2011).

As mucin may be a predominant oligosaccharide supplier in the human gut, several intestinal bacteria such as *Akkermansia muciniphila*, *Bacteroides thetaiotaomicron*, *Bifidobacterium* spp. or *Ruminococcus* spp. have glycosidases in order to degrade the mucus layer and use mucin oligosaccharides for their own metabolism (Katayama *et al.*, 2005; Ouwerkerk *et al.*, 2013). This mucin degradation leads to a colonization advantage competing for the mucosal niche. Conversely, intestinal *E. coli* pathovars and other enteric pathogens most likely penetrate the mucus layer by promoting mucin degradation with proteases such as Hap in *Vibrio cholerae* (Silva *et al.*, 2003), Pic in *Shigella flexneri* (Henderson *et al.*, 1999), StcE and EpeA in enterohemorrhagic *E. coli* (Leyton *et al.*, 2003; Grys *et al.*, 2005), Tsh in avian pathogenic *E. coli* (APEC) (Dozois *et al.*, 2000) and cysteine proteases in *Entamoeba histolytica* (Lidell *et al.*, 2006; Moncada *et al.*, 2006). However, these enzymes have not been previously found in AIEC strains associated with CD.

In this study, we report a new protease secreted by AIEC that contributed to gut colonization in a murine model by enhancing the spread of bacteria through the mucus layer and the adhesion to IECs. Furthermore, several chemical factors in the ileal environment activated the transcription of *vat-AIEC*, strengthening the role of the Vat-AIEC protein in the intestinal colonization.

## Results

### *AIEC strain LF82 exhibits a mucinolytic activity mediated by the vat-AIEC gene*

Because the mucin-degrading proteases that were previously characterized in *S. flexneri* and *E. coli* are secreted proteins (Henderson *et al.*, 1999; Stathopoulos *et al.*, 1999), the culture supernatants of AIEC reference strain LF82, which was isolated from a chronic ileal lesion of a patient with CD (Darfeuille-Michaud *et al.*, 2004; Barnich *et al.*, 2007; Barnich and Darfeuille-Michaud, 2010), were investigated for mucinolytic activity. Unlike the supernatant collected from a culture of laboratory *E. coli* strain K12, the strain LF82 supernatant induced clear zones of mucin lysis in an agarose-mucin medium (Fig. 1A), as observed in strain EcoR42 supernatant known to contain the Hbp mucinase (Restieri *et al.*, 2007). This result suggested that CD-associated AIEC reference strain LF82 produces a secreted factor that degrades mucins.

The amino acid sequences of the already known mucinases were not found in the genomic sequence of *E. coli* LF82 (Miquel *et al.*, 2010). However, we found that a large ORF of 4131 nucleotides (accession no. CAP74833) encodes a protein that is 49% and 79% similar to Pic and Hbp mucinases respectively. The ORF was predicted to encode a primary translation product of 1376 amino acids with a molecular mass of 148 019 Da. This product displays significant similarity to the autotransporter proteins that harbour a trypsin-like serine protease domain and are clustered in the SPATE family (serine protease autotransporters of the *Enterobacteriaceae*) (Yen *et al.*, 2008). This protein is 97% similar to Vat protein, which exhibits vacuolating activity in cultured cells but not proteolytic activity (Parreira and Gyles, 2003). Analysis of the amino acid sequence reveals an extended leading peptide signal sequence (amino acid 1 to 55), an N-terminal α-domain (amino acids 56 to 1099) and a C-terminal β-domain (amino acids 1100 to 1376), as observed in other SPATE proteins. The α-domain, which usually harbours the biological activity of SPATEs, shares 65.6% and 48% identity with the N-terminal domains of Hbp and Pic, respectively, and 96% with the Vat one. This domain harbours the conserved sequence $G_{258}$DSGS$P_{263}$, which is thought to contain active serine (S260) (Yen *et al.*, 2008), and the residues most likely responsible for active serine activation (His-130 and Asp-158). These suggest that this enzyme is an active serine hydrolase. It is worthy to note that this catalytic domain is different from that of Vat (ATSGSP). The β-domain shares 99% and 100% identity with the C-terminal domains of Hbp and Pic, which are known to be

© 2015 John Wiley & Sons Ltd, *Cellular Microbiology*

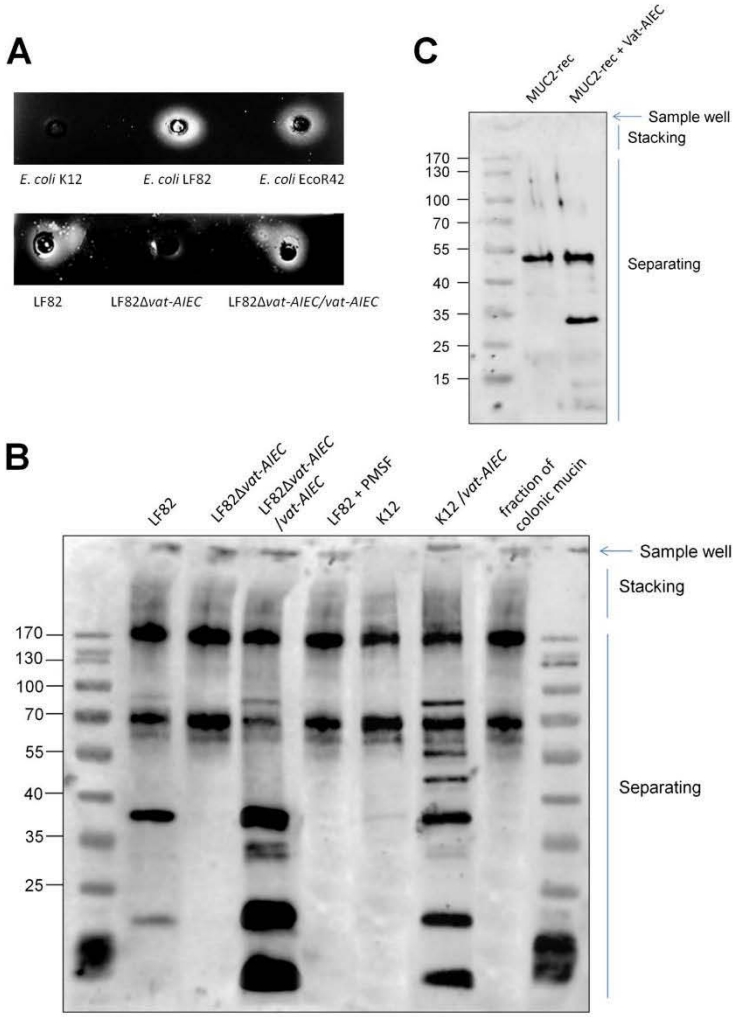



**Fig. 1.** A. Bacterial culture supernatants were tested on a mucin from porcine stomach-supplemented agarose gel staining with amido black. Zones of mucin lysis were observed with culture supernatant of AIEC reference strain LF82, avian pathogenic *Escherichia coli* EcoR42 (which express the Hbp mucinase) and LF82Δ*vat-AIEC*/*vat-AIEC* (transcomplemented strain), unlike culture supernatants of *E. coli* K12 and LF82Δ*vat-AIEC*. These results were obtained from three independent experiments.
B. Reduced colonic mucin extracted from mice gut was incubated overnight with filtered culture supernatants of AIEC LF82 reference strain, LF82Δ*vat-AIEC*, LF82Δ*vat-AIEC*/*vat-AIEC*, LF82 treated with phenylmethylsulfonyl fluoride (PMSF), *E. coli* K12 strain and *E. coli* K12 strain complemented with *vat-AIEC* gene (K12/*vat-AIEC*). Mucin degradation was observed by Western blotting using an anti-MUC2 antibody. PMSF is a common serine protease inhibitor.
C. Recombinant human MUC2 was incubated overnight with cleaved fraction of Vat-AIEC. Mucin degradation was analysed by Western blotting using an anti-MUC2 antibody. MUC2-rec, recombinant human MUC2.

outer membrane translocator domains (Yen *et al.*, 2008; Tajima *et al.*, 2010). The α- and β-domains are joined by a highly conserved sequence that was previously identified as a cleavage site (residues N1099 and N1100) (Fig. S1). These data suggest that LF82 encodes a SPATE-type autotransporter that allows the secretion of an active serine hydrolase.

In order to investigate the mucinolytic activity of this autotransporter named Vat-AIEC (Vat from adherent-invasive *E. coli*), we generated a LF82 isogenic mutant (LF82Δ*vat-AIEC*) and the corresponding transcomplemented mutant (LF82Δ*vat-AIEC*/*vat-AIEC*). Culture supernatants of mutant LFΔ*vat-AIEC* did not yield zones of clearing on crude mucus-forming gels unlike the culture supernatants of LF82 and LF82Δ*vat-AIEC*/*vat-AIEC* (Fig. 1A). Supernatant of LF82 culture degraded non-reduced intestinal mucin extracts as shown by analysis using a Periodic Acid Schiff (PAS) gel (Fig. S2). We then examined whether the Vat-AIEC protein has the ability to hydrolyze intestinal Muc2 by Western blotting. Supernatants of culture of LF82, LF82Δ*vat-AIEC*/*vat-AIEC* and complemented K12/*vat-AIEC* were able to degrade murine intestinal mucins unlike those of LF82Δ*vat-AIEC* and K12 cultures (Fig. 1B). LF16 and LF73, *vat-AIEC*-carrying AIEC strains, also cleaved mucins whereas LF25 and 6259, AIEC strains without the *vat-AIEC* gene failed to degrade mucins (Fig. S2). The C-terminal domain of recombinant human MUC2 was hydrolysed by the cleaved fraction of Vat-AIEC (Fig. 1C). The addition of phenylmethylsulfonyl fluoride,

© 2015 John Wiley & Sons Ltd, *Cellular Microbiology*

a serine protease inhibitor, impaired mucin degradation (Fig. 1B). Moreover, by using site-directed mutagenesis, the putative active site residue at position 260 was changed to isoleucine. No mucinolytic activity against intestinal mucus was observed with this mutant (Fig. S2). These results strongly suggest that the active domain of Vat-AIEC is released outside the cells as reported for Hbp (Otto *et al.*, 1998) and is responsible for the mucinolytic activity. Altogether, our results suggest that the Vat-AIEC protein is an enzyme with a serine protease catalytic domain, responsible of the mucinase activity of LF82 strain.

To determine whether the *vat-AIEC* gene was preferentially associated with AIEC, 100 *E. coli* strains isolated from CD patients (Darfeuille-Michaud *et al.*, 2004; Dreux *et al.*, 2013) and 45 *E. coli* strains from healthy people were screened by PCR amplification. The *vat-AIEC* gene was significantly overrepresented in AIEC isolates (42.6%, $n = 32/75$) in comparison with non-AIEC isolates of CD patients (12%, $n = 3/25$, $P = 0.007$) or healthy people (15.6%, $n = 7/45$, $P = 0.002$). *E. coli* strains were segregated into four major phylogenetic groups, designated 'A', 'B1', 'B2' and 'D' (Clermont *et al.*, 2000). Pathogenic strains belong to group B2 and, to a lesser extent, group D, with more frequent virulence-factor genes than group A and group B1 strains (Picard *et al.*, 1999). Among the 100 *E. coli* strains from CD patients analysed, 97% of *vat-AIEC*-harbouring *E. coli* belonged to phylogenetic group B2 and 3% belonged to phylogroup D. Overall, these results showed that the *vat-AIEC* gene is strongly associated with B2 AIEC strains. This is in accordance with the results of Sepehri *et al.* (2011) that the B2 + D phylotype is significantly associated with inflammation as it is associated with serine protease autotransporters.

*Vat-AIEC contributes to the gut colonization ability of AIEC LF82 in mice*

We hypothesized that Vat-AIEC mucinolytic activity may favour the gut colonization by AIEC. For this purpose, CEABAC10 transgenic mice were orally challenged with LF82 or LF82Δ*vat-AIEC*, and the number of LF82 and LF82Δ*vat-AIEC* bacteria in the faeces or associated with the intestinal mucosa was evaluated 3 days post-infection. Quantification of bacteria in stool samples revealed one-log reduction in LF82Δ*vat-AIEC* mutant compared with wild-type LF82 [$6.1 \times 10^4$ versus $5.7 \times 10^5$ colony-forming units (cfu) per gramme of faeces, $P = 0.0039$, Fig. 2A]. Likewise, the ability of LF82Δ*vat-AIEC* mutants to colonize the ileal tissue of CEABAC10 mice was significantly affected after 3 days of infection (4.9 versus 17.2 cfu mg$^{-1}$ of tissue, $P = 0.014$, Fig. 2A). This difference between strains with and without *vat-AIEC* was much more pronounced in the colon (79.8 *versus* 6.2 cfu mg$^{-1}$ of tissue, $P = 0.0091$, Fig. 2A), possibly because the colon normally has an inner mucus layer impenetrable to bacteria. In addition, confocal microscopy showed the direct contact of LF82 bacteria with colonic epithelial cells (Fig. 2B). The colonization of the colon with LF82 bacteria resulted in a significantly increased secretion of the pro-inflammatory chemokine keratinocyte-derived chemokine and cytokine tumour necrosis factor alpha, respectively, than that with the LF82Δ*vat-AIEC* mutant. Furthermore, an increased level of myeloperoxidase was also observed in mice infected with LF82 compared with that in mice infected with the LF82Δ*vat-AIEC* mutant (Fig. 2C). When CEABAC10 transgenic mice were pretreated with dextran sulfate sodium (0.25%) (Carvalho *et al.*, 2009), no difference of colonization between wild-type and mutant strains was observed (Fig. S3). It has already been shown that bacteria can penetrate the mucus layer even before the development of colitis in mice exposed to 2–5% dextran sulfate sodium, although this layer looks relatively normal with regard to thickness (Johansson *et al.*, 2014). Overall, these results show that Vat-AIEC enhances the ability of AIEC LF82 strain to colonize the intact intestinal mucosa and to induce inflammation.

*Vat-AIEC enhances the spread of LF82 in mucus*

To assess the mechanism by which Vat-AIEC contributes to gut colonization, we first hypothesized that Vat-AIEC-producing AIEC uses mucin as a growth substrate as reported for Pic-producing *E. coli* (Harrington *et al.*, 2009). The growth of wild-type strain LF82 and its isogenic mutant LF82Δ*vat-AIEC* strain was therefore compared in a M9 minimal medium supplemented or not with mucin. No difference of growth was observed (Fig. S4). This indicates that LF82 bacteria do not use mucin as a growth substrate.

Pathogen access to enterocytes is strongly limited by intestinal mucins, which favour their elimination by mechanical processes or the immune system (Dharmani *et al.*, 2009; Johansson and Hansson, 2011; Johansson *et al.*, 2011). We therefore hypothesized that Vat-AIEC enhances the spread of AIEC in close contact to enterocytes through the promotion of mucin barrier degradation. To test our hypothesis, we investigated the ability of LF82 and LF82Δ*vat-AIEC* to penetrate mucus by using a column penetration assay. Unlike LF82Δ*vat-AIEC*, LF82 penetrated through the entire mucus column (fractions 1 to 10, Fig. 3A). At the middle of the column (fractions 5 and 6), LF82Δ*vat-AIEC* was more than two orders of magnitude less efficient at penetrating the mucus column than the wild-type parent strain. We used live-cell imaging to confirm that the difference in the mucus gel column was not caused by decreased mobility in the LF82Δ*vat-AIEC* strain com-

© 2015 John Wiley & Sons Ltd, *Cellular Microbiology*

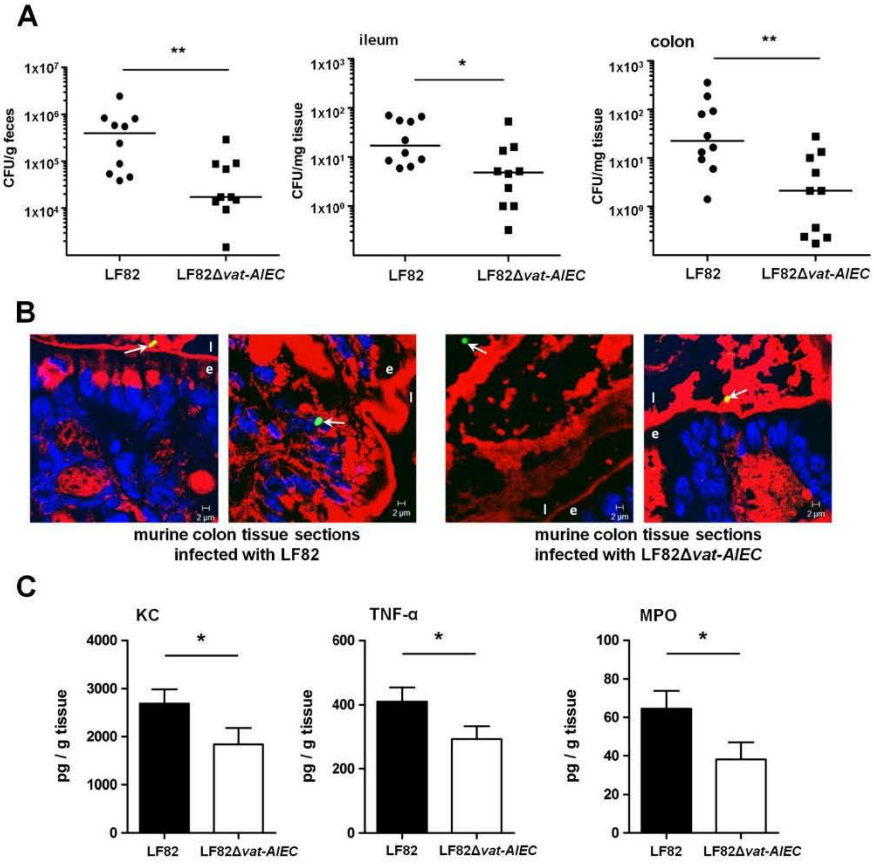



**Fig. 2.** After antibiotics treatment, mice were orally infected with $10^9$ LF82 (round) or $10^9$ LF82Δ*vat-AIEC* (square).
A. The number of LF82 or LF82Δ*vat-AIEC* in faeces collected at day 3 post-infection were counted on TS plates containing amoxicillin. Quantification of ileal and colonic mucosal-associated LF82 or LF82Δ*vat-AIEC* bacteria was also performed 3 days post-infection.
B. Colonic sections of mice infected with LF82 or LF82Δ*vat-AIEC* were stained with anti-*Escherichia coli* O83 antibody (green, bacteria staining), Hoechst (blue, DNA staining) and wheat germ agglutinin (red, mucus and glycocalyx staining) and visualized with confocal microscopy. Bacteria were designed with white arrows. LF82 bacteria were found in contact or inside IECs, whereas LF82Δ*vat-AIEC* bacteria were generally found in the intestinal lumen or inside mucus but not in contact with IECs. Each image is representative of three sections/mouse from 10 mice/group. l, intestinal lumen; e, intestinal epithelium.
C. Colonic keratinocyte-derived chemokine, tumour necrosis factor alpha (TNF-α) and myeloperoxidase (MPO) levels in mice infected with LF82 or LF82Δ*vat-AIEC* were measured by enzyme-linked immunosorbent assay. Statistical analysis were performed with Mann–Whitney test; $^*P < 0.05$; $^{**}P < 0.01$.

pared with LF82. Indeed, track speed mean values were similar between wild-type and mutant strains in medium broth with mucin ($2.6 \pm 2.0$ versus $4.8 \pm 6.3\ \mu m\ s^{-1}$, $P = 0.18$) or without mucin ($14.4 \pm 8.3$ versus $16.9 \pm 3.2\ \mu m\ s^{-1}$, $P = 0.26$).

The viscosity of mucin incubated with LF82 or LF82Δ*vat-AIEC* over 24 h was measured to substantiate this result. LF82 caused a 7.1% reduction in mucin viscosity compared with mucin incubated with LF82Δ*vat-AIEC*, confirming its degradation by Vat-AIEC (with an average of 0.0109 versus 0.0117 Pl, $P < 0.001$) (Fig. 3B). Overall, these results suggest that Vat-AIEC promotes the penetration of LF82 through the mucus layer by altering the gel-forming mucins.

### *Vat-AIEC enhances LF82 access to epithelial cells*

To assess the consequences of Vat-AIEC activity on the adhesion properties of LF82, we infected the mucin-hyperproducing HT29-16E cells (Augeron and Laboisse, 1984) and the closely related HT29 cells, which do not produce significant levels of mucins. We first verified that LF82 and LF82Δ*vat-AIEC* both have type 1 pili, which are necessary for their adhesion to IECs. Compared with the wild-type strain, the LF82Δ*vat-AIEC* mutant exhibited a

© 2015 John Wiley & Sons Ltd, *Cellular Microbiology*

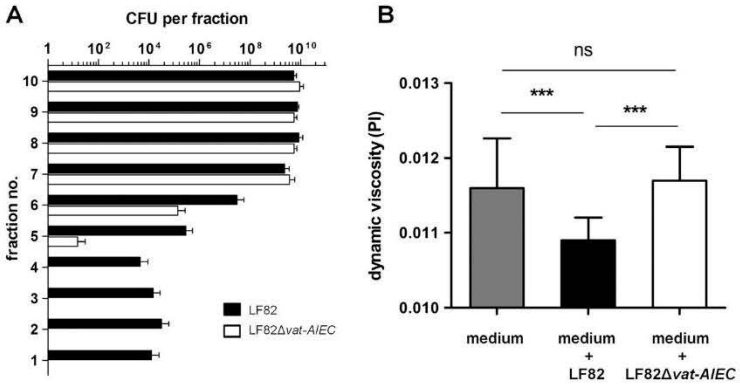



**Fig. 3.** A. Quantification of LF82 (filled bars) and LF82Δ*vat-AIEC* (open bars) in fractions eluted from columns containing gel-forming mucus and inoculated with the bacterial strains (fraction 1, bottom of the column, to 10, top of the column).
B. The viscosity of gel-forming mucins was assessed after 24 h of incubation at 37 °C in presence or absence of LF82 or LF82Δ*vat-AIEC*. Statistical analysis was performed from three independent experiments, and means are shown ± standard deviation. (analysis of variance; ***$P < 0.001$; ns, not significant).

similar ability to agglutinate yeast (agglutination titer of 1/8) as previously described (Dreux *et al.*, 2013), suggesting an expression of functional type 1 pili (see Appendix S1 for detailed method). As shown in Fig. 4A, the wild-type LF82 strain had a significantly higher ability to adhere to HT29-16E cells than the LF82Δ*vat-AIEC* mutant ($P = 0.033$), whereas no difference in adhesion was observed for HT29 cells ($P = 0.052$). Adhesion of LF82 and LF82Δ*vat-AIEC* to IECs was also investigated by confocal microscopy at different post-infection times. The number of adherent bacteria to HT29-16E cells was higher for LF82 than for LF82Δ*vat-AIEC*, suggesting that *vat-AIEC* promotes bacterial access to epithelial cells (Fig. 4B and C). These results were confirmed by electron microscopy analysis of cells at 20 and 60 min post-infection. As shown in Fig. 4D, the wild-type LF82 strain penetrated the mucus to get in contact with HT29-16E cells at 1 h post-infection, whereas the LF82Δ*vat-AIEC* mutant was mucus-embedded. To support these results, we investigated the localization of bacteria *in vivo* using a mouse intestinal loop model infected for 1 h. As shown by immunohistochemistry and electron microscopy, unlike the LF82Δ*vat-AIEC* mutant, LF82 bacteria were mainly found in direct contact with IECs (Fig. 4E and F). Taken together, our data demonstrate that Vat-AIEC improves the ability of AIEC LF82 strain to penetrate the mucus layer to adhere to IECs.

### *Chemical features of the ileum environment activated Vat-AIEC expression*

Adherent-invasive *E. coli* can colonize the ileum of CD patients, and the AIEC reference strain LF82 was isolated from a chronic ileal lesion. Hallmarks of the human ileum environment are an alkaline pH and the presence of bile salts and mucins (Playoust and Isselbacher, 1964; Nugent *et al.*, 2001; McGuckin *et al.*, 2011). We therefore investigated the effect of these chemical features on *vat-AIEC* expression in AIEC strain LF82 by first using a β-galactosidase reporter gene downstream the putative *vat-AIEC* promoter (pRS550-P*vat-AIEC*) and also by quantifying *vat-AIEC* mRNA levels.

pRS550-P*vat-AIEC*-carrying bacteria LF82Δ*vat-AIEC* were grown in a medium buffered at a pH of 6.5 or 7.5, supplemented or not with mucins or bile salts. The pH variation did not significantly modify the β-galactosidase activity (Fig. 5A). However, significant increase in β-galactosidase activity was observed in the presence of mucins and bile salts (1.7-fold and 18.1-fold, respectively, $P < 0.05$; Fig. 5A). Interestingly, the highest increase in activity (32-fold) was obtained under conditions combining mucins, bile salts and a pH value of 7.5 ($P < 0.001$, Fig. 5A), which are chemical features presents in ileum.

The *vat-AIEC* mRNA levels were then assessed in AIEC strain LF82 to substantiate the results with the complete *vat-AIEC* gene. Consistent with β-galactosidase activity results, mucins and bile salts signifcantly increased the *vat-AIEC* mRNA levels unlike the pH (Fig. 5B). A LF82 strain grown in the pH 7.5 medium supplemented with mucins and bile salts exhibited a 6.1-fold increase in *vat-AIEC* mRNA levels when compared with LF82 growth in the pH 6.5 medium that only contained mucins, which corresponds to colonic conditions ($P < 0.001$; Fig. 5B). Altogether, our results suggest that a combination of several conditions found in ileum increases *vat-AIEC* expression.

To confirm these results, adhesion experiments were performed by using mucin-hyperproducing HT29-16E cells infected with AIEC strain LF82 grown at pH 6.5 without bile salts or pH 7.5 with bile salts. At 3 h post-infection, the bacterial adhesion was 10-fold higher in cell medium with a pH of 7.5 than in cell medium with a pH of 6.5 ($P = 0.0078$, Fig. 5C). It should be noticed that the

© 2015 John Wiley & Sons Ltd, *Cellular Microbiology*

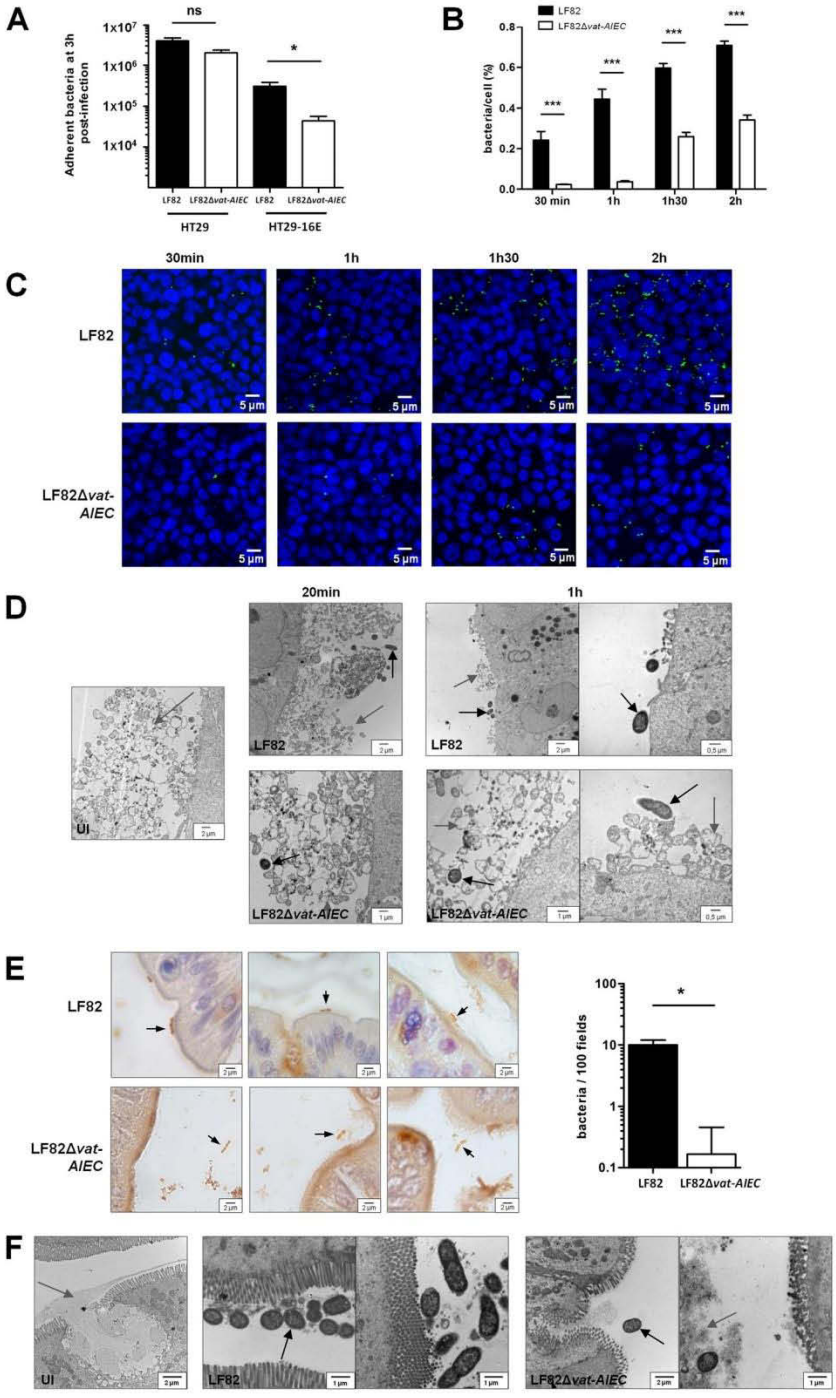



**Fig. 4.** A. Quantification of adherent bacteria to intestinal epithelial cells HT29 or to mucus-hyperproducing intestinal epithelial cells HT29-16E after 3 h of infection. Each value is the mean ± standard error of mean (SEM) of at least three separate experiments. Statistical analysis was performed with a Mann–Whitney test (*$P < 0.05$; ns, not significant).

B. The percentage of adherent LF82 (filled bars) and LF82Δ*vat-AIEC* (open bars) per HT29-16E cell was calculated at different post-infection times. Values are means ± SEM of at least three independent experiments (analysis of variance; ***$P < 0.001$).

C. Adherent LF82 or LF82Δ*vat-AIEC* at different post-infection times were visualized by confocal microscopy.

D. HT29-16E cells were infected for 20 and 60 min with LF82 and LF82Δ*vat-AIEC*, and transmission electron micrograph were taken. Bacteria (black arrows) and mucus (grey arrows) are visualized; UI, uninfected cells.

E. In a murine colonic loop model, paraffin-embedded tissue sections after 1 h of infection with LF82 or LF82Δ*vat-AIEC* were stained with anti-*Escherichia coli* O83 antibody to detect bacteria (arrows). Adherent LF82 (filled bars) and LF82Δ*vat-AIEC* (open bars) were quantified.

F. Intestinal sections of loop model after 1 h of infection with LF82 or LF82Δ*vat-AIEC* were observed by transmission electron micrograph. Bacteria are visualized (arrows); UI, uninfected mice.

© 2015 John Wiley & Sons Ltd, *Cellular Microbiology*

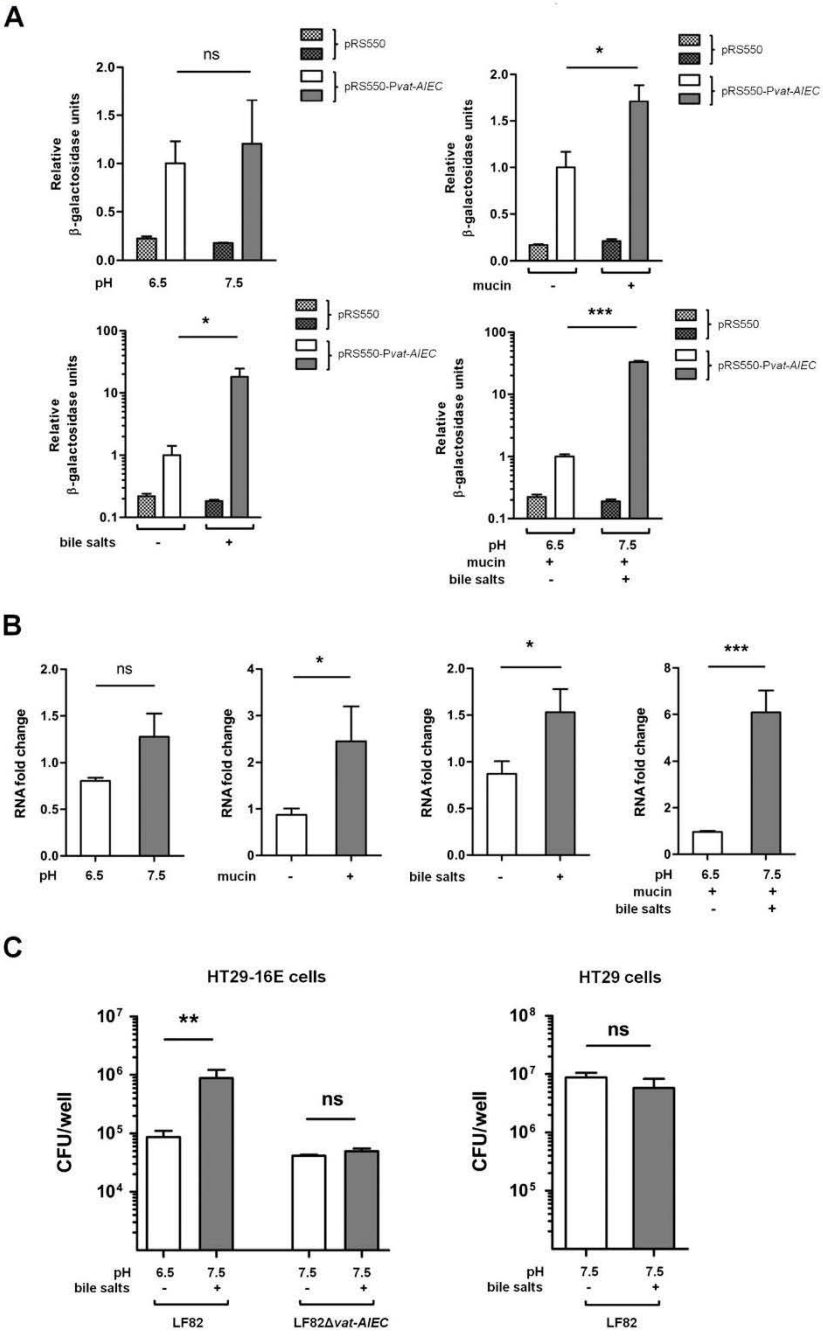



**Fig. 5.** A. Effect of pH, mucin, bile salts and their combination on the activity of the *vat-AIEC* promoter was assessed by β-galactosidase assays in LF82Δ*vat-AIEC*-pRS550 (empty vector) and LF82Δ*vat-AIEC*-pRS550-P*vat-AIEC* (vector containing the *vat-AIEC* promoter) after 6 h of incubation at 37 °C. Each value is the mean ± standard error of mean (SEM) of at least three separate experiments (analysis of variance test; $*P < 0.05$; $***P < 0.001$; ns, not significant).
B. Effect of pH, mucin, bile salts and combined conditions on *vat-AIEC* expression. LF82 wild-type strain was cultured 6 h and *vat-AIEC* mRNA levels were analysed by qRT-PCR. Each value is the mean ± SEM of at least three separate experiments (Mann–Whitney test; $*P < 0.05$; $***P < 0.001$; ns, not significant).
C. Adherent bacteria were quantified by using mucus-hyperproducing HT29-16E cells or HT29 cells after 3 h of infection with culture medium at different pH values. Bacteria were pre-incubated in Dulbecco's modified Eagle's medium at pH 6.5, or pH 7.5 with bile salts. Each value is the mean ± SEM (Mann–Whitney test; $**P < 0.01$; ns, not significant).

incubation of bacteria in medium containing bile salts did not modify AIEC adhesion to HT29 cells. Interestingly, analysis of bile acid composition in patients with CD has revealed increased concentrations of cholic acid (Rutgeerts *et al.*, 1986), which has been used in this study for *vat-AIEC* expression analysis. Taken together, these results indicate that chemical conditions found in the ileum environment could be a signal to increase *Vat-AIEC* expression in AIEC strain LF82 and consequently influences LF82 colonization of intestinal mucosa.

© 2015 John Wiley & Sons Ltd, *Cellular Microbiology*

## Discussion

The mucus covers the mucous membrane and prevents the access of pathogens to the epithelia. Hbp and Pic are serine protease autotransporters that cleave mucins from numerous sources and confers to bacteria a subtle competitive advantage for mucosal colonization (Harrington *et al.*, 2009; Kobayashi *et al.*, 2010). Hbp has been shown to be associated with lethal APEC isolates, and Pic has been found in *S. flexneri* and enteroaggregative *E. coli*. In the present study, we identified a chromosome-mediated gene encoding a protein designated as Vat-AIEC in AIEC reference strain LF82, which is closely similar to the Vat vacuolating toxin from APEC but harbours the catalytic domain of Pic and Hbp serine proteases. Our epidemiologic data revealed that the *vat-AIEC* gene was overrepresented in the AIEC belonging to the B2 phylogenetic subgroup prevalent in biopsy tissues of CD patients (Picard *et al.*, 1999; Kotlowski *et al.*, 2007).

The increased prevalence of AIEC in all CD subtypes reinforces the hypothesis that the microenvironment of CD ileum favours AIEC expansion and that the colon is also a niche colonized by AIEC (Martinez-Medina and Garcia-Gil, 2014). The presence of *vat-AIEC* gene promoted the colonization of mouse intestine by AIEC and thereby enhanced inflammation. The adhesion of CD-associated AIEC to the colonic epithelium following their colonization of gut mucosa requires the penetration of AIEC through the protective mucus barrier, which is impenetrable to bacteria in a physiological context. Vat-AIEC caused a diminution of mucus viscosity and promoted the penetration of AIEC through mucin-based gels suggesting that the proteolytic cleavage of mucins causes disassembly of the oligomerized-mucin macromolecules. Moreover, Vat-AIEC favoured adhesion to IECs recovered of mucus. Consequently, Vat-AIEC-producing AIEC could sneak within the disrupt mucus gel and make contact with the epithelium. This scenario may also apply to strains producing the Hbp and Pic proteases. Mucinases might therefore play a major role in the early stages of AIEC pathogenesis, particularly in the colon where the mucus layer is normally impenetrable to bacteria, as proposed in the pathogenesis of acute enteric pathogens (Leyton *et al.*, 2003; Silva *et al.*, 2003; Lidell *et al.*, 2006; Ruiz-Perez *et al.*, 2011; Kumar *et al.*, 2014).

Bile salts and the intestinal pH are variable environmental signals along the intestinal tract. We observed a strong increase in expression of the *vat-AIEC* gene when bacteria were grown with bile salts. Their presence is a host signal that enteric bacteria encounter as they travel the gastrointestinal tract. It is therefore not surprising that enteric pathogens, including *E. coli*, respond to bile by increasing the expression of virulence factors (Pope *et al.*, 1995; Wibbenmeyer *et al.*, 2002; Prouty *et al.*, 2004; Merritt and Donaldson, 2009; Sahl and Rasko, 2012; Chassaing *et al.*, 2013). However, to our knowledge, *vat-AIEC* is the first example of an autotransporter-encoding gene regulated by bile salts. This regulation resembles the Hap metallo-mucinase of *V. cholerae* (Silva *et al.*, 2003), which is known to colonize the distal ileal segment in a manner similar to that of AIEC (Darfeuille-Michaud *et al.*, 1998). It is notable that the expression of some SPATEs is pH-dependent, with optimum expression at an alkaline pH [pH 7.0–9.0 for EspP (Brunder *et al.*, 1997), pH 8.0 for EpeA (Leyton *et al.*, 2003) and pH 9.0 for Pic (Henderson *et al.*, 1999)]. The terminal ileum pH ranges between 7.3 and 7.7 in healthy adults, and between 7.5 and 8.2 in Crohn's disease patients (Nugent *et al.*, 2001). In this study, we did not observe any significant difference in *vat-AIEC* expression between pH 6.5 and 7.5. However, when the chemical conditions found in the ileum were combined, i.e. a pH of 7.5 with bile salts and mucins, the induction of *vat-AIEC* expression was stronger than that found under colonic conditions, i.e. pH 6.5 and mucins. This suggests that chemical conditions found in the terminal ileum stimulate Vat-AIEC production that promotes crossing of the mucus layer by AIEC. Interestingly, bile salts induced expression of the *lpf* operon coding the long polar fimbriae, which promotes bacterial interaction with Peyer's patches (Chassaing *et al.*, 2013). The Peyer's patches are also covered by mucus, which is strongly secreted by goblet cells in neighbouring villi (Ermund *et al.*, 2013). Hence, chemical features found in the ileum may stimulate *vat-AIEC* and *lfp* expression and enhance the translocation of AIEC strains *via* the M cells of the Peyer's patches.

The expression of mucin-encoding genes in the intestine of CD patients and healthy people is similar (Buisine *et al.*, 2001), but the hydrolysis of MUC2 by Vat-AIEC may be raised in patients with CD. Indeed, Png and colleagues (Png *et al.*, 2010) have reported that in CD intestinal mucosa, there is an increase in relative abundance of the known mucinolytic bacteria, particularly *Ruminococcus gnavus*, which secrete glycosidases capable of removing terminal sugars allowing mucinases to degrade apomucins (Hoskins and Boulding, 1981; Hoskins *et al.*, 1985). Thus, MUC2 degradation by Vat-AIEC-producing AIEC might be promoted in presence of these mucinolytic bacteria. Besides, there is a range of glycosylation abnormalities in CD patients, including reduced length of O-glycans (Bodger *et al.*, 2006; Rhodes *et al.*, 2008), and genetic studies implicate genes involved in glycosylation pathways as CD susceptibility loci (Theodoratou *et al.*, 2014). For example, single nucleotide polymorphisms in the *FUT2* gene, which encodes an α(1,2)fucosyltransferase that catalyses

© 2015 John Wiley & Sons Ltd, *Cellular Microbiology*

addition of terminal α(1,2)fucose residues on mucins and other molecules in mucosal epithelium, have been found to be associated with CD (McGovern *et al.*, 2010). Further studies would need to be undertaken to clarify the role of Vat-AIEC in AIEC implantation in CD patients.

In conclusion, we described in this study a new protease, named Vat-AIEC, that is a contributor to the virulence of AIEC at the first step of infection because it favours the penetration of AIEC through the mucus layer and their adhesion to IECs, whether it is in the ileal or in the colonic subtype of CD. Our study provides a basis for understanding how AIEC can breach the mucosal barrier. An implication of this finding is the development of inhibitors because Vat-AIEC represents a highly accessible antimicrobial target to block AIEC virulence at an early stage of the pathogenicity process.

## Experimental procedures

### *Bacterial strains, plasmids and media*

Bacterial strains, plasmids and primers used in this study are listed in Tables S1 and S2. For *vat-AIEC* induction, *E. coli* strains were grown at 37 °C in medium containing 20% M9 5× broth, 20% lysogeny broth (LB) and 60% Tris buffer with an adjusted pH. Glucose, L-arabinose, porcine mucin and bile salts (cholic acid sodium salt ~50% and deoxycholic acid sodium salt ~50%) (Sigma) were added as required.

### *Construction of isogenic mutant LF82*Δvat-AIEC *and inactive mutant LF82*Δvat-AIEC/vat-AIECS260I

Isogenic mutant of *E. coli* LF82 in which the *vat-AIEC* gene was deleted was generated by using the Red recombinase system described by Datsenko and Chaveroche (Chaveroche *et al.*, 2000; Datsenko and Wanner, 2000) (see Appendix S1 for detailed methods). The corresponding LF82Δ*vat-AIEC* mutant was complemented with *vat-AIEC*. For the experimental procedures, the transcomplemented strain LF82Δ*vat-AIEC*/*vat-AIEC* yielded the same results as the LF82 strain. Inactive mutant LF82Δ*vat-AIEC*/*vat-AIECS260I* was performed by site-directed mutagenesis with the QuikChange™ Site-Directed Mutagenesis Kit II (Stratagene) using the mutagenic primers described in Table S2.

### *Mucinolytic activity, mucin gel penetration assay and viscosity measurements*

The mucin gel method used to detect mucinolytic activity and the mucin gel penetration assay were used as previously described (Henderson *et al.*, 1999; Silva *et al.*, 2003) (see Appendix S1 for detailed methods). For viscosity measurements, bacteria at $5.10^9$ were grown overnight in 25 ml of PBS containing 5% mucin from porcine stomach and 2% bile salts. The measures of viscosity were performed at 37 °C in a constant-temperature chamber by using a falling ball-type viscometer. The viscosity was calculated by using the following equation: $\eta = K \cdot (d_b - d_l) \cdot t$, where $\eta$ is the viscosity in PI, $K$ is the viscometer constant $(\eta_{\text{eau } 37°C})/(1.393 t_{eau})$, $d_b$ is the density of the ball (2.393), $d_l$ is the density of the liquid (g ml$^{-1}$) and $t$ is the time of ball descent. The density of the solution was measured by weighing 1 ml of the given solution at the test temperature.

### *Purification and sequencing of cleaved Vat-AIEC protein*

The LF82 strain was grown for 24 h at 37 °C in M9 medium containing 0.01% glucose. The culture was 0.22 μm filtered and concentrated (1/100 v/v) by filtration using a 10 kDa Amicon® membrane (Merck Millipore). The filtered supernatant was then applied to ion-exchange chromatography onto a HiTrap™ Q Sepharose™ High Performance (Tris-HCl 20 mM, pH 8.5). Precipitated protein was separated by SDS-polyacrylamide gel and stained with Coomassie blue. Band was excised and sequenced by Edman degradation.

### *Muc2 proteolytic cleavage assay*

One-hundred microlitres of culture supernatants filter-sterilized (0.22 μm) or 1 μg of purified fraction of Vat-AIEC were distributed to 1 μg of mucins followed by overnight digestion at 37 °C. Mucins were fractions of extracted mucus from mouse intestine or a fragment of the C-terminal cysteine-rich domain of an unglycosylated recombinant human MUC2, which does not harbour the autocatalytic cleavage site of MUC2 (Creative Biomart, USA). The digestion was terminated by incubating samples with 4× Laemmli sample buffer containing 10% mercaptoethanol at 100 °C for 3 min. Samples were separated on a 4%-stacking and 12%-running SDS-PAGE, transferred to nitrocellulose membrane (GE Healthcare Life Sciences) using a semi-dry blotting apparatus (Bio-Rad), blocked with 1% BSA in Tris-Buffered Saline and Tween 20 (TBS-T) (10 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.05% Tween 20). Membranes were then probed overnight at 4 °C with anti-MUC2 H300 (1:1000 dilution, Santa Cruz). After washes, membranes were incubated with the appropriate HRP-conjugated secondary antibodies (Cell Signaling Technology), and blots were detected using the enhanced chemiluminescence reagent (Bio-Rad).

### *Mouse infection*

For *in vivo* experiments, we used CEABAC10 transgenic mice, which are known to express the human receptor CEACAM6 (Chan and Stanners, 2004; Carvalho *et al.*, 2009). Mice were housed in specific pathogen-free conditions in the animal care facility at the Université d'Auvergne, Clermont-Ferrand, France. Animal protocols were approved by the Committee for Research and Ethical Issues of the CEMEAA (Comité d'Ethique en Matière d'Expérimentation Animale, Auvergne) following international directive 86/609/CEE (no. CE16-09).

Twelve-week-old mice (10 per group, body weight ≈ 26–28 g) were pretreated by administering oral amoxicillin (1 g l$^{-1}$), vancomycin (500 mg l$^{-1}$), metronidazole (500 mg l$^{-1}$) and neomycin (1 g l$^{-1}$) for 3 days and 0.25% dextran sulfate sodium salt

© 2015 John Wiley & Sons Ltd, *Cellular Microbiology*

(Sigma) when needed. The animals were orally challenged with $10^9$ bacteria 24 h after stopping antibiotic treatment. Three days after bacterial infection, fresh faecal pellets (100–200 mg) were collected from the individual mice and resuspended in PBS. After serial dilution, the bacteria were enumerated by plating on trypticase soy (TS) agar medium containing amoxicillin to isolate AIEC LF82 and incubated at 37 °C overnight. Antibiotic susceptibility testing was made on bacteria in a random way to confirm whether they were LF82 or LF82Δ*vat-AIEC*. Three days after infection, the mice were anesthetized with isoflurane and then euthanized by cervical dislocation. Colonization was confirmed by numbering the mucosa-associated AIEC bacteria by homogenizing 0.5 cm of ileum and 1 cm of colon, beginning at 0.5 cm from the cecal junction, in sterile PBS solution. Samples were plated onto TS agar containing amoxicillin and incubated overnight at 37 °C. After fixation with Carnoy's solution to preserve mucus, paraffin-embedded colonic sections were immunofluorescent stained using Hoescht (Sigma), Wheat Germ Agglutinin (Molecular probes) and anti-*E. coli* O83 antibody and analysed by confocal microscopy. Levels of secreted cytokines and myeloperoxidase were quantified by enzyme-linked immunosorbent assay (see Appendix S1 for detailed method).

Bacterial interactions were also studied by using mouse intestinal loops as previously described (Chassaing *et al.*, 2013). The mice were pretreated as described in the previous text. The animals were anesthetized, and their intestines were exteriorized through a midline incision. Their colonic segments (approximately 1 cm) were ligated and inoculated with approximately $5.10^7$ bacteria. After 1 h period, the mice were euthanized. Their colonic loops were analysed next by immunohistochemistry and electron microscopy.

### *In vitro* cell adhesion assays

Human intestinal epithelial HT29 cells and HT29-16E cells (with mucus hyperproduction) were obtained from ATCC and Professor Laboisse (56) respectively. Cells were maintained in an atmosphere containing 5% $CO_2$ at 37 °C in Dulbecco's modified Eagle's medium supplemented with 10% (v/v) fetal calf serum (Lonza), 1% L-glutamine (Life Technologies), 200 U penicillin, 50 mg of streptomycin and 0.25 mg of amphotericin B per litre. For infection experiments, cells were seeded at a density of $2 \times 10^5$ cells per squared centimetre in culture plates (Falcon) for 21 days in order to allow them to produce mucin.

Bacteria were grown overnight in Dulbecco's modified Eagle's medium with adjusted pH and 2% bile salts when needed. Bile salts were used for bacteria growth, but they were not used in cell medium because they induce apoptosis in intestinal cell lines (Hague *et al.*, 1995). For the adhesion assay, cells were infected over a 3 h period at a multiplicity of infection of 10 bacteria per cell with the pH adjusted when needed. At the end of the infection, cells were washed in PBS then lysed with 1% Triton X-100 (Sigma) in deionized water. Finally, the samples were diluted and plated onto TS agar plates to determine the number of cfus corresponding to the total number of cell-associated bacteria.

### *Confocal microscopy*

HT29-16E cells were seeded on sterile glass cover slips. After 21 days of culture, the cells were infected for different periods (30 min, 1 h, 1.5 h and 2 h) with GFP-expressing LF82 or GFP-expressing LF82Δ*vat-AIEC* bacteria at a multiplicity of infection of 10. Hoechst stain (Sigma) was added to the cells before observing them with a Zeiss LSM 510 Meta confocal microscope. For each group, the percentage of bacteria per cell was determined by counting 5000 cells.

### *Promoter expression assay*

The putative *vat-AIEC* promoter (580 pb upstream *vat-AIEC* ORF) was amplified by PCR using BamHIP*vat-AIEC*-F and EcoRIP*vat-AIEC*-R primers (Table S2). The PCR product was ligated into pRS550 plasmid and designated pRS550-P*vat-AIEC*. The galactosidase activity was analysed with a β-galactosidase assay kit (Qiagen) with strains harbouring the pRS550-P*vat-AIEC* construct and grown in medium containing 20% M9 5× broth, 20% LB and 60% Tris buffer at different pH values, in the same medium containing mucin (1%) and/or bile salts (2%). The β-galactosidase activity of each sample was determined by measuring the $OD_{420}$ at 6 h, and the number of bacteria in each sample was calculated by taking the $OD_{600}$ measurement for the Miller unit determination.

### *RNA extraction, reverse transcription and RT-PCR*

*Escherichia coli* strain LF82 was grown at 37 °C for 6 h in a medium containing 20% M9 5× broth, 20% LB and 60% Tris buffer with the pH adjusted and mucin or bile salts as needed. Total RNAs were extracted from bacteria using a Direct-zol RNA MiniPrep kit (Zymo research) and treated with a Turbo DNA-free kit (Ambion) to remove any contaminating genomic DNA. RNAs were reverse transcribed with random hexamers (Invitrogen) and amplified by using specific primers to *vat-AIEC* or *16S*rRNA (Table S2). The amplification of a single expected PCR product was confirmed by electrophoresis on a 2% agarose gel. qRT-PCR was performed using an Eppendorf Realplex. The relative *vat-AIEC* gene expression was normalized to *16S*rRNA by using the $2^{-\Delta\Delta Ct}$ method.

### *Statistical analysis*

Statistical comparisons were performed with a Mann–Whitney nonparametric test. An analysis of variance was performed for experiments with multiple treatment groups, followed by pairwise comparisons with Bonferroni's multiple comparison tests. A *P* value less than 0.05 was considered statistically significant.

© 2015 John Wiley & Sons Ltd, *Cellular Microbiology*

**Acknowledgements**

We thank Sylvine Batista for technical assistance. We also thank Lorraine Gameiro from the CICS and Caroline Vachias from the ICCF platforms (Université d'Auvergne) for assistance with the confocal and electronic microscopy.

**Funding**

We thank Lothar Beutin (Department of Biological Safety, Robert Koch Institut, Berlin, Germany) for kindly providing anti-*Escherichia coli* O83 antibodies.

This work was supported by the Ministère de l'Enseignement supérieur et de la Recherche, the Institut national de la santé et de la recherche médicale (UMR Inserm U1071), the Institut National de la Recherche Agronomique (USC-2018) and CHU de Clermont-Ferrand.

**References**

Augeron, C., and Laboisse, C.L. (1984) Emergence of permanently differentiated cell clones in a human colonic cancer cell line in culture after treatment with sodium butyrate. *Cancer Res* **44**: 3961–3969.

Barnich, N., Boudeau, J., Claret, L., and Darfeuille-Michaud, A. (2003) Regulatory and functional co-operation of flagella and type 1 pili in adhesive and invasive abilities of AIEC strain LF82 isolated from a patient with Crohn's disease. *Mol Microbiol* **48**: 781–794.

Barnich, N., Carvalho, F.A., Glasser, A.-L., Darcha, C., Jantscheff, P., Allez, M., *et al.* (2007) CEACAM6 acts as a receptor for adherent-invasive *E. coli*, supporting ileal mucosa colonization in Crohn disease. *J Clin Invest* **117**: 1566–1574.

Barnich, N., and Darfeuille-Michaud, A. (2007) Adherent-invasive *Escherichia coli* and Crohn's disease. *Curr Opin Gastroenterol* **23**: 16–20.

Barnich, N., and Darfeuille-Michaud, A. (2010) Abnormal CEACAM6 expression in Crohn disease patients favors gut colonization and inflammation by adherent-invasive *E. coli*. *Virulence* **1**: 281–282.

Baumgart, M., Dogan, B., Rishniw, M., Weitzman, G., Bosworth, B., Yantiss, R., *et al.* (2007) Culture independent analysis of ileal mucosa reveals a selective increase in invasive *Escherichia coli* of novel phylogeny relative to depletion of Clostridiales in Crohn's disease involving the ileum. *ISME J* **1**: 403–418.

Bodger, K., Halfvarson, J., Dodson, A.R., Campbell, F., Wilson, S., Lee, R., *et al.* (2006) Altered colonic glycoprotein expression in unaffected monozygotic twins of inflammatory bowel disease patients. *Gut* **55**: 973–977.

Boudeau, J., Barnich, N., and Darfeuille-Michaud, A. (2001) Type 1 pili-mediated adherence of *Escherichia coli* strain LF82 isolated from Crohn's disease is involved in bacterial invasion of intestinal epithelial cells. *Mol Microbiol* **39**: 1272–1284.

Brunder, W., Schmidt, H., and Karch, H. (1997) EspP, a novel extracellular serine protease of enterohaemorrhagic *Escherichia coli* O157:H7 cleaves human coagulation factor V. *Mol Microbiol* **24**: 767–778.

Buisine, M.P., Desreumaux, P., Leteurtre, E., Copin, M.C., Colombel, J.F., Porchet, N., and Aubert, J.P. (2001) Mucin gene expression in intestinal epithelial cells in Crohn's disease. *Gut* **49**: 544–551.

Carvalho, F.A., Barnich, N., Sivignon, A., Darcha, C., Chan, C.H.F., Stanners, C.P., and Darfeuille-Michaud, A. (2009) Crohn's disease adherent-invasive *Escherichia coli* colonize and induce strong gut inflammation in transgenic mice expressing human CEACAM. *J Exp Med* **206**: 2179–2189.

Chan, C.H.F., and Stanners, C.P. (2004) Novel mouse model for carcinoembryonic antigen-based therapy. *Mol Ther J Am Soc Gene Ther* **9**: 775–785.

Chassaing, B., Etienne-Mesmin, L., Bonnet, R., and Darfeuille-Michaud, A. (2013) Bile salts induce long polar fimbriae expression favouring Crohn's disease-associated adherent-invasive *Escherichia coli* interaction with Peyer's patches. *Environ Microbiol* **15**: 355–371.

Chaveroche, M.K., Ghigo, J.M., and d'Enfert, C. (2000) A rapid method for efficient gene replacement in the filamentous fungus *Aspergillus nidulans*. *Nucleic Acids Res* **28** E97.

Clermont, O., Bonacorsi, S., and Bingen, E. (2000) Rapid and simple determination of the *Escherichia coli* phylogenetic group. *Appl Environ Microbiol* **66**: 4555–4558.

Darfeuille-Michaud, A., Boudeau, J., Bulois, P., Neut, C., Glasser, A.-L., Barnich, N., *et al.* (2004) High prevalence of adherent-invasive *Escherichia coli* associated with ileal mucosa in Crohn's disease. *Gastroenterology* **127**: 412–421.

Darfeuille-Michaud, A., Neut, C., Barnich, N., Lederman, E., Martino, P.D., Desreumaux, P., *et al.* (1998) Presence of adherent *Escherichia coli* strains in ileal mucosa of patients with Crohn's disease. *Gastroenterology* **115**: 1405–1413.

Datsenko, K.A., and Wanner, B.L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* **97**: 6640–6645.

Delmas, J., Dalmasso, G., and Bonnet, R. (2015) *Escherichia coli*: the good, the bad and the ugly. Clin Microbiol Open Access 04 Accessed June 18, 2015.

Derrien, M., Passel, M.W., van Bovenkamp, J.H., van de Schipper, R.G., de Vos, W.M., and Dekker, J. (2010) Mucin-bacterial interactions in the human oral cavity and digestive tract. *Gut Microbes* **1**: 254–268.

Dharmani, P., Srivastava, V., Kissoon-Singh, V., and Chadee, K. (2009) Role of intestinal mucins in innate host defense mechanisms against pathogens. *J Innate Immun* **1**: 123–135.

Dozois, C.M., Dho-Moulin, M., Brée, A., Fairbrother, J.M., Desautels, C., and Curtiss, R. (2000) Relationship between the Tsh autotransporter and pathogenicity of avian *Escherichia coli* and localization and analysis of the Tsh genetic region. *Infect Immun* **68**: 4145–4154.

Dreux, N., Denizot, J., Martinez-Medina, M., Mellmann, A., Billig, M., Kisiela, D., *et al.* (2013) Point mutations in FimH adhesin of Crohn's disease-associated adherent-invasive *Escherichia coli* enhance intestinal inflammatory response. *PLoS Pathog* **9** e1003141.

© 2015 John Wiley & Sons Ltd, *Cellular Microbiology*

Ermund, A., Gustafsson, J.K., Hansson, G.C., and Keita, Å.V. (2013) Mucus properties and goblet cell quantification in mouse, rat and human ileal Peyer's patches. PLoS ONE 8 Accessed June 18, 2015.

Gevers, D., Kugathasan, S., Denson, L.A., Vázquez-Baeza, Y., Treuren, W.V., Ren, B., *et al.* (2014) The treatment-naive microbiome in new-onset Crohn's disease. *Cell Host Microbe* **15**: 382–392.

Grys, T.E., Siegel, M.B., Lathem, W.W., and Welch, R.A. (2005) The StcE protease contributes to intimate adherence of enterohemorrhagic *Escherichia coli* O157:H7 to host cells. *Infect Immun* **73**: 1295–1303.

Hague, A., Elder, D.J., Hicks, D.J., and Paraskeva, C. (1995) Apoptosis in colorectal tumour cells: induction by the short chain fatty acids butyrate, propionate and acetate and by the bile salt deoxycholate. *Int J Cancer J Int Cancer* **60**: 400–406.

Harrington, S.M., Sheikh, J., Henderson, I.R., Ruiz-Perez, F., Cohen, P.S., and Nataro, J.P. (2009) The Pic protease of enteroaggregative *Escherichia coli* promotes intestinal colonization and growth in the presence of mucin. *Infect Immun* **77**: 2465–2473.

Henderson, I.R., Czeczulin, J., Eslava, C., Noriega, F., and Nataro, J.P. (1999) Characterization of Pic, a secreted protease of *Shigella flexneri* and enteroaggregative *Escherichia coli*. *Infect Immun* **67**: 5587–5596.

Hoskins, L.C., Agustines, M., McKee, W.B., Boulding, E.T., Kriaris, M., and Niedermeyer, G. (1985) Mucin degradation in human colon ecosystems. Isolation and properties of fecal strains that degrade ABH blood group antigens and oligosaccharides from mucin glycoproteins. *J Clin Invest* **75**: 944–953.

Hoskins, L.C., and Boulding, E.T. (1981) Mucin degradation in human colon ecosystems. *J Clin Invest* **67**: 163–172.

Johansson, M.E.V., Ambort, D., Pelaseyed, T., Schütte, A., Gustafsson, J.K., Ermund, A., *et al.* (2011) Composition and functional role of the mucus layers in the intestine. *Cell Mol Life Sci CMLS* **68**: 3635–3641.

Johansson, M.E.V., Gustafsson, J.K., Holmén-Larsson, J., Jabbar, K.S., Xia, L., Xu, H., *et al.* (2014) Bacteria penetrate the normally impenetrable inner colon mucus layer in both murine colitis models and patients with ulcerative colitis. *Gut* **63**: 281–291.

Johansson, M.E.V., and Hansson, G.C. (2011) Microbiology. *Keeping Bacteria at a Distance Science* **334**: 182–183.

Kaser, A., Zeissig, S., and Blumberg, R.S. (2010) Inflammatory bowel disease. *Annu Rev Immunol* **28**: 573–621.

Katayama, T., Fujita, K., and Yamamoto, K. (2005) Novel bifidobacterial glycosidases acting on sugar chains of mucin glycoproteins. *J Biosci Bioeng* **99**: 457–465.

Kobayashi, R.K., Gaziri, L.C., and Vidotto, M.C. (2010) Functional activities of the Tsh protein from avian pathogenic *Escherichia coli* (APEC) strains. *J Vet Sci* **11**: 315–319.

Kotlowski, R., Bernstein, C.N., Sepehri, S., and Krause, D.O. (2007) High prevalence of *Escherichia coli* belonging to the B2 + D phylogenetic group in inflammatory bowel disease. *Gut* **56**: 669–675.

Kumar, P., Luo, Q., Vickers, T.J., Sheikh, A., Lewis, W.G., and Fleckenstein, J.M. (2014) EatA, an immunogenic protective antigen of enterotoxigenic *Escherichia coli*, degrades intestinal mucin. *Infect Immun* **82**: 500–508.

Leyton, D.L., Sloan, J., Hill, R.E., Doughty, S., and Hartland, E.L. (2003) Transfer region of pO113 from enterohemorrhagic *escherichia coli*: similarity with R64 and identification of a novel plasmid-encoded autotransporter, EpeA. *Infect Immun* **71**: 6307–6319.

Lidell, M.E., Moncada, D.M., Chadee, K., and Hansson, G.C. (2006) *Entamoeba histolytica* cysteine proteases cleave the MUC2 mucin in its C-terminal domain and dissolve the protective colonic mucus gel. *Proc Natl Acad Sci U S A* **103**: 9298–9303.

Man, S.M., Kaakoush, N.O., and Mitchell, H.M. (2011) The role of bacteria and pattern-recognition receptors in Crohn's disease. *Nat Rev Gastroenterol Hepatol* **8**: 152–168.

Martinez-Medina, M., Aldeguer, X., Lopez-Siles, M., González-Huix, F., López-Oliu, C., Dahbi, G., *et al.* (2009) Molecular diversity of *Escherichia coli* in the human gut: new ecological evidence supporting the role of adherent-invasive *E. coli* (AIEC) in Crohn's disease. *Inflamm Bowel Dis* **15**: 872–882.

Martinez-Medina, M., and Garcia-Gil, L.J. (2014) *Escherichia coli* in chronic inflammatory bowel diseases: an update on adherent invasive *Escherichia coli* pathogenicity. *World J Gastrointest Pathophysiol* **5**: 213–227.

Martin, H.M., Campbell, B.J., Hart, C.A., Mpofu, C., Nayar, M., Singh, R., *et al.* (2004) Enhanced *Escherichia coli* adherence and invasion in Crohn's disease and colon cancer. *Gastroenterology* **127**: 80–93.

McGovern, D.P.B., Jones, M.R., Taylor, K.D., Marciante, K., Yan, X., Dubinsky, M., *et al.* (2010) Fucosyltransferase 2 (FUT2) non-secretor status is associated with Crohn's disease. *Hum Mol Genet* **19**: 3468–3476.

McGuckin, M.A., Lindén, S.K., Sutton, P., and Florin, T.H. (2011) Mucin dynamics and enteric pathogens. *Nat Rev Microbiol* **9**: 265–278.

Merritt, M.E., and Donaldson, J.R. (2009) Effect of bile salts on the DNA and membrane integrity of enteric bacteria. *J Med Microbiol* **58**: 1533–1541.

Miquel, S., Martín, R., Rossi, O., Bermúdez-Humarán, L.G., Chatel, J.M., Sokol, H., *et al.* (2013) *Faecalibacterium prausnitzii* and human intestinal health. *Curr Opin Microbiol* **16**: 255–261.

Miquel, S., Peyretaillade, E., Claret, L., Vallée, A. de, Dossat, C., Vacherie, B., *et al.* (2010) Complete genome sequence of Crohn's disease-associated adherent-invasive *E. coli* strain LF82. *PLoS One* **5**: e12714.

Moncada, D., Keller, K., Ankri, S., Mirelman, D., and Chadee, K. (2006) Antisense inhibition of *Entamoeba histolytica* cysteine proteases inhibits colonic mucus degradation. *Gastroenterology* **130**: 721–730.

Mylonaki, M., Rayment, N.B., Rampton, D.S., Hudspith, B.N., and Brostoff, J. (2005) Molecular characterization of rectal mucosa-associated bacterial flora in inflammatory bowel disease. *Inflamm Bowel Dis* **11**: 481–487.

Neut, C., Bulois, P., Desreumaux, P., Membré, J.-M., Lederman, E., Gambiez, L., *et al.* (2002) Changes in the bacterial flora of the neoterminal ileum after ileocolonic resection for Crohn's disease. *Am J Gastroenterol* **97**: 939–946.

Nugent, S.G., Kumar, D., Rampton, D.S., and Evans, D.F. (2001) Intestinal luminal pH in inflammatory bowel disease: possible determinants and implications for therapy with aminosalicylates and other drugs. *Gut* **48**: 571–577.

© 2015 John Wiley & Sons Ltd, *Cellular Microbiology*

Otto, B.R., van Dooren, S.J., Nuijens, J.H., Luirink, J., and Oudega, B. (1998) Characterization of a hemoglobin protease secreted by the pathogenic *Escherichia coli* strain EB1. *J Exp Med* **188**: 1091–1103.

Ouwerkerk, J.P., de Vos, W.M., and Belzer, C. (2013) Glycobiome: bacteria and mucus at the epithelial interface. *Best Pract Res Clin Gastroenterol* **27**: 25–38.

Parreira, V.R., and Gyles, C.L. (2003) A novel pathogenicity island integrated adjacent to the thrW tRNA gene of avian pathogenic *Escherichia coli* encodes a vacuolating autotransporter toxin. *Infect Immun* **71**: 5087–5096.

Pelaseyed, T., Bergström, J.H., Gustafsson, J.K., Ermund, A., Birchenough, G.M.H., Schütte, A., *et al.* (2014) The mucus and mucins of the goblet cells and enterocytes provide the first defense line of the gastrointestinal tract and interact with the immune system. *Immunol Rev* **260**: 8–20.

Picard, B., Garcia, J.S., Gouriou, S., Duriez, P., Brahimi, N., Bingen, E., *et al.* (1999) The link between phylogeny and virulence in *Escherichia coli* extraintestinal infection. *Infect Immun* **67**: 546–553.

Playoust, M.R., and Isselbacher, K.J. (1964) Studies on the transport and metabolism of conjugated bile salts by intestinal mucosa. *J Clin Invest* **43**: 467–476.

Png, C.W., Lindén, S.K., Gilshenan, K.S., Zoetendal, E.G., McSweeney, C.S., Sly, L.I., *et al.* (2010) Mucolytic bacteria with increased prevalence in IBD mucosa augment *in vitro* utilization of mucin by other bacteria. *Am J Gastroenterol* **105**: 2420–2428.

Pope, L.M., Reed, K.E., and Payne, S.M. (1995) Increased protein secretion and adherence to HeLa cells by *Shigella* spp. following growth in the presence of bile salts. *Infect Immun* **63**: 3642–3648.

Prouty, A.M., Brodsky, I.E., Manos, J., Belas, R., Falkow, S., and Gunn, J.S. (2004) Transcriptional regulation of *Salmonella enterica* serovar Typhimurium genes by bile. *FEMS Immunol Med Microbiol* **41**: 177–185.

Restieri, C., Garriss, G., Locas, M.-C., and Dozois, C.M. (2007) Autotransporter-encoding sequences are phylogenetically distributed among *Escherichia coli* clinical isolates and reference strains. *Appl Environ Microbiol* **73**: 1553–1562.

Rhodes, J.M., Campbell, B.J., and Yu, L.-G. (2008) Lectin-epithelial interactions in the human colon. *Biochem Soc Trans* **36**: 1482–1486.

Rolhion, N., and Darfeuille-Michaud, A. (2007) Adherent-invasive *Escherichia coli* in inflammatory bowel disease. *Inflamm Bowel Dis* **13**: 1277–1283.

Ruiz-Perez, F., Wahid, R., Faherty, C.S., Kolappaswamy, K., Rodriguez, L., Santiago, A., *et al.* (2011) Serine protease autotransporters from *Shigella flexneri* and pathogenic *Escherichia coli* target a broad range of leukocyte glycoproteins. *Proc Natl Acad Sci U S A* **108**: 12881–12886.

Rutgeerts, P., Ghoos, Y., Vantrappen, G., and Fevery, J. (1986) Biliary lipid composition in patients with nonoperated Crohn's disease. *Dig Dis Sci* **31**: 27–32.

Sahl, J.W., and Rasko, D.A. (2012) Analysis of global transcriptional profiles of enterotoxigenic *Escherichia coli* isolate E24377A. *Infect Immun* **80**: 1232–1242.

Sepehri, S., Khafipour, E., Bernstein, C.N., Coombes, B.K., Pilar, A.V., Karmali, M., *et al.* (2011) Characterization of *Escherichia coli* isolated from gut biopsies of newly diagnosed patients with inflammatory bowel disease. *Inflamm Bowel Dis* **17**: 1451–1463.

Silva, A.J., Pham, K., and Benitez, J.A. (2003) Haemagglutinin/protease expression and mucin gel penetration in El Tor biotype *Vibrio cholerae*. *Microbiol Read Enlgl* **149**: 1883–1891.

Sokol, H., Pigneur, B., Watterlot, L., Lakhdari, O., Bermúdez-Humarán, L.G., Gratadoux, J.-J., *et al.* (2008) *Faecalibacterium prausnitzii* is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. *Proc Natl Acad Sci* **105**: 16731–16736.

Sokol, H., and Seksik, P. (2010) The intestinal microbiota in inflammatory bowel diseases: time to connect with the host. *Curr Opin Gastroenterol* **26**: 327–331.

Sokol, H., Seksik, P., Furet, J.P., Firmesse, O., Nion-Larmurier, I., Beaugerie, L., *et al.* (2009) Low counts of *Faecalibacterium prausnitzii* in colitis microbiota. *Inflamm Bowel Dis* **15**: 1183–1189.

Stathopoulos, C., Provence, D.L., and Curtiss, R. (1999) Characterization of the avian pathogenic *Escherichia coli* hemagglutinin Tsh, a member of the immunoglobulin a protease-type family of autotransporters. *Infect Immun* **67**: 772–781.

Strober, W., Fuss, I., and Mannon, P. (2007) The fundamental basis of inflammatory bowel disease. *J Clin Invest* **117**: 514–521.

Tajima, N., Kawai, F., Park, S.-Y., and Tame, J.R.H. (2010) A novel intein-like autoproteolytic mechanism in autotransporter proteins. *J Mol Biol* **402**: 645–656.

Theodoratou, E., Campbell, H., Ventham, N.T., Kolarich, D., Pučić-Baković, M., Zoldoš, V., *et al.* (2014) The role of glycosylation in IBD. *Nat Rev Gastroenterol Hepatol* **11**: 588–600.

Wibbenmeyer, J.A., Provenzano, D., Landry, C.F., Klose, K. E., and Delcour, A.H. (2002) *Vibrio cholerae* OmpU and OmpT porins are differentially affected by bile. *Infect Immun* **70**: 121–126.

Willing, B., Halfvarson, J., Dicksved, J., Rosenquist, M., Järnerot, G., Engstrand, L., *et al.* (2009) Twin studies reveal specific imbalances in the mucosa-associated microbiota of patients with ileal Crohn's disease. *Inflamm Bowel Dis* **15**: 653–660.

Xavier, R.J., and Podolsky, D.K. (2007) Unravelling the pathogenesis of inflammatory bowel disease. *Nature* **448**: 427–434.

Yen, Y.T., Kostakioti, M., Henderson, I.R., and Stathopoulos, C. (2008) Common themes and variations in serine protease autotransporters. *Trends Microbiol* **16**: 370–379.

## Supporting information

Additional Supporting Information may found in the online version of this article at the publisher's web-site:

**Table S1.** Bacterial strains and plasmids used in this study.

**Table S2.** Primers used in this study.

**Fig. S1:** A schematic representation of the 1,376 amino acids of Vat-AIEC. Analysis of the amino acid sequence of Vat-AIEC revealed an extended leader peptide signal sequence (amino

© 2015 John Wiley & Sons Ltd, *Cellular Microbiology*

acid 1 to 55), a N-terminal α-domain (amino acids 56 to 1099) and a C-terminal β-domain (amino acids 1100 to 1376), as observed in other SPATE proteins. The predicted molecular mass of each domain is indicated. This data suggested that LF82 encodes a SPATE-type autotransporter that allows the secretion of an active serine hydrolase.

**Fig. S2:** *vat-AIEC*-carrying AIEC cleave Muc2. (A) SDS-PAGE analysis of untreated intestinal mucin from mice and mucin treated with culture supernatant of LF82. The gel is PAS strained, and the band which is cleaved is indicated. (B) Mucin was extracted from mice colon and incubated overnight with culture supernatants of LF73, LF25, LF16 and 6259, four AIEC strains. LF16 and LF73 carry the *vat-AIEC* gene, unlike LF25 and 6259. Mucin degradation was observed with *vat-AIEC*-carrying AIEC by Western blot analysis with anti-Muc2 antibody. (C) Mucin was extracted from mice gut and incubated overnight with culture supernatants of LF82 and LF82Δ*vat-AIEC*/*vat-AIEC*S260I. Mucin degradation was observed with LF82 but not with the inactive mutant.

**Fig. S3:** Effect of Vat-AIEC on mice colonization with dextran sulfate sodium (DSS). Mice were orally infected with $10^9$ LF82 (round) or $10^9$ LF82Δ*vat-AIEC* (square) after pretreatment with 0.25% DSS. Feces were collected three days after oral infection and LF82 or LF82Δ*vat-AIEC* were counted. Quantification of colonic mucosal-associated bacteria was performed three days post-infection. There was no significant difference of colonization between LF82 and LF82Δ*vat-AIEC*. Statistical analysis was performed with Mann-Whitney test. (ns, not significant).

**Fig. S4:** Vat-AIEC did not enhance the growth of LF82 in the presence of mucin. LF82 (round) and LF82Δ*vat-AIEC* (square) were grown in M9 minimal medium containing 0.05% glucose, in which 1% mucin was added after 6 hours of incubation (arrow). Each value is the mean ± SEM of at least three separate experiments (ANOVA, $p > 0.05$).

© 2015 John Wiley & Sons Ltd, *Cellular Microbiology*

**Figure S1: A schematic representation of the 1,376 amino acids of Vat-AIEC.** Analysis of the amino acid sequence of Vat-AIEC revealed an extended leader peptide signal sequence (amino acid 1 to 55), a N-terminal α-domain (amino acids 56 to 1099) and a C-terminal β-domain (amino acids 1100 to 1376), as observed in other SPATE proteins. The predicted molecular mass of each domain is indicated. This data suggested that LF82 encodes a SPATE-type autotransporter that allows the secretion of an active serine hydrolase.

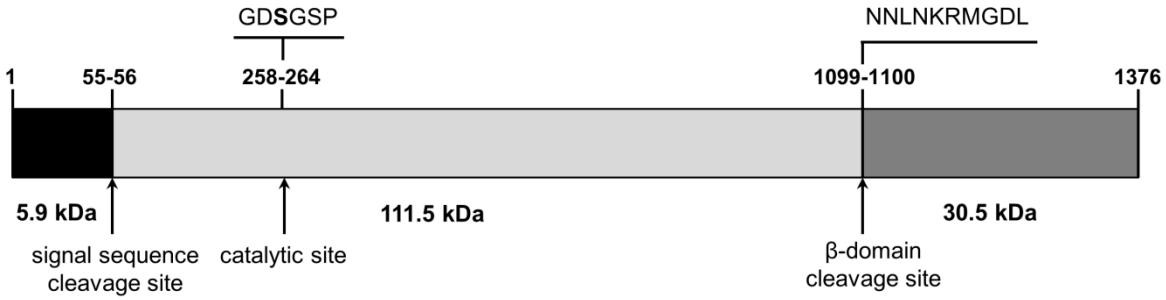


**Figure S2: *vat-AIEC*-carrying AIEC cleave Muc2.** (**A**) SDS-PAGE analysis of untreated intestinal mucin from mice and mucin treated with culture supernatant of LF82. The gel is PAS strained, and the band which is cleaved is indicated. (**B**) Mucin was extracted from mice colon and incubated overnight with culture supernatants of LF73, LF25, LF16 and 6259, four AIEC strains. LF16 and LF73 carry the *vat-AIEC* gene, unlike LF25 and 6259. Mucin degradation was observed with *vat-AIEC*-carrying AIEC by Western blot analysis with anti-Muc2 antibody. (**C**) Mucin was extracted from mice gut and incubated overnight with culture supernatants of LF82 and LF82Δ*vat-AIEC*/*vat-AIEC*S260I. Mucin degradation was observed with LF82 but not with the inactive mutant.

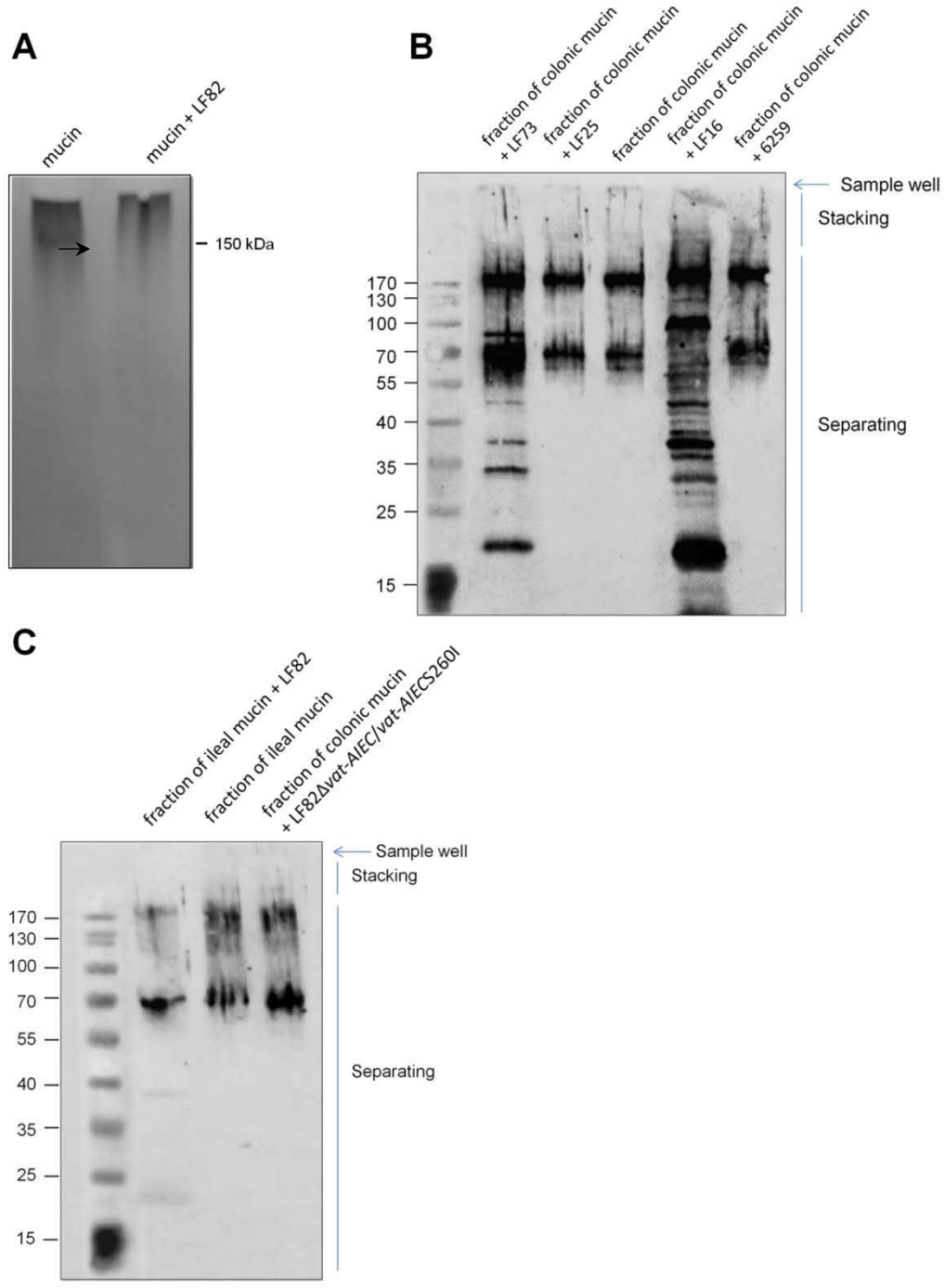


**Fig. S3: Effect of Vat-AIEC on mice colonization with dextran sulfate sodium (DSS).** Mice were orally infected with $10^9$ LF82 (round) or $10^9$ LF82Δ*vat-AIEC* (square) after pretreatment with 0.25% DSS. Feces were collected three days after oral infection and LF82 or LF82Δ*vat-AIEC* were counted. Quantification of colonic mucosal-associated bacteria was performed three days post-infection. There was no significant difference of colonization between LF82 and LF82Δ*vat-AIEC*. Statistical analysis was performed with Mann-Whitney test. (ns, not significant)

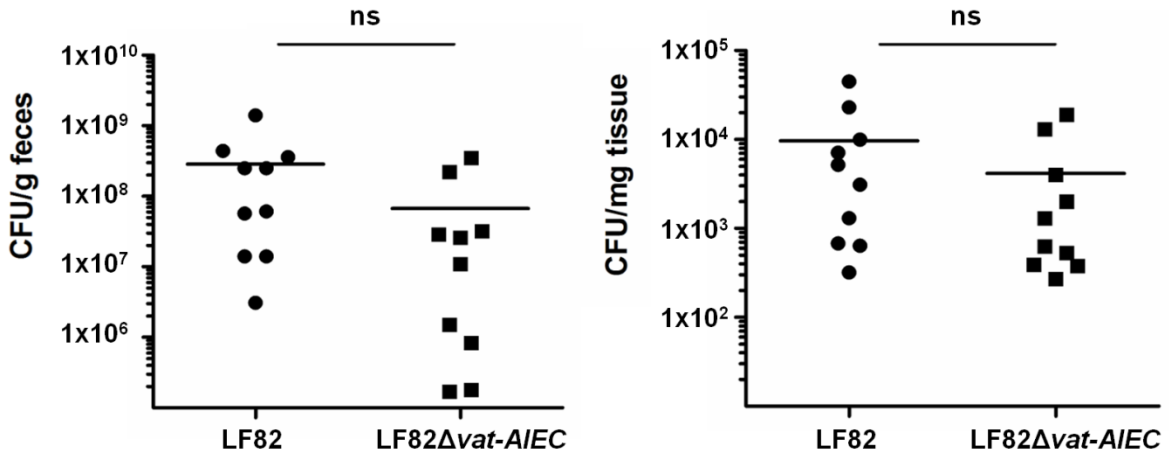


**Figure S4: Vat-AIEC did not enhance the growth of LF82 in the presence of mucin.** LF82 (round) and LF82Δ*vat-AIEC* (square) were grown in M9 minimal medium containing 0.05% glucose, in which 1% mucin was added after 6 hours of incubation (arrow). Each value is the mean ± SEM of at least three separate experiments (ANOVA, $p > 0.05$).

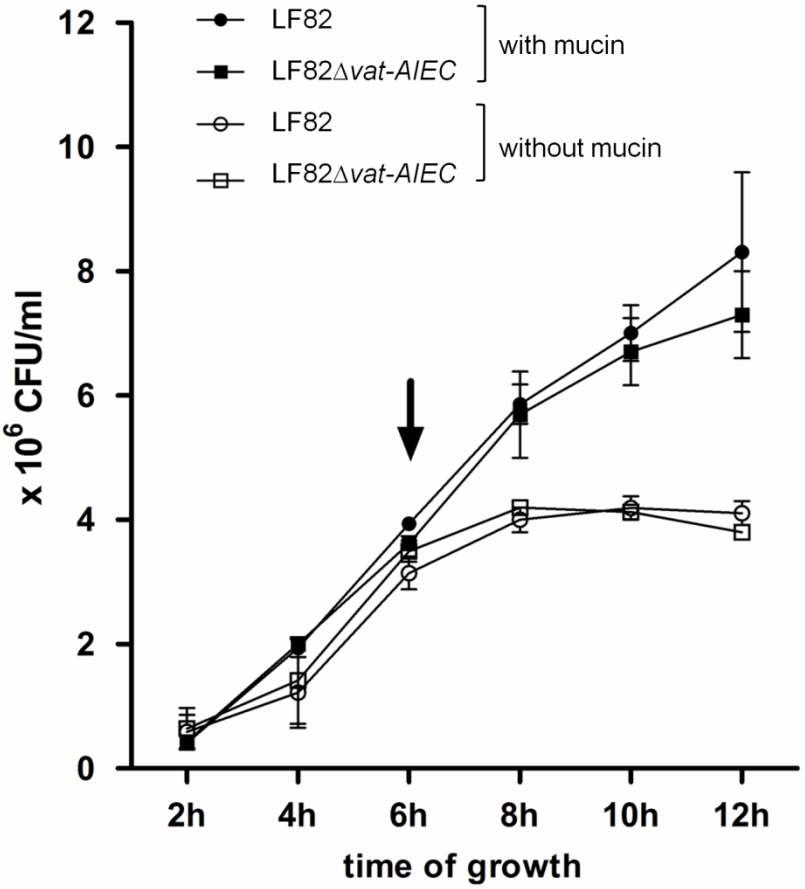


**Table S1:** Bacterial strains and plasmids used in this study

| Strain or plasmid | Relevant characteristics | Reference |
|---|---|---|
| LF82 | AIEC reference strain | (1) |
| LF82Δ*vat-AIEC* | LF82 isogenic mutant deleted for the *vat-AIEC* gene | this study |
| LF82Δ*vat-AIEC*/*vat-AIEC* | LF82Δ*vat-AIEC* containing plasmid pBK-CMV carrying *vat-AIEC* | this study |
| LF82Δ*vat-AIEC*/*vat-AIECS260I* | LF82Δ*vat-AIEC* containing plasmid pBK-CMV carrying *vat-AIECS260I* | this study |
| EcoR42 | APEC strain containing the Tsh mucinase | (2) |
| K12/*vat-AIEC* | *E. coli* K12 strain containing plasmid pBK-CMV carrying *vat-AIEC* | this study |
| LF16 and LF73 | AIEC strains containing the *vat-AIEC* gene | (3) |
| LF25 and 6259 | AIEC strains without the *vat-AIEC* gene | (3) |
| BW25141 | Plasmid pKD4 carrying a kanamycin resistance cassette | (4) |
| pKOBEG | pBAD cloning vector harboring a λ phage redγβα operon; chloramphenicol resistant | (5) |
| pCP20 | Ampicillin and chloramphenicol resistant plasmid that shows temperature-sensitive replication and thermal induction of FLP synthesis | (4) |
| pRS550 | Plasmid harboring β-galactosidase gene | (6) |
| pRS550-P*vat-AIEC* | pRS550 harboring the *vat-AIEC* promoter | this study |
| pGFP | pFPV25.1 plasmid encoding the green fluorescent protein (GFP) | (7) |

## REFERENCES


1. Barnich N, Darfeuille-Michaud A. 2007. Adherent-invasive *Escherichia coli* and Crohn's disease. Curr Opin Gastroenterol 23:16–20.

2. Restieri C, Garriss G, Locas M-C, Dozois CM. 2007. Autotransporter-Encoding Sequences Are Phylogenetically Distributed among *Escherichia coli* Clinical Isolates and Reference Strains. Appl Environ Microbiol 73:1553–1562.

3. Dreux N, Denizot J, Martinez-Medina M, Mellmann A, Billig M, Kisiela D, Chattopadhyay S, Sokurenko E, Neut C, Gower-Rousseau C, Colombel J-F, Bonnet R, Darfeuille-Michaud A, Barnich N. 2013. Point Mutations in FimH Adhesin of Crohn's Disease-Associated Adherent-Invasive *Escherichia coli* Enhance Intestinal Inflammatory Response. PLoS Pathog 9:e1003141.

4. 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.

5. Chaveroche MK, Ghigo JM, d' Enfert C. 2000. A rapid method for efficient gene replacement in the filamentous fungus Aspergillus nidulans. Nucleic Acids Res 28:E97.

6. Simons RW, Houman F, Kleckner N. 1987. Improved single and multicopy lac-based cloning vectors for protein and operon fusions. Gene 53:85–96.

7. Valdivia RH, Falkow S. 1997. Fluorescence-based isolation of bacterial genes expressed within host cells. Science 277:2007–2011.

**Table S2:** Primers used in this study

| Primer | Sequence (5'→3') | PCR product size | Use |
|---|---|---|---|
| LFd*vat-AIEC*-F | TACCATATTGTGGATCGTAATGAACACA GTTCATCTGATCTCCACACACCGTAGGC TGGAGCTGCTTC | 1,574 bp | construction of isogenic mutant |
| LFd*vat-AIEC*-R | AGTTGCGCCATAATCGGCGGCATTCAGC TTATGAAATGTCAGGTCATTCCCATATG AATATCCTCCTTAG | | |
| BamHIVat-AIEC-F | AGCGGATCCATGAATAAAATATACGCTC | 4,130 bp | *vat-AIEC* cloning |
| EcoRIVat-AIEC-R | GGAATTCTCAGAATGAATAACG | | |
| Vat-AIEC-F | CCTCTTTTGGAGAGGCCGGCGACAGT | 204 bp | PCR amplification |
| Vat-AIEC-R | GGGGCGCCGGATGAGGCATTA | | |
| BamHIP*vat-AIEC*-F | AGCGGATCCGTCGAATGATGCCATTCGT A | 475 bp | *vat-AIEC* promoter cloning |
| EcoRIP*vat-AIEC*-R | AACGAATTCAGTAAATGATTCCGGAACG | | |
| Vat-AIECS260I-F | GGAGAGGCCGGCGACATTGGTTCTCCAT TATTTGG | - | construction of inactive mutant |
| Vat-AIECS260I-R | CCAAATAATGGAGAACCAATGTCGCCGG CCTCTCC | | |
| VAT-AIEC-F | ATAGCATCCTGCCCTCTTTTGG | 160 bp | RT-PCR amplification |
| VAT-AIEC-R | CTGTGAGAGAAAACTCTGAGG | | |
| 16S-F | ATGACCAGCCACACTGGAAC | 150 bp | RT-PCR amplification |
| 16S-R | CTTCCTCCCCGCTGAAAGTA | | |

## SUPPLEMENTAL EXPERIMENTAL PROCEDURES

### Bacterial strains and media

AIEC strain LF82 was isolated from a chronic ileal lesion of a patient with CD and belongs to *E. coli* serotype O83:H1 (Darfeuille-Michaud *et al.*, 1998). This strain has a cephalosporinase and can be selected on amoxicillin-containing agar plates. *E. coli* strains were grown overnight at 37°C in lysogeny broth (LB) or in M9 5X containing broth (with 30g of $Na_2HPO_4$, 15g of $KH_2PO_4$, 5g of $NH_4Cl$, 2.5g of NaCl and 15mg of $CaCl_2$ for one liter). Amoxicillin (32 mg/l), kanamycin (50 mg/l) and chloramphenicol (25 mg/l) (Sigma) were added when required.

75 AIEC strains were obtained from CD patients (50 isolated from ileal mucosa and 25 isolated from stools of CD patients), and 25 non-AIEC strains isolated from ileal mucosa of CD patients were also obtained. In these all strains, 46% belonged to *E. coli* phylogenetic group B2, 25% to phylogroup A, 20% to phylogroup D, and 9% to phylogroup B1, as previously reported in the AIEC population isolated from CD patients (Martinez-Medina, Aldeguer, *et al.*, 2009).

### Construction of isogenic mutant LF82Δ*vat-AIEC*

*E. coli* LF82 was transformed with pKOBEG, a plasmid encoding the Red proteins that protect linear DNA from degradation in bacteria. The plasmid was maintained in bacteria at 30°C with 25 mg/l of chloramphenicol and 1 mM of L-arabinose. The Flp recognition target-flanked cassette harboring the kanamycin resistance cassette was generated by PCR from *E. coli* BW25141 with LFd*vat-AIEC*-F and LFd*vat-AIEC*-R primers (Table S2) and High Fidelity Platinum Taq polymerase (Invitrogen) according to the manufacturer's instructions. The PCR products were electroporated in previously glycerol-washed *E. coli* LF82. The resulting LF82Δ*vat-AIEC* isogenic mutants ($Km^R$) were selected on LB agar containing 50 mg/l kanamycin. The replacement of the *vat-AIEC* gene by the kanamycin resistance cassette was confirmed by PCR. The kanamycin resistance cassette was then removed from LF82Δ*vat-AIEC* by the transient expression of the Flp recombinase from the pCP20 plasmid, creating the LF82*Δvat-AIEC* ($Km^S$) strain.

### Transcomplementation of isogenic mutant LF82Δ*vat-AIEC*

The *vat-AIEC* gene was amplified by PCR from *E. coli* LF82 genomic DNA by using BamHIVat-AIEC-F and EcoRIVat-AIEC-R primers (Table S2). The amplified DNA was purified with a NucleoSpin extract kit (Macherey-Nagel), digested with EcoRI and BamHI (New England Biolabs), and ligated to EcoRI-BamHI-digested expression vector pBK-CMV (Agilent Technologies). This construct was electroporated into LF82Δ*vat-AIEC* ($Km^S$) electrocompetent strains and selected on Mueller Hinton agar containing 50 µg/ml kanamycin. The presence of the *vat-AIEC* gene was confirmed by PCR. The construction was checked by double stranded DNA sequencing (GATC biotech).

**Mucinolytic activity assay**

The strains were incubated for 48-72 h at 37°C on plates containing 2.0% agar, 1.0% mucin from porcine stomach (Sigma), and 0.01% glucose in M9 medium. The colonies were suspended in phosphate-buffered saline and harvested by centrifugation at 16,000 g for 60 min at 4°C. The supernatants were filter sterilized (0.22 µm) and incubated for 24 h at 37°C on a medium containing 1.5% agarose and 1% mucin. Zones of clearing were visualized by a staining with a solution containing 0.1% amido black in 3.5 M acetic acid for 20 min, followed by destaining step with the used of 5% acetic acid and 0.5% glycerol for 30 min.

**Mucinolytic activity on SDS-PAGE**

For SDS-PAGE analysis, treated (37°C for 12 h with culture supernatant of LF82) and untreated intestinal mucin specimens from mice were electrophoresed on an SDS–5% polyacrylamide gel. The gel was then incubated in 7.5% acetic acid and 0.5% periodic acid for 1 h and then in Schiff reagent (PAS) overnight. The reaction was stopped by the addition of 5% acetic acid and 40% ethanol.

**Mucin preparation for Muc2 cleavage assay**

Conventional mice (C57BL/6J) were in-house bred. Twenty mice were euthanized; mucosal scrapings from colon or ileum were collected, and the mucins were semi-purified as previously described (Larsson *et al.*, 2009). Briefly, mucosal scrapings were homogenized in extraction guanidinium chloride buffer (extr-GuHCl) (40µL; 6 M with 5 mM EDTA, 0.01 M $NaH_2PO_4$) in the absence of protease inhibitor using a plastic mini pestle. An additional 160 µL extr-GuHCl was added and the homogenized material was stirred gently overnight at 4°C. The solution was centrifuged for 30 min at 14 000 rpm, and the GuHCl soluble and insoluble fractions were separated by collecting the supernatant. To the insoluble fraction containing Muc2, 100µL of extr-GuHCl was added followed by stirring for 4 h. The centrifugation step was repeated with subsequent division of the soluble and the insoluble fractions. Guanidinium chloride for reduction (red-GuHCl) (120 µL, 6 M with 5 mM EDTA, 0.1 M Tris) was added to the insoluble fraction. Samples were transferred to dialysis cups (10 kDa MWCO, Slide-A-Lyzer MINI dialysis cups, Pierce) and dialyzed against water at 4°C overnight or reduced. Reducing agent dithiotreitol was added to the insoluble fractions to a final concentration of 50 mM. The solutions were stirred for 1.5 h at 37°C. Alkylating agent iodoacetamide was added to a final concentration of 125 mM, and the solutions were stirred for 1 h in room temperature in the dark. Reduced samples were dialysed as previously described. Protein content was evaluated by BCA assay (Bio-Rad).

**ELISA**

The amounts of KC and TNF-α released in the culture supernatant of colon specimens after 24h of incubation at 37°C under 5% of $CO_2$, and the amount of myeloperoxidase were measured by ELISA (R&D systems) following the manufacturer's instructions.

**Immunohistochemistry**

After fixation with Carnoy's solution to preserve mucus, paraffin-embedded colonic loops were analyzed as previously described (Nandan *et al.*, 2014). In brief, the colonic loop sections were deparaffinized in xylene, incubated in 3% hydrogen peroxide and rehydrated before heat-induced epitope retrieval with sodium citrate buffer (pH 6.0). The tissue sections were incubated with blocking buffer and then with primary anti-*Escherichia coli* O83 antibody at 4°C overnight. Sections were washed and revealed using the Vectastain ABC kit (Vector Labs) and a DAB detection kit (Invitrogen). The sections were counterstained with Gill's hematoxylin (Vector Labs).

**Electronic microscopy**

Tissue sections were fixed at 4°C for 48 hours with 0.5% PFA and 2% glutaraldehyde in 0.2M sodium cacodylate buffer (pH 7.4), rinsed in the same buffer, and then post-fixed 1 hour in 1% osmium tetroxide at room temperature. Fixation was followed by deshydratation of the tissue in ethanol and acetone, and embedding in Epon resin. Ultrathin sections of approximately 70 nm thickness were cut with an ultra-microtome Leica UC6, stained with uranyl acetate and lead citrate, and examined using transmission electron microscope Hitachi H-7650 with Hamamatsu AMT40 camera.

**Visualization of bacteria mobility with live-cell imaging**

GFP-carrying LF82 and LF82Δ*vat-AIEC* were grown overnight at 37°C without shaking in LB broth. The bacteria were visualized in real-time using a Leica TCS SP5 confocal microscope and track speed mean values were determined ($\mu m.s^{-1}$).

**Yeast agglutination assay**

Commercial baker's yeast (*Saccharomyces cerevisiae*) was suspended in PBS (20 g/L). Bacterial strains were grown overnight at 37°C without agitation in LB broth, washed, and resuspended in PBS at an optical density of 0.5 at 620 nm. The bacterial suspension was diluted and 50 µL was deposited on a 96-well microplate (Greiner). Equal volumes of yeast cell suspension were added to each well. The aggregation was monitored visually, and the titer was recorded as the last dilution of bacteria that gave a positive aggregation reaction.

**Mucin gel penetration assay**

LF82 and LF82Δ*vat-AIEC* that were grown overnight in LB were diluted 1:200 in LB and grown to mid-log phase for 3 hours. The bacteria were pelleted by centrifugation and resuspended in PBS to 10 × $10^9$ cells/ml. A solution containing 10% mucin from porcine stomach and 0.3% agarose in HBSS was loaded into a 1-ml tuberculin syringe fitted with a stopcock, creating a mucous column. A 0.1 ml aliquot of cell suspension was layered onto the mucin by using a separate column for each strain. The columns were incubated for 3 h at 37°C in a vertical position. At the end of the incubation, the stopcocks were removed and 0.1 ml fractions were collected from the button by applying gentle pressure. Each fraction was diluted and plated to LB agar media. The number of CFU per fraction was standardized by the total number of CFU per column (relative cell density).

**Bacterial growth curves**

The growth of the LF82 and LF82Δ*vat-AIEC* strains was observed as previously described (Harrington *et al.*, 2009) by subculturing overnight broth cultures in M9 minimal medium with 0.05% glucose. The optical density at 600 nm and the deduced bacterial quantity were measured at various time points. Porcine mucin 1% was added after 6 hours of culture when needed.

**REFERENCES**


1. Darfeuille-Michaud A, Neut C, Barnich N, Lederman E, Di Martino P, Desreumaux P, Gambiez L, Joly B, Cortot A, Colombel J-F. 1998. Presence of adherent *Escherichia coli* strains in ileal mucosa of patients with Crohn's disease. Gastroenterology 115:1405–1413.
2. Martinez-Medina M, Aldeguer X, Lopez-Siles M, González-Huix F, López-Oliu C, Dahbi G, Blanco JE, Blanco J, Garcia-Gil LJ, Darfeuille-Michaud A. 2009. Molecular diversity of *Escherichia coli* in the human gut: new ecological evidence supporting the role of adherent-invasive E. coli (AIEC) in Crohn's disease. Inflamm Bowel Dis 15:872–882.
3. Larsson JMH, Karlsson H, Sjövall H, Hansson GC. 2009. A complex, but uniform O-glycosylation of the human MUC2 mucin from colonic biopsies analyzed by nanoLC/MSn. Glycobiology 19:756–766.
4. Nandan MO, Ghaleb AM, Liu Y, Bialkowska AB, McConnell BB, Shroyer KR, Robine S, Yang VW. 2014. Inducible intestine-specific deletion of Krüppel-like factor 5 is characterized by a regenerative response in adult mouse colon. Dev Biol 387:191–202.
5. Harrington SM, Sheikh J, Henderson IR, Ruiz-Perez F, Cohen PS, Nataro JP. 2009. The Pic Protease of Enteroaggregative *Escherichia coli* Promotes Intestinal Colonization and Growth in the Presence of Mucin. Infect Immun 77:2465–2473.

## 3. Commentaires de la publication n°1

### ✓ Il existe chez la souche LF82 une activité mucinolytique portée par le gène *vat-AIEC*

Afin d'atteindre les cellules intestinales, les AIEC doivent d'abord franchir la couche de mucus protectrice. Certaines bactéries pathogènes comme *Shigella flexneri* ou certains pathovars d'*E. coli* sécrètent une protéase à activité mucinolytique (Stathopoulos *et al.*, 1999; Henderson *et al.*, 1999). Nous avons donc recherché une activité mucinolytique chez la souche AIEC LF82 à l'aide d'un gel de mucine. Cette activité n'étant pas portée par une mucinase déjà décrite, nous avons retrouvé au sein du génome de la souche LF82 un gène annoté *LF82_p043* (Miquel, Peyretaillade, *et al.*, 2010) codant une sérine protéase, dénommée Vat-AIEC, faisant partie de la classe des autotransporteurs comme la mucinase Pic de *Shigella* ou Hbp d'*E. coli*. La délétion de ce gène par recombinaison homologue a conduit à la perte de l'activité mucinase. Cette dernière a ensuite été restaurée avec l'ajout de ce même gène apporté par un plasmide. Cet autotransporteur est donc bien responsable de l'activité mucinolytique de la souche LF82. De plus, l'activité mucinase retrouvée dans un surnageant de culture suggère que cette protéase est clivée et relarguée dans le milieu extracellulaire.

Cette activité mucinolytique a été ensuite vérifiée par Western Blot à l'aide de surnageants de culture mis en contact avec de la mucine Muc2 purifiée à partir d'intestins de souris. La souche LF82 et le mutant transcomplémenté ont montré une dégradation de cette mucine, contrairement au mutant LF82Δ*vat-AIEC*. Pour vérifier que l'activité mucinase est bien portée par le gène *vat-AIEC* et ne provient pas d'une régulation par Vat-AIEC d'autres gènes chez la souche LF82, nous avons introduit le plasmide porteur du gène *vat-AIEC* dans un *E. coli* K12 et nous avons observé la même dégradation de mucine Muc2. L'ajout de phenylmethylsulfonyl fluoride, un inhibiteur de sérine protéase, inhibe l'activité mucinolytique de la LF82. D'autres souches AIEC porteuses du même gène *vat-AIEC* montrent également une activité mucinolytique sur la mucine Muc2. L'activité mucinase de la LF82 a également été vérifiée sur la protéine MUC2 humaine recombinante.

Toutes ces données montrent que le gène *vat-AIEC* code une sérine protéase responsable d'une activité mucinolytique, en particulier sur Muc2, mucine prédominante au niveau intestinal.

### ✓ La mucinase Vat-AIEC est présente au sein d'un phylogroupe particulier d'*E. coli* AIEC

Le gène *vat-AIEC* est plus fréquemment présent chez les souches AIEC que les non-AIEC, et au sein de ces souches le gène *vat-AIEC* est retrouvé essentiellement au sein du phylogroupe B2 et minoritairement dans le phylogroupe D. La présence de ce gène semble donc être une caractéristique des souches AIEC de phylogroupe B2 chez lesquelles il est retrouvé un ensemble de caractéristiques communes comme par exemple la présence de mutations (S70/N78) au niveau de l'adhésine FimH qui assure une meilleure adhésion des souches aux cellules épithéliales (Dreux *et al.*, 2013). Le phylogroupe B2, avec le phylogroupe D, sont par ailleurs connus pour regrouper la plupart des souches virulentes, et sont également plus fréquemment retrouvés dans les MICI (Picard *et al.*, 1999; Kotlowski *et al.*, 2007)

### ✓ La mucinase Vat-AIEC est impliquée dans la colonisation des cellules intestinales par les souches AIEC

Afin de voir si la protéine Vat-AIEC, de par son activité mucinolytique, favorise la colonisation de la souche LF82, nous avons utilisé un modèle murin. Après 3 jours de traitement antibiotique utilisé pour éliminer la flore résidente, les souris ont été gavées avec la souche sauvage ou la souche mutée. Nous avons ainsi pu montrer que le nombre de bactéries LF82 retrouvées dans les fèces ou associées aux muqueuses intestinales est significativement plus important que celui du mutant. La différence est même plus importante au niveau du colon, sûrement du fait de la couche de mucus qui est plus épaisse à ce niveau-là. Nous avons également montré sur des coupes de tissu la présence de la souche LF82 au contact des cellules épithéliales intestinales, contrairement à la souche délétée du gène *vat-AIEC*. La colonisation du colon de souris par la souche LF82 a conduit également à une augmentation

de la production de cytokines pro-inflammatoires KC et TNF-α et de myélopéroxidases contrairement à la souche mutée.

Ces données montrent que la protéine Vat-AIEC favorise la colonisation de l'intestin par la souche LF82, ce qui conduit à une augmentation de l'inflammation en réponse à cette colonisation.

### ✓ La mucinase Vat-AIEC favorise la pénétration des AIEC au sein du mucus

Pour mieux comprendre quels étaient les mécanismes qui reliaient la présence de la protéine Vat-AIEC et la colonisation de l'intestin des souris, différentes expérimentations ont été entreprises.

La dégradation des mucines par l'enzyme Vat-AIEC pourrait produire des nutriments permettant aux AIEC de se multiplier. Nous avons donc effectué des courbes de croissance en présence de mucine, mais nous n'avons observé aucune différence de croissance entre la souche LF82 et son mutant.

La dégradation de mucines *in vivo* pourrait permettre à la souche LF82 de traverser le mucus et d'atteindre les cellules épithéliales intestinales. Pour vérifier cette hypothèse, nous avons testé la capacité de la souche LF82 et du mutant LF82Δ*vat-AIEC* à pénétrer à travers une colonne constituée de mucines, après avoir vérifié que la délétion du gène *vat-AIEC* n'entrainait aucune perte de mobilité chez la souche LF82. La pénétration du mucus était plus importante avec la souche LF82 qu'avec le mutant.

Nous avons également testé l'effet de Vat-AIEC en mesurant la viscosité d'un milieu constitué de mucines après contact avec la souche LF82 ou le mutant. La viscosité était significativement diminuée dans le milieu en contact avec la bactérie LF82 comparé à celui en contact avec le mutant, suggérant une plus grande facilité pour les bactéries à se faufiler jusqu'aux cellules épithéliales.

### ✓ La mucinase Vat-AIEC favorise l'adhésion et invasion des AIEC

L'impact de la mucinase Vat-AIEC sur la capacité des bactéries à adhérer et envahir les cellules a été observé sur des cellules intestinales en culture qui, à l'état différencié, hyperproduisent du mucus. La souche LF82 a montré une meilleure adhésion et invasion de ces cellules, en comparaison à la souche mutée. Cette différence d'adhésion et d'invasion n'a pas été retrouvée sur des cellules ne produisant pas de mucus. Par ailleurs, le nombre de bactéries adhérentes et invasives était plus important avec des cellules non productrices de mucus, ce qui nous indique que la présence de mucus entrave l'adhésion et l'invasion des cellules par les AIEC, et que dans ces conditions la protéine Vat-AIEC la favorise.

L'adhésion des bactéries aux cellules productrices de mucus a été également objectivée par microscopie confocale et microscopie électronique, qui ont confirmé l'adhésion plus importante de la souche LF82 que du mutant LF82Δ*vat-AIEC*. Enfin, cette meillleure adhésion des souches productrices de Vat-AIEC a été vérifiée *in vivo* sur des anses ligaturées d'intestin de souris par marquage anticorps et microscopie électronique.

L'ensemble de ces données montrent que la protéine Vat-AIEC favorise l'adhésion et l'invasion des bactéries aux cellules lorsque la couche de mucus protectrice est présente.

### ✓ Induction du gène *vat-AIEC* par les conditions physico-chimiques retrouvées au niveau de l'iléon terminal

Les AIEC étant plus fréquemment retrouvées dans les formes iléales de la MC, nous avons voulu savoir si l'environnement physico-chimique iléal avait un impact sur la synthèse de cette protéase.

Pour cela, nous avons cloné le promoteur du gène Vat-AIEC devant un gène rapporteur codant la β-galactosidase, afin de tester l'impact du mucus, du pH et/ou la présence de sels biliaires sur la synthèse de Vat-AIEC. En effet, le pH change tout le long du tractus gastro-intestinal, avec un pH cible autour de 7,5 au niveau de l'iléon, et plutôt de 6,5 au niveau du colon. La couche de mucus est également différente entre l'iléon, où elle est plutôt fine pour permettre une bonne absorption des nutriments nécessaires à l'organisme, et le colon, où elle est plus épaisse. Enfin, la présence de sels biliaires au niveau de l'intestin grêle a un impact important sur la communauté bactérienne, en ralentissant notamment les synthèses bactériennes par des dommages induits sur les membranes cellulaires et sur l'ADN bactérien

et en activant la synthèse de certains facteurs de virulence. Il a donc été supposé qu'ils puissent également jouer un rôle sur la synthèse de la protéine Vat-AIEC.

Une augmentation de l'ARNm du gène *vat-AIEC* a été observée en présence de mucine, de sels biliaires et/ou d'un pH correspondant au pH de l'iléon terminal. Les conditions retrouvées au niveau iléal sont donc toutes les trois favorables à une augmentation de l'expression du gène *vat-AIEC*.

Cette étude a ainsi permis la description d'une nouvelle sérine protéase impliquée dans la virulence des AIEC agissant dans les étapes précoces de l'infection, avant même l'adhésion des bactéries à l'épithélium intestinal des patients atteints de MC. Elle a également permis de mieux comprendre comment les AIEC arrivaient à traverser la barrière de mucus pour atteindre les cellules cibles.

## 4. Résultats complémentaires

### ✓ Régulation de l'expression du gène *vat-AIEC* par les facteurs de transcription Fis et CRP

L'implication de plusieurs paramètres (pH, mucines, sels biliaires) à des temps différents dans l'expression du gène *vat-AIEC* suggère une régulation complexe. L'analyse *in silico* (Münch *et al.*, 2005) du promoteur du gène a identifié des sites de liaison aux facteurs de transcription Fis et CRP, deux facteurs connus pour réguler de nombreux gènes chez les bactéries (**Figure 35**) (Dillon and Dorman, 2010).

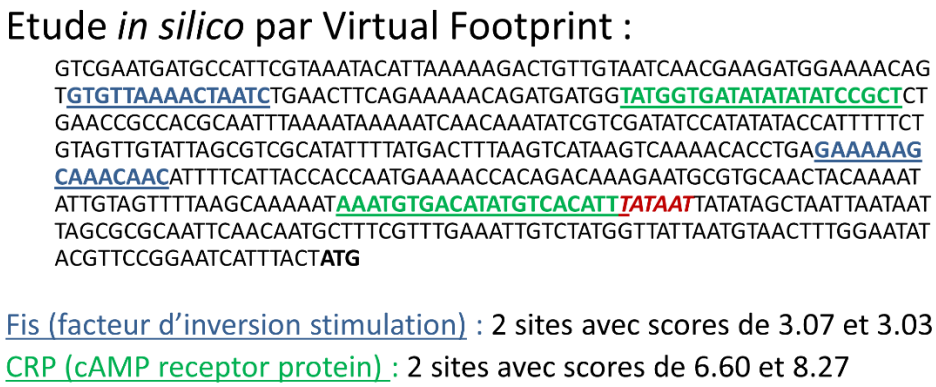



**Figure 35** : Etude *in silico* du promoteur du gène *vat-AIEC*

Afin de vérifier l'implication de ces deux facteurs de transcription dans la régulation du gène *vat-AIEC*, nous avons entrepris la quantification des ARNm de ces facteurs sur les cultures où l'ARNm du gène *vat-AIEC* avait été augmenté. Des résultats préliminaires nous montrent une

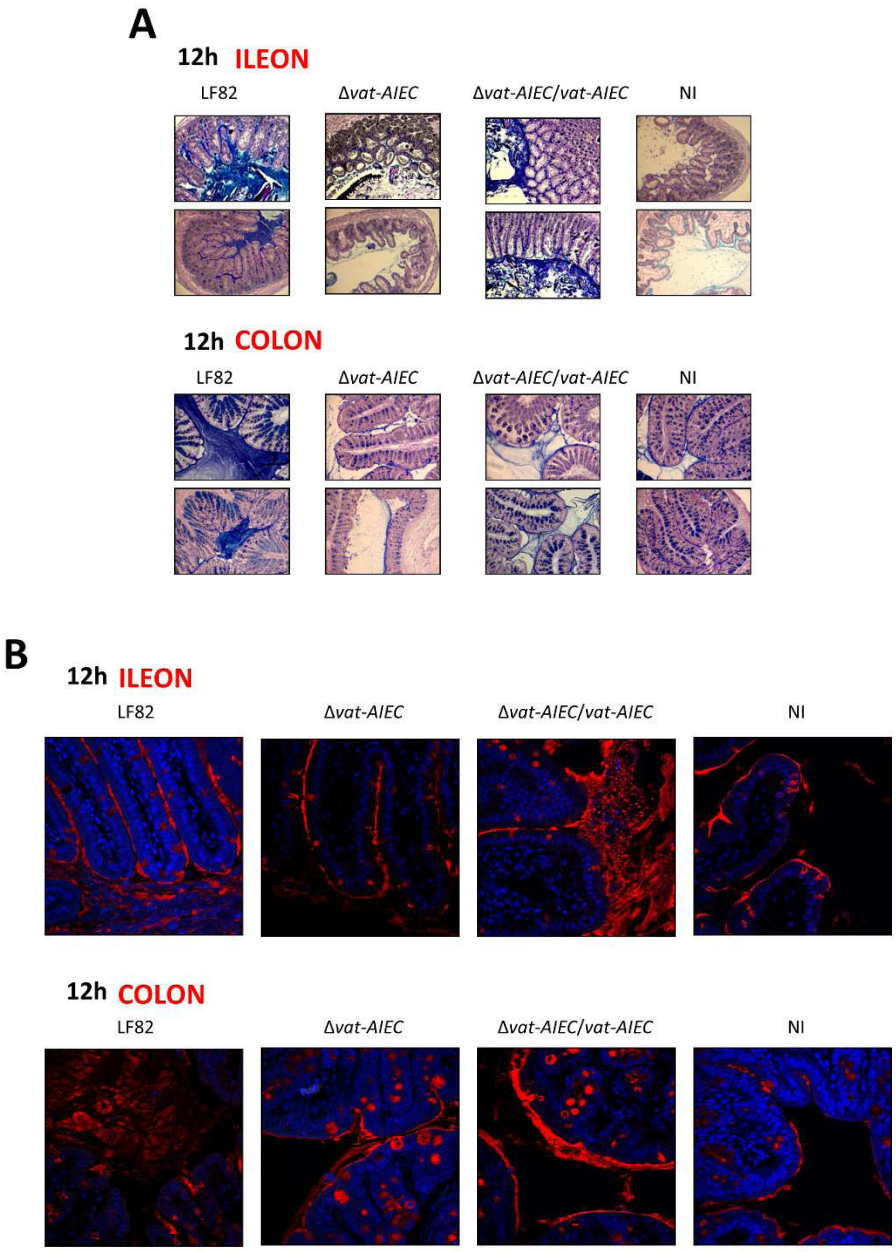



**Figure 36** : Coupes histologiques d'intestin de souris (anses ligaturées) après 12h de contact avec la bactérie. (A) coloration au bleu Alcian, (B) marquage avec un anticorps anti-Muc2 (marquage rouge = anticorps anti-Muc2, marquage bleu = DAPI).

augmentation des ARNm de *fis* qui doivent être confirmés. A l'inverse, les ARNm de *crp* semblent stables ou diminués, en fonction des conditions de culture.

Nous avons également testé deux mutants isogéniques de LF82 : LF82*fis ::Km* (Miquel, Claret, *et al.*, 2010) et LF82Δ*crp,* que nous avons construit, et nous avons étudié l'activité β-galactosidase comme précédemment. En l'absence de ces deux facteurs de transcription, l'activité β-galactosidase diminue, ce qui suggère que ces deux facteurs de transcription sont impliqués dans la régulation, et plus particulièrement dans l'activation du gène *vat-AIEC*.

### ✓ Induction de l'expression des mucines en présence de la mucinase Vat-AIEC

La mucinase Pic, produite par *Shigella flexneri,* induit la production de mucines après 12h d'infection en anses ligaturées de rats (Radhakrishnan *et al.*, 2007; Navarro-Garcia *et al.*, 2010). La mucinase Hbp de *Vibrio cholerae* est également capable d'induire une production de mucus (Bardakchian *et al.*, 2008).

Afin de voir si la mucinase Vat-AIEC était également capable d'induire une hyper-production de mucus, nous avons travaillé *in vivo* sur des anses ligaturées de souris mâles BALB/c au niveau iléal mais également colique, après décontamination digestive à la streptomycine. Ces anses ont été infectées avec les trois souches LF82, LF82Δ*vat-AIEC* (mutant délété du gène *vat-AIEC*) et LF82Δ*vat-AIEC*/*vat-AIEC* (mutant transcomplémenté à l'aide d'un plasmide porteur du gène) et une injection de PBS nous a servi de témoin (anses non infectées : NI). Nous disposions de 6 souris par groupe. Après 12h de contact entre les bactéries et l'intestin, les souris ont été sacrifiées et des échantillons d'intestin (iléon et colon) ont été prélevés. Nous avons ainsi observé une augmentation de la production de mucus dans les anses des souris infectées par LF82 et LF82Δ*vat-AIEC*/*vat-AIEC*, par rapport aux souris infectées par LF82Δ*vat-AIEC* ou non infectées (**Figure 36A**). La principale mucine sécrétée était Muc2 comme visualisée sur les coupes histologiques marquées avec un anticorps anti-Muc2 (**Figure 36B**). Ces données pourraient suggérer que les AIEC en induisant une hyperproduction de mucus favorisent leur implantation au détriment de bactéries non mucinolytiques. Toutefois, il est difficile de conclure sur un modèle d'anses ligaturées utilisé pendant 12 h.

# III.Impact des sels biliaires sur le transcriptome des AIEC

Publication n°2:

Transcriptional response of adherent-invasive *Escherichia coli* to bile salts promoting their implantation in the human gut

Gibold L*, Delmas J*, Faïs T, Sinel C, Cattoir V, Barnich N, Dalmasso G et Bonnet R.

Article prêt à être soumis

(* : Contribution équivalente des auteurs)

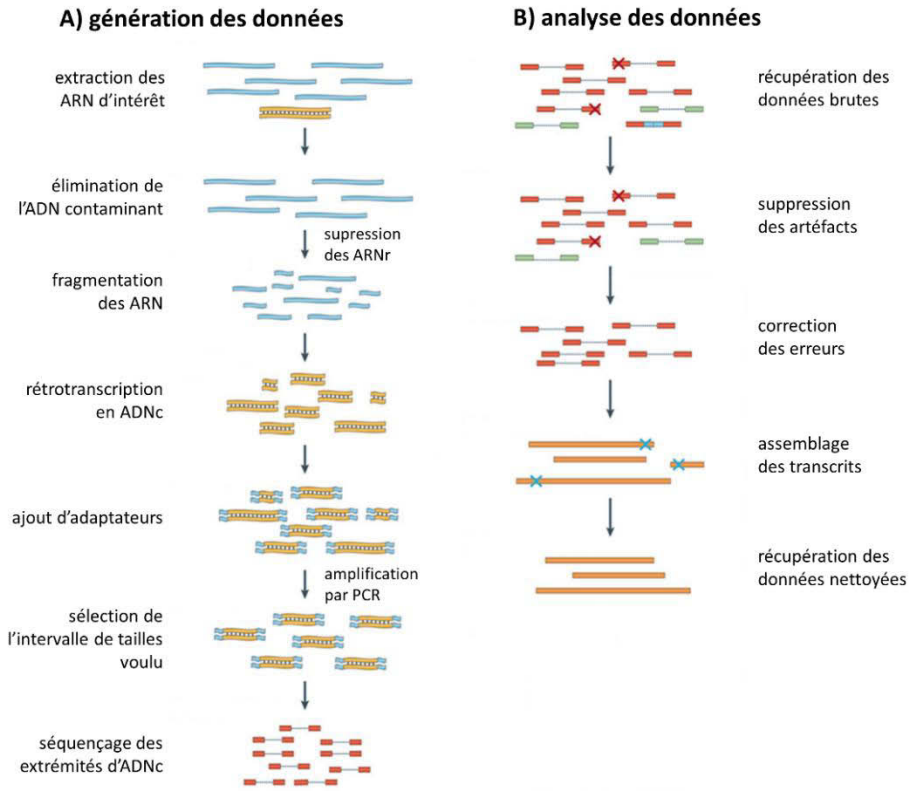



**Figure 37** : Différentes étapes de l'analyse RNA-seq (image modifiée d'après Martin and Wang, 2011)

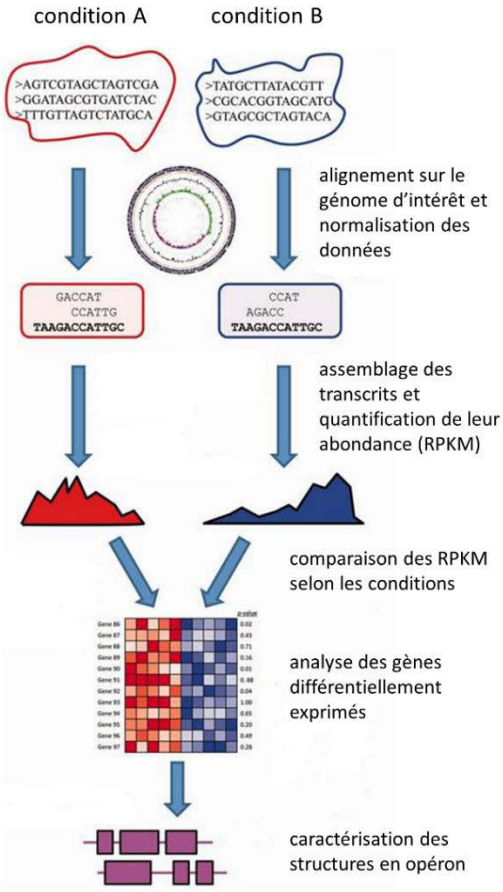



**Figure 38** : Principe de la comparaison des données issues du séquençage des ARN en fonction des conditions de culture (image modifiée d'après McClure *et al.*, 2013).

## 1. Introduction à la publication n°2

Les facteurs micro-environnementaux de l'hôte ont un impact sur la synthèse des facteurs de virulence bactériens. Nous avons déjà démontré que la combinaison d'un pH plutôt basique, la présence de mucines et de sels biliaires, qui correspondent aux conditions retrouvées au niveau de l'iléon, favorise l'expression du gène *vat-AIEC* de la souche LF82.

Dans cette deuxième partie, nous avons étudié l'effet des sels biliaires sur l'expression des gènes de la souche AIEC LF82, dans la mesure où cette bactérie est majoritairement retrouvée au niveau de l'iléon terminal (Darfeuille-Michaud *et al.*, 2004). Pour cela, nous avons utilisé un séquençage à haut débit des ARNm ou RNA-seq en milieu minimum additionné ou non de sels biliaires.

La technique utilisée ici correspond à un séquençage des ARN par une technique de NGS (next generation sequencing) (Voelkerding *et al.*, 2010; Martin and Wang, 2011). Elle a été précédée d'une étape de culture de bactéries en présence ou non de sels biliaires (1%) puis d'une extraction complète des ARN. Après déplétion des ARN ribosomiques bactériens, les ARNm ont été fragmentés et utilisés pour la construction de banques d'ADNc tagués brin spécifique. Les ADNc obtenus ont ensuite été amplifiés par PCR puis séquencés à l'aide d'un Illumina HiSeq et alignés sur le génome connu de la souche LF82 (Miquel, Peyretaillade, *et al.*, 2010) (**Figure 37**).

Une fois les reads nettoyés et alignés sur le génome, une normalisation des données a permis de comparer les différentes conditions de culture (milieu seul et milieu avec sels biliaires). L'abondance de chaque transcrit a été calculée en RPKM (reads per kilobase of transcript per million reads mapped) qui correspond au nombre de fragments du gène d'intérêt rapporté à la longueur du gène d'intérêt et au nombre total de fragments alignés sur le génome. Les différentes conditions ont été ensuite comparées, et les gènes différentiellement exprimés (DEG) en fonction des conditions de culture ont été retrouvés par le calcul du fold change en $Log_2$ et la valeur du *p* associé (**Figure 38**).

L'analyse RNA-seq est différente pour les organismes procaryotes et eucaryotes. Chez les procaryotes, les gènes peuvent être chevauchants, ce qui augmente la difficulté pour déterminer le début et la fin de chaque gène. De plus, certains ARNm procaryotes sont

polycistroniques, c'est-à-dire qu'ils codent pour plusieurs protéines. Plusieurs codons d'initiation et codons stop peuvent se retrouver sur le même ARNm. Enfin, plusieurs promoteurs peuvent commander l'expression d'un gène ou d'un opéron selon différentes conditions (McClure *et al.*, 2013). L'analyse RNA-seq utilisée ici a donc pris en compte ces paramètres. Enfin, dans le but de vérifier la qualité des données, une qRT-PCR a été réalisée en parallèle sur quelques gènes, et devait retrouver les même gènes sur- ou sous-exprimés que l'analyse RNA-seq.

L'analyse RNA-seq a révélé que les sels biliaires ont un effet majeur sur l'expression des gènes de l'AIEC LF82 puisque 40% du génome était dérégulé.

## 2. Publication n°2

# Transcriptional response of adherent-invasive *Escherichia coli* to bile salts promoting their implantation in the human gut

Lucie Gibold*[1,2], Julien Delmas*[1,2#], Tiphanie Faïs[1,2], Clara Sinel[3], Vincent Cattoir[3,4], Nicolas Barnich[2], Guillaume Dalmasso[2], Richard Bonnet[1,2]

Service de Bactériologie, Centre Hospitalo-Universitaire Clermont-Ferrand, France[1]; Microbes, Intestins, Inflammation et Susceptibilité de l'Hôte, INSERM U1071, INRA USC2018, Université d'Auvergne, Clermont-Ferrand, France[2] ; Equipe antibiorésistance EA 4655, Université de Caen Normandie, Caen, France[3] ; Service de bactériologie, Centre Hospitalo-Universitaire Caen, France[4] ;

* these authors contributed equally to this work

#Address correspondence to Julien Delmas

Mailing address:

Laboratoire de Bactériologie, CHU Clermont-Ferrand

58, rue Montalembert

63003 Clermont-Ferrand, France

Phone: (33)473754920

Fax: (33)473754922

jdelmas@chu-clermontferrand.fr.

**KEYWORDS**

*Escherichia coli*, Crohn's Disease, AIEC, bile salts, transcriptome

## Abstract

The etiology of Crohn's disease (CD) involves disorders in host genetic factors and intestinal microbiota. Ileal mucosa of CD patients is often abnormally colonized by adherent-invasive *Escherichia coli* (AIEC). Previous studies have shown an implication of bile salts in the increased expression of genes involved in virulence.

In this study, a global RNA sequencing (RNA-seq) of *E. coli* LF82 has been used to observe the impact of bile salts on the expression of bacterial genes. The results demonstrate the explosive effect of bile salts with a dysregulation of about 40% of the genome, with a global upregulation of genes involved in degradation and downregulation of those implicated in several biosynthesis. We then focused on the impact of bile salts on ethanolamine and 1,2-propanediol utilization by LF82 strain. Our results show that LF82 use ethanolamine as a nitrogen source and propanediol as a carbon source, which can favor their colonization in the gut compared to the other bacteria. We also studied virulence genes expression in the presence of bile salts. They increase the expression of all of the four pathogenicity islands and several other virulence factors. Furthermore, we noticed an increased expression of genes implicated in biofilm formation.

These results improve the understanding of the global regulatory network in the presence of bile salts and thus of AIEC implantation in the human gut of CD patients.

## Introduction

Crohn's disease (CD) is an inflammatory bowel disease (IBD) caused by an abnormal inflammatory response to the gut microbiota with disorders in host genetic factors (Strober et al., 2007; Xavier and Podolsky, 2007; Kaser et al., 2010). Changes in mucosa-associated microbiota have been observed in CD patients, with more prevalent adherent-invasive *Escherichia coli* (AIEC) than in healthy subjects (Martin et al., 2004; Darfeuille-Michaud et al., 2004; Kotlowski et al., 2007; Baumgart et al., 2007; Martinez-Medina and Garcia-Gil, 2014). These bacteria strongly adhere and invade intestinal epithelial cells (IECs), survive within macrophages, move into the deep tissues and activate immune cells inducing inflammatory cytokine secretion (Darfeuille-Michaud et al., 1998; Glasser et al., 2001; Eaves-Pyles et al., 2008).

Several physicochemical conditions found in the gut could have an impact on the present bacteria, especially on their virulence factors (Gupta and Chowdhury, 1997; Harjai *et al.*, 2005; Hamner *et al.*, 2013; Sengupta *et al.*, 2014; Hicks *et al.*, 2015). Bile is a digestive secretion produced by the liver and stored in gallbladder. It includes bile acids for approximately 50%, and also cholesterol, phospholipids and the biliverdin pigment. Bile plays a major role in the emulsification and solubilization of lipids and promotes cholesterol excretion (Hofmann, 1999). The ability to act as detergents also confers potent antimicrobial properties on bile salts. They are able to disrupt cellular integrity with modifications in cell permeability and induce an oxidative stress leading to bacterial DNA damages and SOS response, which can involve cell growth arrest and therefore cell death (Gunn, 2000; Begley *et al.*, 2005; Merritt and Donaldson, 2009).

Thus, bile acids allow the body to protect itself against pathogens. However, bile acids can upregulate virulence factors of some bacterial pathogen species like *Salmonella*, *Shigella* or *Campylobacter*, leading their colonization in the gut (Pope et al., 1995; Doig et al., 1996; Guo

et al., 1997; Gunn, 2000; Crawford et al., 2008; Faherty et al., 2012). Some virulence factors of AIEC such as long polar fimbriae (lpf), GipA and Vat-AIEC, are modulated in the presence of some chemical conditions found in the ileum, especially bile salts (Chassaing et al., 2013; Gibold et al., 2015; Vazeille et al., 2016).

Several clinical studies have highlighted changes in the bile composition of CD patients with an increased level of biliary concentration of bilirubin (Marks et al., 1977; Pereira et al., 2003; Lapidus et al., 2006). Furthermore, CD patients suffer from bile acid malabsorption which lead to bile acids release into the colon (Pereira et al., 2003).

In this study we used the method of deep sequencing-based transcriptome analysis (RNA-seq) to monitor the levels of all transcripts in LF82 grown in the presence and the absence of bile salts, in order to identify bile salts-responsive genes. Analysis of the transcriptome provides key elements of the role of environmental features on AIEC colonization in the human gut. We uncovered several repressions of genes involved in protein syntheses and in contrast upregulation of genes involved in metabolisms, revealing the role of bile salts in bacterial nutrition. Our results also point toward a contribution of the bile salts in controlling the expression of some virulence factors during AIEC implantation in the human gut.

## Material and methods

### Bacterial strain and growth conditions for RNA-seq and qRT-PCR

*E. coli* AIEC strain LF82 was isolated from a chronic ileal lesion of a patient with CD and belongs to *E. coli* serotype O83:H1 (Darfeuille-Michaud et al., 1998) (see Table S1). This strain has a cephalosporinase and can be selected on amoxicillin-containing agar plates. For RNA-seq and qRT-PCR for absolute quantification, LF82 was grown for 24 h at 37°C in M9 minimal medium containing $Na_2HPO_4$ (34 mM), $KH_2PO_4$ (22 mM), $NH_4Cl$ (19 mM), NaCl (8 mM), $CaCl_2$ (0.03 mM) and supplemented with D-glucose (1 g/L) in a 20 mM Tris buffer

with a pH = 7.5, in the presence or absence of 1% of bile salts (cholic acid sodium salt 50%, deoxycholic acid sodium salt 50%, Sigma). Each assay was performed in triplicate.

For qRT- PCR of the *eut* genes, the growth medium used is a modified M9-bile medium (see below) supplemented with glucose (1g/L) and ethanolamine hydrochloride (3 mM). Gene expression has been measured after 8 hours of culture.

**RNA extraction**

Total RNAs were extracted from bacteria using a Direct-zol RNA MiniPrep kit (Zymo research) and treated with a Turbo DNA-free kit (Ambion) to remove any contaminating genomic DNA. DNase-treated RNA samples were purified with RNA clean and concentrator-25 (Zymoresearch). For RNA-seq analysis, they then were quantified using a NanoDrop 1000 spectrophotometer (Thermo Scientific). The integrity of the results (RNA integrity number [RIN]) was assessed using an Agilent 2100 bioanalyzer. A Ribo-zero Magnetic kit for bacteria (Epicentre) was used according to the manufacturer's recommendations to remove the 23S and 16S rRNA from the total RNA samples, and the samples were then purified using RNA clean and concentrator-5 (Zymoresearch). To evaluate the degree of rRNA depletion, the samples were analyzed using an Agilent 2100 bioanalyzer and RNA-seq analysis was made by the genomic ProfileXpert plateform (Claude Bernard University, Lyon, France).

**RNA-seq data analysis**

The RNA-seq data analysis was performed using the R software, Bioconductor packages including edgeR and the SARTools package developed at PF2 - Institut Pasteur (Varet *et al.*, 2016). Briefly, reads were mapped against the genomic sequence of *E. coli* LF82 (GenBank accession n° NC_011993). Reads mapped to several positions and reads mapped to rRNA were removed from further analysis. The number of reads overlapping each gene based on GenBank annotation was recorded. Reads from replicate samples were pooled, and the

number of reads per gene was normalized according to the total number of reads in each library and the gene size. To normalize the expression of genes in different RNA-seq samples, values corresponding to the number of reads per kilobase per million mapped (RPKM) were calculated as follows: (number of reads for the gene x $10^9$) / (total number of reads x size of the gene). A *p*-value adjustment for differentially expressed genes (DEG) was performed to take into account multiple testing and control the false positive rate to a chosen level $\alpha < 0.001$. Products of differentially expressed genes were classified by functional category according to metabolism pathways of the Ecocyc database (http://ecocyc.org) (Karp *et al.*, 2014). RKPM values for each gene were plotted and visualized as a circle using the Circos program (Krzywinski *et al.*, 2009).

**Reverse transcription and qPCR**

For qRT-PCR, extracted RNAs were quantified using a NanoPhotometer P-Class (Implen) and then reverse transcribed with random hexamers (Invitrogen) and NxGen M-MuLV Reverse Transcriptase (Lucigen) as the manufacturer's recommendations. RNA were then amplified by using specific primers to *prpB*, *folX*, *pfkA*, *cfa*, *LF_715, eut* genes or *16S*rRNA (see Table S2). The amplification of a single expected PCR product was confirmed by electrophoresis on a 2% agarose gel. qRT-PCR were performed using an Eppendorf Realplex with the following program: 95°C for 5 min and 40 cycles of 95°C for 15s, 60°C for 20s and 72°C for 40s. For absolute qRT-PCR, transcript levels in each sample were determined by absolute quantification using serial dilutions of PCR products as previously described (Yoder-Himes *et al.*, 2009). Each experiment was performed in triplicate. For relative qRT-PCR, each gene expression was normalized to *16SrRNA* by using the $2^{-\Delta\Delta Ct}$ method.

**Construction of isogenic mutants**

*E. coli* LF82 was transformed with pKOBEG, a plasmid encoding the Red proteins that protect linear DNA from degradation in bacteria. The plasmid was maintained in bacteria at

30°C with 25 mg/l of chloramphenicol and 1 mM of L-arabinose. The Flp recognition target-flanked cassette harboring the kanamycin resistance cassette was generated by PCR from *E. coli* BW25141 with d-eutB-F/d-eutB-R and d-eutE-F/d-eutE-R primers (Table S2) and High Fidelity Platinum Taq polymerase (Invitrogen) according to the manufacturer's instructions. The PCR products were electroporated in previously glycerol-washed *E. coli* LF82. The resulting LF82Δ*eutB* and LF82Δ*eutE* isogenic mutants ($Km^R$) were selected on LB agar containing 50 mg/L kanamycin. The replacement of the *eut* genes by the kanamycin resistance cassette was confirmed by PCR. The kanamycin resistance cassette was then removed from LF82Δ*eutB* by the transient expression of the Flp recombinase from the pCP20 plasmid, creating the LF82*ΔeutB* ($Km^S$) strain.

**Transcomplementation of isogenic mutant LF82Δ*eutB***

The *eutB* gene was amplified by PCR from *E. coli* LF82 genomic DNA by using eutBEcoR1-F and eutBBamH1-R primers (Table S2). The amplified DNA was purified with a NucleoSpin extract kit (Macherey-Nagel), digested with EcoRI and BamHI (New England Biolabs), and ligated to EcoRI-BamHI-digested expression vector pBK-CMV (Agilent Technologies). This construct was electroporated into LF82Δ*eutB* ($Km^S$) electrocompetent strains and selected on Mueller Hinton agar containing 50 mg/L kanamycin. The presence of the *eutB* gene was confirmed by PCR. The construction was checked by double stranded DNA sequencing (GATC biotech, Germany).

**Bacterial growth curves**

The ability of *E. coli* LF82 to grow with bile salts and ethanolamine (EA) was tested on minimal media as previously described (Bertin *et al.*, 2011). A modified M9 minimal medium containing $Na_2HPO_4$ (48 mM), $KH_2PO_4$ (22 mM), and NaCl (8.5 mM) was supplemented with $MgSO_4$ (1 mM), $CaCl_2$ (0.1 mM), Vitamin B12 (cyanocobalamin) (150 nM), vitamin B1 (5 mg/L) and trace metals (0.1 µM $ZnSO_4$, 0.045 µM $FeSO_4$, 0.2 µM $Na_2Se_2O_3$, 0.2 µM

167 $Na_2MoO_4$, 2 µM $MnSO_4$, 0.1 µM $CuSO_4$, 3 µM $CoCl_2$ and 0.1 µM $NiSO_4$) (2 ml/L) (mM9).
168 This media was supplemented with bile salts (1%) when needed (mM9b).

169 Utilization of EA as a nitrogen source was investigated by addition of EA hydrochloride (3
170 mM) and glucose (1g/L) to the mM9b medium (mM9b-EA). To test the capacity of LF82 to
171 use EA as the sole carbon source, the mM9b medium was supplemented with $NH_4Cl$ (20 mM)
172 and EA hydrochloride (3 mM). All chemicals were from Sigma. For each condition, three LB
173 broth cultures were each started from a single colony and grown overnight at 37°C with
174 aeration. Cells were then pelleted by centrifugation, resuspended in medium with EA as the
175 sole nitrogen or carbon source and diluted 50-fold in the corresponding media. Cultures were
176 then incubated with shaking at 37°C and the growth was monitored in three parallel cultures
177 by following optical density (λ=600 nm). For the other growth curves, mM9b-EA was used.

178 To test the utilization of 1,2-propane-diol and/or citrate as carbon source, LF82 was grown
179 overnight at 37°C in a brain-heart infusion medium with aeration. Cells were then diluted
180 250-fold in a M9 minimal medium containing bile salts (1%) supplemented with citrate (6.8
181 mM) or the mM9b medium supplemented with propane-diol (4 mM).

182 **Competition experiments**

183 Precultures of *E. coli* strains (inoculated from a single colony) were incubated in lysogeny
184 broth (LB) with the appropriate antibiotic (amoxicillin 32 mg/L for LF82 and kanamycin 50
185 mg/L for LFΔ*eutB*). The bacterial cultures were diluted 50-fold in LB broth and grown
186 overnight at 37°C without shaking. mM9b medium previously used was supplemented with
187 glucose (1g/L) and EA hydrochloride (3 mM) and inoculated with approximately $5.10^7$
188 bacteria per milliliter of each of the two strains tested in competition assays and incubated at
189 37°C without shaking. At each time point, the co-culture was diluted in PBS and spotted on
190 trypticase soy agar plates with and without antibiotic. Co-cultures of LF82/LFΔ*eutB* were
191 spotted on plates without antibiotic and on plates containing kanamycin (50 mg/L). Co-

192 cultures of LF82 or LFΔ*eutB*/MG1655 were spotted on plates without antibiotic and on plates
193 containing amoxicillin (32 mg/L). The plates were then incubated overnight at 37°C before
194 counting colony-forming unit (CFU). The CFU count of the antibiotic-sensitive strain for each
195 competition assay was calculated by subtracting the number of CFU resistant to amoxicillin or
196 kanamycin from the number of CFU counted on agar plate without antibiotic. Each
197 experiment was replicated three times and a competitive index (CI) was calculated.

198 **Mouse infection**

199 For *in vivo* experiments, we used C57BL/6 mice, which were housed in specific pathogen-free
200 conditions in the animal care facility at the Université d'Auvergne, Clermont-Ferrand, France.
201 Animal protocols were approved by the Committee for Research and Ethical Issues of the
202 CEMEAA (Comité d'Ethique en Matière d'Expérimentation Animale, Auvergne) following
203 international directive 86/609/CEE (n°CE16-09).

204 Ten twelve-week-old mice (body weight ≈26–28g) were pretreated by administering oral
205 amoxicillin (1 g/L), vancomycin (500 mg/L), metronidazole (500 mg/L) and neomycin (1
206 g/L) for four days, and 3% dextran sulfate sodium salt (Sigma) for the last day. 24h after
207 stopping antibiotic treatment, the animals were orally challenged with $10^9$ bacteria (50% LF82
208 - 50% LFΔ*eutB*). Three days after the last bacterial infection, fresh fecal pellets (100–200mg)
209 were collected from the individual mice and resuspended in PBS. After serial dilution, the
210 bacteria were enumerated by plating on Trypticase Soy (TS) agar medium containing
211 amoxicillin to isolate the two bacteria and amoxicillin + kanamycin to isolate LFΔ*eutB*. The
212 plates were then incubated overnight at 37°C before counting colony-forming unit (CFU).
213 The CFU count of the LF82 strain for each mouse was calculated by subtracting the number
214 of CFU resistant to kanamycin from the number of CFU counted on agar plate with only
215 amoxicillin. Three days after infection, the mice were anesthetized with isoflurane and then
216 euthanized by cervical dislocation. Colonization of the two strains was studied by numbering

the mucosa-associated AIEC bacteria by homogenizing 0.5 cm of ileum and 1 cm of colon, beginning at 0.5 cm from the cecal junction, in sterile PBS solution. Samples were plated onto TS agar containing amoxicillin or amoxicillin + kanamycin and incubated overnight at 37°C.

**Motility assay**

Motility was evaluated as previously described (Lane *et al.*, 2005). Briefly, soft-agar plates (1% tryptone, 0.5% NaCl, 0.25% agar) were used with or without 1% of bile salts, which were prepared the day prior to use and left at room temperature overnight. After culture of the bacteria incubated with or without 1% of bile salts at 37°C up to the exponential phase, 2µL of the culture was inoculated into the middle of the soft-agar plates using a sterile inoculating needle (care was taken not to touch the bottom of the plate so as to avoid possible twitching motility). After 16/24h of incubation at 30-37°C, motile strains were identified as those with the largest diameter.

**Biofilm formation assay**

Biofilm assays were used as previously described with some modifications (O'Toole and Kolter, 1998; Hennequin *et al.*, 2012). Briefly, 6 µL of 3 hours of culture in mM9-EA medium supplemented or not with 1% of bile salts was inoculated into 144 µL of the same medium in a 96-well culture-treated polystyrene microtiter plate (Nunc). Wells filled with growth medium alone were included as negative controls. After 4h30 of incubation at 37°C, surface-adherent biofilm formation was measured by staining bound cells for 15 min with a 0.5% (w/v) aqueous solution of crystal violet. After rinsing with distilled water, the bound dye was released from the stained cells using 95% ethanol, and OD at 540 nm was determined.

### Autoaggregation assay

After an overnight culture at 37°C in M9 minimal medium supplemented with glucose (1g/L) and bile salts (1%) when needed, cells were diluted in the same medium and grown at 37°C to the exponential phase. 20µL of each culture was Gram's stained and visualized by microscopy.

### Statistical analysis

Statistical comparisons were performed with a Mann-Whitney nonparametric test. An analysis of variance (ANOVA) was performed for experiments with multiple treatment groups, followed by pairwise comparisons with Bonferroni's multiple comparison tests. A $p$ value less than 0.05 was considered statistically significant. For RNA-seq, $p$ values $< 0.001$ were considered as significant.

## Results

### Transcriptome of LF82 determined by RNA-seq

The AIEC reference strain LF82 was grown for 24h at 37°C in a minimal media with or without bile salts. We used RNA-seq to compare transcriptomes in order to understand the regulatory networks that control gene expression in AIEC strains during their ileal colonization. The growth kinetics of the culture grown with bile salts was more slowly with an optical density not exceeding 0.5. Total RNA was isolated from LF82 at the stationary phase and, following rRNA depletion, the remaining RNA was reverse transcribed into cDNA, fragmented and sequenced using Illumina HiSeq 2500 technology.

A total of 53,900,605 reads from minimal medium and 56,327,631 reads from bile salts added medium were obtained for each cDNA library. Among them, 50,353,549 (93.4%) and 53,246,854 (94.5%) reads mapped to the genome of *Escherichia coli* LF82 which as a total size of 4,773,108 pb. The median numbers of reads per region were 462 (without bile salts)

and 747 (with bile salts), with a coverage of 98.4% and 98.6% of the 4376 coding DNA sequences (CDS) of *E. coli* LF82 represented by at least one single read.

**Transcription profile within the two media and validation by quantitative RT-PCR**

The absolute and relative distributions of reads in the two media in the annotated genes of the LF82 strain are shown in Fig.1A (see also Table S3 in the supplemental material). A dendrogram used to represent all the samples shows replicates grouping and a difference between the biological conditions (Fig.S1A). Box plots and volcano plot are also represented to show the quality of the normalization and the differential expression genes (see Fig.S1). To assess the reliability of RNA-seq in determining the relative abundances of individual transcripts in absence and presence of bile salts, we determined by real-time quantitative reverse transcription-polymerase chain reaction the mRNA levels of three upregulated genes (*prpB*, *eutB* and *LF82_715*), two downregulated genes (*pfkA* and *cfa*) and one unchanged gene (*folX*). The data of these genes gave a $r2$ value of 0.9708, with slightly lower values by qRT-PCR than by RNA-seq (Fig.1B; see Table S4 in the supplemental material).

**Differentially expressed genes during bile salts exposition**

The analysis of RNA-seq data and comparisons showed an explosive effect of bile salts with 1796 differentially expressed genes (DEG) including 1138 genes showing an increase and 658 a decrease in mRNA concentrations, representing about 40% of the genome. Among these DEG, 517 genes (29%) coded for proteins of unknown function.

After a classification into functional categories, we observed 255 DEG on the 722 genes involved in metabolism (35%), including 182 upregulated and 73 downregulated genes. In the presence of bile salts, expression of genes coding proteins involved in degradation, utilization

and assimilation of compounds was significantly induced relative to the other overexpressed genes involved in metabolism (37% *versus* 16.8%; $p<0.001$). More particularly, the expression of genes involved in degradation of amines, alcohols, carboxylates and secondary metabolism were overexpressed for 47% of them (Fig.2A, B and C). Conversely, bile salts induce a repression of genes involved in biosynthesis relative to the other underexpressed genes involved in metabolism (13.2% *versus* 6.1%; $p<0.001$) (Fig.2D and E).

**Induction of ethanolamine metabolism by bile salts**

Ethanolamine, a major component of the cell membranes of bacteria and animals in the form of phosphatidylethanolamine, can be a source of carbon and/or nitrogen for many bacterial species. The genetic information required for the degradation and utilization of EA by *E. coli* is encoded by 17 genes in the *eut* operon. In the presence of bile salts, all of these genes were significantly up-regulated in LF82 (Table 1). EutS, EutM, EutN, EutL and EutK are shell proteins constituting a carboxysome, an organelle required to conserve volatile metabolites and to concentrate the EA catabolic enzymes. Ethanolamine ammonia lyase, encoded by the genes *eutB* and *eutC*, converts EA to ammonia and acetaldehyde. This process requires the cofactor adenosylcobalamin, which is produced from cobalamin by the corrinoid cobalamin adenosyltransferase protein encoded by *eutT*. While ammonia can serve as a cellular source of reduced nitrogen, the acetaldehyde is further converted to acetyl-CoA by an aldehyde oxidoreductase encoded by *eutE*, and enters the carbon pool of the cell. In the case of acetyl-CoA accumulation, acetaldehyde can be converted to alcohol by another oxidoreductase encoded by *eutG*. The ability to use ethanolamine could contribute to the pathogenesis of LF82 in an indirect pathway by providing source of carbon and/or nitrogen that promotes successful colonization of the intestine. To evaluate the ability of AIEC to utilize ethanolamine as a nitrogen source, EA was added along with 0.1% (w/v) glucose to minimal

312 medium. To test the ability of the bacterium to use EA as a carbon source, EA and $NH_4Cl$
313 were added to minimal medium. LF82 failed to utilize EA as a carbon source but was able to
314 use EA as the sole nitrogen source. The mutant defective for EutB (LF82Δ*eutB)* was unable to
315 grow in M9 medium with EA as the sole nitrogen source, while complementation *in trans*
316 with the *eutB* gene (LF82Δ*eutB/*p*eutB*) restored a bacterial growth. In contrast, the mutant
317 LF82Δ*eutE* showed growth phenotype similar to those of the wild-type strain. These results
318 clearly demonstrated that the AIEC strain LF82 can use EA as a nitrogen source and that a
319 functional ethanolamine ammonia lyase including EutB was required for the release of
320 ammonia (Fig.3).

321 The *eut* operon is found within *E. coli* species. The growth of the AIEC strain LF82 and the
322 commensal *E. coli* strain MG1655 in mM9-EA supplemented with bile salts was evaluated in
323 comparison with the maximal of growth of each strain in LB medium. In these conditions, the
324 growth of MG1655 was very low compared to those of the LF82 strain (Fig.4A). To argue
325 this difference of growth, the ratio of mRNA level of several genes of *eut* operon was
326 measured in the both strains in the early phase of growth. *eut* genes were found up-regulated
327 in the AIEC strain LF82 relative to the MG1655 strain (Fig.4B). Particularly, a very high
328 induction of *eutT, eutH, eutS and eutQ* genes was observed. These data suggest that AIEC
329 strain LF82 degrade EA more efficiently than commensal *E. coli.* The utilization of EA could
330 provide a nutritional advantage for AIEC. To test this hypothesis, competitive assays were
331 performed between LF82, its isogenic *eutB* mutant and the commensal strain MG1655. LF82
332 was able to grow in the presence of EA much more effectively than its isogenic mutant or the
333 MG1655 strain (Fig.4C). Interestingly, competitive index values close to 1 were obtained
334 when LF82Δ*eutB* and MG1655 strains were co-incubated, indicating that EA present in mM9-
335 EA supplemented in bile salts was poorly used by commensal *E. coli.* AIEC have been
336 reported to extensively colonize the ileum of CD patients. To analyze the capacity of the

facultative anaerobes bacteria of ileum to catabolize EA, 6 strains (*Lactobacillus plantarum, Lactococcus lactis, Enterococcus faecium* and three strains of *Enterococcus faecalis*) were inoculated in mM9-EA with bile salts. No growth was observed (data not shown).

We next evaluated the capacity of AIEC to use EA as a nitrogen source to colonize the ileum, which contain high bile concentration. Mice were orally challenged with an equal mixture of the LF82 strain and the LF82Δ*eutB* isogenic mutant as previously described in the study of EA utilization by *Salmonella* (Thiennimitr *et al.*, 2011). Quantification of bacteria in stool samples, ileal and colon tissues revealed approximatively one-log reduction in LF82Δ*eutB*-AIEC mutant compared with wild-type LF82 (Fig.5). Analysis of the competitive index of wild-type LF82 bacteria grown compared with isogenic mutant grown revealed a median competitive index of 19, 15 and 11 for stools, ileum and colon, respectively.

Collectively, these results indicate that AIEC LF82 gains a competitive advantage by using ethanolamine as a nitrogen source in the ileum.

**Bile salts active methylcitrate pathway**

We searched carbon sources that may be relevant to the intestine colonization by AIEC. The analysis of RNA-seq data has revealed interactions between the pathways required for the degradation of 1,2-propanediol, propionate and citrate. These data indicated that 1,2-PD utilization genes (*pdu*) are induced in the presence of bile salts (Table 2) suggesting the formation of propionyl-CoA. The *prpE* gene encoding the enzyme that converts propionate to propionyl-CoA was also significantly up-regulated (3.3 fold). The propionyl-CoA is the precursor of 2-methylcitrate, the coactivator needed for transcriptional activation of the *prpBCDE* operon by the PrpR activator protein. We found that the expression of methylcitrate pathway genes was strongly up-regulated (3.2 to 4.5 fold) except for the *acnB* gene which has

361 a constitutive expression (Table 3). The methylcitrate cycle converts propionyl-CoA to
362 pyruvate and succinate. Thus, bile salts may favor formation of pyruvate which is a key
363 metabolite located at the intersection of several major metabolic pathways of living cells, such
364 as glycolysis, citric acid cycle and gluconeogenesis, and can be converted to fatty acid,
365 alanine or ethanol after oxidative decarboxylation to acetyl-coenzyme A (Mazumdar *et al.*,
366 2010; Clomburg and Gonzalez, 2010). The genes encoding citrate lyase (*citD, citE* and *citF*)
367 were also up-regulated (Table 3). The citrate lyase converts citrate to oxaloacetate. This latter
368 might be condensed with propionyl-CoA by the PrpC enzyme to go into the 2-methylcitric
369 acid cycle since the *prpC* gene has a 3.6-fold increased mRNA level.
370 To further establish the relevance of the catabolic genes identified by RNA-seq, we cultured
371 AIEC LF82 with 1,2-PD or citrate as carbon source. LF82 was able to grow in the presence of
372 1,2-PD as sole carbon source, but not citrate (Fig. 6). Thus, the methylcitrate pathway could
373 be used by AIEC strains to provide a carbon source.
374
375 **Bile salts increase expression of genes encoding virulence factors**
376 In the LF82 strain, four pathogenicity islands (PAI) and several virulence genes have been
377 identified (Miquel *et al.*, 2010). On these 148 genes found here, 80 were differentially
378 expressed, with an upregulation of 69 genes (Table S3B). Over half the PAIs genes showed
379 significant increase of their expression in the presence of bile salts (Table 4). Among them,
380 many genes of the PAI I and III, which encode T6SS, were significantly up-regulated. These
381 secretion systems are known to inject effectors into eukaryotic cells and bacteria thereby
382 interrupting their own cell functions (Schwarz *et al.*, 2010). The PAI II genes encoding the
383 yersiniabactin siderophore were also 2 to 3.5-fold up-regulated. In fact, the ability to acquire
384 and use iron appears to be important to the AIEC pathotype (Dogan *et al.*, 2014).
385 Furthermore, all the genes the PAI IV were increased in the presence of bile salts. PAI IV

386 contains gene clusters encoding group 2 capsule, which is a well-established virulence factor,
387 often acting by protecting the cell from opsonophagocytosis and complement-mediated killing
388 (Whitfield, 2006).
389 Interestingly, we also found that the *ibeA* gene, which encoding an invasin associated with
390 AIEC invasion process and macrophage survival, was 3.3-fold up-regulated in the presence of
391 bile salts (Cieza *et al.*, 2015).
392
393 **Bile increase genes involved in biofilm formation**
394 The *E. coli* capsule has been shown to play an indirect role in biofilms by shielding of
395 bacterial surface adhesin (Schembri *et al.*, 2004) and biofilm formation is a phenotypic feature
396 of AIEC (Martinez-Medina *et al.*, 2009). Thus, we analyzed the genes expression involved in
397 biofilm formation. An important feature in biofilm development of many bacterial pathogens
398 is a mucoid-like substance known as exopolysaccharide (EPS) or extracellular matrix (Danese
399 *et al.*, 2000). The EPS matrices are composed of polysaccharides such as colanic acid or
400 cellulose (Beloin *et al.*, 2008). The *wza* and *wca g*enes involved in colanic acid synthesis
401 (Stevenson *et al.*, 1996) were almost all significantly 2.1 to 3.3-fold up-regulated under bile
402 stress conditions. The genes encoding the phosphomannomutase CpsG, the GDP-mannose
403 4,6-deshydratase Gmd and the GDP-L-fucose synthetase Fcl, which are important for the
404 formation of the fucose precursor for colanic acid production, were upregulated under bile
405 stress conditions (2.6, 2.3 and 2-fold respectively). The cellulose biosynthesis genes (*yhj*)
406 were in part also up-regulated. Furthermore, evidence was supporting the participation of curli
407 fimbriae in the biofilm formation (Prigent-Combaret *et al.*, 2000). Curli fibers are adhesive
408 surface fibers that bond several host extracellular matrix and contact phase proteins which
409 have been implicated in the internalization of bacteria by eukaryotic cells (Gophna *et al.*,
410 2001). A high expression of curli fibers may cause bacterial autoaggregation. In this study, the

411 *csgA* and *csgB* genes encoding the structural components of curli fimbriae were significantly
412 up-regulated. In parallel, we observed increased mRNA for *auf yeh* and *yfc* genes which are
413 likely involved in adhesion. Besides, the inhibition of bacterial motility appears to be essential
414 in biofilm formation (Guttenplan and Kearns, 2013). We noticed here a decrease in the
415 expression of the *flhDC* operon in presence of bile salts, which was bound to an inhibition of
416 bacterial mobility (Fig.S2). On the other way, indole has been described as a potential
417 extracellular signal that has an effect on biofilm development depending on the conditions
418 used (Di Martino *et al.*, 2003). Indole is produced by the degradation of tryptophane, a
419 reaction performed by tryptophanase encoded by the *tnaA* gene (Newton and Snell, 1964).
420 The *tnaA* gene was 2.45-fold up-regulated at the mRNA level upon bile salts exposure. The
421 *tnaB* gene encoding the tryptophan-specific transporter TnaB and *acrE* and *acrF* genes
422 encoding the AcrEF multidrug efflux pump that exports indole outside the cells were 1 to 2.3-
423 fold up-regulated.
424 Considering all of these data, we tested the ability of LF82 to form biofilms in minimal
425 medium with or without bile salts. As expected, AIEC LF82 formed thicker biofilms in the
426 presence of bile salts (Fig.7A). Furthermore, the presence of bile salts induced an
427 autoaggregation of bacteria (Fig.7B). Thus, bile salts play important roles in colonization and
428 biofilm formation of AIEC strains in the gut.
429
430 **Bile decrease expression of genes involved in resistance to bile**
431 It is well known that bile salts are protecting us by having harmful effects to bacteria. Bile
432 salts activate *sulA* gene which induce SOS response in *Escherichia coli* and inhibit bacterial
433 growth (Jones and Holland, 1985). However, *sulA* seems to be significantly downregulated in
434 LF82 strain grown with bile salts. Enteric pathogens have developed mechanisms in order to
435 survive and grow in the gut despite the presence of bile salts. Several bacteria have efflux

436 pumps to evict bile salts from the cells. The most frequent efflux pumps are encoded by
437 *acrAB* and *emrAB* genes, which are curiously decreased here. Furthermore, *rob* gene which
438 activate *acrAB* (Rosenberg *et al.*, 2003) and *tol* genes, which are also involved in bile
439 resistance with the AcrAB efflux pump system, are not dysregulated. Another efflux pump
440 system implicated in bile resistance is encoded by *mdtABCD* gene cluster (Baranova and
441 Nikaido, 2002). In the LF82 strain, this gene cluster is slightly increased with a significant
442 raise of 2 of the 4 genes. The lipopolysaccharide (LPS) of bacteria is known to be another
443 way to resist to bile salts (Gunn, 2000; Begley *et al.*, 2005). Within the genes implicated in
444 LPS biosynthesis, we have observed in LF82 a diminution in the expression of genes involved
445 in antigen-O biosynthesis and in of *rfa* genes which mediate core region of LPS synthesis
446 (Schnaitman *et al.*, 1991). Another way to resist to bile salts is to raise the phospholipid
447 biosynthesis to help maintain integrity of the cell envelope (Hamon *et al.*, 2011). In fact, the
448 cytoplasmic membrane of *cfa*-mutants *Escherichia coli* strains displayed increased overall
449 permeability (Shabala and Ross, 2008). However, the expression of *cfa* gene is also reduced
450 in LF82 in the presence of bile salts. Finally, *Escherichia coli* have a transcriptional regulator
451 called RpoS which favors bacterial resistance to harmful environment, especially bile salts
452 (Chen *et al.*, 2010). The RpoS-encoded gene is also significantly decreased in LF82.
453

454 ## Discussion

455 Ileal mucosa of CD patients is often abnormally colonized by AIEC. When they infect the
456 intestine, they are exposed to bile salts which are present at high molar concentrations (0.1 to
457 1.3 %) (Northfield and McColl, 1973). In the present study, we used high throughput deep
458 sequencing of mRNA transcripts (termed RNAseq) to study the transcriptome of AIEC strain
459 LF82 in the presence or the absence of bile salts. Specific studies in AIEC have shown that
460 bile salts can enhance expression of virulence genes (Chassaing *et al.*, 2013; Gibold *et al.*,

2015). However, this is the first report to provide evidence that bile salts have very important repercussions on the expression of virulence factors and on the metabolism of the AIEC strain LF82.

Many genes of all PAIs of the LF82 strain were up-regulated in the presence of bile salts, including genes involved in the synthesis of two T6SS, the yersiniabactin siderophore and the group 2 capsule. Bile salts also increased the transcription of another virulence gene encoding the IbeA invasine, which is involved in AIEC invasion process and macrophage survival. This virulence factor is predominant among AIEC (Conte *et al.*, 2014). These results were not necessarily expected since exposure to bile salts is not associated with a global increase of virulence factors in enteropathogen bacteria. For example, bile treatment does not increase Shiga-like toxin and *LEE* mRNA levels in *E. coli* O157:H7 (Hamner *et al.*, 2013). Nevertheless, it is well established that pathogenic bacteria such as *Listeria monocytogenes*, *Pseudomonas aeruginosa*, *Klebsiella pneumoniae* and *Vibrio cholerae* can form biofilm when they are exposed to bile (Hung *et al.*, 2006; Begley *et al.*, 2009; Hennequin and Forestier, 2009; Reen *et al.*, 2012). In the current study, we observed the increase of gene transcription involved in biosynthesis of fibers, such as curli that display direct interactions with the substratum and form interbacterial bundles, and EPS such as cellulose and colanic acid that contribute to the biofilm architecture and allow the formation of voluminous biofilms (Prigent-Combaret *et al.*, 2000; Serra *et al.*, 2013). The bacterial aggregation and the increase of the biofilm thickness in the presence of bile salts were confirmed by phenotypic test. AIEC strains are known to have stronger biofilm formation abilities than other *E. coli* strains isolated from the intestinal mucosa (Chassaing and Darfeuille-Michaud, 2013). This present work suggests that the enhanced biofilm formation in response to bile may be associated with the colonization and persistence of AIEC strains in the ileum.

Vat-AIEC mucinase and Lpf are virulence factors associated with AIEC strains. Their mRNA levels remained unchanged in the presence or absence of bile. These results differ from previous work of our laboratory which show that bile salts promoted activation of Vat-AIEC and Lpf transcription when bacteria are exposed to a concentration of 2% of bile salts (Chassaing *et al.*, 2013; Gibold *et al.*, 2015). This discrepancy may be due to the setting of exposure time of the bacteria or the concentration of bile acids used. In the present study, we preferred to use a lower bile salts concentration. Furthermore, in our previous work, the *vat-AIEC* transcription was studied at mid-log phase. Supplemental tests show that *vat-AIEC* transcription is not increased in the presence of bile salts at 24h, suggesting a feedback control mechanism (Fig.S3). Hence, it could conceivably be hypothesized that bile salts induce differential gene activation time dependent.

Another important finding was that 1,2-propanediol and ethanolamine constitute a carbon and a nitrogen source, respectively, for the AIEC strain LF82. Indeed, the both metabolic pathways, 1,2-propandiol and ethanolamine degradation, which are coenzyme B12-dependent and involve bacterial microcompartments, were stimulated under bile stress conditions as revealed by RNA-seq data analysis and growth curves in minimal medium.

The 1,2-PD degradation supports growth by providing ATP, an electron sink, and a three-carbon intermediate (propionyl-coenzyme A) that is degraded to pyruvate and succinate *via* the methylcitrate pathway which genes were up-regulated in the presence of bile salts. The *pdu* operon is significantly correlated with the AIEC pathotype (Dogan *et al.*, 2014). The presence of *pduC*, which is significantly up-regulated in the current study, is correlated with increased cellular invasion and persistence *in vitro* of AIEC strains. The *pdu* operon is prevalent among enteropathogenic bacteria, such as *Salmonella*, *Yersinia*, *Listeria*, and *Clostridium perfringens* (Lawhon *et al.*, 2003; Korbel *et al.*, 2005; Thomson *et al.*, 2006). 1,2-

PD is a major product of the anaerobic degradation of rhamnose and fucose, which are common sugars in plant cell walls, bacterial exopolysaccharides, and the glycoconjugates of intestinal epithelial cells (Obradors *et al.*, 1988). Hence, 1,2-PD degradation might give to AIEC a competitive advantage relative to other bacteria, including other *E. coli* that cannot use fucose or propanediol.

Ethanolamine is abundant in the intestinal tract due to the turnover and exfoliation of enterocytes and bacterial cells (Garsin, 2010) and the substrate concentration can be higher in intestinal inflammation (Thiennimitr *et al.*, 2011). In this study, we demonstrated for the first time that EA can be used as a source of nitrogen by AIEC strain LF82. EA utilization in the presence of bile salts gave a competitive growth advantage over *eutB* mutant strain, and it was shown to catabolize EA much more avidly than commensal *E. coli* MG1655 strain and other facultative anaerobes bacteria of ileum. There are similarities between these results and those described by Bertin *et al.* that shows that the O157:H7 EHEC strain EDL933 use EA as a nitrogen source in the bovine intestinal content and it is poorly metabolized by endogenous microbiota of the bovine small intestine, including commensal *E. coli* (Bertin *et al.*, 2011). The ability of AIEC to utilize EA as a source of nitrogen may provide AIEC an important competitive advantage in the gut. Furthermore, genes involved in the citrate degradation were also up regulated. All these degradation pathways (citrate, methylcitrate and ethanolamine) can lead to the formation of acetate and pyruvate that might participate to the formation of the side chains of colanic acid (Fig.S4).

We also analyzed the genes known to be involved in bacterial resistance to bile (Gunn, 2000; Begley *et al.*, 2005). Surprisingly, all of these genes were downregulated. This inconsistency might be due to the kinetic expression of these genes. Indeed, their expression may be increased rapidly after bile exposure then decreased after a control loop in the

stationary growth phase. Another explanation is that bile salts concentration used here is not sufficient for activating of these genes. Indeed, a concentration of 3-5% of bile salts seems to be necessary for genes activation (Hernández *et al.*, 2012).

Our findings in addition to previous study on AIEC and bile salts revealed that characteristics associated with AIEC pathovar (lpf, Vat-AIEC, IbeA invasine, biofilm formation, yersiniabactin, capsular material and *pdu* operon) can be up-regulated by bile salts suggesting that AIEC strains may use bile as an environmental signal to colonize the ileum (Chassaing *et al.*, 2013; Conte *et al.*, 2014; Gibold *et al.*, 2015; Zhang *et al.*, 2015).

In summary, the findings presented here indicate that AIEC LF82 responds to bile through wide-ranging changes in gene transcription. Many up-regulated genes encode virulence factors involved in intestinal colonization. Our study reveals that 1,2-propanediol and ethanolamine utilization bestows a competitive advantage on AIEC strains that are metabolically capable of their degradation. To conclude, the current study provides insight into how AIEC adapt to niches, such as ileum.

## Acknowledgements

We thank Sylvine Batista, Anne-Sophie Marinelli and Virginie Bonnin for technical assistance. We also thank GM. Weinstock (University of Texas) for kindly providing the MG1655 strain.

**Funding**: This work was supported by the Ministère de l'Enseignement supérieur et de la Recherche, the Institut national de la santé et de la recherche médicale (UMR Inserm U1071), the Institut National de la Recherche Agronomique (USC-2018), and CHU de Clermont-Ferrand.

561

562

563 ## References


564 Baranova, N., and Nikaido, H. (2002) The BaeSR Two-Component Regulatory System
565 Activates Transcription of the yegMNOB (mdtABCD) Transporter Gene Cluster in
566 *Escherichia coli* and Increases Its Resistance to Novobiocin and Deoxycholate. *J*
567 *Bacteriol* **184**: 4168–4176.

568 Barnich, N., and Darfeuille-Michaud, A. (2007) Adherent-invasive *Escherichia coli* and
569 Crohn's disease. *Curr Opin Gastroenterol* **23**: 16–20.

570 Baumgart, M., Dogan, B., Rishniw, M., Weitzman, G., Bosworth, B., Yantiss, R., *et al.*
571 (2007) Culture independent analysis of ileal mucosa reveals a selective increase in
572 invasive *Escherichia coli* of novel phylogeny relative to depletion of Clostridiales in
573 Crohn's disease involving the ileum. *ISME J* **1**: 403–418.

574 Begley, M., Gahan, C.G.M., and Hill, C. (2005) The interaction between bacteria and bile.
575 *FEMS Microbiol Rev* **29**: 625–651.

576 Begley, M., Kerr, C., and Hill, C. (2009) Exposure to bile influences biofilm formation by
577 *Listeria monocytogenes*. *Gut Pathog* **1**: 11.

578 Beloin, C., Roux, A., and Ghigo, J.-M. (2008) *Escherichia coli* biofilms. *Curr Top Microbiol*
579 *Immunol* **322**: 249–289.

580 Bertin, Y., Girardeau, J.P., Chaucheyras-Durand, F., Lyan, B., Pujos-Guillot, E., Harel, J., and
581 Martin, C. (2011) Enterohaemorrhagic *Escherichia coli* gains a competitive advantage by
582 using ethanolamine as a nitrogen source in the bovine intestinal content. *Environ*
583 *Microbiol* **13**: 365–377.

584 Chan, C.H.F., and Stanners, C.P. (2004) Novel mouse model for carcinoembryonic antigen-
585 based therapy. *Mol Ther J Am Soc Gene Ther* **9**: 775–785.

586 Chassaing, B., and Darfeuille-Michaud, A. (2013) The σE pathway is involved in biofilm
587 formation by Crohn's disease-associated adherent-invasive *Escherichia coli*. *J Bacteriol*
588 **195**: 76–84.

589 Chassaing, B., Etienne-Mesmin, L., Bonnet, R., and Darfeuille-Michaud, A. (2013) Bile salts
590 induce long polar fimbriae expression favouring Crohn's disease-associated adherent-
591 invasive *Escherichia coli* interaction with Peyer's patches. *Environ Microbiol* **15**: 355–
592 371.

593 Chassaing, B., Garénaux, E., Carriere, J., Rolhion, N., Guérardel, Y., Barnich, N., *et al.*
594 (2015) Analysis of the σE Regulon in Crohn's Disease-Associated *Escherichia coli*
595 Revealed Involvement of the waaWVL Operon in Biofilm Formation. *J Bacteriol* **197**:
596 1451–1465.

597 Chaveroche, M.K., Ghigo, J.M., and Enfert, C. d' (2000) A rapid method for efficient gene
598 replacement in the filamentous fungus *Aspergillus nidulans*. *Nucleic Acids Res* **28**: E97.

599 Chen, W.-L., Oliver, J.D., and Wong, H. (2010) Adaptation of *Vibrio vulnificus* and an rpoS
600 mutant to bile salts. *Int J Food Microbiol* **140**: 232–238.

601 Cieza, R.J., Hu, J., Ross, B.N., Sbrana, E., and Torres, A.G. (2015) The IbeA Invasin of
602 Adherent-Invasive *Escherichia coli* Mediates Interaction with Intestinal Epithelia and
603 Macrophages. *Infect Immun* **83**: 1904–1918.

604 Clomburg, J.M., and Gonzalez, R. (2010) Biofuel production in *Escherichia coli*. *Appl*
605 *Microbiol Biotechnol* **86**: 419–434.

606 Conte, M.P., Longhi, C., Marazzato, M., Conte, A.L., Aleandri, M., Lepanto, M.S., *et al.*
607 (2014) Adherent-invasive *Escherichia coli* (AIEC) in pediatric Crohn's disease patients:
608 phenotypic and genetic pathogenic features. *BMC Res Notes* **7**
609 http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4210564/. Accessed July 3, 2016.

610 Crawford, R.W., Gibson, D.L., Kay, W.W., and Gunn, J.S. (2008) Identification of a Bile-
611 Induced Exopolysaccharide Required for Salmonella Biofilm Formation on Gallstone
612 Surfaces. *Infect Immun* **76**: 5341–5349.

613 Danese, P.N., Pratt, L.A., and Kolter, R. (2000) Exopolysaccharide Production Is Required
614 for Development of *Escherichia coli* K-12 Biofilm Architecture. *J Bacteriol* **182**: 3593–
615 3596.

616 Darfeuille-Michaud, A., Boudeau, J., Bulois, P., Neut, C., Glasser, A.-L., Barnich, N., *et al.*
617 (2004) High prevalence of adherent-invasive *Escherichia coli* associated with ileal
618 mucosa in Crohn's disease. *Gastroenterology* **127**: 412–421.

619 Darfeuille-Michaud, A., Neut, C., Barnich, N., Lederman, E., Di Martino, P., Desreumaux, P.,
620 *et al.* (1998) Presence of adherent *Escherichia coli* strains in ileal mucosa of patients with
621 Crohn's disease. *Gastroenterology* **115**: 1405–1413.

622 Datsenko, K.A., and Wanner, B.L. (2000) One-step inactivation of chromosomal genes in
623 *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* **97**: 6640–6645.

624 Di Martino, P., Fursy, R., Bret, L., Sundararaju, B., and Phillips, R.S. (2003) Indole can act as
625 an extracellular signal to regulate biofilm formation of *Escherichia coli* and other indole-
626 producing bacteria. *Can J Microbiol* **49**: 443–449.

627 Dogan, B., Suzuki, H., Herlekar, D., Sartor, R.B., Campbell, B.J., Roberts, C.L., *et al.* (2014)
628 Inflammation-associated adherent-invasive *Escherichia coli* are enriched in pathways for
629 use of propanediol and iron and M-cell translocation. *Inflamm Bowel Dis* **20**: 1919–1932.

630 Doig, P., Yao, R., Burr, D.H., Guerry, P., and Trust, T.J. (1996) An environmentally regulated
631 pilus-like appendage involved in *Campylobacter* pathogenesis. *Mol Microbiol* **20**: 885–
632 894.

633 Eaves-Pyles, T., Allen, C.A., Taormina, J., Swidsinski, A., Tutt, C.B., Jezek, G.E., *et al.*
634 (2008) *Escherichia coli* isolated from a Crohn's disease patient adheres, invades, and
635 induces inflammatory responses in polarized intestinal epithelial cells. *Int J Med*
636 *Microbiol IJMM* **298**: 397–409.

637 Faherty, C.S., Redman, J.C., Rasko, D.A., Barry, E.M., and Nataro, J.P. (2012) Shigella
638 flexneri effectors OspE1 and OspE2 mediate induced adherence to the colonic epithelium
639 following bile salts exposure. *Mol Microbiol* **85**: 107–121.

640 Garsin, D.A. (2010) Ethanolamine Utilization in Bacterial Pathogens: Roles and Regulation.
641 *Nat Rev Microbiol* **8**: 290–295.

642 Gibold, L., Garenaux, E., Dalmasso, G., Gallucci, C., Cia, D., Mottet-Auselo, B., *et al.* (2015)
643 The Vat-AIEC protease promotes crossing of the intestinal mucus layer by Crohn's
644 disease-associated *Escherichia coli*. *Cell Microbiol* ;18(5):617-31.

645 Glasser, A.-L., Boudeau, J., Barnich, N., Perruchot, M.-H., Colombel, J.-F., and Darfeuille-
646 Michaud, A. (2001) Adherent Invasive *Escherichia coli* Strains from Patients with
647 Crohn's Disease Survive and Replicate within Macrophages without Inducing Host Cell
648 Death. *Infect Immun* **69**: 5529–5537.

649 Gophna, U., Barlev, M., Seijffers, R., Oelschlager, T.A., Hacker, J., and Ron, E.Z. (2001)
650 Curli Fibers Mediate Internalization of *Escherichia coli* by Eukaryotic Cells. *Infect*
651 *Immun* **69**: 2659–2665.

652 Gunn, J.S. (2000) Mechanisms of bacterial resistance and response to bile. *Microbes Infect* **2**:
653 907–913.

654 Guo, L., Lim, K.B., Gunn, J.S., Bainbridge, B., Darveau, R.P., Hackett, M., and Miller, S.I.
655 (1997) Regulation of Lipid A Modifications by *Salmonella typhimurium* Virulence Genes
656 phoP-phoQ. *Science* **276**: 250–253.

657 Gupta, S., and Chowdhury, R. (1997) Bile affects production of virulence factors and motility
658 of *Vibrio cholerae*. *Infect Immun* **65**: 1131–1134.

659 Guttenplan, S.B., and Kearns, D.B. (2013) Regulation of flagellar motility during biofilm
660 formation. *FEMS Microbiol Rev* **37**: 849–871.

661 Hamner, S., McInnerney, K., Williamson, K., Franklin, M.J., and Ford, T.E. (2013) Bile Salts
662 Affect Expression of *Escherichia coli* O157:H7 Genes for Virulence and Iron
663 Acquisition, and Promote Growth under Iron Limiting Conditions. *PLoS ONE* **8**: e74647.

664 Hamon, E., Horvatovich, P., Izquierdo, E., Bringel, F., Marchioni, E., Aoudé-Werner, D., and
665 Ennahar, S. (2011) Comparative proteomic analysis of *Lactobacillus plantarum* for the
666 identification of key proteins in bile tolerance. *BMC Microbiol* **11**: 63.

667 Harjai, K., Khandwaha, R.K., Mittal, R., Yadav, V., Gupta, V., and Sharma, S. (2005) Effect
668 of pH on production of virulence factors by biofilm cells of *Pseudomonas aeruginosa.*
669 *Folia Microbiol (Praha)* **50**: 99–102.

670 Hennequin, C., Aumeran, C., Robin, F., Traore, O., and Forestier, C. (2012) Antibiotic
671 resistance and plasmid transfer capacity in biofilm formed with a CTX-M-15-producing
672 *Klebsiella pneumoniae* isolate. *J Antimicrob Chemother* **67**: 2123–2130.

673 Hennequin, C., and Forestier, C. (2009) oxyR, a LysR-Type Regulator Involved in *Klebsiella*
674 *pneumoniae* Mucosal and Abiotic Colonization. *Infect Immun* **77**: 5449–5457.

675 Hernández, S.B., Cota, I., Ducret, A., Aussel, L., and Casadesús, J. (2012) Adaptation and
676 Preadaptation of *Salmonella enterica* to Bile. *PLoS Genet* **8**
677 http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3261920/. Accessed July 3, 2016.

678 Hicks, K.G., Delbecq, S.P., Sancho-Vaello, E., Blanc, M.-P., Dove, K.K., Prost, L.R., *et al.*
679 (2015) Acidic pH and divalent cation sensing by PhoQ are dispensable for systemic
680 salmonellae virulence. *eLife* **4** http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4473727/.
681 Accessed May 30, 2016.

682 Hofmann, A.F. (1999) Bile Acids: The Good, the Bad, and the Ugly. *News Physiol Sci Int J*
683 *Physiol Prod Jointly Int Union Physiol Sci Am Physiol Soc* **14**: 24–29.

684 Hung, D.T., Zhu, J., Sturtevant, D., and Mekalanos, J.J. (2006) Bile acids stimulate biofilm
685 formation in *Vibrio cholerae*. *Mol Microbiol* **59**: 193–201.

686 Jones, C., and Holland, I.B. (1985) Role of the SulB (FtsZ) protein in division inhibition
687 during the SOS response in *Escherichia coli*: FtsZ stabilizes the inhibitor SulA in
688 maxicells. *Proc Natl Acad Sci U S A* **82**: 6045–6049.

689 Karp, P.D., Weaver, D., Paley, S., Fulcher, C., Kubo, A., Kothari, A., *et al.* (2014) The
690 EcoCyc Database. *EcoSal Plus* **6**
691 http://www.asmscience.org/content/journal/ecosalplus/10.1128/ecosalplus.ESP-0009-
692 2013. Accessed June 29, 2016.

693 Kaser, A., Zeissig, S., and Blumberg, R.S. (2010) Inflammatory bowel disease. *Annu Rev*
694 *Immunol* **28**: 573–621.

695 Korbel, J.O., Doerks, T., Jensen, L.J., Perez-Iratxeta, C., Kaczanowski, S., Hooper, S.D., *et al.*
696 (2005) Systematic Association of Genes to Phenotypes by Genome and Literature
697 Mining. *PLoS Biol* **3** http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1073694/. Accessed
698 July 3, 2016.

699 Kotlowski, R., Bernstein, C.N., Sepehri, S., and Krause, D.O. (2007) High prevalence of
700 *Escherichia coli* belonging to the B2+D phylogenetic group in inflammatory bowel
701 disease. *Gut* **56**: 669–675.

702 Krzywinski, M., Schein, J., Birol, İ., Connors, J., Gascoyne, R., Horsman, D., *et al.* (2009)
703 Circos: An information aesthetic for comparative genomics. *Genome Res* **19**: 1639–1645.

704 Lane, M.C., Lockatell, V., Monterosso, G., Lamphier, D., Weinert, J., Hebel, J.R., *et al.*
705 (2005) Role of Motility in the Colonization of Uropathogenic *Escherichia coli* in the
706 Urinary Tract. *Infect Immun* **73**: 7644–7656.

707 Lapidus, A., Akerlund, J.-E., and Einarsson, C. (2006) Gallbladder bile composition in
708 patients with Crohn ’s disease. *World J Gastroenterol WJG* **12**: 70–74.

709 Lawhon, S.D., Frye, J.G., Suyemoto, M., Porwollik, S., McClelland, M., and Altier, C. (2003)
710 Global regulation by CsrA in *Salmonella typhimurium*. *Mol Microbiol* **48**: 1633–1645.

711 Marks, J.W., Conley, D.R., Capretta, T.L., Bonorris, G.G., Chung, A., Coyne, M.J., and
712 Schoenfield, L.J. (1977) Gallstone prevalence and biliary lipid composition in
713 inflammatory bowel disease. *Am J Dig Dis* **22**: 1097–1100.

714 Martin, H.M., Campbell, B.J., Hart, C.A., Mpofu, C., Nayar, M., Singh, R., *et al.* (2004)
715 Enhanced *Escherichia coli* adherence and invasion in Crohn’s disease and colon cancer.
716 *Gastroenterology* **127**: 80–93.

717 Martinez-Medina, M., and Garcia-Gil, L.J. (2014) *Escherichia coli* in chronic inflammatory
718 bowel diseases: An update on adherent invasive *Escherichia coli* pathogenicity. *World J*
719 *Gastrointest Pathophysiol* **5**: 213–227.

720 Martinez-Medina, M., Naves, P., Blanco, J., Aldeguer, X., Blanco, J.E., Blanco, M., *et al.*
721 (2009) Biofilm formation as a novel phenotypic feature of adherent-invasive *Escherichia*
722 *coli* (AIEC). *BMC Microbiol* **9**: 202.

723 Mazumdar, S., Clomburg, J.M., and Gonzalez, R. (2010) *Escherichia coli* Strains Engineered
724 for Homofermentative Production of d-Lactic Acid from Glycerol. *Appl Environ*
725 *Microbiol* **76**: 4327–4336.

726 Merritt, M.E., and Donaldson, J.R. (2009) Effect of bile salts on the DNA and membrane
727 integrity of enteric bacteria. *J Med Microbiol* **58**: 1533–1541.

728 Miquel, S., Peyretaillade, E., Claret, L., Vallée, A. de, Dossat, C., Vacherie, B., *et al.* (2010)
729 Complete Genome Sequence of Crohn's Disease-Associated Adherent-Invasive *E. coli*
730 Strain LF82. *PLoS ONE* **5**: e12714.

731 Newton, W.A., and Snell, E.E. (1964) Catalytic properties of tryptophanase, a multifunctional
732 pyridoxal phosphate enzyme. *Proc Natl Acad Sci U S A* **51**: 382–389.

733 Northfield, T.C., and McColl, I. (1973) Postprandial concentrations of free and conjugated
734 bile acids down the length of the normal human small intestine. *Gut* **14**: 513–518.

735 Obradors, N., Badía, J., Baldomà, L., and Aguilar, J. (1988) Anaerobic metabolism of the L-
736 rhamnose fermentation product 1,2-propanediol in *Salmonella typhimurium*. *J Bacteriol*
737 **170**: 2159–2162.

738 O'Toole, G.A., and Kolter, R. (1998) Initiation of biofilm formation in *Pseudomonas*
739 *fluorescens* WCS365 proceeds via multiple, convergent signalling pathways: a genetic
740 analysis. *Mol Microbiol* **28**: 449–461.

741 Pereira, S.P., Bain, I.M., Kumar, D., and Dowling, R.H. (2003) Bile composition in
742 inflammatory bowel disease: ileal disease and colectomy, but not colitis, induce
743 lithogenic bile. *Aliment Pharmacol Ther* **17**: 923–933.

744 Pope, L.M., Reed, K.E., and Payne, S.M. (1995) Increased protein secretion and adherence to
745 HeLa cells by Shigella spp. following growth in the presence of bile salts. *Infect Immun*
746 **63**: 3642–3648.

747 Pos, K.M., Dimroth, P., and Bott, M. (1998) The *Escherichia coli* Citrate Carrier CitT: a
748 Member of a Novel Eubacterial Transporter Family Related to the 2-

749 Oxoglutarate/Malate Translocator from Spinach Chloroplasts. *J Bacteriol* **180**: 4160–
750 4165.

751 Prigent-Combaret, C., Prensier, G., Le Thi, T.T., Vidal, O., Lejeune, P., and Dorel, C. (2000)
752 Developmental pathway for biofilm formation in curli-producing *Escherichia coli* strains:
753 role of flagella, curli and colanic acid. *Environ Microbiol* **2**: 450–464.

754 Reen, F.J., Woods, D.F., Mooij, M.J., Adams, C., and O'Gara, F. (2012) Respiratory
755 Pathogens Adopt a Chronic Lifestyle in Response to Bile. *PLoS ONE* **7**
756 http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3458808/. Accessed July 3, 2016.

757 Rosenberg, E.Y., Bertenthal, D., Nilles, M.L., Bertrand, K.P., and Nikaido, H. (2003) Bile
758 salts and fatty acids induce the expression of *Escherichia coli* AcrAB multidrug efflux
759 pump through their interaction with Rob regulatory protein. *Mol Microbiol* **48**: 1609–
760 1619.

761 Schembri, M.A., Dalsgaard, D., and Klemm, P. (2004) Capsule Shields the Function of Short
762 Bacterial Adhesins. *J Bacteriol* **186**: 1249–1257.

763 Schnaitman, C.A., Parker, C.T., Klena, J.D., Pradel, E.L., Pearson, N.B., Sanderson, K.E., and
764 MacClachlan, P.R. (1991) Physical maps of the rfa loci of *Escherichia coli* K-12 and
765 Salmonella typhimurium. *J Bacteriol* **173**: 7410–7411.

766 Schwarz, S., Hood, R.D., and Mougous, J.D. (2010) What is type VI secretion doing in all
767 those bugs? *Trends Microbiol* **18**: 531–537.

768 Sengupta, C., Ray, S., and Chowdhury, R. (2014) Fine tuning of virulence regulatory
769 pathways in enteric bacteria in response to varying bile and oxygen concentrations in the
770 gastrointestinal tract. *Gut Pathog* **6**: 38.

771 Serra, D.O., Richter, A.M., and Hengge, R. (2013) Cellulose as an Architectural Element in
772 Spatially Structured *Escherichia coli* Biofilms. *J Bacteriol* **195**: 5540–5554.

773 Shabala, L., and Ross, T. (2008) Cyclopropane fatty acids improve *Escherichia coli* survival
774 in acidified minimal media by reducing membrane permeability to H+ and enhanced
775 ability to extrude H+. *Res Microbiol* **159**: 458–461.

776 Stevenson, G., Andrianopoulos, K., Hobbs, M., and Reeves, P.R. (1996) Organization of the
777 *Escherichia coli* K-12 gene cluster responsible for production of the extracellular
778 polysaccharide colanic acid. *J Bacteriol* **178**: 4885–4893.

779 Strober, W., Fuss, I., and Mannon, P. (2007) The fundamental basis of inflammatory bowel
780 disease. *J Clin Invest* **117**: 514–521.

Thiennimitr, P., Winter, S.E., Winter, M.G., Xavier, M.N., Tolstikov, V., Huseby, D.L., *et al.* (2011) Intestinal inflammation allows Salmonella to use ethanolamine to compete with the microbiota. *Proc Natl Acad Sci* **108**: 17480–17485.

Thomson, N.R., Howard, S., Wren, B.W., Holden, M.T.G., Crossman, L., Challis, G.L., *et al.* (2006) The Complete Genome Sequence and Comparative Genome Analysis of the High Pathogenicity Yersinia enterocolitica Strain 8081. *PLoS Genet* **2** http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1698947/. Accessed July 3, 2016.

Varet, H., Brillet-Guéguen, L., Coppée, J.-Y., and Dillies, M.-A. (2016) SARTools: A DESeq2- and EdgeR-Based R Pipeline for Comprehensive Differential Analysis of RNA-Seq Data. *PLOS ONE* **11**: e0157022.

Vazeille, E., Chassaing, B., Buisson, A., Dubois, A., Vallée, A. de, Billard, E., *et al.* (2016) GipA Factor Supports Colonization of Peyer's Patches by Crohn's Disease-associated *Escherichia coli*. *Inflamm Bowel Dis* **22**: 68–81.

Whitfield, C. (2006) Biosynthesis and Assembly of Capsular Polysaccharides in *Escherichia coli*. *Annu Rev Biochem* **75**: 39–68.

Xavier, R.J., and Podolsky, D.K. (2007) Unravelling the pathogenesis of inflammatory bowel disease. *Nature* **448**: 427–434.

Yoder-Himes, D.R., Chain, P.S.G., Zhu, Y., Wurtzel, O., Rubin, E.M., Tiedje, J.M., and Sorek, R. (2009) Mapping the *Burkholderia cenocepacia* niche response via high-throughput sequencing. *Proc Natl Acad Sci U S A* **106**: 3976–3981.

Zhang, Y., Rowehl, L., Krumsiek, J.M., Orner, E.P., Shaikh, N., Tarr, P.I., *et al.* (2015) Identification of Candidate Adherent-Invasive *E. coli* Signature Transcripts by Genomic/Transcriptomic Analysis. *PLoS ONE* **10** http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4509574/. Accessed March 17, 2016.

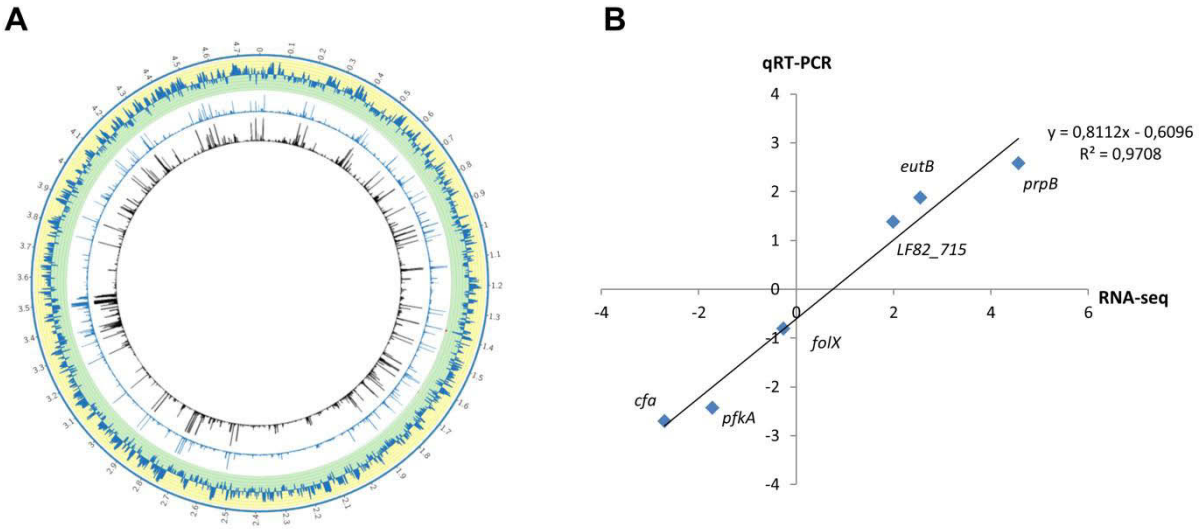
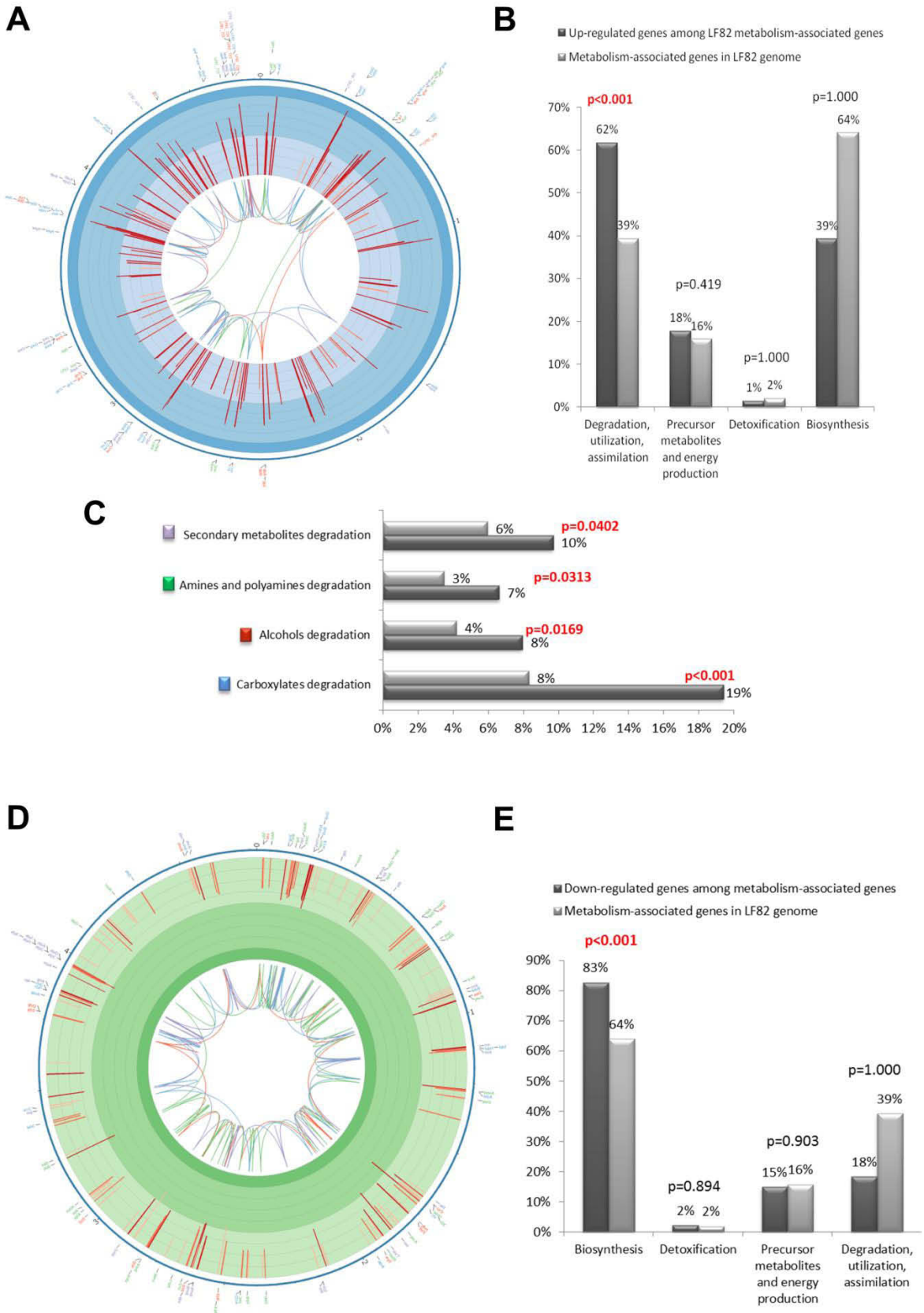



**Figure 1: Analysis and validation of RNA-seq experiments**.

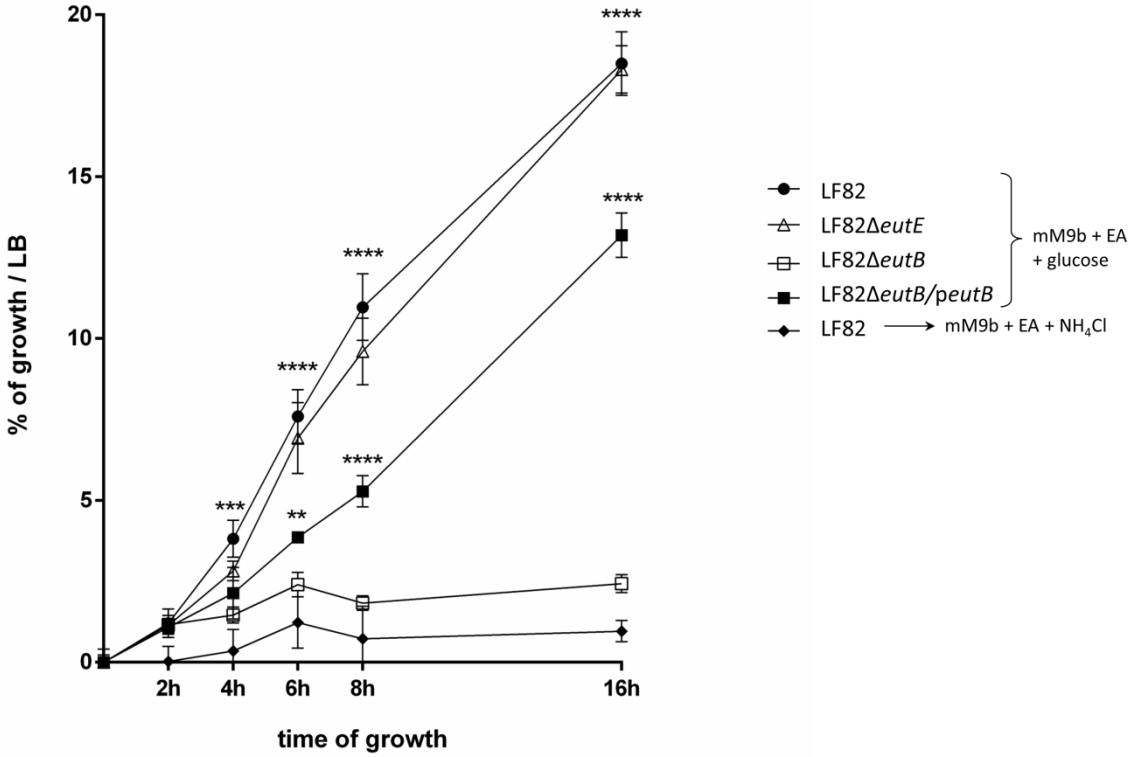
(**A**) Global analysis of transcript levels in *E. coli* LF82 by RNA-seq. The gray circular lines represent a $\log_2 = 1$ (i.e., RPKM bile salts containing medium/ minimal medium ratio values) of each gene within overexpressed genes (yellow background) and underexpressed genes (green background). Black and blue circles on white background correspond to the expression of each gene (represented as RPKM values) in bacteria grown under minimal condition and bile salts containing medium, respectively. The outermost circle represents the 4,773,108 bp of the LF82 genome. (**B**) Validation of RNA-seq results by qRT-PCR on selected genes. Mean $\log_2$ ratios of values determined in the qRT-PCR experiments are plotted against the mean $\log_2$ ratios of values determined in RNA-seq experiments (see Table S2 in the supplemental material).


A
B
Up-regulated genes among LF82 metabolism-associated genes
Metabolism-associated genes in LF82 genome
p<0.001
p=1.000
p=0.419
p=1.000
62%
39%
18%
16%
1%
2%
39%
64%
70%
60%
50%
40%
30%
20%
10%
0%
Degradation, utilization, assimilation
Precursor metabolites and energy production
Detoxification
Biosynthesis
C
Secondary metabolites degradation
Amines and polyamines degradation
Alcohols degradation
Carboxylates degradation
6%
10%
p=0.0402
3%
7%
p=0.0313
4%
8%
p=0.0169
8%
19%
p<0.001
0%
2%
4%
6%
8%
10%
12%
14%
16%
18%
20%
D
E
Down-regulated genes among metabolism-associated genes
Metabolism-associated genes in LF82 genome
p<0.001
83%
64%
p=0.894
2%
2%
p=0.903
15%
16%
p=1.000
18%
39%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
Biosynthesis
Detoxification
Precursor metabolites and energy production
Degradation, utilization, assimilation

822
823

**Figure 2: Analysis of the expression of the LF82 genes involved in metabolism**

(**A**) Global analysis of increased transcript levels of genes involved in metabolism in *E. coli* LF82 with connections into each metabolic pathway. The gray circular lines represent a $\log_2$ = 0.5 (i.e., RPKM bile salts containing medium/ minimal medium ratio values) of each gene within overexpressed genes. (**B**) Repartition of the overexpressed genes in the different metabolic pathways. The percentage of increased transcript levels in every pathway (dark grey) and the percentage of LF82 genes involved in every metabolic pathway (light grey) are represented. (**C**) Repartition of the overexpressed genes in degradation pathways. The percentage of increased transcript levels in degradation pathways (dark grey) and the percentage of LF82 genes involved in each pathway (light grey) are represented. A $p < 0.05$ is considered as significant. (**D**) Global analysis of decreased transcript levels of genes involved in metabolism in *E. coli* LF82 with connections into each metabolic pathway. The gray circular lines represent a $\log_2 = 0.5$ (i.e., RPKM bile salts containing medium/ minimal medium ratio values) of each gene within underexpressed genes. (**E**) Repartition of the underexpressed genes in the different metabolic pathways. The percentage of decreased transcript levels in every pathway (dark grey) and the percentage of LF82 genes involved in every metabolic pathway (light grey) are represented. A $p < 0.05$ is considered as significant.




842

843 **Figure 3: Bacterial growth curves of LF82 and *eut* mutants in mM9b.**

844 All the strains were incubated at 37°C with shaking in mM9b with ethanolamine (EA) and
845 glucose or $NH_4Cl$. LF82 was incubated in mM9b with EA as the sole source of carbon (filled
846 rhombus) and mM9b with EA as the sole source of nitrogen (filled circles); LF82Δ*eutB* (open
847 squares), LF82Δ*eutE* (open triangles) and LF82Δ*eutB*/p*eutB* (filled squares) were incubated
848 in EA as the sole source of nitrogen. Values are means ± SEM of at least three independent
849 experiments (ANOVA; **$p$<0.01; ***$p$<0.001; ****$p$ <0.0001).

850

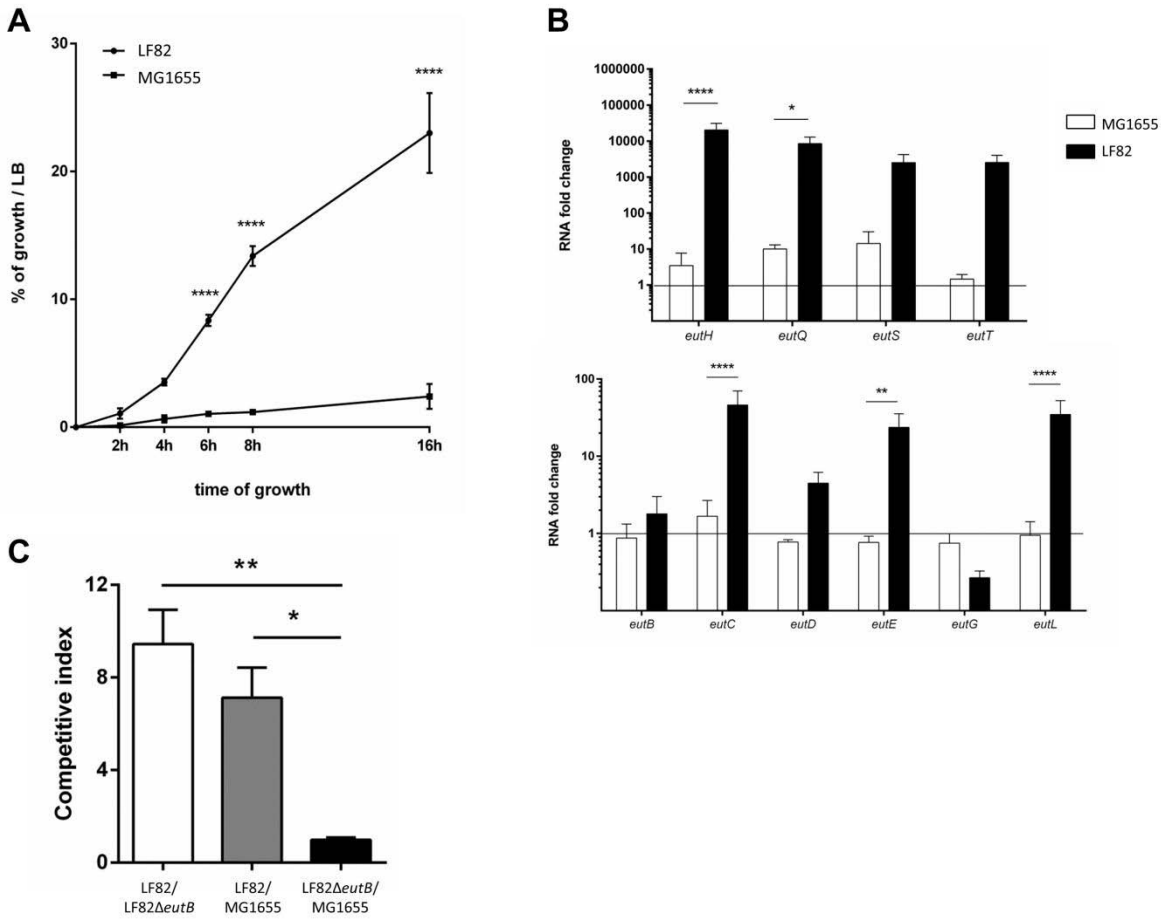
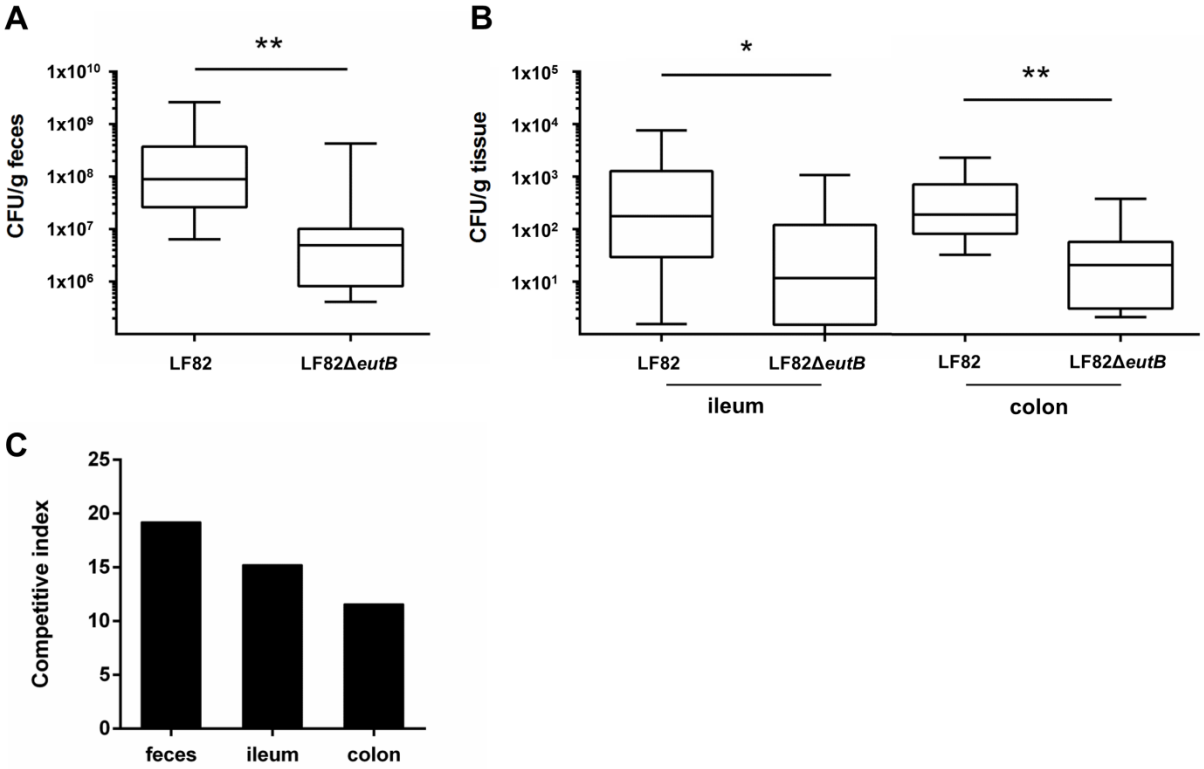



**Figure 4: EA utilization by AIEC provides them an advantage of growth.**

(**A**) Bacterial growth curves of LF82 and MG1655 in mM9b-EA. (**B**) Relative expression level of the *eut* genes. The ration of mRNA level of each gene was measured in the LF82 strain and in MG1655 strain in mM9b-EA after 8 hours of culture. (**C**) Competitive index between LF82, LF82Δ*eutB* and/or MG1655 in minimal medium supplemented with glucose, bile salts and ethanolamine. Values are means ± SD of at least three independent experiments. Statistical analysis was performed with a Mann-Whitney test or ANOVA test for multiple comparaison; * $p<0.05$; ** $p<0.01$; **** $p<0.0001$.



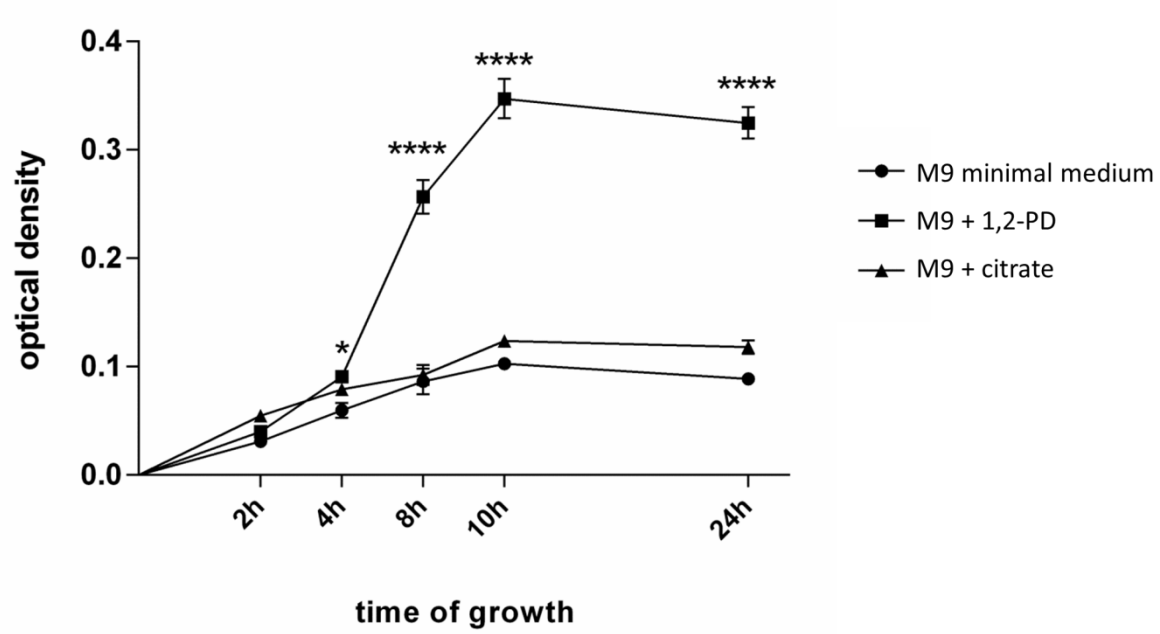

862

863

864 **Figure 5: *eutB* favors AIEC gut colonization in mice.**

865 Mice were orally co-infected with $0.5x10^9$ LF82 and $0.5x10^9$ LF82Δ*eutB*. (**A**) Feces were
866 collected three days post-infection and LF82 and LF82Δ*eutB* were counted. (**B**)
867 Quantification of ileal and colonic mucosal-associated LF82 and LF82Δ*eutB* bacteria was
868 performed three days post-infection. Statistical analysis was performed with a Mann-Whitney
869 test; *, $p < 0.05$; **, $p < 0.01$. (**C**) Median competitive index between LF82 and LF82Δ*eutB*
870 in feces, ileum and colon collected three days post-infection.

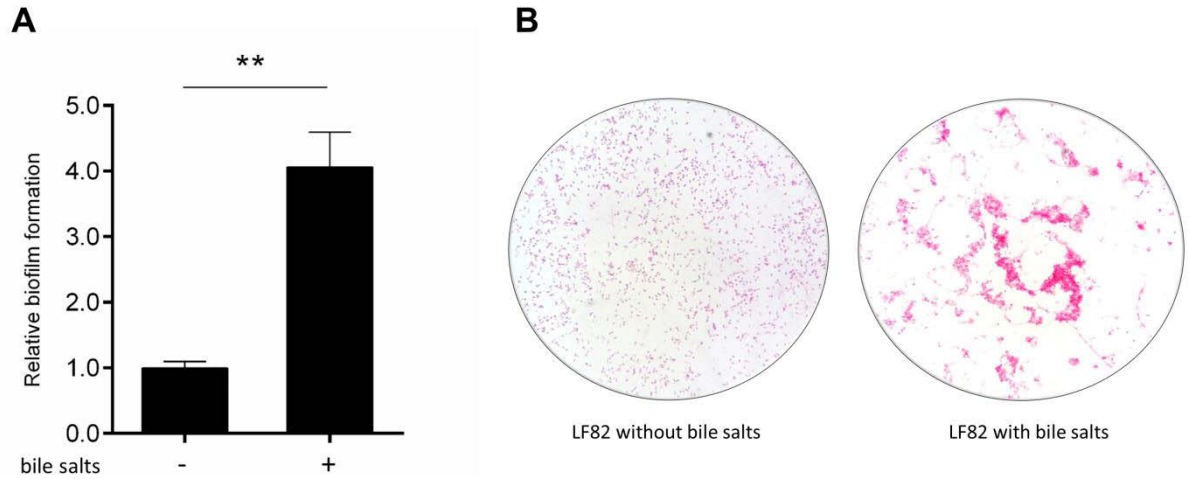
871




872

873 **Figure 6: 1.2-PD utilization by AIEC provides them an advantage of growth**.

874 LF82 was able to grow in the presence of 1,2-PD as sole carbon source, but not citrate
875 Bacterial growth curves of LF82 in M9 minimal medium containing bile salts and 1,2-PD or
876 citrate. Values are means ± SEM of at least three independent experiments (ANOVA;
877 *$p$<0.05; ****$p$ <0.0001).




878
879

880 **Figure 7: Bile salts induce biofilm formation and autoaggregation of bacteria.**

881 (**A**) Biofilm formation in the presence of bile salts. Optical density after biofilm staining has
882 been measured. Values are means ± SD of at least three independent experiments. (**B**)
883 Autoaggregation of bacteria from at least three independent experiments visualized by
884 microscopy after Gram's staining (1000x magnification).

885

886 **Table 1:** Variations of the expression of *eut* genes

| locus_tag | name | gene size (bp) | RNA-seq values (minimal medium) RPKM | RNA-seq values (bile salts) RPKM | Fold change ($Log_2$) | p value |
|---|---|---|---|---|---|---|
| LF82_0580 | eutA | 1403 | 36 | 223 | **2,65** | 2,151E-13 |
| LF82_0581 | eutB | 1361 | 61 | 358 | **2,55** | 1,818E-17 |
| LF82_0582 | eutC | 887 | 47 | 286 | **2,60** | 5,747E-17 |
| LF82_0583 | eutD | 1016 | 108 | 357 | **1,72** | 2,746E-11 |
| LF82_0584 | eutE | 1403 | 14 | 149 | **3,38** | 1,849E-16 |
| LF82_0585 | eutG | 1187 | 35 | 168 | **2,28** | 2,691E-11 |
| LF82_0586 | eutH | 1226 | 106 | 514 | **2,27** | 1,172E-15 |
| LF82_0587 | eutJ | 836 | 13 | 147 | **3,502** | 8,290E-18 |
| LF82_0588 | eutK | 500 | 37 | 94 | **1,34** | 9,218E-05 |
| LF82_0589 | eutL | 659 | 30 | 126 | **2,079** | 1,825E-09 |
| LF82_0590 | eutM | 335 | 55 | 177 | **1,69** | 1,128E-05 |
| LF82_0591 | eutN | 287 | 5 | 62 | **3,61** | 1,249E-12 |
| LF82_0592 | eutP | 479 | 39 | 166 | **2,08** | 7,469E-10 |
| LF82_0593 | eutQ | 701 | 24 | 192 | **3,009** | 1,052E-12 |
| LF82_0594 | eutR | 1052 | 142 | 308 | **1,121** | 3,639E-05 |
| LF82_0595 | eutS | 335 | 20 | 76 | **1,962** | 8,018E-07 |
| LF82_0596 | eutT | 803 | 13 | 83 | **2,68** | 6,423E-10 |

887

**Table 2:** Variations of the expression of *pdu* genes

| locus_tag | product | gene size (bp) | RNA-seq values (minimal medium) RPKM | RNA-seq values (bile salts) RPKM | Fold change ($Log_2$) | p value |
|---|---|---|---|---|---|---|
| LF82_p331 | propanediol utilization protein:polyhedral bodies | 812 | 71 | 203 | **1,51** | 3,202E-07 |
| LF82_p332 | propanediol utilization protein: glycerol dehydratase large subunit | 1664 | 251 | 532 | **1,08** | 3,910E-06 |
| LF82_p333 | propanediol utilization protein: diol dehydratase medium subunit | 668 | 110 | 236 | **1,11** | 5,478E-04 |
| LF82_p334 | propanediol utilization protein: diol dehydratase small subunit | 518 | 39 | 118 | **1,62** | 7,898E-06 |
| LF82_p335 | propanediol utilization protein: diol dehydratase reactivation | 1832 | 185 | 494 | **1,42** | 8,661E-08 |
| LF82_p336 | propanediol utilization protein: diol dehydratase reactivation | 350 | 71 | 174 | **1,292** | 4,274E-05 |
| LF82_p337 | propanediol utilization protein: polyhedral bodies | 275 | 42 | 135 | **1,694** | 2,362E-07 |
| LF82_p338 | propanediol utilization protein: polyhedral bodies | 422 | 16 | 59 | **1,89** | 1,644E-06 |
| LF82_p339 | propanediol utilization protein | 632 | 53 | 134 | **1,353** | 5,949E-05 |
| LF82_p340 | propanediol utilization protein | 491 | 78 | 233 | **1,591** | 2,260E-07 |
| LF82_p341 | propanediol utilization protein: polyhedral bodies | 275 | 48 | 90 | 0,91 | 1,456E-02 |
| LF82_p342 | propanediol utilization protein: B12 related | 1007 | 106 | 261 | **1,30** | 6,990E-06 |
| LF82_p344 | propanediol utilization protein: propanol dehydrogenase | 1112 | 81 | 211 | **1,394** | 2,872E-06 |
| LF82_p345 | Propanediol utilization protein | 1343 | 125 | 416 | **1,738** | 3,516E-09 |
| LF82_p346 | propanediol utilization protein: polyhedral bodies | 554 | 55 | 157 | **1,509** | 1,503E-05 |
| LF82_p347 | propanediol utilization protein: polyhedral bodies | 350 | 95 | 102 | 0,107 | 7,703E-01 |
| LF82_p348 | propanediol utilization protein | 443 | 188 | 152 | -0,31 | 2,703E-01 |

**Table 3:** Variations of the expression of genes involved in methylcitrate pathway

| locus_tag | name | product | gene size (bp) | RNA-seq values (minimal medium) RPKM | RNA-seq values (bile salts) RPKM | Fold change ($Log_2$) | p value |
|---|---|---|---|---|---|---|---|
| LF82_0012 | aceA | Isocitrate lyase | 1304 | 2632 | 2978 | **0,179** | 6,231E-01 |
| LF82_0016 | aceK | Isocitrate dehydrogenase kinase/phosphatase | 1724 | 1774 | 2189 | 0,304 | 2,885E-01 |
| LF82_0307 | citC | [Citrate [pro-3S]-lyase] ligase | 1058 | 326 | 588 | **0,85** | 8,689E-04 |
| LF82_0308 | citD | Citrate lyase acyl carrier protein | 296 | 5 | 50 | **3,18** | 3,731E-10 |
| LF82_0309 | citE | Citrate lyase beta subunit | 908 | 48 | 332 | **2,80** | 1,628E-17 |
| LF82_0310 | citF | Citrate lyase alpha chain | 1532 | 55 | 362 | **2,72** | 8,791E-16 |
| LF82_0311 | citG | 2-(5"-triphosphoribosyl)-3'-dephosphocoenzyme-A | 878 | 32 | 183 | **2,49** | 5,951E-14 |
| LF82_0312 | citT | Citrate carrier | 1463 | 271 | 732 | **1,434** | 2,076E-09 |
| LF82_0313 | citX | Apo-citrate lyase phosphoribosyl-dephospho-CoA | 551 | 5 | 59 | **3,414** | 1,042E-13 |
| LF82_p629 | LF82_p629 | citrate permease | 1316 | 352 | 828 | **1,234** | 6,277E-06 |
| LF82_0875 | gltA | Citrate synthase | 1283 | 15175 | 7479 | **-1,02** | 7,423E-04 |
| LF82_1086 | icd | Isocitrate dehydrogenase [NadP] | 1250 | 30473 | 17987 | **-0,76** | 1,944E-02 |
| LF82_1743 | prpR | Propionate catabolism operon regulatory protein | 1586 | 192 | 435 | **1,19** | 1,071E-04 |
| LF82_1739 | prpB | Methylisocitrate lyase | 890 | 131 | 3094 | **4,561** | 1,828E-37 |
| LF82_1740 | prpC | 2-methylcitrate synthase | 1169 | 106 | 1313 | **3,624** | 3,179E-30 |
| LF82_1741 | prpD | 2-methylcitrate dehydratase | 1451 | 66 | 609 | **3,213** | 2,654E-26 |
| LF82_1742 | prpE | Propionate--CoA ligase | 1886 | 80 | 786 | **3,30** | 5,999E-33 |

**Table 4:** Variations of the expression of genes involved in virulence

| locus_tag | name | product | gene size (bp) | RNA-seq values (minimal medium) RPKM | RNA-seq values (bile salts) RPKM | Fold change (Log2) | p value | type of virulence gene |
|---|---|---|---|---|---|---|---|---|
| ESCLF0245 | ESCLF0245 | orf | 179 | 289 | 127 | **-1,19** | 1,932E-04 | PAI-I/T6SS |
| LF82_0244 | borD | Lipoprotein bor homolog from lambdoid prophage DLP12 | 293 | 1029 | 1279 | 0,31 | 2,985E-01 | other |
| LF82_0360 | csgA | Major curlin subunit | 455 | 74 | 206 | **1,48** | 4,550E-06 | pili |
| LF82_0361 | csgB | Minor curlin subunit | 455 | 67 | 138 | **1,055** | 8,281E-04 | pili |
| LF82_0362 | csgC | curli production protein csgC | 332 | 232 | 208 | -0,15 | 6,320E-01 | pili |
| LF82_0642 | fepC | Ferric enterobactin transport ATP-binding protein fepC | 815 | 534 | 494 | -0,113 | 6,451E-01 | other |
| LF82_0643 | fepD | Ferric enterobactin transport system permease protein | 1004 | 67 | 166 | **1,305** | 1,949E-04 | other |
| LF82_0645 | fepG | Ferric enterobactin transport system permease protein | 992 | 98 | 270 | **1,466** | 5,091E-07 | other |
| LF82_0650 | fhuA | | 2243 | NA | NA | NA | NA | other |
| LF82_0651 | fhuB | Ferrichrome transport system permease protein fhuB | 1982 | 375 | 706 | **0,91** | 7,489E-05 | other |
| LF82_0652 | fhuC | Ferrichrome transport ATP-binding protein fhuC | 797 | 209 | 350 | 0,742 | 1,563E-03 | other |
| LF82_0653 | fhuD | Ferrichrome-binding periplasmic protein | 890 | 161 | 243 | 0,597 | 3,542E-02 | other |
| LF82_0658 | fimA | Type-1 fimbrial protein, A chain | 548 | 7094 | 21306 | **1,587** | 1,812E-06 | pili |
| LF82_0659 | fimB | Type 1 fimbriae regulatory protein fimB | 602 | 413 | 428 | 0,05 | 8,898E-01 | pili |
| LF82_0660 | fimC | chaperone protein fimC | 725 | 2840 | 3333 | 0,231 | 4,971E-01 | pili |
| LF82_0661 | fimD | Outer membrane usher protein fimD | 2636 | 3624 | 6328 | 0,804 | 9,777E-03 | pili |
| LF82_0662 | fimE | Type 1 fimbriae regulatory protein fimE | 596 | 809 | 450 | **-0,845** | 1,584E-04 | pili |
| LF82_0663 | fimF | Protein fimF | 530 | 206 | 415 | **1,008** | 1,064E-04 | pili |
| LF82_0664 | fimG | Protein fimG | 503 | 208 | 528 | **1,34** | 4,303E-08 | pili |
| LF82_0665 | fimH | Protein fimH | 902 | 574 | 920 | 0,681 | 5,124E-03 | pili |
| LF82_0666 | fimI | Fimbrin-like protein fimI | 539 | 3699 | 5088 | 0,46 | 2,335E-01 | pili |
| LF82_1503 | nlpI | Lipoprotein nlpI | 884 | 61411 | 21074 | **-1,54** | 2,140E-04 | other |
| LF82_1559 | ompA | OMP A | 1040 | 117890 | 28810 | **-2,03** | 2,059E-07 | OMP |
| LF82_1560 | ompC | OMP C | 1127 | 7553 | 12146 | 0,685 | 2,604E-02 | OMP |
| LF82_1700 | ppdD | Prepilin peptidase-dependent protein D | 440 | 116 | 128 | 0,14 | 6,745E-01 | pili |
| LF82_2474 | yadC | fimbrial-like protein yadC | 1262 | 1607 | 1611 | 0,004 | 9,912E-01 | pili |
| LF82_2480 | yadK | orf | 596 | 103 | 271 | 1,392 | 1,260E-01 | pili |
| LF82_2481 | yadL | orf | 608 | 92 | 160 | 0,81 | 3,809E-03 | pili |
| LF82_2505 | yafT | lipoprotein yafT | 785 | 10365 | 1545 | **-2,75** | 1,150E-15 | PAI-I/T6SS |
| LF82_2821 | ydeP | Protein ydeP | 2279 | 568 | 862 | 0,601 | 1,532E-02 | pili |
| LF82_2822 | ydeQ | fimbrial-like protein ydeQ | 914 | 113 | 250 | **1,148** | 5,470E-04 | pili |
| LF82_2823 | ydeR | fimbrial-like protein ydeR | 503 | 339 | 421 | 0,315 | 2,302E-01 | pili |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| LF82_2824 | ydes | fimbrial-like protein ydes | 530 | 181 | 323 | 0,83 | 7,688E-03 | pili |
| LF82_2988 | yehA | orf | 1034 | 351 | 465 | 0,41 | 1,382E-01 | pili |
| LF82_2989 | yehB | outer membrane usher protein yehB | 2480 | 243 | 822 | **1,761** | 2,152E-11 | pili |
| locus_tag | name | product | gene size (bp) | RNA-seq values (minimal medium) RPKM | RNA-seq values (bile salts) RPKM | **Fold change (Log2)** | p value | type of virulence gene |
| LF82_2990 | yehC | fimbrial chaperone yehC | 674 | 27 | 223 | **3,02** | 7,172E-18 | pili |
| LF82_2991 | yehD | orf | 542 | 28 | 124 | **2,14** | 5,677E-11 | pili |
| LF82_2992 | yehE | orf | 281 | 438 | 452 | 0,04 | 8,737E-01 | pili |
| LF82_3097 | yfgl | Lipoprotein yfgl | 1178 | 5402 | 1840 | **-1,554** | 4,347E-08 | OMP |
| LF82_p013 | LF82_p013 | orf | 479 | 1233 | 805 | -0,62 | 9,764E-03 | PAI-I/T6SS |
| LF82_p014 | LF82_p014 | orf | 1358 | 457 | 523 | 0,195 | 4,563E-01 | PAI-I/T6SS |
| LF82_p015 | LF82_p015 | orf | 3488 | 544 | 1134 | **1,061** | 9,424E-06 | PAI-I/T6SS |
| LF82_p016 | LF82_p016 | orf | 1442 | 56 | 264 | **2,226** | 1,606E-11 | PAI-I/T6SS |
| LF82_p017 | LF82_p017 | orf | 743 | 15 | 55 | **1,842** | 1,353E-06 | PAI-I/T6SS |
| LF82_p018 | LF82_p018 | orf | 2762 | 105 | 595 | **2,501** | 3,437E-18 | PAI-I/T6SS |
| LF82_p019 | LF82_p019 | orf | 764 | 62 | 209 | **1,751** | 9,261E-09 | PAI-I/T6SS |
| LF82_p020 | LF82_p020 | orf | 1346 | 115 | 374 | **1,693** | 8,700E-10 | PAI-I/T6SS |
| LF82_p021 | LF82_p021 | orf | 524 | 51 | 247 | **2,279** | 7,122E-13 | PAI-I/T6SS |
| LF82_p022 | LF82_p022 | orf | 1292 | 48 | 291 | **2,598** | 1,824E-16 | PAI-I/T6SS |
| LF82_p023 | LF82_p023 | orf | 1049 | 81 | 299 | **1,877** | 1,288E-11 | PAI-I/T6SS |
| LF82_p024 | LF82_p024 | orf | 1841 | 1681 | 1284 | -0,389 | 1,292E-01 | PAI-I/T6SS |
| LF82_p025 | LF82_p025 | orf | 425 | 83 | 133 | 0,688 | 3,854E-02 | PAI-I/T6SS |
| LF82_p026 | LF82_p026 | orf | 1484 | 276 | 553 | **1,003** | 2,139E-05 | PAI-I/T6SS |
| LF82_p027 | LF82_p027 | orf | 503 | 266 | 385 | 0,536 | 3,588E-02 | PAI-I/T6SS |
| LF82_p028 | LF82_p028 | orf | 518 | 144 | 190 | 0,407 | 1,324E-01 | PAI-I/T6SS |
| LF82_p029 | LF82_p029 | orf | 2342 | 535 | 1322 | **1,31** | 8,033E-07 | PAI-I/T6SS |
| LF82_p030 | LF82_p030 | orf | 899 | 326 | 666 | **1,03** | 5,724E-05 | PAI-I/T6SS |
| LF82_p031 | LF82_p031 | orf | 665 | 261 | 332 | 0,34 | 2,194E-01 | PAI-I/T6SS |
| LF82_p032 | LF82_p032 | receptor protein | 650 | 780 | 526 | -0,57 | 2,263E-02 | PAI-I/T6SS |
| LF82_p033 | LF82_p033 | orf | 1982 | 192 | 1082 | **2,50** | 2,050E-17 | PAI-I/T6SS |
| LF82_p034 | LF82_p034 | orf | 1046 | 514 | 692 | 0,43 | 9,951E-02 | PAI-I/T6SS |
| LF82_p035 | LF82_p035 | orf | 1439 | 1290 | 1369 | 0,09 | 7,537E-01 | PAI-I/T6SS |
| LF82_p036 | LF82_p036 | orf | 290 | 394 | 451 | 0,20 | 5,257E-01 | PAI-I/T6SS |
| LF82_p037 | LF82_p037 | yhhI transposase | 485 | 789 | 667 | -0,24 | 3,086E-01 | PAI-I/T6SS |
| LF82_p038 | LF82_p038 | orf | 503 | 309 | 366 | 0,25 | 3,879E-01 | PAI-I/T6SS |
| LF82_p039 | LF82_p039 | orf | 512 | 292 | 402 | 0,46 | 4,845E-02 | PAI-I/T6SS |
| LF82_p043 | LF82_p043 | Haemoglobin protease | 4130 | 3127 | 2512 | -0,316 | 2,394E-01 | other |
| LF82_p085 | GipA | Peyer's patch-specific virulence factor GipA | 1130 | 3565 | 3096 | -0,204 | 5,573E-01 | other |
| LF82_p184 | LF82_p184 | SitD protein | 857 | 920 | 842 | -0,126 | 6,203E-01 | other |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| LF82_p185 | LF82_p185 | SitC protein | 857 | 1064 | 995 | -0,097 | 6,951E-01 | other |
| LF82_p186 | LF82_p186 | SitB protein | 827 | 1412 | 747 | **-0,92** | 1,331E-04 | other |
| LF82_p294 | LF82_p294 | Prophage P4 integrase | 1043 | 3520 | 1397 | **-1,33** | 3,190E-05 | PAI-II/yersinia path island |
| LF82_p296 | LF82_p296 | orf | 1280 | 193 | 224 | 0,22 | 4,766E-01 | PAI-II/yersinia path island |

| locus_tag | name | product | gene size (bp) | RNA-seq values (minimal medium) RPKM | RNA-seq values (bile salts) RPKM | Fold change (Log2) | p value | type of virulence gene |
|---|---|---|---|---|---|---|---|---|
| LF82_p297 | LF82_p297 | ABC transporter protein | 1802 | 394 | 358 | -0,14 | 6,662E-01 | PAI-II/yersinia path island |
| LF82_p298 | LF82_p298 | inner membrane ABC-transporter | 1802 | 785 | 469 | -0,74 | 1,600E-02 | PAI-II/yersinia path island |
| LF82_p299 | LF82_p299 | Arac type regulator | 959 | 271 | 307 | 0,19 | 5,537E-01 | PAI-II/yersinia path island |
| LF82_p300 | LF82_p300 | peptide synthetase-like protein | 6125 | 383 | 961 | **1,33** | 1,637E-06 | PAI-II/yersinia path island |
| LF82_p301 | LF82_p301 | HMWP1 nonribosomal peptide/polyketide synthase | 9491 | 254 | 1207 | **2,25** | 3,826E-18 | PAI-II/yersinia path island |
| LF82_p302 | LF82_p302 | orf | 1100 | 30 | 206 | **2,793** | 4,647E-15 | PAI-II/yersinia path island |
| LF82_p303 | LF82_p303 | YbtT protein | 788 | 36 | 174 | **2,29** | 7,254E-12 | PAI-II/yersinia path island |
| LF82_p304 | LF82_p304 | PAI-II/yersinia bactin siderophore biosynthetic protein | 1577 | 48 | 207 | **2,095** | 3,255E-10 | PAI-II/yersinia path island |
| LF82_p305 | LF82_p305 | pesticin receptor | 2021 | 4483 | 746 | **-2,59** | 2,506E-20 | PAI-II/yersinia path island |
| LF82_p306 | LF82_p306 | orf | 749 | 300 | 250 | -0,26 | 3,826E-01 | PAI-II/yersinia path island |
| LF82_p307 | LF82_p307 | orf | 1469 | 1117 | 807 | -0,47 | 3,788E-02 | PAI-II/yersinia path island |
| LF82_p308 | LF82_p308 | orf | 449 | 150 | 181 | 0,28 | 3,499E-01 | PAI-II/yersinia path island |
| LF82_p366 | LF82_p366 | orf | 857 | 683 | 608 | -0,169 | 4,454E-01 | pili |
| LF82_p367 | LF82_p367 | fimbrial-like protein yfcP | 446 | 40 | 149 | **1,89** | 1,704E-08 | pili |
| LF82_p368 | LF82_p368 | Yfc fimbriae subunit YfcQ | 479 | 39 | 99 | **1,33** | 1,296E-04 | pili |
| LF82_p369 | LF82_p369 | orf | 503 | 28 | 47 | 0,74 | 8,409E-02 | pili |
| LF82_p370 | LF82_p370 | periplasmic chaperone YfcS | 752 | 100 | 261 | **1,38** | 5,212E-06 | pili |
| LF82_p371 | LF82_p371 | outer membrane usher YfcU | 2654 | 980 | 1178 | 0,27 | 3,152E-01 | pili |
| LF82_p372 | LF82_p372 | Yfc fimbriae subunit YfcV | 566 | 429 | 143 | **-1,58** | 2,635E-08 | pili |
| LF82_p374 | LF82_p374 | RatA-like protein | 8051 | 5137 | 5158 | 0,01 | 9,879E-01 | other |
| LF82_p426 | LF82_p426 | orf | 500 | 142 | 228 | 0,68 | 1,725E-02 | PAI-III/T6SS |
| LF82_p427 | LF82_p427 | orf | 1538 | 155 | 473 | **1,61** | 1,909E-09 | PAI-III/T6SS |
| LF82_p428 | LF82_p428 | orf | 1337 | 45 | 253 | **2,49** | 1,584E-15 | PAI-III/T6SS |
| LF82_p429 | LF82_p429 | orf | 665 | 22 | 178 | **3,03** | 2,723E-14 | PAI-III/T6SS |
| LF82_p430 | LF82_p430 | orf | 1652 | 32 | 273 | **3,107** | 4,958E-19 | PAI-III/T6SS |
| LF82_p431 | LF82_p431 | Secreted protein Hcp | 491 | 91 | 274 | **1,59** | 1,995E-08 | PAI-III/T6SS |
| LF82_p432 | LF82_p432 | ClpB protein | 2636 | 55 | 440 | **3,01** | 5,135E-20 | PAI-III/T6SS |
| LF82_p433 | LF82_p433 | orf | 2474 | 128 | 744 | **2,54** | 8,126E-19 | PAI-III/T6SS |
| LF82_p434 | LF82_p434 | orf | 824 | 1133 | 929 | -0,29 | 2,269E-01 | PAI-III/T6SS |
| LF82_p435 | LF82_p435 | orf | 1976 | 1426 | 1421 | 0,00 | 9,894E-01 | PAI-III/T6SS |
| LF82_p436 | LF82_p436 | orf | 272 | 72 | 317 | **2,133** | 3,567E-10 | PAI-III/T6SS |
| LF82_p437 | LF82_p437 | orf | 1433 | 57 | 139 | **1,28** | 3,881E-04 | PAI-III/T6SS |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| LF82_p438 | LF82_p438 | orf | 3338 | 417 | 900 | **1,11** | 5,987E-07 | PAI-III/T6SS |
| LF82_p439 | LF82_p439 | orf | 1610 | 83 | 287 | **1,79** | 1,018E-10 | PAI-III/T6SS |
| LF82_p440 | LF82_p440 | orf | 1256 | 525 | 888 | **0,76** | 8,336E-04 | PAI-III/T6SS |
| LF82_p441 | LF82_p441 | orf | 443 | 254 | 274 | 0,11 | 7,182E-01 | PAI-III/T6SS |
| LF82_p442 | LF82_p442 | orf | 1760 | 259 | 860 | **1,73** | 4,158E-10 | PAI-III/T6SS |
| LF82_p443 | LF82_p443 | orf | 1079 | 101 | 233 | **1,20** | 1,118E-05 | PAI-III/T6SS |

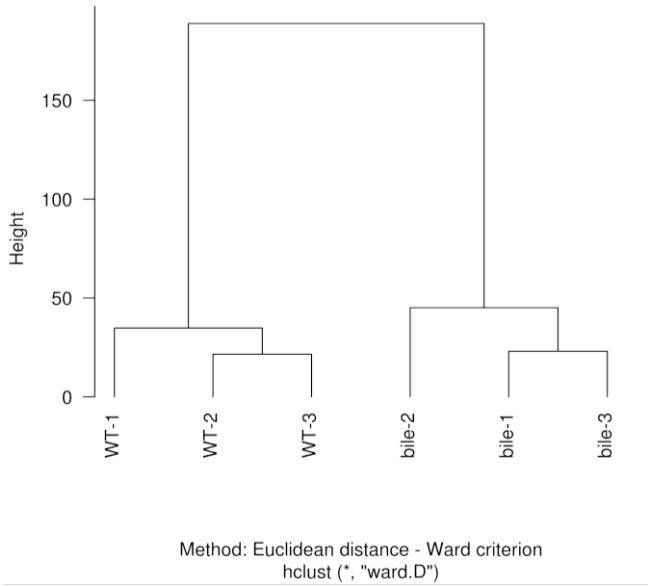
| locus_tag | name | product | gene size (bp) | RNA-seq values (minimal medium) RPKM | RNA-seq values (bile salts) RPKM | **Fold change (Log2)** | p value | type of virulence gene |
|---|---|---|---|---|---|---|---|---|
| LF82_p444 | LF82_p444 | orf | 536 | 84 | 424 | **2,34** | 2,445E-13 | PAI-III/T6SS |
| LF82_p445 | LF82_p445 | orf | 428 | 34 | 178 | **2,40** | 1,491E-12 | PAI-III/T6SS |
| LF82_p446 | LF82_p446 | orf | 1376 | 447 | 632 | 0,50 | 3,048E-02 | PAI-III/T6SS |
| LF82_p447 | LF82_p447 | 2-hydroxy acid dehydrogenase | 947 | 634 | 640 | 0,01 | 9,605E-01 | PAI-III/T6SS |
| LF82_p461 | LF82_p461 | orf | 983 | 140 | 680 | **2,273** | 8,694E-21 | PAI-IV/group 2 capsul |
| LF82_p462 | LF82_p462 | Polysialic acid capsule synthesis protein KpsE, Capsule polysaccharide export inner-membrane protein kpsE | 1148 | 187 | 629 | **1,75** | 1,511E-15 | PAI-IV/group 2 capsul |
| LF82_p463 | LF82_p463 | KpsD protein | 1676 | 260 | 802 | **1,625** | 2,746E-14 | PAI-IV/group 2 capsul |
| LF82_p464 | LF82_p464 | 3-deoxy-manno-octulosonatecytidylyltransferase(EC 27738) | 740 | 54 | 296 | **2,443** | 1,120E-17 | PAI-IV/group 2 capsul |
| LF82_p465 | LF82_p465 | KpsC protein | 2027 | 324 | 1002 | **1,627** | 4,229E-14 | PAI-IV/group 2 capsul |
| LF82_p466 | LF82_p466 | KpsS protein | 1259 | 246 | 916 | **1,90** | 4,083E-18 | PAI-IV/group 2 capsul |
| LF82_p467 | LF82_p467 | Transposase within prophage | 1124 | 212 | 547 | **1,37** | 2,025E-08 | PAI-IV/group 2 capsul |
| LF82_p468 | LF82_p468 | orf | 983 | 222 | 544 | **1,29** | 5,044E-06 | PAI-IV/group 2 capsul |
| LF82_p469 | LF82_p469 | orf | 1166 | 333 | 773 | **1,22** | 8,374E-07 | PAI-IV/group 2 capsul |
| LF82_p470 | LF82_p470 | orf | 1190 | 430 | 709 | **0,72** | 6,277E-04 | PAI-IV/group 2 capsul |
| LF82_p471 | LF82_p471 | orf | 1250 | 212 | 476 | **1,17** | 8,775E-04 | PAI-IV/group 2 capsul |
| LF82_p472 | LF82_p472 | glycerol-3-phosphate cytidyltransferase | 395 | 137 | 490 | **1,842** | 5,837E-14 | PAI-IV/group 2 capsul |
| LF82_p473 | LF82_p473 | KpsT protein | 674 | 119 | 280 | **1,233** | 1,382E-05 | PAI-IV/group 2 capsul |
| LF82_p474 | LF82_p474 | KpsM protein | 776 | 250 | 498 | **1,00** | 6,626E-05 | PAI-IV/group 2 capsul |
| LF82_p509 | LF82_p509 | Yqi fimbriae subunit YgiL | 563 | 129 | 206 | 0,69 | 1,835E-02 | pili |
| LF82_p510 | LF82_p510 | outer membrane usher protein YqiG | 2522 | 1748 | 1134 | -0,62 | 8,810E-03 | pili |
| LF82_p511 | LF82_p511 | fimbrial chaperone yqiH | 749 | 779 | 620 | -0,33 | 1,175E-01 | pili |
| LF82_p512 | LF82_p512 | Yqi fimbrial adhesin | 1049 | 3757 | 1483 | **-1,34** | 1,502E-06 | pili |
| LF82_p527 | LF82_p527 | fimbrial adhesin | 1130 | 190 | 558 | **1,552** | 4,169E-09 | pili |
| LF82_p528 | LF82_p528 | fimbrial chaperone | 749 | 116 | 377 | **1,71** | 1,416E-09 | pili |
| LF82_p529 | LF82_p529 | minor fimbrial subunit precursor | 515 | 129 | 277 | **1,103** | 1,024E-04 | pili |
| LF82_p530 | LF82_p530 | minor fimbrial subunit precursor | 563 | 187 | 481 | **1,361** | 4,641E-07 | pili |
| LF82_p531 | LF82_p531 | Outer membrane usher protein AufC | 2594 | 422 | 1586 | **1,91** | 1,131E-11 | pili |
| LF82_p532 | LF82_p532 | Auf fimbiral chaperone 1 | 752 | 324 | 487 | 0,589 | 1,275E-02 | pili |
| LF82_p533 | LF82_p533 | | 704 | 882 | 621 | -0,51 | 2,276E-02 | pili |
| LF82_p541 | LF82_p541 | Outer membrane hemin receptor chuA | 1982 | 845 | 1089 | 0,367 | 1,896E-01 | OMP |
| LF82_p548 | LF82_p548 | major fimbrial subunit | 524 | 180 | 294 | 0,71 | 1,020E-02 | pili |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| LF82_p549 | LF82_p549 | fimbriae | 1052 | 1495 | 1122 | -0,414 | 1,111E-01 | pili |
| LF82_p550 | LF82_p550 | Outer membrane usher protein lpfC precursor | 2531 | 1297 | 1444 | 0,154 | 5,427E-01 | pili |
| LF82_p551 | LF82_p551 | Fimbrial chaperone protein | 686 | 96 | 193 | 1,00 | 1,393E-03 | pili |
| LF82_p552 | LF82_p552 | fimbrial-like protein | 524 | 244 | 188 | -0,37 | 1,943E-01 | pili |
| LF82_p734 | LF82_p734 | Invasion protein IbeA | 1370 | 135 | 678 | **2,33** | 5,930E-14 | other |

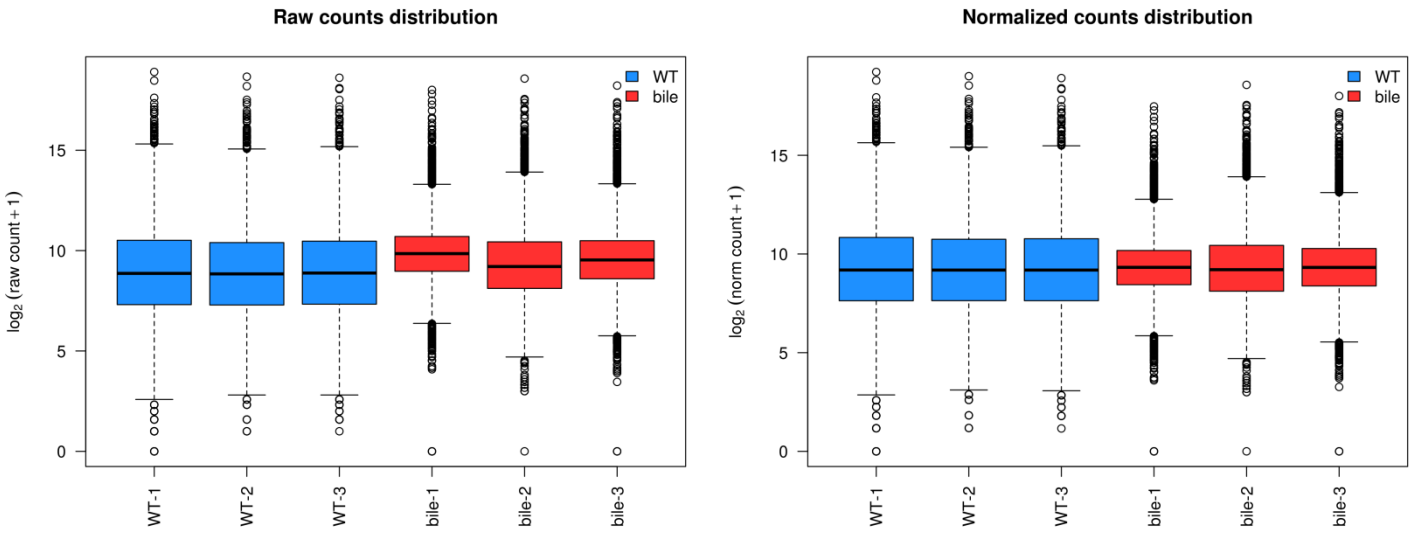
## Supplemental material

**Figure S1: Global data retrieved from RNA-seq. A**) Dendrogram built upon the Ward criteria which show variability in raw data within the experiment conditions. **B)** Boxplots of raw (left) and normalized (right) data to assess the quality of the normalization process. **C**) Volcano plot (bile salts containing medium *versus* medium) representing the log of the adjusted p value as a function of the log ratio of differential expression.

**A**)




**B**)



**C)**

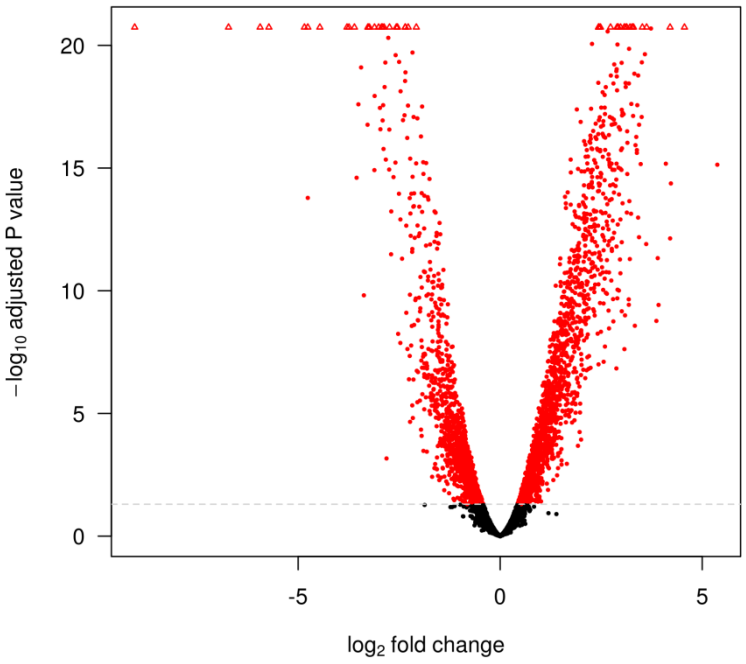



**Figure S2: Bile salts induce a diminution of bacterial motility.**

LF82 was grown overnight at 37°C in a medium supplemented or not with bile salts and then inoculated into soft agar plates. The motility of LF82 strain was altered with the addition of bile salts in the medium.

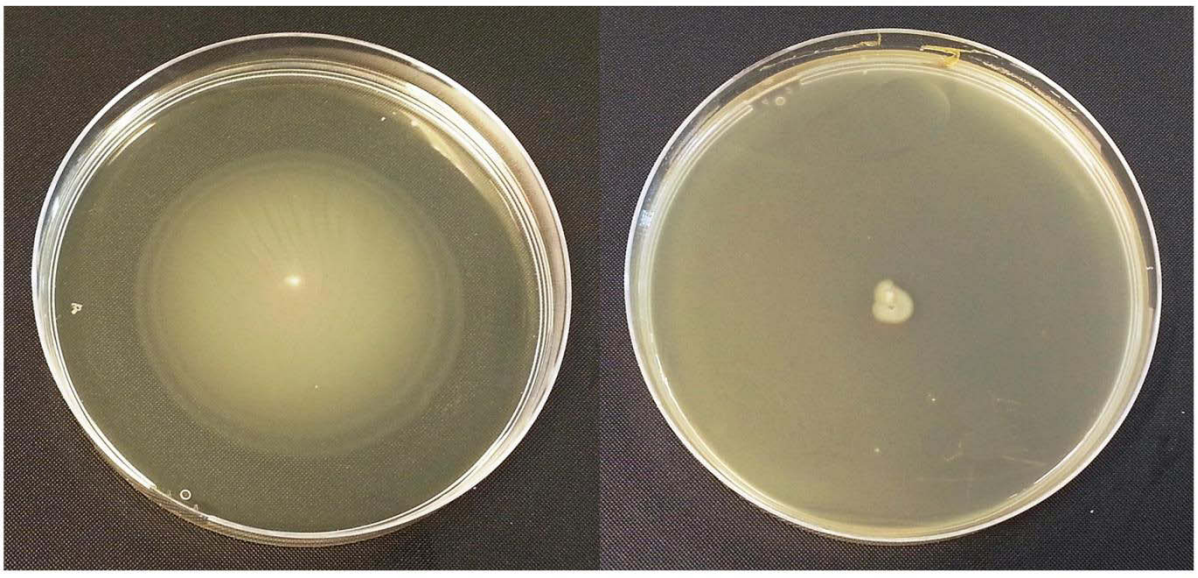


**Figure S3**: **Bile salts induce increased expression of *vat-AIEC* in early growth phase**.

Effect of bile salts on *vat-AIEC* expression. LF82 wild-type strain was cultured for 2, 6, 10 and 24 hours and *vat-AIEC* mRNA levels were analyzed by qRT-PCR. Each value is the mean ± SEM of at least three separate experiments (Mann-Whitney test; *, $P$ <0.05; **, $P$ <0.01).

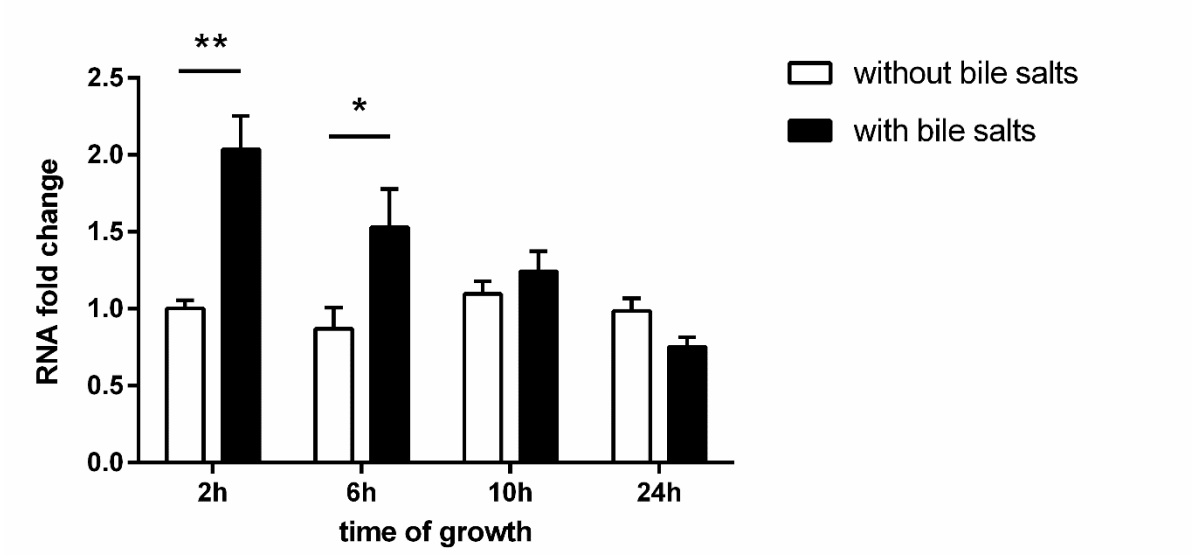
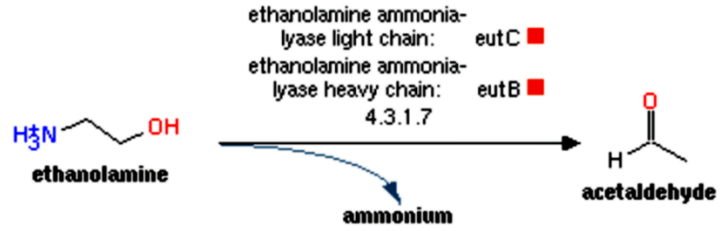



**Figure S4: Bile salts induce variations of gene expression corresponding to metabolic pathways.** For the different pathways, such as ethanolamine utilization (**A**), 2-methylcitrate cycle (**B**) and colanic acid building blocks biosynthesis (**C**), an overview of the reaction from http://www.biocyc.org/ is shown. The overexpressed genes are highlighted with little red squares and the corresponding gene cluster are represented below each metabolic reaction.

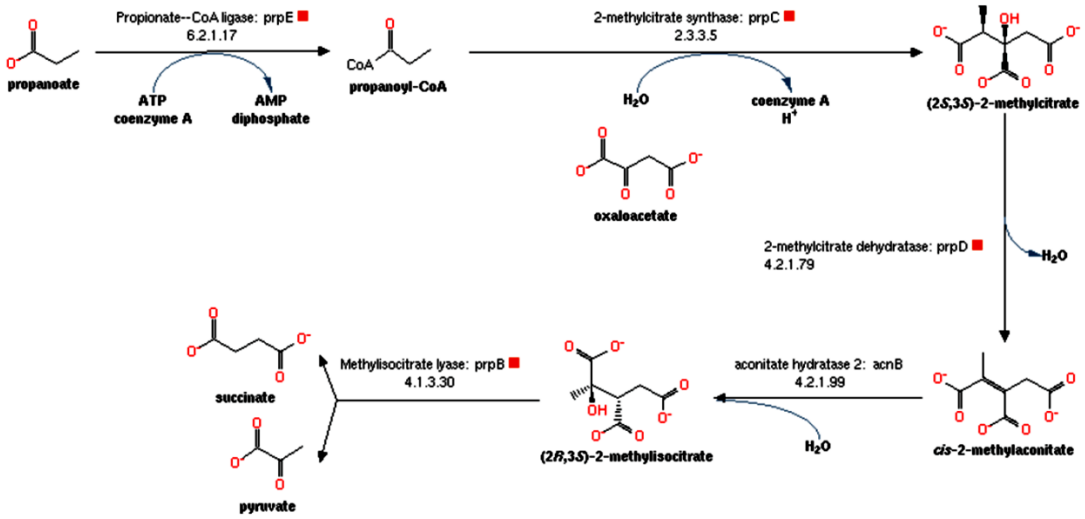
**A**



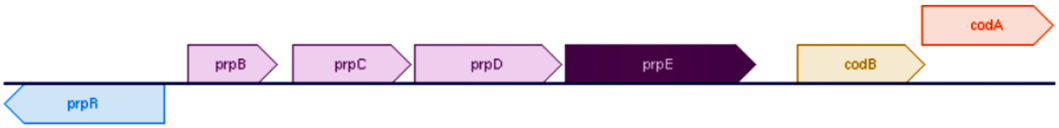

Gene Local Context (not to scale -- see Genome Browser for correct scale)

Transcription Unit

**B**

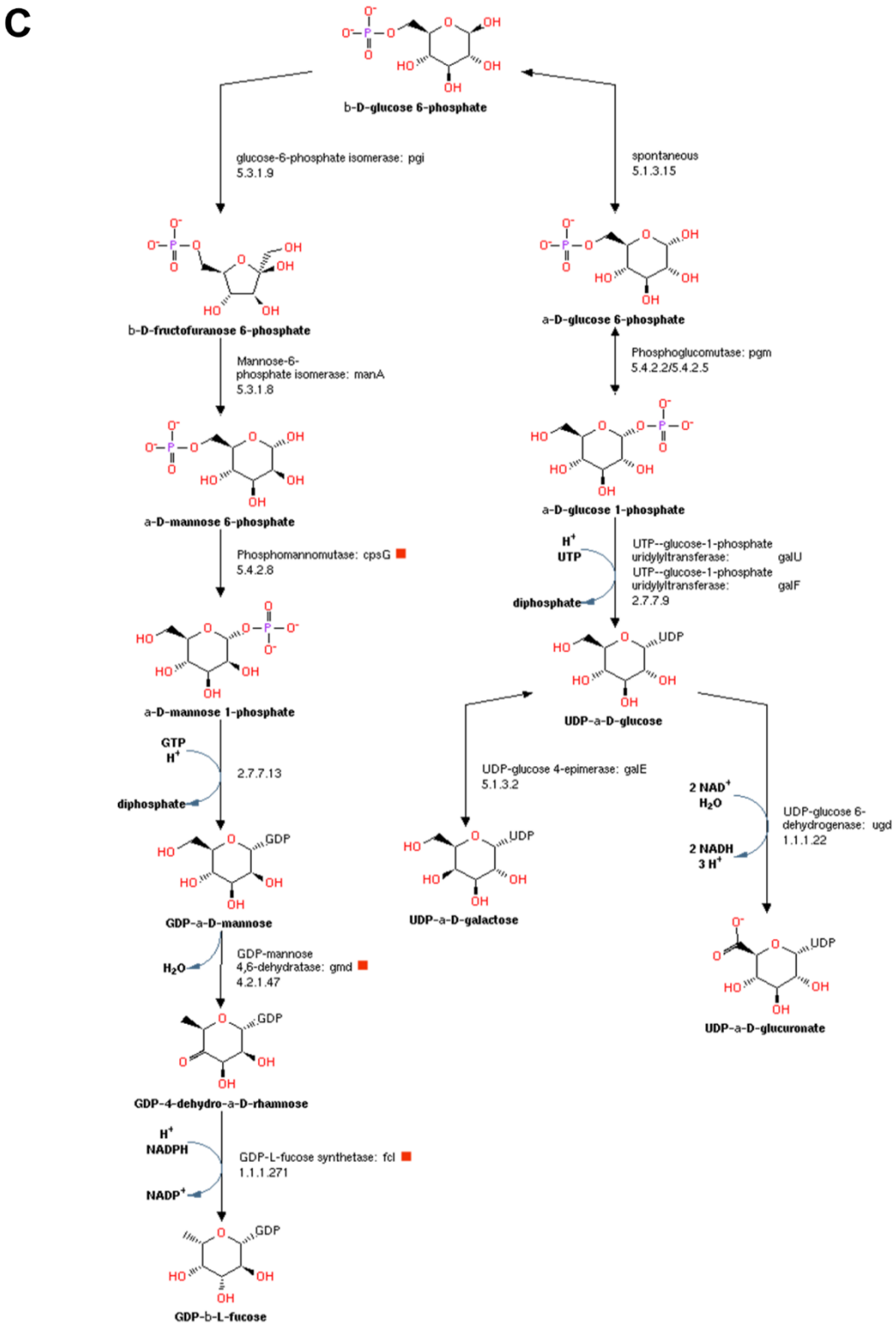



Gene Local Context (not to scale -- see Genome Browser for correct scale)




Transcription Unit

C




**Gene Local Context (not to scale -- see Genome Browser for correct scale)**

**Transcription Unit**

**Table S1:** Bacterial strains and plasmids used in this study

| Strain or plasmid | Relevant characteristics | Reference |
|---|---|---|
| LF82 | AIEC reference strain | (1) |
| LF82Δ*eutB* | LF82 isogenic mutant deleted for the *eutB* gene | this study |
| LF82Δ*eutE* | LF82 isogenic mutant deleted for the *eutE* gene | this study |
| LF82Δ*eutB*/p*eutB* | LF82Δ*eutB* containing plasmid pBK-CMV carrying *eutB* | this study |
| MG1655 | *Escherichia coli* K-12 serotype OR:H48:K- | Laboratory stock |
| BW25141 | Plasmid pKD4 carrying a kanamycin resistance cassette | (2) |
| pKOBEG | pBAD cloning vector harboring a λ phage redγβα operon; chloramphenicol resistant | (3) |
| pCP20 | Ampicillin and chloramphenicol resistant plasmid that shows temperature-sensitive replication and thermal induction of FLP synthesis | (2) |

**REFERENCES**

1. Barnich N, Darfeuille-Michaud A. Adherent-invasive Escherichia coli and Crohn's disease. Curr Opin Gastroenterol. janv 2007;23(1):16‑20.

2. 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. 6 juin 2000;97(12):6640‑5.

3. Chaveroche MK, Ghigo JM, d'Enfert C. A rapid method for efficient gene replacement in the filamentous fungus Aspergillus nidulans. Nucleic Acids Res. 15 nov 2000;28(22):E97.

**Table S2:** Primers used in this study

| Primer | Sequence (5'→3') | PCR product size | Use |
|---|---|---|---|
| d-eutB-F | GTGCGGTGACGGTGAAATCACTCGCATTTCCTTCCTGAGGGAACGACTTGTAGGCTGGAGCTGCTTC | - | construction of isogenic mutant |
| d-eutB-R | TACAATTTCTTCAATCTGTTTTTGATCCATGATGTGTTATCTCCGCGTCACATATGAATATCCTCCTTAG | | |
| d-eutE-F | GAGGTGGTGTCTGGCGGTCAGGTCATTTTCCACAAATAAGGCAGAACATCGTAGGCTGGAGCTGCTTC | - | construction of isogenic mutant |
| d-eutE-R | TTGCAGACGTGGAGTGAGCCATTGTTCATCGTGCGCCATCGGTTACTCCTCATATGAATATCCTCCTTAG | | |
| eutBEcoR1-F | GGAATTCATGAAACTAAAGACCAC | 1,362 bp | eutB cloning |
| eutBBamH1-R | CGGGATCCCGGGATCCTCAGAAGAACA | | |
| eutB-F | AGTATCGCCGCGCAAATCTA | 151 bp | RT-PCR amplification |
| eutB-R | CCCTGGGTCGGAATGTTGAA | | |
| eutC-F | GCAGGTCGTTATTTCTGATG | 228 bp | RT-PCR amplification |
| eutC-R | GCTTTCTGACTGCCCCAACC | | |
| eutD-F | GTTTTTCGGATGCGTTAGA | 198 bp | RT-PCR amplification |
| eutD-R | CCAGCGATGAGCAAATTCTT | | |
| eutE-F | TGAAACTGACCGCAGAGCAG | 206 bp | RT-PCR amplification |
| eutE-R | TCAGTTCGGTCACGGCAAAC | | |
| eutG-F | GCTGGTTTCTGACATCTCTG | 213 bp | RT-PCR amplification |
| eutG-R | CCTGGCCGTTAAAGGCATC | | |
| eutH-F | CTGGTGGTCTGGTAGCGAT | 206 bp | RT-PCR amplification |
| eutH-R | CAAGACCGAGGGTGATCAAT | | |
| eutL-F | GCGTACCGGTTCTTATCTCT | 191 bp | RT-PCR amplification |
| eutL-R | TGGCTACCGGTTAAAAATGC | | |
| eutQ-F | TAGGCTTGACTGATCTCGT | 196 bp | RT-PCR amplification |
| eutQ-R | GCCTTTCGGGATAAACATCA | | |
| eutS-F | CGGCGACTCATGGATAAGA | 193 bp | RT-PCR amplification |
| eutS-R | GCTTTCAACGCTAAATCAC | | |

| | | | |
|---|---|---|---|
| eutT-F | GTGCGTAGCAGATTCAGC | 279 bp | RT-PCR amplification |
| eutT-R | TTCGCGCTGTTCTCG | | |
| **Primer** | **Sequence (5'→3')** | **PCR product size** | **Use** |
| prpB-F | TGGATGCGAAAACCGATCCT | 149 bp | RT-PCR amplification |
| prpB-R | TGGCGAGTTCGGTAATTGCT | | |
| folX-F | CCACTACCCCGCCGATAAAG | 174 bp | RT-PCR amplification |
| folX-R | TAGCATACGTCACCCAGTG | | |
| pfkA-F | TCGTGGCGATTACCGAACAT | 178 bp | RT-PCR amplification |
| pfkA -R | CAGCAGCAGATCGATAGCGT | | |
| cfa-F | TTTGCGGTGGTGGATCGTAA | 157 bp | RT-PCR amplification |
| cfa -R | TCGCTGGACTGAGCAATCTG | | |
| LF_715-F | GGGAGAGCCAAAAGAAGCCT | 182 bp | RT-PCR amplification |
| LF_715-R | CAGTCCTTTGAGGCCGTAGG | | |
| 16S-F | ATGACCAGCCACACTGGAAC | 150 bp | RT-PCR amplification |
| 16S-R | CTTCCTCCCCGCTGAAAGTA | | |

**Table S4:** qRT-PCR validation of selected genes

| Gene number | Gene name | Product name | qRT-PCR fold change ($Log_2$) | RNA-seq fold change ($Log_2$) |
|---|---|---|---|---|
| LF82_0581 | *eutB* | ethanolamine ammonia-lyase heavy chain | 1.882 | 2.55 |
| LF82_1739 | *prpB* | Methylisocitrate lyase | 2.59 | 4.561 |
| LF82_0729 | *folX* | D-erythro-7,8-dihydroneopterin triphosphate epimerase | -0.802 | -0.262 |
| LF82_1619 | *pfkA* | 6-phosphofructokinase isozyme 1 | -2.43 | -1.725 |
| LF82_0286 | *cfa* | Cyclopropane-fatty-acyl-phospholipid synthase | -2.70 | -2.70 |
| LF82_p715 | LF82_p715 | Ornithine carbamoyltransferase chain I | 1.383 | 1.99 |

## 3. Commentaires de la publication n°2

### ✓ L'analyse RNA-seq a montré que les données obtenues étaient de bonne qualité

L'analyse bioinformatique du RNA-seq nous a permis de montrer que la qualité de nos séquences (ou « reads ») était satisfaisante, avec un nombre de reads alignés sur le génome important, une normalisation correcte des données et un dendrogramme retrouvant bien nos triplicats à l'intérieur de chaque condition de culture.

Par ailleurs, la qRT-PCR a permis de confirmer la bonne qualité de ces données, avec un coefficient de corrélation $r2$ de 0.9708 sur les 6 gènes testés.

### ✓ Les sels biliaires induisent chez les AIEC une surexpression des gènes impliqués dans les catabolismes et une répression des gènes impliqués dans les biosynthèses

Les sels biliaires ont un énorme impact sur l'expression des gènes, avec environ 40% de gènes différentiellement exprimés par rapport au milieu sans sels biliaires. Parmi eux, nous avons étudié l'expression des 722 gènes de la souche LF82 impliqués dans les voies métaboliques. Nous avons ainsi pu montrer que les sels biliaires activent les gènes impliqués dans les voies de dégradation, avec 37% de gènes surexprimés dans ces voies contre 16,8% de gènes surexprimés toutes voies confondues, avec plus particulièrement une activation du métabolisme des amines (comme par exemple l'éthanolamine), des carboxylates (comme le 1,2-propanediol, le propionate ou le citrate), des alcools et des métabolites secondaires, avec 47% de gènes surexprimés.

En parallèle de l'activation de certaines voies métaboliques, les sels biliaires éteignent des gènes impliqués dans les biosynthèses, avec 13.2% de gènes sous-exprimés dans ces voies contre 6,1% de gènes sous-exprimés au total des gènes impliqués dans les voies métaboliques.

### ✓ Les sels biliaires activent l'utilisation de l'éthanolamine par les AIEC

Au sein des gènes impliqués dans les différentes voies de dégradation, l'opéron *eut* qui code différentes protéines en charge de la dégradation de l'éthanolamine, est surexprimé en présence de sels biliaires. Différentes conditions de croissance de la souche LF82 ainsi que l'utilisation de mutants isogéniques ont montré que LF82 ne semble pas être capable d'utiliser l'éthanolamine comme source de carbone, mais peut l'utiliser comme source d'azote. Des courbes de compétition entre la LF82 et le *E. coli* K12 MG1655 ont par ailleurs montré une surexpression des gènes *eut* chez la souche LF82 comparativement à la souche MG1655, et que l'utilisation de l'éthanolamine permet à la souche LF82 de croître aux dépens des autres bactéries. A partir d'un test de compétition en modèle murin, nous avons pu confirmer que la dégradation de l'éthanolamine par la souche LF82 lui confère un avantage de colonisation au niveau de l'intestin de souris. L'éthanolamine est abondant dans le tube digestif car il est le principal constituant des membranes cellulaires où il est trouvé sous forme de phosphatidyléthanolamine. Il a déjà été démontré qu'il est utilisé par la souche d'EHEC O157:H7 EDL933 comme une source d'azote dans le contenu instestinal des bovins alors qu'il est peu métabolisé par la flore endogène (Bertin *et al.*, 2011). La dégradation de l'éthanolamine par les AIEC pourrait donc participer à la colonisation iléale des patients atteints de MC.

### ✓ Les sels biliaires activent la voie du méthylcitrate chez les AIEC

L'analyse RNA-seq a également montré une surexpression des gènes impliqués dans la voie du méthylcitrate. Pour vérifier l'activation de cette voie par la souche LF82, nous avons réalisé des courbes de croissance et ainsi démontré l'utilisation du 1,2 propanediol, mais pas du citrate, comme source de carbone par la souche AIEC LF82.

Ces données ont montré que l'éthanolamine et le 1,2-propanediol constituent respectivement, pour la souche LF82, une source d'azote et une source de carbone, toutes deux présentes au niveau intestinal.

### ✓ Les sels biliaires activent la formation de biofilm

L'activation des voies métaboliques (méthylcitrate, éthanolamine), en plus de constituer une source de nutriments et d'énergie pour la bactérie, peut conduire à la formation d'acétate et de pyruvate, qui pourraient eux-même participer à la formation d'acide colanique, constituant principal des biofilms.

Nous avons également montré une surexpression de gènes impliqués dans la biosynthèse de protéines curli et d'exopolysaccharides comme la cellulose et l'acide colanique, impliqués dans la formation de biofilm. La capacité de la souche LF82 à produire du biofilm en présence de sels biliaires a également été confirmée *in vitro*.

L'augmentation de la formation de biofilm en présence de sels biliaires pourrait être associée à une meilleure colonisation et la persistance des AIEC au niveau de l'iléon des patients atteints de MC.

### ✓ Les sels biliaires activent des facteurs de virulence des AIEC

Notre étude a révélé que la présence de sels biliaires a des répercussions importantes sur le transcriptome de la souche LF82, et notamment sur les gènes impliqués dans la virulence de la souche.

En effet, sur les 148 gènes connus impliqués dans la virulence de la souche LF82, 69 ont montré une surexpression en présence de sels biliaires. Parmi eux, environ la moitié des gènes appartenant aux ilots de pathogénicité de la souche LF82 sont surexprimés. Ces gènes sont responsables de la synthèse de deux SST6, un sidérophore et une capsule.

Les SST6 sont des systèmes permettant l'injection de composés bactériens dans des cellules cibles eucaryotes ou procaryotes cibles, afin de perturber leur homéostasie. Ces systèmes pourraient donc favoriser la colonisation des AIEC dans l'épithélium digestif aux dépens d'autres bactéries. Le sidérophore permet quant à lui la captation et l'utilisation du Fer par les bactéries. Enfin, la capsule est un autre facteur de virulence impliqué dans l'échappement des AIEC au système immunitaire (Whitfield, 2006) et dans la formation de biofilm (Schembri *et al.*, 2004).

Le gène codant l'invasine ibeA, impliquée dans l'invasion et la survie en macrophages des AIEC, ainsi que dans la réponse inflammatoire, est également surexprimé en présence de sels biliaires.

En réponse à la bile, la souche LF82 est donc capable de surexprimer un certain nombre de gènes, et notamment ceux codant des facteurs de virulence impliqués dans la colonisation de l'intestin par les bactéries AIEC, ce qui peut en partie expliquer le tropisme plutôt iléal de cette bactérie.

# IV. Discussion – Perspectives

Les lésions des patients atteints de MC sont anormalement colonisées par des souches d'*E. coli* appartenant au pathovar AIEC (Darfeuille-Michaud *et al.*, 2004). Cette colonisation met en jeu plusieurs facteurs de virulence des AIEC, tels que leur pouvoir d'adhésion et d'invasion des cellules épithéliales intestinales, de colonisation de la muqueuse et la formation de biofilm (Boudeau *et al.*, 1999; Barnich *et al.*, 2007; Martinez-Medina, Naves, *et al.*, 2009; Carvalho *et al.*, 2009).

L'apport de notre travail est d'avoir (i) décrit un nouveau facteur de virulence des AIEC impliqué dans la dégradation et le franchissement du mucus intestinal (ii) montré l'effet majeur des sels biliaires sur la sur-expression des facteurs de virulence des AIEC (iii) et mis en évidence des voies métaboliques qui pourraient favoriser l'implantation des AIEC au niveau intestinal.

## ✓ Purification et cristallisation de la protéine Vat-AIEC

Suite à ce travail, dans le but de mieux comprendre la fonction de la protéase Vat-AIEC, il pourrait s'avérer intéressant de déterminer sa structure tridimensionnelle. Cette protéine étant un autotransporteur, la production de cette protéine entière à grande échelle dans un système bactérien risque fort de conduire à la mort des cellules bactériennes, du fait d'un trop grand nombre de pores insérés dans la membrane. Nous avons cloné le gène codant la partie enzymatique de la protéine Vat-AIEC dans un plasmide sous contrôle d'un promoteur fort dans des bactéries *E. coli* BL21, dans le but d'hyperproduire ce fragment et le purifier. Suite à cette purification, une cristallisation de la partie enzymatique pourra être réalisée au laboratoire en servant de l'expérience acquise sur les β-lactamases. Une collaboration pourra être entreprise en cas de difficulté dans la réalisation d'une de ces étapes.

Les données recueillies devraient permettre de mieux comprendre le mécanisme hydrolytique de l'enzyme Vat-AIEC. L'analyse et la comparaison de cette structure avec celle de Hbp pourra nous aider à mettre en évidence les éléments structuraux de la poche catalytique

essentiels à la reconnaissance et à l'hydrolyse des mucines. Des mutants pourront par la suite être générés par mutagénèse dirigée pour confirmer l'importance de ces éléments. Ces données enrichiront les connaissances sur les mucinases et plus largement sur les sérines protéases.

### ✓ Recherche d'inhibiteurs de la protéine Vat-AIEC

L'étude des éléments structuraux essentiels à l'hydrolyse des mucines à partir de la structure obtenue de Vat-AIEC ou de celle déjà connue de Hbp devraient permettre le développement d'inhibiteurs intéressants. Un docking de petites molécules pourra être réalisé, c'est-à-dire la recherche d'inhibiteurs en fonction du calcul de leur affinité pour la poche catalytique de la protéase (Vat-AIEC ou Hbp). L'activité de ces inhibiteurs serait ensuite vérifiée *in vitro* par visualisation en Western Blot de la dégradation de la mucine, et sur des cultures de cellules HT29-16E productrices de mucus.

L'utilisation d'inhibiteurs pourrait ainsi bloquer la virulence des AIEC à un stade précoce de la maladie et ainsi limiter l'invasion de l'épithélium intestinal, dans la mesure où Vat-AIEC constitue une cible bactérienne facile à atteindre car présente à la surface des bactéries ou dans le milieu extracellulaire.

### ✓ Rôle de Vat-AIEC chez les *E. coli* uropathogènes

Nous avons déterminé la prévalence du gène *vat-AIEC* au sein des *E. coli*. Nous avons ainsi pu montrer que la mucinase Vat-AIEC est fréquemment retrouvée au niveau de l'espèce *E. coli* puisqu'elle est retrouvée chez 21 % des souches de la collection EcoR et 21 % des souches de *E. coli* isolées de selles de personnes saines. Elle est beaucoup plus fréquemment associée aux souches uropathogènes (50 % des souches), ce qui laisse supposer que cette mucinase est importante dans la physiopathologie de l'infection urinaire. Une autre mucinase, Pic, décrite chez *Escherichia coli*, est également retrouvée parmi les souches responsables de pyélonéphrites (30 % des souches) (Heimer *et al.*, 2004), renforcant le rôle potentiel de ces enzymes dans la virulence des UPEC. Les cellules épithéliales de la vessie étant recouvertes

d'une faible couche de mucines et de dérivés glycosaminés, on peut envisager que ces enzymes permettent la dégradation des mucines urinaires et favorisent l'adhésion des souches uropathogènes aux cellules épithéliales, comme cela avait été montré au niveau digestif avec les souches entéropathogènes. Des études pourraient être réalisées afin de tester l'hydrolyse des dérivés glycosaminés (héparane sulfate, chondroïtine sulfate) par les souches uropathogènes productrices de Vat-AIEC. De plus, il a été démontré que certaines mucinases chez *Shigella* et *E. coli* hydrolysaient des glycoprotéines (CD43, CD44, CD45 et CD93), dites « mucin-like », présentes à la surface des leucocytes ; cette hydrolyse affecte l'activation des polynucléaires neutrophiles et l'inflammation, suggérant un rôle immunomodulateur de ces protéines (Ruiz-Perez *et al.*, 2011; Ayala-Lujan *et al.*, 2014).

L'obtention du mutant délété du gène *vat-AIEC* dans la souche de référence uropathogène UPEC 536 permettra d'étudier dans un modèle d'infection urinaire murin l'implication de cette protéase dans la pathologie.

### ✓ Nouvelles propriétés pro-inflammatoires de la LF82

Les résultats de l'étude du transcriptome de la souche LF82 en présence de sels biliaires nous a permis de mieux comprendre pourquoi les AIEC ont un tropisme iléal préférentiel et quelles sont les voies métaboliques éventuellement mises en jeu pour faciliter l'implantation de ces bactéries dans l'intestin. Nous avons montré une surexpression des gènes impliqués dans l'utilisation de certains composés comme le propionate.

Le propionate est connu pour avoir des propriétés anti-inflammatoires (Tedelind *et al.*, 2007; Vinolo *et al.*, 2011). Il est produit par la flore endogène (*Bacteroïdetes* par exemple). La dégradation du propionate par la souche LF82 pourrait donc favoriser l'inflammation de l'intestin inflammé. Il serait donc intéressant de confirmer l'utilisation du propionate par la souche LF82 *in vitro,* puis dans un modèle murin en dosant ce dernier dans les selles et si possible dans des fragments d'intestin. L'utilisation de souris infectées ou non par la souche LF82, et traitées par DSS et propionate, devraient nous permettre d'établir si un réel rôle existe entre dégradation du propionate et exarcébation de l'inflammation.

## ✓ Impact de la cholestyramine sur la LF82

Nous avons, dans la deuxième publication, démontré que les acides biliaires jouent un rôle important dans l'activation des facteurs de virulence des AIEC notamment les ilots de pathogénicité et la formation de biofilm.

La cholestyramine est un traitement chélateur des acides biliaires, utilisé principalement dans les hypercholestérolémies, mais qui peut également être utilisé chez les patients atteints de MC comme traitement symptomatique lors des phases de poussée, quand les lésions iléales conduisent à une malabsorption des acides biliaires et provoquent une diarrhée. Cependant, l'analyse RNA-seq montre un effet délétère important des sels biliaires, avec une activation de la virulence de la souche LF82, pouvant conduire à une colonisation plus importante de cette dernière au niveau de l'épithélium intestinal des patients atteints de MC.

Il serait donc intéressant de regarder l'effet de la cholestyramine sur la colonisation de la LF82 dans un modèle murin. En effet, un tel effet pourrait peut-être conduire à une utilisation plus large de ce médicament.

## ✓ Impact de la mucine sur le transcriptome de la souche LF82

La mucine a déjà montré un effet sur l'expression de la mucinase Vat-AIEC, mais elle pourrait moduler l'expression d'autres gènes. De la même manière que nous avons étudié l'effet des sels biliaires sur l'expression des gènes de la souche LF82, il serait intéressant d'étudier le transcriptome de cette souche en présence de mucines intestinales.

# Références bibliographiques

# Bibliographie


Abraham, C., and Cho, J.H. (2009) Inflammatory Bowel Disease. *N Engl J Med* **361**: 2066–2078.

Abreu, M.T. (2010) Toll-like receptor signalling in the intestinal epithelium: how bacterial recognition shapes intestinal function. *Nat Rev Immunol* **10**: 131–144.

Agrawal, B., Gendler, S.J., and Longenecker, B.M. (1998) The biological role of mucins in cellular interactions and immune regulation: prospects for cancer immunotherapy. *Mol Med Today* **4**: 397–403.

Aguiar Vallim, T.Q. de, Tarling, E.J., and Edwards, P.A. (2013) Pleiotropic roles of bile acids in metabolism. *Cell Metab* **17**: 657–669.

Agus, A., Massier, S., Darfeuille-Michaud, A., Billard, E., and Barnich, N. (2014) Understanding Host-Adherent-Invasive *Escherichia coli* Interaction in Crohn's Disease: Opening Up New Therapeutic Strategies. *BioMed Res Int* **2014** http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4279263/. Accessed April 3, 2015.

Aksoy, N., and Akinci, O.F. (2004) Mucin Macromolecules in Normal, Adenomatous, and Carcinomatous Colon: Evidence for the Neotransformation. *Macromol Biosci* **4**: 483–496.

Al-Hasani, K., Henderson, I.R., Sakellaris, H., Rajakumar, K., Grant, T., Nataro, J.P., *et al.* (2000) The sigA Gene Which Is Borne on the she Pathogenicity Island of *Shigella flexneri* 2a Encodes an Exported Cytopathic Protease Involved in Intestinal Fluid Accumulation. *Infect Immun* **68**: 2457–2463.

Allen, A., Leonard, A.J., and Sellers, L.A. (1988) The mucus barrier. Its role in gastroduodenal mucosal protection. *J Clin Gastroenterol* **10 Suppl 1**: S93-98.

Amer, A.O., Byrne, B.G., and Swanson, M.S. (2005) Macrophages Rapidly Transfer Pathogens from Lipid Raft Vacuoles to Autophagosomes. *Autophagy* **1**: 53–58.

Amiot, A., and Peyrin-Biroulet, L. (2015) Current, new and future biological agents on the horizon for the treatment of inflammatory bowel diseases. *Ther Adv Gastroenterol* **8**: 66–82.

Ananthakrishnan, A.N., Khalili, H., Konijeti, G.G., Higuchi, L.M., Silva, P. de, Korzenik, J.R., *et al.* (2013) A prospective study of long-term intake of dietary fiber and risk of Crohn's disease and ulcerative colitis. *Gastroenterology* **145**: 970–977.

Anderson, D.M., and Frank, D.W. (2012) Five Mechanisms of Manipulation by Bacterial Effectors: A Ubiquitous Theme. *PLoS Pathog* **8** http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3426537/. Accessed May 8, 2016.

Andersson, R.E., Olaison, G., Tysk, C., and Ekbom, A. (2003) Appendectomy is followed by increased risk of Crohn's disease. *Gastroenterology* **124**: 40–46.

Andrianifahanana, M., Moniaux, N., and Batra, S.K. (2006) Regulation of mucin expression: Mechanistic aspects and implications for cancer and inflammatory diseases. *Biochim Biophys Acta BBA - Rev Cancer* **1765**: 189–222.

Ann M. Stock, Victoria L. Robinson, and Goudreau, and P.N. (2000) Two-Component Signal Transduction. *Annu Rev Biochem* **69**: 183–215.

Arnold, G.L., Beaves, M.R., Pryjdun, V.O., and Mook, W.J. (2002) Preliminary study of ciprofloxacin in active Crohn's disease. *Inflamm Bowel Dis* **8**: 10–15.

Arumugam, M., Raes, J., Pelletier, E., Le Paslier, D., Yamada, T., Mende, D.R., *et al.* (2011) Enterotypes of the human gut microbiome. *Nature* **473**: 174–180.

Ashgar, S.S.A., Oldfield, N.J., Wooldridge, K.G., Jones, M.A., Irving, G.J., Turner, D.P.J., and Ala'Aldeen, D.A.A. (2007) CapA, an autotransporter protein of *Campylobacter jejuni*, mediates association with human epithelial cells and colonization of the chicken gut. *J Bacteriol* **189**: 1856–1865.

Atuma, C., Strugala, V., Allen, A., and Holm, L. (2001) The adherent gastrointestinal mucus gel layer: thickness and physical state in vivo. *Am J Physiol Gastrointest Liver Physiol* **280**: G922-929.

Averboukh, F., Ziv, Y., Kariv, Y., Zmora, O., Dotan, I., Klausner, J.M., *et al.* (2011) Colorectal carcinoma in inflammatory bowel disease: a comparison between Crohn's and ulcerative colitis. *Colorectal Dis* **13**: 1230–1235.

Ayala-Lujan, J.L., Vijayakumar, V., Gong, M., Smith, R., Santiago, A.E., and Ruiz-Perez, F. (2014) Broad spectrum activity of a lectin-like bacterial serine protease family on human leukocytes. *PloS One* **9**: e107920.

Baert, F., Noman, M., Vermeire, S., Van Assche, G., D' Haens, G., Carbonez, A., and Rutgeerts, P. (2003) Influence of Immunogenicity on the Long-Term Efficacy of Infliximab in Crohn's Disease. *N Engl J Med* **348**: 601–608.

Baldus, S.E., and Hanisch, F.G. (2000) Biochemistry and pathological importance of mucin-associated antigens in gastrointestinal neoplasia. *Adv Cancer Res* **79**: 201–248.

Bandzar, S., Gupta, S., and Platt, M.O. (2013) Crohn's disease: A review of treatment options and current research. *Cell Immunol* **286**: 45–52.

Baranova, N., and Nikaido, H. (2002) The BaeSR Two-Component Regulatory System Activates Transcription of the yegMNOB (mdtABCD) Transporter Gene Cluster in *Escherichia coli* and Increases Its Resistance to Novobiocin and Deoxycholate. *J Bacteriol* **184**: 4168–4176.

Bardakchian, E.A., Monakhova, E.V., Kharlanova, N.G., Sayamov, S.R., and Pisanov, R.V. (2008) Ultrastructural changes in the small intestine of suckling mice, caused by *Vibrio cholerae* hemagglutinin/protease. *Bull Exp Biol Med* **145**: 490–494.

Barnard, T.J., Dautin, N., Lukacik, P., Bernstein, H.D., and Buchanan, S.K. (2007) Autotransporter structure reveals intra-barrel cleavage followed by conformational changes. *Nat Struct Mol Biol* **14**: 1214–1220.

Barnes, M.J., and Powrie, F. (2009) Regulatory T cells reinforce intestinal homeostasis. *Immunity* **31**: 401–411.

Barnhart, M.M., and Chapman, M.R. (2006) Curli Biogenesis and Function. *Annu Rev Microbiol* **60**: 131–147.

Barnich, N., Boudeau, J., Claret, L., and Darfeuille-Michaud, A. (2003) Regulatory and functional co-operation of flagella and type 1 pili in adhesive and invasive abilities of AIEC strain LF82 isolated from a patient with Crohn's disease. *Mol Microbiol* **48**: 781–794.

Barnich, N., Carvalho, F.A., Glasser, A.-L., Darcha, C., Jantscheff, P., Allez, M., *et al.* (2007) CEACAM6 acts as a receptor for adherent-invasive *E. coli*, supporting ileal mucosa colonization in Crohn disease. *J Clin Invest* **117**: 1566–1574.

Barnich, N., and Darfeuille-Michaud, A. (2007a) Adherent-invasive *Escherichia coli* and Crohn's disease. *Curr Opin Gastroenterol* **23**: 16–20.

Barnich, N., and Darfeuille-Michaud, A. (2007b) Role of bacteria in the etiopathogenesis of inflammatory bowel disease. *World J Gastroenterol WJG* **13**: 5571–5576.

Baron, S.F., and Hylemon, P.B. (1995) Expression of the bile acid-inducible NADH:flavin oxidoreductase gene of *Eubacterium* sp. VPI 12708 in *Escherichia coli*. *Biochim Biophys Acta* **1249**: 145–154.

Barreau, F., Meinzer, U., Chareyre, F., Berrebi, D., Niwa-Kawakita, M., Dussaillant, M., *et al.* (2007) CARD15/NOD2 Is Required for Peyer's Patches Homeostasis in Mice. *PLoS ONE* **2** http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1885825/. Accessed March 4, 2016.

Barrett, J.C., Hansoul, S., Nicolae, D.L., Cho, J.H., Duerr, R.H., Rioux, J.D., *et al.* (2008) Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nat Genet* **40**: 955–962.

Baumgart, D.C., and Sandborn, W.J. (2012) Crohn's disease. *The Lancet* **380**: 1590–1605.

Baumgart, M., Dogan, B., Rishniw, M., Weitzman, G., Bosworth, B., Yantiss, R., *et al.* (2007) Culture independent analysis of ileal mucosa reveals a selective increase in invasive *Escherichia coli* of novel phylogeny relative to depletion of Clostridiales in Crohn's disease involving the ileum. *ISME J* **1**: 403–418.

Bäumler, A.J., Tsolis, R.M., and Heffron, F. (1996) The lpf fimbrial operon mediates adhesion of *Salmonella typhimurium* to murine Peyer's patches. *Proc Natl Acad Sci U S A* **93**: 279–283.

Bearson, B.L., Wilson, L., and Foster, J.W. (1998) A Low pH-Inducible, PhoPQ-Dependent Acid Tolerance Response Protects *Salmonella typhimurium* against Inorganic Acid Stress. *J Bacteriol* **180**: 2409–2417.

Beaven, S.W., and Abreu, M.T. (2004) Biomarkers in inflammatory bowel disease. *Curr Opin Gastroenterol* **20**: 318–327.

Begley, M., Gahan, C.G.M., and Hill, C. (2005) The interaction between bacteria and bile. *FEMS Microbiol Rev* **29**: 625–651.

Begley, M., Kerr, C., and Hill, C. (2009) Exposure to bile influences biofilm formation by *Listeria monocytogenes*. *Gut Pathog* **1**: 11.

Begue, B., Dumant, C., Bambou, J. c., Beaulieu, J. f., Chamaillard, M., Hugot, J. p., *et al.* (2006) Microbial induction of CARD15 expression in intestinal epithelial cells via toll-like receptor 5 triggers an antibacterial response loop. *J Cell Physiol* **209**: 241–252.

Behari, J., Stagon, L., and Calderwood, S.B. (2001) pepA, a Gene Mediating pH Regulation of Virulence Genes in *Vibrio cholerae*. *J Bacteriol* **183**: 178–188.

Belley, A., and Chadee, K. (1999) Prostaglandin E2 stimulates rat and human colonic mucin exocytosis via the EP4 receptor. *Gastroenterology* **117**: 1352–1362.

Beloin, C., Roux, A., and Ghigo, J.-M. (2008) *Escherichia coli* biofilms. *Curr Top Microbiol Immunol* **322**: 249–289.

Benitez, J.-M., Meuwis, M.-A., Reenaers, C., Kemseke, C.V., Meunier, P., and Louis, E. (2013) Role of endoscopy, cross-sectional imaging and biomarkers in Crohn's disease monitoring. *Gut* **62**: 1806–1816.

Benjamin, M.M., and Datta, A.R. (1995) Acid tolerance of enterohemorrhagic *Escherichia coli*. *Appl Environ Microbiol* **61**: 1669–1672.

Benjelloun-Touimi, Z., Sansonetti, P.J., and Parsot, C. (1995) SepA, the major extracellular protein of *Shigella flexneri*: autonomous secretion and involvement in tissue invasion. *Mol Microbiol* **17**: 123–135.

Benson, A., Pifer, R., Behrendt, C.L., Hooper, L.V., and Yarovinsky, F. (2009) Gut commensal bacteria direct a protective immune response against the human pathogen *Toxoplasma gondii*. *Cell Host Microbe* **6**: 187–196.

Benz, I., and Schmidt, M.A. (2001) Glycosylation with heptose residues mediated by the aah gene product is essential for adherence of the AIDA-I adhesin. *Mol Microbiol* **40**: 1403–1413.

Berg, R.D. (1996) The indigenous gastrointestinal microflora. *Trends Microbiol* **4**: 430–435.

Bergthorsson, U., and Ochman, H. (1998) Distribution of chromosome length variation in natural isolates of *Escherichia coli*. *Mol Biol Evol* **15**: 6–16.

Berks, B.C., Sargent, F., and Palmer, T. (2000) The Tat protein export pathway. *Mol Microbiol* **35**: 260–274.

Bernalier-Donadille, A. (2010) Fermentative metabolism by the human gut microbiota. *Gastroentérologie Clin Biol* **34**, **Supplement 1**: S16–S22.

Bernstein, L.H., Frank, M.S., Brandt, L.J., and Boley, S.J. (1980) Healing of perineal Crohn's disease with metronidazole. *Gastroenterology* **79**: 599.

Bertin, Y., Girardeau, J.P., Chaucheyras-Durand, F., Lyan, B., Pujos-Guillot, E., Harel, J., and Martin, C. (2011) Enterohaemorrhagic *Escherichia coli* gains a competitive advantage by using ethanolamine as a nitrogen source in the bovine intestinal content. *Environ Microbiol* **13**: 365–377.

Bertók, L. (2004) Bile acids in physico-chemical host defence. *Pathophysiology* **11**: 139–145.

Best, W.R. (2006) Predicting the Crohn's disease activity index from the Harvey-Bradshaw Index. *Inflamm Bowel Dis* **12**: 304–310.

Best, W.R., Becktel, J.M., Singleton, J.W., and Kern, F. (1976) Development of a Crohn's disease activity index. National Cooperative Crohn's Disease Study. *Gastroenterology* **70**: 439–444.

Bettelli, E., and Kuchroo, V.K. (2005) IL-12– and IL-23–induced T helper cell subsets. *J Exp Med* **201**: 169–171.

Binder, V., and Orholm, M. (1996) Familial occurrence and inheritance studies in inflammatory bowel disease. *Neth J Med* **48**: 53–56.

Bingen, E., Picard, B., Brahimi, N., Mathy, S., Desjardins, P., Elion, J., and Denamur, E. (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.

Björkstén, B., Sepp, E., Julge, K., Voor, T., and Mikelsaar, M. (2001) Allergy development and the intestinal microflora during the first year of life. *J Allergy Clin Immunol* **108**: 516–520.

Blichfeldt, P., Blomhoff, J.P., Myhre, E., and Gjone, E. (1978) Metronidazole in Crohn's Disease. *Scand J Gastroenterol* **13**: 123–127.

Bodger, K., Halfvarson, J., Dodson, A.R., Campbell, F., Wilson, S., Lee, R., *et al.* (2006) Altered colonic glycoprotein expression in unaffected monozygotic twins of inflammatory bowel disease patients. *Gut* **55**: 973–977.

Booth, I.R. (1985) Regulation of cytoplasmic pH in bacteria. *Microbiol Rev* **49**: 359.

Borody, T.J., and Khoruts, A. (2012) Fecal microbiota transplantation and emerging applications. *Nat Rev Gastroenterol Hepatol* **9**: 88–96.

Borzutzky, A., Fried, A., Chou, J., Bonilla, F.A., Kim, S., and Dedeoglu, F. (2010) NOD2-associated diseases: Bridging innate immunity and autoinflammation. *Clin Immunol* **134**: 251–261.

Bossuyt, X. (2006) Serologic Markers in Inflammatory Bowel Disease. *Clin Chem* **52**: 171–181.

Boudeau, J., Barnich, N., and Darfeuille-Michaud, A. (2001) Type 1 pili-mediated adherence of *Escherichia coli* strain LF82 isolated from Crohn's disease is involved in bacterial invasion of intestinal epithelial cells. *Mol Microbiol* **39**: 1272–1284.

Boudeau, J., Glasser, A.-L., Masseret, E., Joly, B., and Darfeuille-Michaud, A. (1999) Invasive Ability of an *Escherichia coli* Strain Isolated from the Ileal Mucosa of a Patient with Crohn's Disease. *Infect Immun* **67**: 4499.

Bouma, G., and Strober, W. (2003) The immunological and genetic basis of inflammatory bowel disease. *Nat Rev Immunol* **3**: 521–533.

Bower, J.M., Eto, D.S., and Mulvey, M.A. (2005) Covert Operations of Uropathogenic *Escherichia coli* within the Urinary Tract. *Traffic Cph Den* **6**: 18–31.

Brain, O., Owens, B.M.J., Pichulik, T., Allan, P., Khatamzas, E., Leslie, A., *et al.* (2013) The Intracellular Sensor NOD2 Induces MicroRNA-29 Expression in Human Dendritic Cells to Limit IL-23 Release. *Immunity* **39**: 521–536.

Brand, S., Staudinger, T., Schnitzler, F., Pfennig, S., Hofbauer, K., Dambacher, J., *et al.* (2005) The role of Toll-like receptor 4 Asp299Gly and Thr399Ile polymorphisms and CARD15/NOD2 mutations in the susceptibility and phenotype of Crohn's disease. *Inflamm Bowel Dis* **11**: 645–652.

Breese, E.J., Michie, C.A., Nicholls, S.W., Murch, S.H., Williams, C.B., Domizio, P., *et al.* (1994) Tumor necrosis factor alpha-producing cells in the intestinal mucosa of children with inflammatory bowel disease. *Gastroenterology* **106**: 1455–1466.

Brest, P., Corcelle, E.A., Cesaro, A., Chargui, A., Belaïd, A., Klionsky, D.J., *et al.* (2010) Autophagy and Crohn's Disease: At the Crossroads of Infection, Inflammation, Immunity, and Cancer. *Curr Mol Med* **10**: 486–502.

Brest, P., Lapaquette, P., Souidi, M., Lebrigand, K., Cesaro, A., Vouret-Craviari, V., *et al.* (2011) A synonymous variant in IRGM alters a binding site for miR-196 and causes deregulation of IRGM-dependent xenophagy in Crohn's disease. *Nat Genet* **43**: 242–245.

Bringer, M.-A., Billard, E., Glasser, A.-L., Colombel, J.-F., and Darfeuille-Michaud, A. (2012) Replication of Crohn's disease-associated AIEC within macrophages is dependent on TNF-α secretion. *Lab Invest* **92**: 411–419.

Brooke, B., Hoffmann, D.C., and Swarbrick, E.T. (1969) Azathioprine for Crohn's disease. *The Lancet* **294**: 612–614.

Brooks, J., and Watson, A.J.M. (2015) The Enteric Virome in Inflammatory Bowel Disease. *Gastroenterology* **149**: 1120–1121.

Brunder, W., Schmidt, H., and Karch, H. (1997) EspP, a novel extracellular serine protease of enterohaemorrhagic *Escherichia coli* O157:H7 cleaves human coagulation factor V. *Mol Microbiol* **24**: 767–778.

Bry, L., Falk, P., Huttner, K., Ouellette, A., Midtvedt, T., and Gordon, J.I. (1994) Paneth cell differentiation in the developing intestine of normal and transgenic mice. *Proc Natl Acad Sci* **91**: 10335–10339.

Buffie, C.G., Bucci, V., Stein, R.R., McKenney, P.T., Ling, L., Gobourne, A., *et al.* (2014) Precision microbiome reconstitution restores bile acid mediated resistance to *Clostridium difficile*. *Nature* **517**: 205–208.

Buisine, M.P., Desreumaux, P., Leteurtre, E., Copin, M.C., Colombel, J.F., Porchet, N., and Aubert, J.P. (2001) Mucin gene expression in intestinal epithelial cells in Crohn's disease. *Gut* **49**: 544–551.

Bulet, P., Stöcklin, R., and Menin, L. (2004) Anti-microbial peptides: from invertebrates to vertebrates. *Immunol Rev* **198**: 169–184.

Burger, D., and Travis, S. (2011) Conventional Medical Management of Inflammatory Bowel Disease. *Gastroenterology* **140**: 1827–1837.e2.

Burisch, J., Pedersen, N., Cukovic-Cavka, S., Turk, N., Kaimakliotis, I., Duricova, D., *et al.* (2014) Environmental factors in a population-based inception cohort of inflammatory bowel disease patients in Europe — An ECCO-EpiCom study. *J Crohns Colitis* **8**: 607–616.

Cadwell, K., Liu, J., Brown, S.L., Miyoshi, H., Loh, J., Lennerz, J., *et al.* (2008) A unique role for autophagy and Atg16L1 in Paneth cells in murine and human intestine. *Nature* **456**: 259–263.

Calkins, B.M. (1989) A meta-analysis of the role of smoking in inflammatory bowel disease. *Dig Dis Sci* **34**: 1841–1854.

Cambronne, E.D., and Roy, C.R. (2006) Recognition and Delivery of Effector Proteins into Eukaryotic Cells by Bacterial Secretion Systems. *Traffic* **7**: 929–939.

Cario, E. (2005) Bacterial interactions with cells of the intestinal mucosa: Toll-Like Receptors and NOD2. *Gut* **54**: 1182–1193.

Cario, E., and Podolsky, D.K. (2000) Differential Alteration in Intestinal Epithelial Cell Expression of Toll-Like Receptor 3 (TLR3) and TLR4 in Inflammatory Bowel Disease. *Infect Immun* **68**: 7010–7017.

Carlos, C., Pires, M.M., Stoppe, N.C., Hachich, E.M., Sato, M.I.Z., Gomes, T.A.T., *et al.* (2010) *Escherichia coli* phylogenetic group determination and its application in the identification of the major animal source of fecal contamination. *BMC Microbiol* **10**: 161.

Carvalho, F.A., Barnich, N., Sauvanet, P., Darcha, C., Gelot, A., and Darfeuille-Michaud, A. (2008) Crohn's disease-associated *Escherichia coli* LF82 aggravates colitis in injured mouse colon via signaling by flagellin. *Inflamm Bowel Dis* **14**: 1051–1060.

Carvalho, F.A., Barnich, N., Sivignon, A., Darcha, C., Chan, C.H.F., Stanners, C.P., and Darfeuille-Michaud, A. (2009) Crohn's disease adherent-invasive *Escherichia coli* colonize and induce strong gut inflammation in transgenic mice expressing human CEACAM. *J Exp Med* **206**: 2179–2189.

Castanie-Cornet, M.-P., Penfound, T.A., Smith, D., Elliott, J.F., and Foster, J.W. (1999) Control of Acid Resistance in *Escherichia coli*. *J Bacteriol* **181**: 3525–3535.

Cattoir, V., Narasimhan, G., Skurnik, D., Aschard, H., Roux, D., Ramphal, R., *et al.* (2013) Transcriptional response of mucoid *Pseudomonas aeruginosa* to human respiratory mucus. *mBio* **3**: e00410-412.

CDU-HGE (2014) *Les fondamentaux de la pathologie digestive*. Elsevier-Masson, . http://www.snfge.org/content/les-fondamentaux-de-la-pathologie-digestive. Accessed July 23, 2015.

Celli, J.P., Turner, B.S., Afdhal, N.H., Keates, S., Ghiran, I., Kelly, C.P., *et al.* (2009) *Helicobacter pylori* moves through mucus by reducing mucin viscoelasticity. *Proc Natl Acad Sci U S A* **106**: 14321–14326.

Cerf-Bensussan, N., and Gaboriau-Routhiau, V. (2010) The immune system and the gut microbiota: friends or foes? *Nat Rev Immunol* **10**: 735–744.

Chagneau, C., Heyde, M., Alonso, S., Portalier, R., and Laloi, P. (2001) External-pH-Dependent Expression of the Maltose Regulon and ompF Gene in *Escherichia coli* Is Affected by the Level of Glycerol Kinase, Encoded by glpK. *J Bacteriol* **183**: 5675–5683.

Chan, C.H.F., and Stanners, C.P. (2004) Novel mouse model for carcinoembryonic antigen-based therapy. *Mol Ther J Am Soc Gene Ther* **9**: 775–785.

Charrier, L., and Merlin, D. (2006) The oligopeptide transporter hPepT1: gateway to the innate immune response. *Lab Invest* **86**: 538–546.

Chassaing, B., and Darfeuille–Michaud, A. (2011) The Commensal Microbiota and Enteropathogens in the Pathogenesis of Inflammatory Bowel Diseases. *Gastroenterology* **140**: 1720–1728.e3.

Chassaing, B., and Darfeuille-Michaud, A. (2013) The σE pathway is involved in biofilm formation by Crohn's disease-associated adherent-invasive *Escherichia coli*. *J Bacteriol* **195**: 76–84.

Chassaing, B., Etienne-Mesmin, L., Bonnet, R., and Darfeuille-Michaud, A. (2013) Bile salts induce long polar fimbriae expression favouring Crohn's disease-associated adherent-invasive *Escherichia coli* interaction with Peyer's patches. *Environ Microbiol* **15**: 355–371.

Chassaing, B., Koren, O., Carvalho, F.A., Ley, R.E., and Gewirtz, A.T. (2014) AIEC Pathobiont Instigates Chronic colitis in Susceptible Hosts by Altering Microbiota composition. *Gut* **63**: 1069–1080.

Chassaing, B., Rolhion, N., Vallée, A. de, Salim, S.Y., Prorok-Hamon, M., Neut, C., *et al.* (2011) Crohn disease–associated adherent-invasive *E. coli* bacteria target mouse and human Peyer's patches via long polar fimbriae. *J Clin Invest* **121**: 966–975.

Chatterjee, A., Chaudhuri, S., Saha, G., Gupta, S., and Chowdhury, R. (2004) Effect of Bile on the Cell Surface Permeability Barrier and Efflux System of *Vibrio cholerae*. *J Bacteriol* **186**: 6809–6814.

Chaveroche, M.K., Ghigo, J.M., and Enfert, C. d' (2000) A rapid method for efficient gene replacement in the filamentous fungus *Aspergillus nidulans*. *Nucleic Acids Res* **28**: E97.

Chen, W.-L., Oliver, J.D., and Wong, H. (2010) Adaptation of *Vibrio vulnificus* and an rpoS mutant to bile salts. *Int J Food Microbiol* **140**: 232–238.

Chen, Y., Zhao, Y.H., Kalaslavadi, T.B., Hamati, E., Nehrke, K., Le, A.D., *et al.* (2004) Genome-wide search and identification of a novel gel-forming mucin MUC19/Muc19 in glandular tissues. *Am J Respir Cell Mol Biol* **30**: 155–165.

Cheville, A.M., Arnold, K.W., Buchrieser, C., Cheng, C.M., and Kaspar, C.W. (1996) rpoS regulation of acid, heat, and salt tolerance in *Escherichia coli* O157:H7. *Appl Environ Microbiol* **62**: 1822–1824.

Chiba, M., Komatsu, M., Iizuka, M., Masamune, O., Hoshina, S., and Kong, M. (1998) Microbiology of the intestinal lymph follicle: a clue to elucidate causative microbial agent(s) in Crohn's disease. *Med Hypotheses* **51**: 421–427.

Chuang, A.Y., Chuang, J.C., Zhai, Z., Wu, F., and Kwon, J.H. (2014) NOD2 expression is regulated by microRNAs in colonic epithelial HCT116 cells. *Inflamm Bowel Dis* **20**: 126–135.

Ciccocioppo, R., Bernardo, M.E., Sgarella, A., Maccario, R., Avanzini, M.A., Ubezio, C., *et al.* (2011) Autologous bone marrow-derived mesenchymal stromal cells in the treatment of fistulising Crohn's disease. *Gut* **60**: 788–798.

Cieza, R.J., Hu, J., Ross, B.N., Sbrana, E., and Torres, A.G. (2015) The IbeA Invasin of Adherent-Invasive *Escherichia coli* Mediates Interaction with Intestinal Epithelia and Macrophages. *Infect Immun* **83**: 1904–1918.

Claes, A.-K., Zhou, J.Y., and Philpott, D.J. (2015) NOD-Like Receptors: Guardians of Intestinal Mucosal Barriers. *Physiology* **30**: 241–250.

Clements, A., Young, J.C., Constantinou, N., and Frankel, G. (2012) Infection strategies of enteric pathogenic *Escherichia coli*. *Gut Microbes* **3**: 71–87.

Clerici, M., Cassinotti, A., Onida, F., Trabattoni, D., Annaloro, C., Della Volpe, A., *et al.* (2011) Immunomodulatory effects of unselected haematopoietic stem cells autotransplantation in refractory Crohn's disease. *Dig Liver Dis* **43**: 946–952.

Clermont, O., Christenson, J.K., Denamur, E., and Gordon, D.M. (2013) The Clermont *Escherichia coli* phylo-typing method revisited: improvement of specificity and detection of new phylo-groups. *Environ Microbiol Rep* **5**: 58–65.

Clevers, H. (2013) The Intestinal Crypt, A Prototype Stem Cell Compartment. *Cell* **154**: 274–284.

Clomburg, J.M., and Gonzalez, R. (2010) Biofuel production in *Escherichia coli*. *Appl Microbiol Biotechnol* **86**: 419–434.

Coleman, R., and Holdsworth, G. (1976) The release of membrane components prior to haemolysis during extraction of intact eryghrocytes with bile salts. *Biochim Biophys Acta* **426**: 776–780.

Coleman, R., Lowe, P.J., and Billington, D. (1980) Membrane lipid composition and susceptibility to bile salt damage. *Biochim Biophys Acta BBA - Biomembr* **599**: 294–300.

Collins, S.M., Surette, M., and Bercik, P. (2012) The interplay between the intestinal microbiota and the brain. *Nat Rev Microbiol* **10**: 735–742.

Colombel, J.-F., Cortot, A., and Kruiningen, H.J. van (2001) Antibiotics in Crohn's disease. *Gut* **48**: 647–647.

Colombel, J.-F., Lémann, M., Cassagnou, M., Bouhnik, Y., Duclos, B., Dupas, J.-L., *et al.* (1999) A controlled trial comparing ciprofloxacin with mesalazine for the treatment of active Crohn's disease. *Am J Gastroenterol* **94**: 674–678.

Colombel, J.F., Sandborn, W.J., Reinisch, W., Mantzaris, G.J., Kornbluth, A., Rachmilewitz, D., *et al.* (2010) Infliximab, Azathioprine, or Combination Therapy for Crohn's Disease. *N Engl J Med* **362**: 1383–1395.

Conte, M.P., Longhi, C., Marazzato, M., Conte, A.L., Aleandri, M., Lepanto, M.S., *et al.* (2014) Adherent-invasive *Escherichia coli* (AIEC) in pediatric Crohn's disease patients: phenotypic and genetic pathogenic features. *BMC Res Notes* **7** http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4210564/. Accessed July 3, 2016.

Conte, M.P., Schippa, S., Zamboni, I., Penta, M., Chiarini, F., Seganti, L., *et al.* (2006) Gut-associated bacterial microbiota in paediatric patients with inflammatory bowel disease. *Gut* **55**: 1760–1767.

Cooney, R., Baker, J., Brain, O., Danis, B., Pichulik, T., Allan, P., *et al.* (2010) NOD2 stimulation induces autophagy in dendritic cells influencing bacterial handling and antigen presentation. *Nat Med* **16**: 90–97.

Corfield, A.P., Wagner, S.A., Clamp, J.R., Kriaris, M.S., and Hoskins, L.C. (1992) Mucin degradation in the human colon: production of sialidase, sialate O-acetylesterase, N-acetylneuraminate lyase, arylesterase, and glycosulfatase activities by strains of fecal bacteria. *Infect Immun* **60**: 3971–3978.

Cosnes, J., Gower–Rousseau, C., Seksik, P., and Cortot, A. (2011) Epidemiology and Natural History of Inflammatory Bowel Diseases. *Gastroenterology* **140**: 1785–1794.e4.

Cotter, S.E., Surana, N.K., Grass, S., and St Geme, J.W. (2006) Trimeric autotransporters require trimerization of the passenger domain for stability and adhesive activity. *J Bacteriol* **188**: 5400–5407.

Coutte, L., Willery, E., Antoine, R., Drobecq, H., Locht, C., and Jacob-Dubuisson, F. (2003) Surface anchoring of bacterial subtilisin important for maturation function. *Mol Microbiol* **49**: 529–539.

Crawford, R.W., Gibson, D.L., Kay, W.W., and Gunn, J.S. (2008) Identification of a Bile-Induced Exopolysaccharide Required for *Salmonella* Biofilm Formation on Gallstone Surfaces. *Infect Immun* **76**: 5341–5349.

Crohn, B.B., Ginzburg, L., and Oppenheimer, G.D. (2000) Regional ileitis: a pathologic and clinical entity. 1932. *Mt Sinai J Med N Y* **67**: 263–268.

Crosnier, C., Stamataki, D., and Lewis, J. (2006) Organizing cell renewal in the intestine: stem cells, signals and combinatorial control. *Nat Rev Genet* **7**: 349–359.

Cross, C., Halliwell, B., and Allen, A. (1984) Antioxidant protection: a function of tracheobronchial and gastrointestinal mucus. *The Lancet* **323**: 1328–1330.

Croxen, M.A., and Finlay, B.B. (2010) Molecular mechanisms of *Escherichia coli* pathogenicity. *Nat Rev Microbiol* **8**: 26–38.

Croxen, M.A., Law, R.J., Scholz, R., Keeney, K.M., Wlodarska, M., and Finlay, B.B. (2013) Recent Advances in Understanding Enteric Pathogenic *Escherichia coli*. *Clin Microbiol Rev* **26**: 822–880.

Cui, B., Feng, Q., Wang, H., Wang, M., Peng, Z., Li, P., *et al.* (2015) Fecal microbiota transplantation through mid-gut for refractory Crohn's disease: Safety, feasibility, and efficacy trial results. *J Gastroenterol Hepatol* **30**: 51–58.

Culp, D.J., Latchney, L.R., Fallon, M.A., Denny, P.A., Denny, P.C., Couwenhoven, R.I., and Chuang, S. (2004) The gene encoding mouse Muc19: cDNA, genomic organization and relationship to Smgc. *Physiol Genomics* **19**: 303–318.

Cummings, J.H. (1981) Short chain fatty acids in the human colon. *Gut* **22**: 763–779.

Cummings, J.H., Pomare, E.W., Branch, W.J., Naylor, C.P., and Macfarlane, G.T. (1987) Short chain fatty acids in human large intestine, portal, hepatic and venous blood. *Gut* **28**: 1221–1227.

Curkovic, I., Egbring, M., and Kullak-Ublick, G.A. (2013) Risks of Inflammatory Bowel Disease Treatment with Glucocorticosteroids and Aminosalicylates. *Dig Dis* **31**: 368–373.

D'Aldebert, E., Biyeyeme Bi Mve, M., Mergey, M., Wendum, D., Firrincieli, D., Coilly, A., et al. (2009) Bile Salts Control the Antimicrobial Peptide Cathelicidin Through Nuclear Receptors in the Human Biliary Epithelium. Gastroenterology 136: 1435–1443.

Dalmasso, G., Cougnoux, A., Delmas, J., Darfeuille-Michaud, A., and Bonnet, R. (2014) The bacterial genotoxin colibactin promotes colon tumor growth by modifying the tumor microenvironment. *Gut Microbes* **5**: 675–680.

Dalmasso, G., Nguyen, H.T.T., Charrier-Hisamuddin, L., Yan, Y., Laroui, H., Demoulin, B., *et al.* (2010) PepT1 mediates transport of the proinflammatory bacterial tripeptide l-Ala-γ-d-Glu-meso-DAP in intestinal epithelial cells. *Am J Physiol - Gastrointest Liver Physiol* **299**: G687–G696.

Dalmasso, G., Nguyen, H.T.T., Ingersoll, S.A., Ayyadurai, S., Laroui, H., Charania, M.A., *et al.* (2011) The PepT1–NOD2 Signaling Pathway Aggravates Induced Colitis in Mice. *Gastroenterology* **141**: 1334–1345.

Danese, P.N., Pratt, L.A., and Kolter, R. (2000) Exopolysaccharide Production Is Required for Development of *Escherichia coli* K-12 Biofilm Architecture. *J Bacteriol* **182**: 3593–3596.

Darfeuille-Michaud, A., Boudeau, J., Bulois, P., Neut, C., Glasser, A.-L., Barnich, N., *et al.* (2004) High prevalence of adherent-invasive *Escherichia coli* associated with ileal mucosa in Crohn's disease. *Gastroenterology* **127**: 412–421.

Darfeuille-Michaud, A., Neut, C., Barnich, N., Lederman, E., Di Martino, P., Desreumaux, P., *et al.* (1998) Presence of adherent *Escherichia coli* strains in ileal mucosa of patients with Crohn's disease. *Gastroenterology* **115**: 1405–1413.

Daszak, P., Purcell, M., Lewin, J., Dhillon, A.P., Pounder, R.E., and Wakefield, A.J. (1997) Detection and comparative analysis of persistent measles virus infection in Crohn's disease by immunogold electron microscopy. *J Clin Pathol* **50**: 299–304.

Datsenko, K.A., and Wanner, B.L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* **97**: 6640–6645.

Dautin, N. (2010) Serine Protease Autotransporters of *Enterobacteriaceae* (SPATEs): Biogenesis and Function. *Toxins* **2**: 1179–1206.

Dautin, N., Barnard, T.J., Anderson, D.E., and Bernstein, H.D. (2007) Cleavage of a bacterial autotransporter by an evolutionarily convergent autocatalytic mechanism. *EMBO J* **26**: 1942–1952.

Delmas, J., Dalmasso, G., and Bonnet, R. (2015) *Escherichia coli*: The Good, the Bad and the Ugly. *Clin Microbiol Open Access* **4** http://www.esciencecentral.org/journals/escherichia-coli-the-good-the-bad-and-the-ugly-2327-5073-1000195.php?aid=50935. Accessed June 18, 2015.

Deretic, V. (2005) Autophagy in innate and adaptive immunity. *Trends Immunol* **26**: 523–528.

Derrien, M., Passel, M.W. van, Bovenkamp, J.H. van de, Schipper, R.G., Vos, W.M. de, and Dekker, J. (2010) Mucin-bacterial interactions in the human oral cavity and digestive tract. *Gut Microbes* **1**: 254–268.

Desilets, M., Deng, X., Rao, C., Ensminger, A.W., Krause, D.O., Sherman, P.M., and Gray-Owen, S.D. (2016) Genome-based Definition of an Inflammatory Bowel Disease-associated Adherent-Invasive *Escherichia coli* Pathovar. *Inflamm Bowel Dis* **22**: 1–12.

Desreumaux, P., Brandt, E., Gambiez, L., Emilie, D., Geboes, K., Klein, O., *et al.* (1997) Distinct cytokine patterns in early and chronic ileal lesions of Crohn's disease. *Gastroenterology* **113**: 118–126.

Desseyn, J.-L., Aubert, J.-P., Seuningen, I.V., Porchet, N., and Laine, A. (1997) Genomic Organization of the 3′ Region of the Human Mucin GeneMUC5B. *J Biol Chem* **272**: 16873–16883.

Desvaux, M., Hébraud, M., Talon, R., and Henderson, I.R. (2009) Secretion and subcellular localizations of bacterial proteins: a semantic awareness issue. *Trends Microbiol* **17**: 139–145.

Desvaux, M., Parham, N.J., and Henderson, I.R. (2004) The autotransporter secretion system. *Res Microbiol* **155**: 53–60.

Desvaux, M., Parham, N.J., and Henderson, I.R. (2004) Type V protein secretion: simplicity gone awry? *Curr Issues Mol Biol* **6**: 111–124.

D'Haens, G., Van Deventer, S., Van Hogezand, R., Chalmers, D., Kothe, C., Baert, F., *et al.* (1999) Endoscopic and histological healing with infliximab anti–tumor necrosis factor antibodies in Crohn's disease: A European multicenter trial. *Gastroenterology* **116**: 1029–1034.

Dharmani, P., Srivastava, V., Kissoon-Singh, V., and Chadee, K. (2009) Role of Intestinal Mucins in Innate Host Defense Mechanisms against Pathogens. *J Innate Immun* **1**: 123–135.

Di Martino, P., Fursy, R., Bret, L., Sundararaju, B., and Phillips, R.S. (2003) Indole can act as an extracellular signal to regulate biofilm formation of *Escherichia coli* and other indole-producing bacteria. *Can J Microbiol* **49**: 443–449.

Diamond, G., Beckloff, N., Weinberg, A., and Kisich, K.O. (2009) The Roles of Antimicrobial Peptides in Innate Host Defense. *Curr Pharm Des* **15**: 2377–2392.

Dibb-Fuller, M.P., Allen-Vercoe, E., Thorns, C.J., and Woodward, M.J. (1999) Fimbriae- and flagella-mediated association with and invasion of cultured epithelial cells by *Salmonella* enteritidis. *Microbiology* **145**: 1023–1031.

Dictionnaire médical de l'Académie de Médecine http://dictionnaire.academie-medecine.fr/. Accessed July 4, 2016.

Dillon, S.C., and Dorman, C.J. (2010) Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. *Nat Rev Microbiol* **8**: 185–195.

Dilworth, M.J., and Glenn, A.R. (1999) Problems of adverse pH and bacterial strategies to combat it. *Novartis Found Symp* **221**: 4–14, NaN-18.

Dior, M., Delagrèverie, H., Duboc, H., Jouet, P., Coffin, B., Brot, L., *et al.* (2016) Interplay between bile acid metabolism and microbiota in irritable bowel syndrome. *Neurogastroenterol Motil Off J Eur Gastrointest Motil Soc* **28**: 1330–1340.

Dogan, B., Suzuki, H., Herlekar, D., Sartor, R.B., Campbell, B.J., Roberts, C.L., *et al.* (2014) Inflammation-associated adherent-invasive *Escherichia coli* are enriched in pathways for use of propanediol and iron and M-cell translocation. *Inflamm Bowel Dis* **20**: 1919–1932.

Doig, P., Yao, R., Burr, D.H., Guerry, P., and Trust, T.J. (1996) An environmentally regulated pilus-like appendage involved in *Campylobacter* pathogenesis. *Mol Microbiol* **20**: 885–894.

Dozois, C.M., Dho-Moulin, M., Brée, A., Fairbrother, J.M., Desautels, C., and Curtiss, R. (2000) Relationship between the Tsh autotransporter and pathogenicity of avian *Escherichia coli* and localization and analysis of the Tsh genetic region. *Infect Immun* **68**: 4145–4154.

Dreux, N., Denizot, J., Martinez-Medina, M., Mellmann, A., Billig, M., Kisiela, D., *et al.* (2013) Point Mutations in FimH Adhesin of Crohn's Disease-Associated Adherent-Invasive *Escherichia coli* Enhance Intestinal Inflammatory Response. *PLoS Pathog* **9**: e1003141.

Dridi, B., Raoult, D., and Drancourt, M. (2011) Archaea as emerging organisms in complex human microbiomes. *Anaerobe* **17**: 56–63.

Duboc, H., Rajca, S., Rainteau, D., Benarous, D., Maubert, M.-A., Quervain, E., *et al.* (2013) Connecting dysbiosis, bile-acid dysmetabolism and gut inflammation in inflammatory bowel diseases. *Gut* **62**: 531–539.

Duboc, H., Taché, Y., and Hofmann, A.F. (2014) The bile acid TGR5 membrane receptor: From basic research to clinical application. *Dig Liver Dis* **46**: 302–312.

Duerr, R.H., Taylor, K.D., Brant, S.R., Rioux, J.D., Silverberg, M.S., Daly, M.J., *et al.* (2006) A Genome-Wide Association Study Identifies IL23R as an Inflammatory Bowel Disease Gene. *Science* **314**: 1461–1463.

Eade, C.R., Hung, C.-C., Bullard, B., Gonzalez-Escobedo, G., Gunn, J.S., and Altier, C. (2016) Bile Acids Function Synergistically to Repress Invasion Gene Expression in *Salmonella* by Destabilizing the Invasion Regulator HilD. *Infect Immun* IAI.00177-16.

Eaves-Pyles, T., Allen, C.A., Taormina, J., Swidsinski, A., Tutt, C.B., Jezek, G.E., *et al.* (2008) *Escherichia coli* isolated from a Crohn's disease patient adheres, invades, and induces inflammatory responses in polarized intestinal epithelial cells. *Int J Med Microbiol IJMM* **298**: 397–409.

Eckburg, P.B., Bik, E.M., Bernstein, C.N., Purdom, E., Dethlefsen, L., Sargent, M., *et al.* (2005) Diversity of the Human Intestinal Microbial Flora. *Science* **308**: 1635–1638.

Economou, M., and Pappas, G. (2008) New global map of Crohn's disease: Genetic, environmental, and socioeconomic correlations. *Inflamm Bowel Dis* **14**: 709–720.

Ekbom, A., Adami, H.-O., Helmick, C.G., Jonzon, A., and Zack, M.M. (1990) Perinatal Risk Factors for Inflammatory Bowel Disease: A Case-Control Study. *Am J Epidemiol* **132**: 1111–1119.

Elliott, S.J., Sperandio, V., Girón, J.A., Shin, S., Mellies, J.L., Wainwright, L., *et al.* (2000) The Locus of Enterocyte Effacement (LEE)-Encoded Regulator Controls Expression of Both LEE- and Non-LEE-Encoded Virulence Factors in Enteropathogenic and Enterohemorrhagic *Escherichia coli*. *Infect Immun* **68**: 6115–6126.

Ellis, R.D., Goodlad, J.R., Limb, G.A., Powell, J.J., Thompson, R.P.H., and Punchard, N.A. (2014) Activation of nuclear factor kappa B in Crohn's disease. *Inflamm Res* **47**: 440–445.

Ermund, A., Gustafsson, J.K., Hansson, G.C., and Keita, Å.V. (2013) Mucus Properties and Goblet Cell Quantification in Mouse, Rat and Human Ileal Peyer's Patches. *PLoS ONE* **8** http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3865249/. Accessed June 18, 2015.

Escherich, T. (1989) The Intestinal Bacteria of the Neonate and Breast-Fed Infant. *Rev Infect Dis* **11**: 352–356.

Eskelinen, E.-L. (2005) Maturation of Autophagic Vacuoles in Mammalian Cells. *Autophagy* **1**: 1–10.

Evans, J.M., Morris, L.S., and Marchesi, J.R. (2013) The gut microbiome: the role of a virtual organ in the endocrinology of the host. *J Endocrinol* **218**: R37–R47.

Faherty, C.S., Redman, J.C., Rasko, D.A., Barry, E.M., and Nataro, J.P. (2012) *Shigella flexneri* effectors OspE1 and OspE2 mediate induced adherence to the colonic epithelium following bile salts exposure. *Mol Microbiol* **85**: 107–121.

Fallingborg, J. (1999) Intraluminal pH of the human gastrointestinal tract. *Dan Med Bull* **46**: 183–196.

Favier, C., Neut, C., Mizon, C., Cortot, A., Colombel, J.F., and Mizon, J. (1997) Fecal β-D-Galactosidase Production and Bifidobacteria Are Decreased in Crohn's Disease. *Dig Dis Sci* **42**: 817–822.

Feagan, B.G., Rochon, J., Fedorak, R.N., Irvine, E.J., Wild, G., Sutherland, L., *et al.* (1995) Methotrexate for the Treatment of Crohn's Disease. *N Engl J Med* **332**: 292–297.

Feehily, C., and Karatzas, K. a. g. (2013) Role of glutamate metabolism in bacterial responses towards acid and other stresses. *J Appl Microbiol* **114**: 11–24.

Fekkes, P., and Driessen, A.J.M. (1999) Protein Targeting to the Bacterial Cytoplasmic Membrane. *Microbiol Mol Biol Rev* **63**: 161.

Feller, M., Huwiler, K., Stephan, R., Altpeter, E., Shang, A., Furrer, H., *et al.* (2007) *Mycobacterium avium* subspecies paratuberculosis and Crohn's disease: a systematic review and meta-analysis. *Lancet Infect Dis* **7**: 607–613.

Filloux, A., Hachani, A., and Bleves, S. (2008) The bacterial type VI secretion machine: yet another player for protein transport across membranes. *Microbiol Read Engl* **154**: 1570–1583.

Fink, D.L., Cope, L.D., Hansen, E.J., and Geme, J.W. (2001) The Hemophilus influenzae Hap autotransporter is a chymotrypsin clan serine protease and undergoes autoproteolysis via an intermolecular mechanism. *J Biol Chem* **276**: 39492–39500.

Fiocca, R., Necchi, V., Sommi, P., Ricci, V., Telford, J., Cover, T.L., and Solcia, E. (1999) Release of *Helicobacter pylori* vacuolating cytotoxin by both a specific secretion pathway and budding of outer membrane vesicles. Uptake of released toxin and vesicles by gastric epithelium. *J Pathol* **188**: 220–226.

Fiorucci, S., Cipriani, S., Baldelli, F., and Mencarelli, A. (2010) Bile acid-activated receptors in the treatment of dyslipidemia and related disorders. *Prog Lipid Res* **49**: 171–185.

Floyd, D.N., Langham, S., Séverac, H.C., and Levesque, B.G. (2015) The economic and quality-of-life burden of Crohn's disease in Europe and the United States, 2000 to 2013: a systematic review. *Dig Dis Sci* **60**: 299–312.

Franchimont, D., Vermeire, S., El Housni, H., Pierik, M., Van Steen, K., Gustot, T., *et al.* (2004) Deficient host-bacteria interactions in inflammatory bowel disease? The toll-like receptor (TLR)-4 Asp299gly polymorphism is associated with Crohn's disease and ulcerative colitis. *Gut* **53**: 987–992.

Franke, A., McGovern, D.P.B., Barrett, J.C., Wang, K., Radford-Smith, G.L., Ahmad, T., *et al.* (2010) Meta-Analysis Increases to 71 the Tally of Confirmed Crohn's Disease Susceptibility Loci. *Nat Genet* **42**: 1118–1125.

Fraser, A.G. (2003) Methotrexate: first-line or second-line immunomodulator? *Eur J Gastroenterol Hepatol* **15**: 225–231.

Fronzes, R., Christie, P.J., and Waksman, G. (2009) The structural biology of type IV secretion systems. *Nat Rev Microbiol* **7**: 703–714.

Fujimura, Y., Kamoi, R., and Iida, M. (1996) Pathogenesis of aphthoid ulcers in Crohn's disease: correlative findings by magnifying colonoscopy, electron microscopy, and immunohistochemistry. *Gut* **38**: 724–732.

Fujisawa, T., and Mori, M. (1996) Influence of bile salts on β-glucuronidase activity of intestinal bacteria. *Lett Appl Microbiol* **22**: 271–274.

Gale, E.F., and Epps, H.M. (1942) The effect of the pH of the medium during growth on the enzymic activities of bacteria (*Escherichia coli* and *Micrococcus lysodeikticus*) and the biological significance of the changes produced. *Biochem J* **36**: 600–618.

Garlatti, M., and Barouki, R. (2002) Le stress du réticulum endoplasmique : adaptation et toxicité. *médecine/sciences* **18**: 585–594.

Garsin, D.A. (2010) Ethanolamine Utilization in Bacterial Pathogens: Roles and Regulation. *Nat Rev Microbiol* **8**: 290–295.

Gaudier, E., and Hoebler, C. (2006) [Physiological role of mucins in the colonic barrier integrity]. *Gastroentérologie Clin Biol* **30**: 965–974.

Gawarzewski, I., Smits, S.H.J., Schmitt, L., and Jose, J. (2013) Structural comparison of the transport units of type V secretion systems. *Biol Chem* **394**: 1385–1398.

Gent, A.E., Hellier, M.D., Grace, R.H., Swarbrick, E.T., and Coggon, D. (1994) Inflammatory bowel disease and domestic hygiene in infancy. *The Lancet* **343**: 766–767.

Gerbe, F., Legraverend, C., and Jay, P. (2012) The intestinal epithelium tuft cells: specification and function. *Cell Mol Life Sci* **69**: 2907–2917.

Gevers, D., Kugathasan, S., Denson, L.A., Vázquez-Baeza, Y., Van Treuren, W., Ren, B., *et al.* (2014) The treatment-naive microbiome in new-onset Crohn's disease. *Cell Host Microbe* **15**: 382–392.

Ghaisas, S., Maher, J., and Kanthasamy, A. (2016) Gut microbiome in health and disease: Linking the microbiome–gut–brain axis and environmental factors in the pathogenesis of systemic and neurodegenerative diseases. *Pharmacol Ther* **158**: 52–62.

Ghosh, S., Goldin, E., Gordon, F.H., Malchow, H.A., Rask-Madsen, J., Rutgeerts, P., *et al.* (2003) Natalizumab for Active Crohn's Disease. *N Engl J Med* **348**: 24–32.

Giaffer, M.H., Holdsworth, C.D., and Duerden, B.I. (1992) Virulence properties of *Escherichia coli* strains isolated from patients with inflammatory bowel disease. *Gut* **33**: 646–650.

Giampapa, C.S., Abraham, S.N., Chiang, T.M., and Beachey, E.H. (1988) Isolation and characterization of a receptor for type 1 fimbriae of *Escherichia coli* from guinea pig erythrocytes. *J Biol Chem* **263**: 5362–5367.

Gibold, L., Garenaux, E., Dalmasso, G., Gallucci, C., Cia, D., Mottet-Auselo, B., *et al.* (2015) The Vat-AIEC protease promotes crossing of the intestinal mucus layer by Crohn's disease-associated *Escherichia coli*. *Cell Microbiol* 18(5):617-31.

Gibson, P.R., and Shepherd, S.J. (2005) Personal view: food for thought – western lifestyle and susceptibility to Crohn's disease. The FODMAP hypothesis. *Aliment Pharmacol Ther* **21**: 1399–1409.

Girardin, S.E., Boneca, I.G., Viala, J., Chamaillard, M., Labigne, A., Thomas, G., *et al.* (2003) Nod2 Is a General Sensor of Peptidoglycan through Muramyl Dipeptide (MDP) Detection. *J Biol Chem* **278**: 8869–8872.

Girón, J.A., Torres, A.G., Freer, E., and Kaper, J.B. (2002) The flagella of enteropathogenic *Escherichia coli* mediate adherence to epithelial cells. *Mol Microbiol* **44**: 361–379.

Glasser, A.-L., Boudeau, J., Barnich, N., Perruchot, M.-H., Colombel, J.-F., and Darfeuille-Michaud, A. (2001) Adherent Invasive *Escherichia coli* Strains from Patients with Crohn's Disease Survive and Replicate within Macrophages without Inducing Host Cell Death. *Infect Immun* **69**: 5529–5537.

Godet, P.G., May, G.R., and Sutherland, L.R. (1995) Meta-analysis of the role of oral contraceptive agents in inflammatory bowel disease. *Gut* **37**: 668–673.

Gohlke, U., Pullan, L., McDevitt, C.A., Porcelli, I., Leeuw, E. de, Palmer, T., *et al.* (2005) The TatA component of the twin-arginine protein transport system forms channel complexes of variable diameter. *Proc Natl Acad Sci U S A* **102**: 10482–10486.

Goldberg, M.B., Bârzu, O., Parsot, C., and Sansonetti, P.J. (1993) Unipolar localization and ATPase activity of IcsA, a *Shigella flexneri* protein involved in intracellular movement. *J Bacteriol* **175**: 2189–2196.

Gómez Zavaglia, A., Kociubinski, G., Pérez, P., Disalvo, E., and De Antoni, G. (2002) Effect of bile on the lipid composition and surface properties of bifidobacteria. *J Appl Microbiol* **93**: 794–799.

Gonzales, L., Ali, Z.B., Nygren, E., Wang, Z., Karlsson, S., Zhu, B., *et al.* (2013) Alkaline pH Is a signal for optimal production and secretion of the heat labile toxin, LT in enterotoxigenic *Escherichia coli* (ETEC). *PloS One* **8**: e74069.

Gophna, U., Barlev, M., Seijffers, R., Oelschlager, T.A., Hacker, J., and Ron, E.Z. (2001) Curli Fibers Mediate Internalization of *Escherichia coli* by Eukaryotic Cells. *Infect Immun* **69**: 2659–2665.

Gordon, D.M., and Cowling, A. (2003) The distribution and genetic structure of *Escherichia coli* in Australian vertebrates: host and geographic effects. *Microbiology* **149**: 3575–3586.

Gouridis, G., Karamanou, S., Gelis, I., Kalodimos, C.G., and Economou, A. (2009) Signal peptides are allosteric activators of the protein translocase. *Nature* **462**: 363–367.

Goyal, P., Krasteva, P.V., Van Gerven, N., Gubellini, F., Van den Broeck, I., Troupiotis-Tsaïlaki, A., *et al.* (2014) Structural and mechanistic insights into the bacterial amyloid secretion channel CsgG. *Nature* **advance online publication** http://www.nature.com.gate2.inist.fr/nature/journal/vaop/ncurrent/full/nature13768.html. Accessed September 19, 2014.

Gradel, K.O., Nielsen, H.L., Schønheyder, H.C., Ejlertsen, T., Kristensen, B., and Nielsen, H. (2009) Increased Short- and Long-Term Risk of Inflammatory Bowel Disease After *Salmonella* or *Campylobacter* Gastroenteritis. *Gastroenterology* **137**: 495–501.

Grenier, D., and Mayrand, D. (1987) Functional characterization of extracellular vesicles produced by *Bacteroides gingivalis*. *Infect Immun* **55**: 111–117.

Grijpstra, J., Arenas, J., Rutten, L., and Tommassen, J. (2013) Autotransporter secretion: varying on a theme. *Res Microbiol* **164**: 562–582.

Groschwitz, K.R., and Hogan, S.P. (2009) Intestinal Barrier Function: Molecular Regulation and Disease Pathogenesis. *J Allergy Clin Immunol* **124**: 3–22.

Gross, V., Andus, T., Leser, H.G., Roth, M., and Schölmerich, J. (1991) Inflammatory mediators in chronic inflammatory bowel diseases. *Klin Wochenschr* **69**: 981–987.

Grys, T.E., Siegel, M.B., Lathem, W.W., and Welch, R.A. (2005) The StcE protease contributes to intimate adherence of enterohemorrhagic *Escherichia coli* O157:H7 to host cells. *Infect Immun* **73**: 1295–1303.

Guaní-Guerra, E., Santos-Mendoza, T., Lugo-Reyes, S.O., and Terán, L.M. (2010) Antimicrobial peptides: General overview and clinical implications in human health and disease. *Clin Immunol* **135**: 1–11.

Guarner, F., and Malagelada, J.-R. (2003) Gut flora in health and disease. *The Lancet* **361**: 512–519.

Guerry, P. (2007) *Campylobacter* flagella: not just for motility. *Trends Microbiol* **15**: 456–461.

Gullberg, E., and Söderholm, J.D. (2006) Peyer's Patches and M Cells as Potential Sites of the Inflammatory Onset in Crohn's Disease. *Ann N Y Acad Sci* **1072**: 218–232.

Gum, J.R., Hicks, J.W., Toribara, N.W., Siddiki, B., and Kim, Y.S. (1994) Molecular cloning of human intestinal mucin (MUC2) cDNA. Identification of the amino terminus and overall sequence similarity to prepro-von Willebrand factor. *J Biol Chem* **269**: 2440–2446.

Gunn, J.S. (2000) Mechanisms of bacterial resistance and response to bile. *Microbes Infect* **2**: 907–913.

Guo, L., Lim, K.B., Gunn, J.S., Bainbridge, B., Darveau, R.P., Hackett, M., and Miller, S.I. (1997) Regulation of Lipid A Modifications by *Salmonella typhimurium* Virulence Genes phoP-phoQ. *Science* **276**: 250–253.

Gupta, S., and Chowdhury, R. (1997) Bile affects production of virulence factors and motility of *Vibrio cholerae*. *Infect Immun* **65**: 1131–1134.

Gutierrez, M.G., Master, S.S., Singh, S.B., Taylor, G.A., Colombo, M.I., and Deretic, V. (2004) Autophagy Is a Defense Mechanism Inhibiting BCG and *Mycobacterium tuberculosis* Survival in Infected Macrophages. *Cell* **119**: 753–766.

Guttenplan, S.B., and Kearns, D.B. (2013) Regulation of flagellar motility during biofilm formation. *FEMS Microbiol Rev* **37**: 849–871.

Hacker, J., and Carniel, E. (2001) Ecological fitness, genomic islands and bacterial pathogenicity. *EMBO Rep* **2**: 376–381.

Halfvarson, J., Bodin, L., Tysk, C., Lindberg, E., and Järnerot, G. (2003) Inflammatory bowel disease in a Swedish twin cohort: a long-term follow-up of concordance and clinical characteristics. *Gastroenterology* **124**: 1767–1773.

Halme, L., Paavola-Sakki, P., Turunen, U., Lappalainen, M., Färkkilä, M., and Kontula, K. (2006) Family and twin studies in inflammatory bowel disease. *World J Gastroenterol WJG* **12**: 3668–3672.

Hamner, S., McInnerney, K., Williamson, K., Franklin, M.J., and Ford, T.E. (2013) Bile Salts Affect Expression of *Escherichia coli* O157:H7 Genes for Virulence and Iron Acquisition, and Promote Growth under Iron Limiting Conditions. *PLoS ONE* **8**: e74647.

Hamon, E., Horvatovich, P., Izquierdo, E., Bringel, F., Marchioni, E., Aoudé-Werner, D., and Ennahar, S. (2011) Comparative proteomic analysis of *Lactobacillus plantarum* for the identification of key proteins in bile tolerance. *BMC Microbiol* **11**: 63.

Hampe, J., Franke, A., Rosenstiel, P., Till, A., Teuber, M., Huse, K., *et al.* (2007) A genome-wide association scan of nonsynonymous SNPs identifies a susceptibility variant for Crohn disease in ATG16L1. *Nat Genet* **39**: 207–211.

Hansen, R., Thomson, J.M., El-Omar, E.M., and Hold, G.L. (2010) The role of infection in the aetiology of inflammatory bowel disease. *J Gastroenterol* **45**: 266–276.

Hansson, G.C. (2012) Role of mucus layers in gut infection and inflammation. *Curr Opin Microbiol* **15**: 57–62.

Hardy, J., Francis, K.P., DeBoer, M., Chu, P., Gibbs, K., and Contag, C.H. (2004) Extracellular Replication of *Listeria monocytogenes* in the Murine Gall Bladder. *Science* **303**: 851–853.

Harjai, K., Khandwaha, R.K., Mittal, R., Yadav, V., Gupta, V., and Sharma, S. (2005) Effect of pH on production of virulence factors by biofilm cells of *Pseudomonas aeruginosa*. *Folia Microbiol (Praha)* **50**: 99–102.

Harrington, S.M., Sheikh, J., Henderson, I.R., Ruiz-Perez, F., Cohen, P.S., and Nataro, J.P. (2009) The Pic Protease of Enteroaggregative *Escherichia coli* Promotes Intestinal Colonization and Growth in the Presence of Mucin. *Infect Immun* **77**: 2465–2473.

Heijtz, R.D., Wang, S., Anuar, F., Qian, Y., Björkholm, B., Samuelsson, A., *et al.* (2011) Normal gut microbiota modulates brain development and behavior. *Proc Natl Acad Sci* **108**: 3047–3052.

Heimer, S.R., Rasko, D.A., Lockatell, C.V., Johnson, D.E., and Mobley, H.L.T. (2004) Autotransporter Genes pic and tsh Are Associated with *Escherichia coli* Strains That Cause Acute Pyelonephritis and Are Expressed during Urinary Tract Infection. *Infect Immun* **72**: 593–597.

Heithoff, D.M., Conner, C.P., Hentschel, U., Govantes, F., Hanna, P.C., and Mahan, M.J. (1999) Coordinate Intracellular Expression of *Salmonella* Genes Induced during Infection. *J Bacteriol* **181**: 799–807.

Henderson, I.R., Czeczulin, J., Eslava, C., Noriega, F., and Nataro, J.P. (1999) Characterization of pic, a secreted protease of *Shigella flexneri* and enteroaggregative *Escherichia coli*. *Infect Immun* **67**: 5587–5596.

Henderson, I.R., and Nataro, J.P. (2001) Virulence Functions of Autotransporter Proteins. *Infect Immun* **69**: 1231–1243.

Henderson, I.R., Navarro-Garcia, F., Desvaux, M., Fernandez, R.C., and Ala'Aldeen, D. (2004) Type V Protein Secretion Pathway: the Autotransporter Story. *Microbiol Mol Biol Rev* **68**: 692–744.

Henderson, I.R., Navarro-Garcia, F., and Nataro, J.P. (1998) The great escape: structure and function of the autotransporter proteins. *Trends Microbiol* **6**: 370–378.

Hennequin, C., Aumeran, C., Robin, F., Traore, O., and Forestier, C. (2012) Antibiotic resistance and plasmid transfer capacity in biofilm formed with a CTX-M-15-producing *Klebsiella pneumoniae* isolate. *J Antimicrob Chemother* **67**: 2123–2130.

Hennequin, C., and Forestier, C. (2009) oxyR, a LysR-Type Regulator Involved in *Klebsiella pneumoniae* Mucosal and Abiotic Colonization. *Infect Immun* **77**: 5449–5457.

Hernández, S.B., Cota, I., Ducret, A., Aussel, L., and Casadesús, J. (2012) Adaptation and Preadaptation of *Salmonella enterica* to Bile. *PLoS Genet* **8** http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3261920/. Accessed July 3, 2016.

Hetz, C., Thielen, P., Matus, S., Nassif, M., Court, F., Kiffin, R., *et al.* (2009) XBP-1 deficiency in the nervous system protects against amyotrophic lateral sclerosis by increasing autophagy. *Genes Dev* **23**: 2294–2306.

Heyde, M., and Portalier, R. (1987) Regulation of major outer membrane porin proteins of *Escherichia coli* K 12 by pH. *Mol Gen Genet MGG* **208**: 511–517.

Hicks, K.G., Delbecq, S.P., Sancho-Vaello, E., Blanc, M.-P., Dove, K.K., Prost, L.R., *et al.* (2015) Acidic pH and divalent cation sensing by PhoQ are dispensable for systemic *Salmonella*e virulence. *eLife* **4** http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4473727/. Accessed May 30, 2016.

Hirst, T.R., Sanchez, J., Kaper, J.B., Hardy, S.J., and Holmgren, J. (1984) Mechanism of toxin secretion by *Vibrio cholerae* investigated in strains harboring plasmids that encode heat-labile enterotoxins of *Escherichia coli*. *Proc Natl Acad Sci U S A* **81**: 7752–7756.

Hoekstra, D., Laan, J.W. van der, Leij, L. de, and Witholt, B. (1976) Release of outer membrane fragments from normally growing *Escherichia coli*. *Biochim Biophys Acta BBA - Biomembr* **455**: 889–899.

Hofmann, A.F. (1999) Bile Acids: The Good, the Bad, and the Ugly. *News Physiol Sci Int J Physiol Prod Jointly Int Union Physiol Sci Am Physiol Soc* **14**: 24–29.

Hollingsworth, M.A., and Swanson, B.J. (2004) Mucins in cancer: protection and control of the cell surface. *Nat Rev Cancer* **4**: 45–60.

Homer, C.R., Richmond, A.L., Rebert, N.A., Achkar, J.-P., and McDonald, C. (2010) ATG16L1 and NOD2 interact in an autophagy-dependent, anti-bacterial pathway implicated in Crohn's disease pathogenesis. *Gastroenterology* **139**: 1630–1641.e2.

Hommes, D.W., Duijvestein, M., Zelinkova, Z., Stokkers, P.C.F., Ley, M.H., Stoker, J., *et al.* (2011) Long-term follow-up of autologous hematopoietic stem cell transplantation for severe refractory Crohn's disease. *J Crohns Colitis* **5**: 543–549.

Hommes, D.W., Oldenburg, B., Bodegraven, A.A. van, Hogezand, R.A. van, Jong, D.J. de, Romberg-Camps, M.J.L., *et al.* (2006) Guidelines for treatment with infliximab for Crohn's disease. *Neth J Med* **64**: 219–229.

Hooper, L.V., Littman, D.R., and Macpherson, A.J. (2012) Interactions between the microbiota and the immune system. *Science* **336**: 1268–1273.

Hooper, L.V., and Macpherson, A.J. (2010) Immune adaptations that maintain homeostasis with the intestinal microbiota. *Nat Rev Immunol* **10**: 159–169.

Horiuchi, S., Inagaki, Y., Okamura, N., Nakaya, R., and Yamamoto, N. (1992) Type 1 Pili Enhance the Invasion of *Salmonella* braenderup and *Salmonella typhimurium* to HeLa Cells. *Microbiol Immunol* **36**: 593–602.

Hsieh, Y.-C., Liang, S.-M., Tsai, W.-L., Chen, Y.-H., Liu, T.-Y., and Liang, C.-M. (2003) Study of Capsular Polysaccharide from *Vibrio parahaemolyticus*. *Infect Immun* **71**: 3329–3336.

Hugdahl, M.B., Beery, J.T., and Doyle, M.P. (1988) Chemotactic behavior of *Campylobacter jejuni*. *Infect Immun* **56**: 1560–1566.

Hugot, J.-P., Alberti, C., Berrebi, D., Bingen, E., and Cézard, J.-P. (2003) Crohn's disease: the cold chain hypothesis. *The Lancet* **362**: 2012–2015.

Hugot, J.-P., Chamaillard, M., Zouali, H., Lesage, S., Cézard, J.-P., Belaiche, J., *et al.* (2001) Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature* **411**: 599–603.

Hugot, J.P., Laurent-Puig, P., Gower-Rousseau, C., Caillat-Zucman, S., Beaugerie, L., Dupas, J.-L., *et al.* (1994) Linkage analyses of chromosome 6 loci, including HLA, in familial aggregations of Crohn disease. *Am J Med Genet* **52**: 207–213.

Humbert, L., Maubert, M.A., Wolf, C., Duboc, H., Mahé, M., Farabos, D., *et al.* (2012) Bile acid profiling in human biological samples: comparison of extraction procedures and application to normal and cholestatic patients. *J Chromatogr B Analyt Technol Biomed Life Sci* **899**: 135–145.

Hung, D.T., Zhu, J., Sturtevant, D., and Mekalanos, J.J. (2006) Bile acids stimulate biofilm formation in *Vibrio cholerae*. *Mol Microbiol* **59**: 193–201.

Huttner, K.M., and Bevins, C.L. (1999) Antimicrobial Peptides as Mediators of Epithelial Host Defense. *Pediatr Res* **45**: 785–794.

Ieva, R., Tian, P., Peterson, J.H., and Bernstein, H.D. (2011) Sequential and spatially restricted interactions of assembly factors with an autotransporter beta domain. *Proc Natl Acad Sci U S A* **108**: E383-391.

Inagaki, T., Moschetta, A., Lee, Y.-K., Peng, L., Zhao, G., Downes, M., *et al.* (2006) Regulation of antibacterial defense in the small intestine by the nuclear bile acid receptor. *Proc Natl Acad Sci U S A* **103**: 3920–3925.

Iontcheva, I., Oppenheim, F.G., Offner, G.D., and Troxler, R.F. (2000) Molecular mapping of statherin- and histatin-binding domains in human salivary mucin MG1 (MUC5B) by the yeast two-hybrid system. *J Dent Res* **79**: 732–739.

Ivanov, I.I., Atarashi, K., Manel, N., Brodie, E.L., Shima, T., Karaoz, U., *et al.* (2009) Induction of intestinal Th17 cells by segmented filamentous bacteria. *Cell* **139**: 485–498.

Jacob-Dubuisson, F., Locht, C., and Antoine, R. (2001) Two-partner secretion in Gram-negative bacteria: a thrifty, specific pathway for large virulence proteins. *Mol Microbiol* **40**: 306–313.

Jacquot, A., Sakamoto, C., Razafitianamaharavo, A., Caillet, C., Merlin, J., Fahs, A., *et al.* (2014) Dynamic Modulation of Fimbrial Extension and FimH-Mannose Binding Force on Live Bacteria Under pH Changes: A Molecular Atomic Force Microscopy Analysis. *J Biomed Nanotechnol* **10**: 3361–3372.

Jani, A.J., and Cotter, P.A. (2010) Type VI Secretion: Not Just for Pathogenesis Anymore. *Cell Host Microbe* **8**: 2–6.

Jawhara, S., and Poulain, D. (2007) Saccharomyces boulardii decreases inflammation and intestinal colonization by *Candida albicans* in a mouse model of chemically-induced colitis. *Med Mycol* **45**: 691–700.

Jayamani, E., and Mylonakis, E. (2014) Effector triggered manipulation of host immune response elicited by different pathotypes of *Escherichia coli*. *Virulence* **5**: 733–739.

Jess, T., Simonsen, J., Nielsen, N.M., Jørgensen, K.T., Bager, P., Ethelberg, S., and Frisch, M. (2011) Enteric *Salmonella* or *Campylobacter* infections and the risk of inflammatory bowel disease. *Gut* **60**: 318–324.

Johansson, M.E.V. (2012) Fast Renewal of the Distal Colonic Mucus Layers by the Surface Goblet Cells as Measured by In Vivo Labeling of Mucin Glycoproteins. *PLoS ONE* **7** http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3398881/. Accessed December 15, 2015.

Johansson, M.E.V., Ambort, D., Pelaseyed, T., Schütte, A., Gustafsson, J.K., Ermund, A., *et al.* (2011) Composition and functional role of the mucus layers in the intestine. *Cell Mol Life Sci CMLS* **68**: 3635–3641.

Johansson, M.E.V., Gustafsson, J.K., Holmén-Larsson, J., Jabbar, K.S., Xia, L., Xu, H., *et al.* (2014) Bacteria penetrate the normally impenetrable inner colon mucus layer in both murine colitis models and patients with ulcerative colitis. *Gut* **63**: 281–291.

Johansson, M.E.V., and Hansson, G.C. (2011) Microbiology. Keeping bacteria at a distance. *Science* **334**: 182–183.

Johansson, M.E.V., Larsson, J.M.H., and Hansson, G.C. (2011) The two mucus layers of colon are organized by the MUC2 mucin, whereas the outer layer is a legislator of host-microbial interactions. *Proc Natl Acad Sci U S A* **108 Suppl 1**: 4659–4665.

Johansson, M.E.V., Phillipson, M., Petersson, J., Velcich, A., Holm, L., and Hansson, G.C. (2008) The inner of the two Muc2 mucin-dependent mucus layers in colon is devoid of bacteria. *Proc Natl Acad Sci U S A* **105**: 15064–15069.

Johnson, J.R., Delavari, P., Kuskowski, M., and Stell, A.L. (2001) Phylogenetic Distribution of Extraintestinal Virulence-Associated Traits in *Escherichia coli*. *J Infect Dis* **183**: 78–88.

Johnson, J.R., and Russo, T.A. (2002) Extraintestinal pathogenic *Escherichia coli* : "The other bad E coli ." *J Lab Clin Med* **139**: 155–162.

Johnson, T.L., Abendroth, J., Hol, W.G.J., and Sandkvist, M. (2006) Type II secretion: from structure to function. *FEMS Microbiol Lett* **255**: 175–186.

Jones, C., and Holland, I.B. (1985) Role of the SulB (FtsZ) protein in division inhibition during the SOS response in *Escherichia coli*: FtsZ stabilizes the inhibitor SulA in maxicells. *Proc Natl Acad Sci U S A* **82**: 6045–6049.

Jong, W.S.P., Hagen-Jongman, C.M. ten, Ruijter, E., Orru, R.V.A., Genevaux, P., and Luirink, J. (2010) YidC is involved in the biogenesis of the secreted autotransporter hemoglobin protease. *J Biol Chem* **285**: 39682–39690.

Joo, L.M., Macfarlane-Smith, L.R., and Okeke, I.N. (2007) Error-Prone DNA Repair System in Enteroaggregative *Escherichia coli* Identified by Subtractive Hybridization. *J Bacteriol* **189**: 3793–3803.

Joossens, S., Reinisch, W., Vermeire, S., Sendid, B., Poulain, D., Peeters, M., *et al.* (2002) The value of serologic markers in indeterminate colitis: A prospective follow-up study. *Gastroenterology* **122**: 1242–1247.

Jostins, L., Ripke, S., Weersma, R.K., Duerr, R.H., McGovern, D.P., Hui, K.Y., *et al.* (2012) Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* **491**: 119–124.

Jung, H.C., Eckmann, L., Yang, S.K., Panja, A., Fierer, J., Morzycka-Wroblewska, E., and Kagnoff, M.F. (1995) A distinct array of proinflammatory cytokines is expressed in human colon epithelial cells in response to bacterial invasion. *J Clin Invest* **95**: 55–65.

Junker, M., Schuster, C.C., McDonnell, A.V., Sorg, K.A., Finn, M.C., Berger, B., and Clark, P.L. (2006) Pertactin beta-helix folding mechanism suggests common themes for the secretion and folding of autotransporter proteins. *Proc Natl Acad Sci U S A* **103**: 4918–4923.

Kadurugamuwa, J.L., and Beveridge, T.J. (1999) Membrane vesicles derived from *Pseudomonas aeruginosa* and *Shigella flexneri* can be integrated into the surfaces of other Gram-negative bacteria. *Microbiology* **145**: 2051–2060.

Kajava, A.V., and Steven, A.C. (2006) The turn of the screw: Variations of the abundant β-solenoid motif in passenger domains of Type V secretory proteins. *J Struct Biol* **155**: 306–315.

Kakkanat, A., Totsika, M., Schaale, K., Duell, B.L., Lo, A.W., Phan, M.-D., *et al.* (2015) The role of H4 flagella in *Escherichia coli* ST131 virulence. *Sci Rep* **5** http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4637896/. Accessed May 14, 2016.

Kalla, R., Ventham, N.T., Kennedy, N.A., Quintana, J.F., Nimmo, E.R., Buck, A.H., and Satsangi, J. (2015) MicroRNAs: new players in IBD. *Gut* **64**: 504–513.

Kallinowski, F., Wassmer, A., Hofmann, M.A., Harmsen, D., Heesemann, J., Karch, H., *et al.* (1998) Prevalence of enteropathogenic bacteria in surgically treated chronic inflammatory bowel disease. *Hepatogastroenterology* **45**: 1552–1558.

Kalliomäki, M., Kirjavainen, P., Eerola, E., Kero, P., Salminen, S., and Isolauri, E. (2001) Distinct patterns of neonatal gut microflora in infants in whom atopy was and was not developing. *J Allergy Clin Immunol* **107**: 129–134.

Kanamaru, S. (2009) Structural similarity of tailed phages and pathogenic bacterial secretion systems. *Proc Natl Acad Sci U S A* **106**: 4067–4068.

Kaper, J.B., Nataro, J.P., and Mobley, H.L. (2004) Pathogenic *Escherichia coli*. *Nat Rev Microbiol* **2**: 123–140.

Karp, P.D., Weaver, D., Paley, S., Fulcher, C., Kubo, A., Kothari, A., *et al.* (2014) The EcoCyc Database. *EcoSal Plus* **6** http://www.asmscience.org/content/journal/ecosalplus/10.1128/ecosalplus.ESP-0009-2013. Accessed June 29, 2016.

Kaser, A., Lee, A.-H., Franke, A., Glickman, J.N., Zeissig, S., Tilg, H., *et al.* (2008) XBP1 links ER stress to intestinal inflammation and confers genetic risk for human inflammatory bowel disease. *Cell* **134**: 743–756.

Kaser, A., Zeissig, S., and Blumberg, R.S. (2010) Inflammatory bowel disease. *Annu Rev Immunol* **28**: 573–621.

Katayama, T., Fujita, K., and Yamamoto, K. (2005) Novel bifidobacterial glycosidases acting on sugar chains of mucin glycoproteins. *J Biosci Bioeng* **99**: 457–465.

Kawai, T., and Akira, S. (2006) TLR signaling. *Cell Death Differ* **13**: 816–825.

Kawakubo, M., Ito, Y., Okimura, Y., Kobayashi, M., Sakura, K., Kasama, S., *et al.* (2004) Natural Antibiotic Function of a Human Gastric Mucin Against *Helicobacter pylori* Infection. *Science* **305**: 1003–1006.

Keighley, M.R., Arabi, Y., Dimock, F., Burdon, D.W., Allan, R.N., and Alexander-Williams, J. (1978) Influence of inflammatory bowel disease on intestinal microflora. *Gut* **19**: 1099–1104.

Keita, Å., Salim, S., Jiang, T., Yang, P.-C., Franzén, L., Söderkvist, P., *et al.* (2008) Increased uptake of non-pathogenic *E. coli* via the follicle-associated epithelium in longstanding ileal Crohn's disease. *J Pathol* **215**: 135–144.

Kelsall, B. (2008) Recent progress in understanding the phenotype and function of intestinal dendritic cells and macrophages. *Mucosal Immunol* **1**: 460–469.

Khan, K.J., Ullman, T.A., Ford, A.C., Abreu, M.T., Abadir, A., Marshall, J.K., *et al.* (2011) Antibiotic Therapy in Inflammatory Bowel Disease: A Systematic Review and Meta-Analysis. *Am J Gastroenterol* **106**: 661–673.

Khan, S., Mian, H., Sandercock, L., Chirgadze, N., and Pai, E. (2011) Crystal Structure of the Passenger Domain of the *Escherichia coli* Autotransporter EspP. *J Mol Biol* **413**: 985–1000.

Khoruts, A., Rank, K.M., Newman, K.M., Viskocil, K., Vaughn, B.P., Hamilton, M.J., and Sadowsky, M.J. (2016) Inflammatory Bowel Disease Affects the Outcome of Fecal Microbiota Transplantation for Recurrent *Clostridium difficile* Infection. *Clin Gastroenterol Hepatol Off Clin Pract J Am Gastroenterol Assoc* .

Kim, J.C., Yoon, J.W., Kim, C.-H., Park, M.-S., and Cho, S.-H. (2012) Repression of flagella motility in enterohemorrhagic *Escherichia coli* O157:H7 by mucin components. *Biochem Biophys Res Commun* **423**: 789–792.

Kleessen, B., Kroesen, A.J., Buhr, H.J., and Blaut, M. (2002) Mucosal and Invading Bacteria in Patients with Inflammatory Bowel Disease Compared with Controls. *Scand J Gastroenterol* **37**: 1034–1041.

Klemm, P., Vejborg, R.M., and Sherlock, O. (2006) Self-associating autotransporters, SAATs: functional and structural similarities. *Int J Med Microbiol IJMM* **296**: 187–195.

Klumpp, J., and Fuchs, T.M. (2007) Identification of novel genes in genomic islands that contribute to *Salmonella typhimurium* replication in macrophages. *Microbiology* **153**: 1207–1220.

Knowles, T.J., Scott-Tucker, A., Overduin, M., and Henderson, I.R. (2009) Membrane protein architects: the role of the BAM complex in outer membrane protein assembly. *Nat Rev Microbiol* **7**: 206–214.

Kobayashi, R.K., Gaziri, L.C., and Vidotto, M.C. (2010) Functional activities of the Tsh protein from avian pathogenic *Escherichia coli* (APEC) strains. *J Vet Sci* **11**: 315–319.

Kobayashi, T., Okamoto, S., Hisamatsu, T., Kamada, N., Chinen, H., Saito, R., *et al.* (2008) IL23 differentially regulates the Th1/Th17 balance in ulcerative colitis and Crohn's disease. *Gut* **57**: 1682–1689.

Koenig, J.E., Spor, A., Scalfone, N., Fricker, A.D., Stombaugh, J., Knight, R., *et al.* (2011) Succession of microbial consortia in the developing infant gut microbiome. *Proc Natl Acad Sci U S A* **108**: 4578–4585.

Köhler, C.-D., and Dobrindt, U. (2011) What defines extraintestinal pathogenic *Escherichia coli*? *Int J Med Microbiol IJMM* **301**: 642–647.

Konisky, J. (1982) Colicins and other Bacteriocins with Established Modes of Action. *Annu Rev Microbiol* **36**: 125–144.

Korbel, J.O., Doerks, T., Jensen, L.J., Perez-Iratxeta, C., Kaczanowski, S., Hooper, S.D., *et al.* (2005) Systematic Association of Genes to Phenotypes by Genome and Literature Mining. *PLoS Biol* **3** http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1073694/. Accessed July 3, 2016.

Kotlowski, R., Bernstein, C.N., Sepehri, S., and Krause, D.O. (2007) High prevalence of *Escherichia coli* belonging to the B2+D phylogenetic group in inflammatory bowel disease. *Gut* **56**: 669–675.

Kreisel, W., Potthoff, K., Bertz, H., Schmitt-Graeff, A., Ruf, G., Rasenack, J., and Finke, J. (2003) Complete remission of Crohn's disease after high-dose cyclophosphamide and autologous stem cell transplantation. *Bone Marrow Transplant* **32**: 337–340.

Kristoffersen, S.M., Ravnum, S., Tourasse, N.J., Økstad, O.A., Kolstø, A.-B., and Davies, W. (2007) Low Concentrations of Bile Salts Induce Stress Responses and Reduce Motility in *Bacillus cereus* ATCC 14570. *J Bacteriol* **189**: 5302–5313.

Krzywinski, M., Schein, J., Birol, İ., Connors, J., Gascoyne, R., Horsman, D., *et al.* (2009) Circos: An information aesthetic for comparative genomics. *Genome Res* **19**: 1639–1645.

Kuehn, M.J., and Kesty, N.C. (2005) Bacterial outer membrane vesicles and the host–pathogen interaction. *Genes Dev* **19**: 2645–2655.

Kufe, D.W. (2009) Mucins in cancer: function, prognosis and therapy. *Nat Rev Cancer* **9**: 874–885.

Küster, W., Pascoe, L., Purrmann, J., Funk, S., Majewski, F., Opitz, J.M., and Reynolds, J.F. (1989) The genetics of Crohn disease: Complex segregation analysis of a family study with 265 patients with Crohn disease and 5,387 relatives. *Am J Med Genet* **32**: 105–108.

Laass, M.W., Roggenbuck, D., and Conrad, K. (2014) Diagnosis and classification of Crohn's disease. *Autoimmun Rev* **13**: 467–471.

Labbé, A., Ganopolsky, J.G., Martoni, C.J., Prakash, S., and Jones, M.L. (2014) Bacterial bile metabolising gene abundance in Crohn's, ulcerative colitis and type 2 diabetes metagenomes. *PloS One* **9**: e115175.

Laffineur, G., Lescut, D., Vincent, P., Quandalle, P., Wurtz, A., and Colombel, J.F. (1992) [Bacterial translocation in Crohn disease]. *Gastroentérologie Clin Biol* **16**: 777–781.

Landers, C.J., Cohavy, O., Misra, R., Yang, H., Lin, Y., Braun, J., and Targan, S.R. (2002) Selected loss of tolerance evidenced by Crohn's disease–associated immune responses to auto- and microbial antigens. *Gastroenterology* **123**: 689–699.

Lane, M.C., Lockatell, V., Monterosso, G., Lamphier, D., Weinert, J., Hebel, J.R., *et al.* (2005) Role of Motility in the Colonization of Uropathogenic *Escherichia coli* in the Urinary Tract. *Infect Immun* **73**: 7644–7656.

Lapaquette, P., Bringer, M.-A., and Darfeuille-Michaud, A. (2012) Defects in autophagy favour adherent-invasive *Escherichia coli* persistence within macrophages leading to increased pro-inflammatory response. *Cell Microbiol* **14**: 791–807.

Lapaquette, P., Glasser, A.-L., Huett, A., Xavier, R.J., and Darfeuille-Michaud, A. (2010) Crohn's disease-associated adherent-invasive *E. coli* are selectively favoured by impaired autophagy to replicate intracellularly. *Cell Microbiol* **12**: 99–113.

Lapidus, A., Akerlund, J.-E., and Einarsson, C. (2006) Gallbladder bile composition in patients with Crohn 's disease. *World J Gastroenterol WJG* **12**: 70–74.

Lapidus, A., and Einarsson, C. (1998) Bile composition in patients with ileal resection due to Crohn's disease. *Inflamm Bowel Dis* **4**: 89–94.

Larsson, J.M.H., Karlsson, H., Sjövall, H., and Hansson, G.C. (2009) A complex, but uniform O-glycosylation of the human MUC2 mucin from colonic biopsies analyzed by nanoLC/MSn. *Glycobiology* **19**: 756–766.

Lawhon, S.D., Frye, J.G., Suyemoto, M., Porwollik, S., McClelland, M., and Altier, C. (2003) Global regulation by CsrA in *Salmonella typhimurium*. *Mol Microbiol* **48**: 1633–1645.

Lecointre, G., Rachdi, L., Darlu, P., and Denamur, E. (1998) *Escherichia coli* molecular phylogeny using the incongruence length difference test. *Mol Biol Evol* **15**: 1685–1695.

Lee, E.C. (1984) Aim of surgical treatment of Crohn's disease. *Gut* **25**: 217–222.

Leo, J.C., Grin, I., and Linke, D. (2012) Type V secretion: mechanism(s) of autotransport through the bacterial outer membrane. *Phil Trans R Soc B* **367**: 1088–1101.

Lepage, P., Colombet, J., Marteau, P., Sime-Ngando, T., Doré, J., and Leclerc, M. (2008) Dysbiosis in inflammatory bowel disease: a role for bacteriophages? *Gut* **57**: 424–425.

Lepage, P., Seksik, P., Sutren, M., Cochetière, M.-F. de la, Jian, R., Marteau, P., and Doré, J. (2005) Biodiversity of the mucosa-associated microbiota is stable along the distal digestive tract in healthy individuals and patients with IBD. *Inflamm Bowel Dis* **11**: 473–480.

Levine, B., and Deretic, V. (2007) Unveiling the roles of autophagy in innate and adaptive immunity. *Nat Rev Immunol* **7**: 767–777.

Lewis, J.D. (2011) The Utility of Biomarkers in the Diagnosis and Therapy of Inflammatory Bowel Disease. *Gastroenterology* **140**: 1817–1826.e2.

Leyton, D.L., Rossiter, A.E., and Henderson, I.R. (2012) From self sufficiency to dependence: mechanisms and factors important for autotransporter biogenesis. *Nat Rev Microbiol* **10**: 213–225.

Leyton, D.L., Sloan, J., Hill, R.E., Doughty, S., and Hartland, E.L. (2003) Transfer Region of pO113 from Enterohemorrhagic *Escherichia coli*: Similarity with R64 and Identification of a Novel Plasmid-Encoded Autotransporter, EpeA. *Infect Immun* **71**: 6307–6319.

Li, C.-M., Adler, K.B., and Cheng, P.-W. (1998) Mucin Biosynthesis: Molecular Cloning and Expression of Bovine Lung Mucin Core 2 N-Acetylglucosaminyltransferase cDNA. *Am J Respir Cell Mol Biol* **18**: 343–352.

Li, Q., Wang, C., Tang, C., He, Q., Li, N., and Li, J. (2014) Dysbiosis of Gut Fungal Microbiota is Associated With Mucosal Inflammation in Crohn's Disease. *J Clin Gastroenterol* **48**: 513–523.

Liew, A., O'Brien, T., and Egan, L. (2014) Mesenchymal Stromal Cell Therapy for Crohn's Disease. *Dig Dis* **32**: 50–60.

Lin, J., Sahin, O., Overbye Michel, L., and Zhang, Q. (2003) Critical Role of Multidrug Efflux Pump CmeABC in Bile Resistance and In Vivo Colonization of *Campylobacter jejuni*. *Infect Immun* **71**: 4250–4259.

Lindler, L.E., Klempner, M.S., and Straley, S.C. (1990) Yersinia pestis pH 6 antigen: genetic, biochemical, and virulence characterization of a protein involved in the pathogenesis of bubonic plague. *Infect Immun* **58**: 2569–2577.

Linke, D., Riess, T., Autenrieth, I.B., Lupas, A., and Kempf, V.A.J. (2006) Trimeric autotransporter adhesins: variable structure, common function. *Trends Microbiol* **14**: 264–270.

Liu, B., Gulati, A.S., Cantillana, V., Henry, S.C., Schmidt, E.A., Daniell, X., *et al.* (2013) Irgm1-deficient mice exhibit Paneth cell abnormalities and increased susceptibility to acute intestinal inflammation. *Am J Physiol - Gastrointest Liver Physiol* **305**: G573–G584.

Liu, J.Z., Sommeren, S. van, Huang, H., Ng, S.C., Alberts, R., Takahashi, A., *et al.* (2015) Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nat Genet* **47**: 979–986.

Liu, X., and Ferenci, T. (2001) An analysis of multifactorial influences on the transcriptional control of ompF and ompC porin expression under nutrient limitation. *Microbiology* **147**: 2981–2989.

Liu, Y., Van Kruiningen, H.J., West, A.B., Cartun, R.W., Cortot, A., and Colombel, J.-F. (1995) Immunocytochemical evidence of *Listeria*, Escherichia coil, and *Streptococcus* antigens in Crohn's disease. *Gastroenterology* **108**: 1396–1404.

Loftus Jr, E.V. (2004) Clinical epidemiology of inflammatory bowel disease: incidence, prevalence, and environmental influences. *Gastroenterology* **126**: 1504–1517.

Louis, E., Collard, A., Oger, A., Degroote, E., El Yafi, F.A.N., and Belaiche, J. (2001) Behaviour of Crohn's disease according to the Vienna classification: changing pattern over the course of the disease. *Gut* **49**: 777–782.

Louis, E., Irving, P., and Beaugerie, L. (2014) Use of azathioprine in IBD: modern aspects of an old drug. *Gut* **63**: 1695–1699.

Loveless, B.J., and Saier, M.H. (1997) A novel family of channel-forming, autotransporting, bacterial virulence factors. *Mol Membr Biol* **14**: 113–123.

Lozupone, C.A., Stombaugh, J.I., Gordon, J.I., Jansson, J.K., and Knight, R. (2012) Diversity, stability and resilience of the human gut microbiota. *Nature* **489**: 220–230.

Lund, P., Tramonti, A., and Biase, D.D. (2014) Coping with low pH: molecular strategies in neutralophilic bacteria. *FEMS Microbiol Rev* **38**: 1091–1125.

M. Orholm, K.O.K., V. Binder, T.I.A. Sørensen, L.P. Rasmussen (2000) Concordance of Inflammatory Bowel Disease among Danish Twins: Results of a Nationwide Study. *Scand J Gastroenterol* **35**: 1075–1081.

Mackie, R.I., Sghir, A., and Gaskins, H.R. (1999) Developmental microbial ecology of the neonatal gastrointestinal tract. *Am J Clin Nutr* **69**: 1035s–1045s.

Macpherson, A., Khoo, U.Y., Forgacs, I., Philpott-Howard, J., and Bjarnason, I. (1996) Mucosal antibodies in inflammatory bowel disease are directed against intestinal bacteria. *Gut* **38**: 365–375.

Madara, J.L. (1982) Cup cells: structure and distribution of a unique class of epithelial cells in guinea pig, rabbit, and monkey small intestine. *Gastroenterology* **83**: 981–994.

Madi, A., Lakhdari, O., Blottière, H.M., Guyard-Nicodème, M., Le Roux, K., Groboillot, A., *et al.* (2010) The clinical *Pseudomonas fluorescens* MFN1032 strain exerts a cytotoxic effect on epithelial intestinal cells and induces Interleukin-8 via the AP-1 signaling pathway. *BMC Microbiol* **10**: 215.

Manichanh, C., Borruel, N., Casellas, F., and Guarner, F. (2012) The gut microbiota in IBD. *Nat Rev Gastroenterol Hepatol* **9**: 599–608.

Manichanh, C., Rigottier-Gois, L., Bonnaud, E., Gloux, K., Pelletier, E., Frangeul, L., *et al.* (2006) Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach. *Gut* **55**: 205–211.

Marchesi, J.R. (2011) Human distal gut microbiome. *Environ Microbiol* **13**: 3088–3102.

Marks, J.W., Conley, D.R., Capretta, T.L., Bonorris, G.G., Chung, A., Coyne, M.J., and Schoenfield, L.J. (1977) Gallstone prevalence and biliary lipid composition in inflammatory bowel disease. *Am J Dig Dis* **22**: 1097–1100.

Martin, H.M., Campbell, B.J., Hart, C.A., Mpofu, C., Nayar, M., Singh, R., *et al.* (2004) Enhanced *Escherichia coli* adherence and invasion in Crohn's disease and colon cancer. *Gastroenterology* **127**: 80–93.

Martin, J.A., and Wang, Z. (2011) Next-generation transcriptome assembly. *Nat Rev Genet* **12**: 671–682.

Martinez-Medina, M., Aldeguer, X., Lopez-Siles, M., González-Huix, F., López-Oliu, C., Dahbi, G., *et al.* (2009) Molecular diversity of *Escherichia coli* in the human gut: new ecological evidence supporting the role of adherent-invasive *E. coli* (AIEC) in Crohn's disease. *Inflamm Bowel Dis* **15**: 872–882.

Martinez-Medina, M., and Garcia-Gil, L.J. (2014) *Escherichia coli* in chronic inflammatory bowel diseases: An update on adherent invasive *Escherichia coli* pathogenicity. *World J Gastrointest Pathophysiol* **5**: 213–227.

Martinez-Medina, M., Mora, A., Blanco, M., López, C., Alonso, M.P., Bonacorsi, S., *et al.* (2009) Similarity and Divergence among Adherent-Invasive *Escherichia coli* and Extraintestinal Pathogenic *E. coli* Strains. *J Clin Microbiol* **47**: 3968–3979.

Martinez-Medina, M., Naves, P., Blanco, J., Aldeguer, X., Blanco, J.E., Blanco, M., *et al.* (2009) Biofilm formation as a novel phenotypic feature of adherent-invasive *Escherichia coli* (AIEC). *BMC Microbiol* **9**: 202.

Martini, G.A., and Brandes, J.W. (1976) Increased consumption of refined carbohydrates in patients with Crohn's disease. *Klin Wochenschr* **54**: 367–371.

Marzban, H., Del Bigio, M.R., Alizadeh, J., Ghavami, S., Zachariah, R.M., and Rastegar, M. (2015) Cellular commitment in the developing cerebellum. *Front Cell Neurosci* **8**: 450.

Massironi, S., Zilli, A., Cavalcoli, F., Conte, D., and Peracchi, M. (2016) Chromogranin A and other enteroendocrine markers in inflammatory bowel disease. *Neuropeptides* http://www.sciencedirect.com/science/article/pii/S0143417916000032. Accessed February 14, 2016.

Masson, L., and Holbein, B.E. (1985) Influence of nutrient limitation and low pH on serogroup B *Neisseria meningitidis* capsular polysaccharide levels: correlation with virulence for mice. *Infect Immun* **47**: 465–471.

Maurer, L.M., Yohannes, E., Bondurant, S.S., Radmacher, M., and Slonczewski, J.L. (2005) pH Regulates Genes for Flagellar Motility, Catabolism, and Oxidative Stress in *Escherichia coli* K-12. *J Bacteriol* **187**: 304–319.

Mazar, J., and Cotter, P.A. (2006) Topology and maturation of filamentous haemagglutinin suggest a new model for two-partner secretion. *Mol Microbiol* **62**: 641–654.

Mazumdar, S., Clomburg, J.M., and Gonzalez, R. (2010) *Escherichia coli* Strains Engineered for Homofermentative Production of d-Lactic Acid from Glycerol. *Appl Environ Microbiol* **76**: 4327–4336.

McCarroll, S.A., Huett, A., Kuballa, P., Chilewski, S.D., Landry, A., Goyette, P., *et al.* (2008) Deletion polymorphism upstream of IRGM associated with altered IRGM expression and Crohn's disease. *Nat Genet* **40**: 1107–1112.

McClure, R., Balasubramanian, D., Sun, Y., Bobrovskyy, M., Sumby, P., Genco, C.A., *et al.* (2013) Computational analysis of bacterial RNA-Seq data. *Nucleic Acids Res* **41**: e140–e140.

McCracken, V.J., and Lorenz, R.G. (2001) The gastrointestinal ecosystem: a precarious alliance among epithelium, immunity and microbiota. *Cell Microbiol* **3**: 1–11.

McDermott, A.J., and Huffnagle, G.B. (2014) The microbiome and regulation of mucosal immunity. *Immunology* **142**: 24–31.

McGovern, D.P.B., Jones, M.R., Taylor, K.D., Marciante, K., Yan, X., Dubinsky, M., *et al.* (2010) Fucosyltransferase 2 (FUT2) non-secretor status is associated with Crohn's disease. *Hum Mol Genet* **19**: 3468–3476.

McGovern, D.P.B., Rotter, J.I., Mei, L., Haritunians, T., Landers, C., Derkowski, C., *et al.* (2009) Genetic epistasis of IL23/IL17 pathway genes in Crohn's disease. *Inflamm Bowel Dis* **15**: 883–889.

McGuckin, M.A., Lindén, S.K., Sutton, P., and Florin, T.H. (2011) Mucin dynamics and enteric pathogens. *Nat Rev Microbiol* **9**: 265–278.

Meconi, S., Vercellone, A., Levillain, F., Payré, B., Al Saati, T., Capilla, F., *et al.* (2007) Adherent-invasive *Escherichia coli* isolated from Crohn's disease patients induce granulomas in vitro. *Cell Microbiol* **9**: 1252–1261.

Meng, G., Spahich, N., Kenjale, R., Waksman, G., and St Geme, J.W. (2011) Crystal structure of the *Haemophilus influenzae* Hap adhesin reveals an intercellular oligomerization mechanism for bacterial aggregation. *EMBO J* **30**: 3864–3874.

Merlin, D., Si-Tahar, M., Sitaraman, S.V., Eastburn, K., Williams, I., Liu, X., *et al.* (2001) Colonic epithelial hPepT1 expression occurs in inflammatory bowel disease: Transport of bacterial peptides influences expression of MHC class 1 molecules. *Gastroenterology* **120**: 1666–1679.

Merritt, M.E., and Donaldson, J.R. (2009) Effect of bile salts on the DNA and membrane integrity of enteric bacteria. *J Med Microbiol* **58**: 1533–1541.

Meuwis, M.-A., Fillet, M., Geurts, P., Seny, D. de, Lutteri, L., Chapelle, J.-P., *et al.* (2007) Biomarker discovery for inflammatory bowel disease, using proteomic serum profiling. *Biochem Pharmacol* **73**: 1422–1433.

Miquel, S., Claret, L., Bonnet, R., Dorboz, I., Barnich, N., and Darfeuille-Michaud, A. (2010) Role of decreased levels of Fis histone-like protein in Crohn's disease-associated adherent invasive *Escherichia coli* LF82 bacteria interacting with intestinal epithelial cells. *J Bacteriol* **192**: 1832–1843.

Miquel, S., Martín, R., Rossi, O., Bermúdez-Humarán, L.G., Chatel, J.M., Sokol, H., *et al.* (2013) *Faecalibacterium prausnitzii* and human intestinal health. *Curr Opin Microbiol* **16**: 255–261.

Miquel, S., Peyretaillade, E., Claret, L., Vallée, A. de, Dossat, C., Vacherie, B., *et al.* (2010) Complete Genome Sequence of Crohn's Disease-Associated Adherent-Invasive *E. coli* Strain LF82. *PLoS ONE* **5**: e12714.

Miyamoto, H., Tanaka, T., Kitamoto, N., Fukuda, Y., and Shimoyama, T. (1995) Detection of immunoreactive antigen, with a monoclonal antibody to measles virus, in tissue from a patient with Crohn's disease. *J Gastroenterol* **30**: 28–33.

Mizuno, T. (1997) Compilation of All Genes Encoding Two-component Phosphotransfer Signal Transducers in the Genome of *Escherichia coli*. *DNA Res* **4**: 161–168.

Mizushima, N. (2007) Autophagy: process and function. *Genes Dev* **21**: 2861–2873.

Molodecky, N.A., Soon, I.S., Rabi, D.M., Ghali, W.A., Ferris, M., Chernoff, G., *et al.* (2012) Increasing Incidence and Prevalence of the Inflammatory Bowel Diseases With Time, Based on Systematic Review. *Gastroenterology* **142**: 46–54.e42.

Moniaux, N., Escande, F., Porchet, N., Aubert, J.P., and Batra, S.K. (2001) Structural organization and classification of the human mucin genes. *Front Biosci J Virtual Libr* **6**: D1192-1206.

Monsén, U., Bernell, O., Johansson, C., and Hellers, G. (1991) Prevalence of Inflammatory Bowel Disease among Relatives of Patients with Crohn's Disease. *Scand J Gastroenterol* **26**: 302–306.

Morgan, X.C., Segata, N., and Huttenhower, C. (2013) Biodiversity and functional genomics in the human microbiome. *Trends Genet* **29**: 51–58.

Morson, B.C. (1972) The early histological lesion of Crohn's disease. *Proc R Soc Med* **65**: 71–72.

Mosli, M.H., Zou, G., Garg, S.K., Feagan, S.G., MacDonald, J.K., Chande, N., *et al.* (2015) C-Reactive Protein, Fecal Calprotectin, and Stool Lactoferrin for Detection of Endoscopic Activity in Symptomatic Inflammatory Bowel Disease Patients: A Systematic Review and Meta-Analysis. *Am J Gastroenterol* **110**: 802–819.

Mougous, J.D., Cuff, M.E., Raunser, S., Shen, A., Zhou, M., Gifford, C.A., *et al.* (2006) A Virulence Locus of *Pseudomonas aeruginosa* Encodes a Protein Secretion Apparatus. *Science* **312**: 1526–1530.

Mow, W.S., Vasiliauskas, E.A., Lin, Y.-C., Fleshner, P.R., Papadakis, K.A., Taylor, K.D., *et al.* (2004) Association of antibody responses to microbial antigens and complications of small bowel Crohn's disease. *Gastroenterology* **126**: 414–424.

Münch, R., Hiller, K., Grote, A., Scheer, M., Klein, J., Schobert, M., and Jahn, D. (2005) Virtual Footprint and PRODORIC: an integrative framework for regulon prediction in prokaryotes. *Bioinformatics* **21**: 4187–4189.

Musso, G., Gambino, R., and Cassader, M. (2011) Interactions Between Gut Microbiota and Host Metabolism Predisposing to Obesity and Diabetes. *Annu Rev Med* **62**: 361–380.

Nakagawa, I., Amano, A., Mizushima, N., Yamamoto, A., Yamaguchi, H., Kamimoto, T., *et al.* (2004) Autophagy Defends Cells Against Invading Group A Streptococcus. *Science* **306**: 1037–1040.

Nakagome, S., Mano, S., Kozlowski, L., Bujnicki, J.M., Shibata, H., Fukumaki, Y., *et al.* (2012) Crohn's Disease Risk Alleles on the NOD2 Locus Have Been Maintained by Natural Selection on Standing Variation. *Mol Biol Evol* **29**: 1569–1585.

Nakamura, R.M., Matsutani, M., and Barry, M. (2003) Advances in clinical laboratory tests for inflammatory bowel disease. *Clin Chim Acta* **335**: 9–20.

Nandan, M.O., Ghaleb, A.M., Liu, Y., Bialkowska, A.B., McConnell, B.B., Shroyer, K.R., *et al.* (2014) Inducible intestine-specific deletion of Krüppel-like factor 5 is characterized by a regenerative response in adult mouse colon. *Dev Biol* **387**: 191–202.

Naser, S.A., Ghobrial, G., Romero, C., and Valentine, J.F. (2004) Culture of *Mycobacterium avium* subspecies paratuberculosis from the blood of patients with Crohn's disease. *The Lancet* **364**: 1039–1044.

Nava, G.M., and Stappenbeck, T.S. (2011) Diversity of the autochthonous colonic microbiota. *Gut Microbes* **2**: 99–104.

Navarro-Garcia, F., Gutierrez-Jimenez, J., Garcia-Tovar, C., Castro, L.A., Salazar-Gonzalez, H., and Cordova, V. (2010) Pic, an Autotransporter Protein Secreted by Different Pathogens in the *Enterobacteriaceae* Family, Is a Potent Mucus Secretagogue. *Infect Immun* **78**: 4101–4109.

Neut, C., Bulois, P., Desreumaux, P., Membré, J.-M., Lederman, E., Gambiez, L., *et al.* (2002) Changes in the bacterial flora of the neoterminal ileum after ileocolonic resection for Crohn's disease. *Am J Gastroenterol* **97**: 939–946.

Neutra, M., and Leblond, C.P. (1966) Synthesis of the carbohydrate of mucus in the golgi complex as shown by electron microscope radioautography of goblet cells from rats injected with glucose-H3. *J Cell Biol* **30**: 119–136.

Newton, W.A., and Snell, E.E. (1964) Catalytic properties of tryptophanase, a multifunctional pyridoxal phosphate enzyme. *Proc Natl Acad Sci U S A* **51**: 382–389.

Nguyen, H.T.T., Dalmasso, G., Müller, S., Carrière, J., Seibold, F., and Darfeuille-Michaud, A. (2014) Crohn's disease-associated adherent invasive *Escherichia coli* modulate levels of microRNAs in intestinal epithelial cells to reduce autophagy. *Gastroenterology* **146**: 508–519.

Nguyen, H.T.T., Lapaquette, P., Bringer, M.-A., and Darfeuille-Michaud, A. (2013) Autophagy and Crohn's disease. *J Innate Immun* **5**: 434–443.

Niessner, M., and Volk, B.A. (1995) Altered Th1/Th2 cytokine profiles in the intestinal mucosa of patients with inflammatory bowel disease as assessed by quantitative reversed transcribed polymerase chain reaction (RT-PCR). *Clin Exp Immunol* **101**: 428–435.

Nigro, G., Rossi, R., Commere, P.-H., Jay, P., and Sansonetti, P.J. (2014) The Cytosolic Bacterial Peptidoglycan Sensor Nod2 Affords Stem Cell Protection and Links Microbes to Gut Epithelial Regeneration. *Cell Host Microbe* **15**: 792–798.

Nikolaus, S., and Schreiber, S. (2007) Diagnostics of Inflammatory Bowel Disease. *Gastroenterology* **133**: 1670–1689.

Noah, T.K., Donahue, B., and Shroyer, N.F. (2011) Intestinal development and differentiation. *Exp Cell Res* **317**: 2702–2710.

Noh, D.O., and Gilliland, S.E. (1993) Influence of Bile on Cellular Integrity and β-Galactosidase Activity of *Lactobacillus acidophilus*1. *J Dairy Sci* **76**: 1253–1259.

Norman, J.M., Handley, S.A., Baldridge, M.T., Droit, L., Liu, C.Y., Keller, B.C., *et al.* (2015) Disease-specific Alterations in the Enteric Virome in Inflammatory Bowel Disease. *Cell* **160**: 447–460.

Northfield, T.C., and McColl, I. (1973) Postprandial concentrations of free and conjugated bile acids down the length of the normal human small intestine. *Gut* **14**: 513–518.

Noverr, M.C., and Huffnagle, G.B. (2005) The "microflora hypothesis" of allergic diseases. *Clin Exp Allergy* **35**: 1511–1520.

Nugent, S.G., Kumar, D., Rampton, D.S., and Evans, D.F. (2001) Intestinal luminal pH in inflammatory bowel disease: possible determinants and implications for therapy with aminosalicylates and other drugs. *Gut* **48**: 571–577.

Nummelin, H., Merckel, M.C., Leo, J.C., Lankinen, H., Skurnik, M., and Goldman, A. (2004) The Yersinia adhesin YadA collagen-binding domain structure is a novel left-handed parallel beta-roll. *EMBO J* **23**: 701–711.

Nyangale, E.P., Mottram, D.S., and Gibson, G.R. (2012) Gut microbial activity, implications for health and disease: the potential role of metabolite analysis. *J Proteome Res* **11**: 5573–5585.

Obradors, N., Badía, J., Baldomà, L., and Aguilar, J. (1988) Anaerobic metabolism of the L-rhamnose fermentation product 1,2-propanediol in *Salmonella typhimurium*. *J Bacteriol* **170**: 2159–2162.

Ofek, I., Mirelman, D., and Sharon, N. (1977) Adherence of *Escherichia coli* to human mucosal cells mediated by mannose receptors. *Nature* **265**: 623–625.

Ogata, H., and Podolsky, D.K. (1997) Trefoil peptide expression and secretion is regulated by neuropeptides and acetylcholine. *Am J Physiol - Gastrointest Liver Physiol* **273**: G348–G354.

Ogura, Y., Bonen, D.K., Inohara, N., Nicolae, D.L., Chen, F.F., Ramos, R., *et al.* (2001) A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. *Nature* **411**: 603–606.

O'Hara, A.M., and Shanahan, F. (2006) The gut flora as a forgotten organ. *EMBO Rep* **7**: 688–693.

Oliver, D.B., and Beckwith, J. (1982) Identification of a new gene (secA) and gene product involved in the secretion of envelope proteins in *Escherichia coli*. *J Bacteriol* **150**: 686–691.

Oliver, D.C., Huang, G., Nodel, E., Pleasance, S., and Fernandez, R.C. (2003) A conserved region within the *Bordetella pertussis* autotransporter BrkA is necessary for folding of its passenger domain. *Mol Microbiol* **47**: 1367–1383.

Oppmann, B., Lesley, R., Blom, B., Timans, J.C., Xu, Y., Hunte, B., *et al.* (2000) Novel p19 Protein Engages IL-12p40 to Form a Cytokine, IL-23, with Biological Activities Similar as Well as Distinct from IL-12. *Immunity* **13**: 715–725.

Orholm, M., Munkholm, P., Langholz, E., Nielsen, O.H., Sφrensen, T.I.A., and Binder, V. (1991) Familial Occurrence of Inflammatory Bowel Disease. *N Engl J Med* **324**: 84–88.

O'Toole, G.A., and Kolter, R. (1998) Initiation of biofilm formation in *Pseudomonas fluorescens* WCS365 proceeds via multiple, convergent signalling pathways: a genetic analysis. *Mol Microbiol* **28**: 449–461.

Ott, C., and Schölmerich, J. (2013) Extraintestinal manifestations and complications in IBD. *Nat Rev Gastroenterol Hepatol* **10**: 585–595.

Otto, B.R., Sijbrandi, R., Luirink, J., Oudega, B., Heddle, J.G., Mizutani, K., *et al.* (2005) Crystal Structure of Hemoglobin Protease, a Heme Binding Autotransporter Protein from Pathogenic *Escherichia coli*. *J Biol Chem* **280**: 17339–17345.

Ouwerkerk, J.P., Vos, W.M. de, and Belzer, C. (2013) Glycobiome: Bacteria and mucus at the epithelial interface. *Best Pract Res Clin Gastroenterol* **27**: 25–38.

Paerregaard, A., Espersen, F., Jensen, O.M., and Skurnik, M. (1991) Interactions between *Yersinia enterocolitica* and rabbit ileal mucus: growth, adhesion, penetration, and subsequent changes in surface hydrophobicity and ability to adhere to ileal brush border membrane vesicles. *Infect Immun* **59**: 253–260.

Palanichamy, J.K., and Rao, D.S. (2014) miRNA dysregulation in cancer: towards a mechanistic understanding. *Front Genet* **5** http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3957189/. Accessed March 6, 2016.

Papillon, E., Bonaz, B., and Fournet, J. (1999) [Short chain fatty acids: effects on gastrointestinal function and therapeutic potential in gastroenterology]. *Gastroentérologie Clin Biol* **23**: 761–769.

Parker, C.T., and Guard-Petter, J. (2001) Contribution of flagella and invasion proteins to pathogenesis of *Salmonella enterica* serovar enteritidis in chicks. *FEMS Microbiol Lett* **204**: 287–291.

Parkes, M., Barrett, J.C., Prescott, N.J., Tremelling, M., Anderson, C.A., Fisher, S.A., *et al.* (2007) Sequence variants in the autophagy gene IRGM and multiple other replicating loci contribute to Crohn's disease susceptibility. *Nat Genet* **39**: 830–832.

Parreira, V.R., and Gyles, C.L. (2003) A Novel Pathogenicity Island Integrated Adjacent to the thrW tRNA Gene of Avian Pathogenic *Escherichia coli* Encodes a Vacuolating Autotransporter Toxin. *Infect Immun* **71**: 5087–5096.

Pavlova, O., Peterson, J.H., Ieva, R., and Bernstein, H.D. (2013) Mechanistic link between barrel assembly and the initiation of autotransporter secretion. *Proc Natl Acad Sci* **110**: E938–E947.

Payne, D., Tatham, D., Williamson, E.D., and Titball, R.W. (1998) The pH 6 Antigen of Yersinia pestis Binds to β1-Linked Galactosyl Residues in Glycosphingolipids. *Infect Immun* **66**: 4545–4548.

Pelaseyed, T., Bergström, J.H., Gustafsson, J.K., Ermund, A., Birchenough, G.M.H., Schütte, A., *et al.* (2014) The mucus and mucins of the goblet cells and enterocytes provide the first defense line of the gastrointestinal tract and interact with the immune system. *Immunol Rev* **260**: 8–20.

Peluso, I., Pallone, F., and Monteleone, G. (2006) Interleukin-12 and Th1 immune response in Crohn's disease: Pathogenetic relevance and therapeutic inplication. *World J Gastroenterol WJG* **12**: 5606–5610.

Pereira, S.P., Bain, I.M., Kumar, D., and Dowling, R.H. (2003) Bile composition in inflammatory bowel disease: ileal disease and colectomy, but not colitis, induce lithogenic bile. *Aliment Pharmacol Ther* **17**: 923–933.

Peterson, J.H., Tian, P., Ieva, R., Dautin, N., and Bernstein, H.D. (2010) Secretion of a bacterial virulence factor is driven by the folding of a C-terminal segment. *Proc Natl Acad Sci U S A* **107**: 17739–17744.

Philpott, D.J., and Viala, J. (2004) Towards an understanding of the role of NOD2/CARD15 in the pathogenesis of Crohn's disease. *Best Pract Res Clin Gastroenterol* **18**: 555–568.

Picard, B., Garcia, J.S., Gouriou, S., Duriez, P., Brahimi, N., Bingen, E., *et al.* (1999) The Link between Phylogeny and Virulence in *Escherichia coli* Extraintestinal Infection. *Infect Immun* **67**: 546–553.

Pironi, L., Bonvicini, F., Gionchetti, P., D'Errico, A., Rizzello, F., Corsini, C., *et al.* (2009) Parvovirus B19 Infection Localized in the Intestinal Mucosa and Associated with Severe Inflammatory Bowel Disease. *J Clin Microbiol* **47**: 1591–1595.

Pohlner, J., Halter, R., Beyreuther, K., and Meyer, T.F. (1987) Gene structure and extracellular secretion of *Neisseria gonorrhoeae* IgA protease. *Nature* **325**: 458–462.

Pope, L.M., Reed, K.E., and Payne, S.M. (1995) Increased protein secretion and adherence to HeLa cells by *Shigella spp*. following growth in the presence of bile salts. *Infect Immun* **63**: 3642–3648.

Prantera, C. (2006) Probiotics for Crohn's disease: what have we learned? *Gut* **55**: 757–759.

Prantera, C., Lochs, H., Campieri, M., Scribano, M.L., Sturniolo, G.C., Castiglione, F., and Cottone, M. (2006) Antibiotic treatment of Crohn's disease: results of a multicentre, double blind, randomized, placebo-controlled trial with rifaximin. *Aliment Pharmacol Ther* **23**: 1117–1125.

Prantera, C., and Scribano, M.L. (2009) Antibiotics and probiotics in inflammatory bowel disease: why, when, and how: *Curr Opin Gastroenterol* **25**: 329–333.

Prantera, C., Zannoni, F., Scribano, M.L., Berto, E., Andreoli, A., Kohn, A., and Luzi, C. (1996) An antibiotic regimen for the treatment of active Crohn's disease: a randomized, controlled clinical trial of metronidazole plus ciprofloxacin. *Am J Gastroenterol* **91**: 328–332.

Prescott, N.J., Dominy, K.M., Kubo, M., Lewis, C.M., Fisher, S.A., Redon, R., *et al.* (2010) Independent and population-specific association of risk variants at the IRGM locus with Crohn's disease. *Hum Mol Genet* **19**: 1828–1839.

Present, D. h. (1999) Review article: the efficacy of infliximab in Crohn's disease — healing of fistulae. *Aliment Pharmacol Ther* **13**: 23–28.

Prigent-Combaret, C., Prensier, G., Le Thi, T.T., Vidal, O., Lejeune, P., and Dorel, C. (2000) Developmental pathway for biofilm formation in curli-producing *Escherichia coli* strains: role of flagella, curli and colanic acid. *Environ Microbiol* **2**: 450–464.

Prouty, A.M., Brodsky, I.E., Manos, J., Belas, R., Falkow, S., and Gunn, J.S. (2004) Transcriptional regulation of *Salmonella enterica* serovar Typhimurium genes by bile. *FEMS Immunol Med Microbiol* **41**: 177–185.

Puertollano, E., Kolida, S., and Yaqoob, P. (2014) Biological significance of short-chain fatty acid metabolism by the intestinal microbiome. *Curr Opin Clin Nutr Metab Care* **17**: 139–144.

Pukatzki, S., Ma, A.T., Revel, A.T., Sturtevant, D., and Mekalanos, J.J. (2007) Type VI secretion system translocates a phage tail spike-like protein into target cells where it cross-links actin. *Proc Natl Acad Sci U S A* **104**: 15508–15513.

Pukatzki, S., Ma, A.T., Sturtevant, D., Krastins, B., Sarracino, D., Nelson, W.C., *et al.* (2006) Identification of a conserved bacterial protein secretion system in *Vibrio cholerae* using the Dictyostelium host model system. *Proc Natl Acad Sci U S A* **103**: 1528–1533.

Pukatzki, S., McAuley, S.B., and Miyata, S.T. (2009) The type VI secretion system: translocation of effectors and effector-domains. *Curr Opin Microbiol* **12**: 11–17.

Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K.S., Manichanh, C., *et al.* (2010) A human gut microbial gene catalog established by metagenomic sequencing. *Nature* **464**: 59–65.

Quinton, J., Sendid, B., Reumaux, D., Duthilleul, P., Cortot, A., Grandbastien, B., *et al.* (1998) Anti-*Saccharomyces cerevisiae* mannan antibodies combined with antineutrophil cytoplasmic autoantibodies in inflammatory bowel disease: prevalence and diagnostic role. *Gut* **42**: 788–791.

Radhakrishnan, P., Halagowder, D., and Devaraj, S.N. (2007) Altered expression of MUC2 and MUC5AC in response to Shigella infection, an in vivo study. *Biochim Biophys Acta BBA - Gen Subj* **1770**: 884–889.

Randow, F., and Seed, B. (2001) Endoplasmic reticulum chaperone gp96 is required for innate immunity but not cell viability. *Nat Cell Biol* **3**: 891–896.

Reen, F.J., Woods, D.F., Mooij, M.J., Adams, C., and O'Gara, F. (2012) Respiratory Pathogens Adopt a Chronic Lifestyle in Response to Bile. *PLoS ONE* **7** http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3458808/. Accessed July 3, 2016.

Reid, G., Gaudier, E., Guarner, F., Huffnagle, G.B., Macklaim, J.M., Munoz, A.M., *et al.* (2010) Responders and non-responders to probiotic interventions. *Gut Microbes* **1**: 200–204.

Reif, S., Klein, I., Lubin, F., Farbstein, M., Hallak, A., and Gilat, T. (1997) Pre-illness dietary factors in inflammatory bowel disease. *Gut* **40**: 754–760.

Reyes, A., Haynes, M., Hanson, N., Angly, F.E., Heath, A.C., Rohwer, F., and Gordon, J.I. (2010) Viruses in the fecal microbiota of monozygotic twins and their mothers. *Nature* **466**: 334–338.

Rhee, S.H., Pothoulakis, C., and Mayer, E.A. (2009) Principles and clinical implications of the brain–gut–enteric microbiota axis. *Nat Rev Gastroenterol Hepatol* **6**: 306–314.

Rhodes, J.M., Campbell, B.J., and Yu, L.-G. (2008) Lectin-epithelial interactions in the human colon. *Biochem Soc Trans* **36**: 1482–1486.

Ridlon, J.M., Kang, D.-J., and Hylemon, P.B. (2006) Bile salt biotransformations by human intestinal bacteria. *J Lipid Res* **47**: 241–259.

Rigas, A., Rigas, B., Glassman, M., Yen, Y.-Y., Lan, S.J., Petridou, E., *et al.* (1993) Breast-feeding and maternal smoking in the etiology of Crohn's disease and ulcerative colitis in childhood. *Ann Epidemiol* **3**: 387–392.

Riley, P.A. (2004) Bacteriophages in autoimmune disease and other inflammatory conditions. *Med Hypotheses* **62**: 493–498.

Rio, M.-C. (2002) Fonctions paradoxales pour TFF1. *médecine/sciences* **18**: 1177–1178.

Rioux, J.D., Xavier, R.J., Taylor, K.D., Silverberg, M.S., Goyette, P., Huett, A., *et al.* (2007) Genome-wide association study identifies five novel susceptibility loci for Crohn's disease and implicates a role for autophagy in disease pathogenesis. *Nat Genet* **39**: 596–604.

Roda, G., Sartini, A., Zambon, E., Calafiore, A., Marocchi, M., Caponi, A., *et al.* (2010) Intestinal epithelial cells in inflammatory bowel diseases. *World J Gastroenterol WJG* **16**: 4264–4271.

Rolhion, N., Barnich, N., Bringer, M.-A., Glasser, A.-L., Ranc, J., Hébuterne, X., *et al.* (2010) Abnormally expressed ER stress response chaperone Gp96 in CD favours adherent-invasive *Escherichia coli* invasion. *Gut* **59**: 1355–1362.

Rolhion, N., Barnich, N., Claret, L., and Darfeuille-Michaud, A. (2005) Strong Decrease in Invasive Ability and Outer Membrane Vesicle Release in Crohn's Disease-Associated Adherent-Invasive *Escherichia coli* Strain LF82 with the yfgL Gene Deleted. *J Bacteriol* **187**: 2286–2296.

Rolhion, N., Carvalho, F.A., and Darfeuille-Michaud, A. (2007) OmpC and the σE regulatory pathway are involved in adhesion and invasion of the Crohn's disease-associated *Escherichia coli* strain LF82. *Mol Microbiol* **63**: 1684–1700.

Rosenberg, E.Y., Bertenthal, D., Nilles, M.L., Bertrand, K.P., and Nikaido, H. (2003) Bile salts and fatty acids induce the expression of *Escherichia coli* AcrAB multidrug efflux pump through their interaction with Rob regulatory protein. *Mol Microbiol* **48**: 1609–1619.

Round, J.L., and Mazmanian, S.K. (2009) The gut microbiome shapes intestinal immune responses during health and disease. *Nat Rev Immunol* **9**: 313–323.

Roush, S., and Slack, F.J. (2008) The let-7 family of microRNAs. *Trends Cell Biol* **18**: 505–516.

Royle, L., Roos, A., Harvey, D.J., Wormald, M.R., Gijlswijk-Janssen, D. van, Redwan, E.-R.M., *et al.* (2003) Secretory IgA N- and O-glycans provide a link between the innate and adaptive immune systems. *J Biol Chem* **278**: 20140–20153.

Ruiz-Perez, F., Henderson, I.R., Leyton, D.L., Rossiter, A.E., Zhang, Y., and Nataro, J.P. (2009) Roles of periplasmic chaperone proteins in the biogenesis of serine protease autotransporters of *Enterobacteriaceae*. *J Bacteriol* **191**: 6571–6583.

Ruiz-Perez, F., and Nataro, J.P. (2014) Bacterial serine proteases secreted by the autotransporter pathway: classification, specificity, and role in virulence. *Cell Mol Life Sci CMLS* **71**: 745–770.

Ruiz-Perez, F., Wahid, R., Faherty, C.S., Kolappaswamy, K., Rodriguez, L., Santiago, A., *et al.* (2011) Serine protease autotransporters from *Shigella flexneri* and pathogenic *Escherichia coli* target a broad range of leukocyte glycoproteins. *Proc Natl Acad Sci U S A* **108**: 12881–12886.

Rupaimoole, R., Calin, G.A., Lopez-Berestein, G., and Sood, A.K. (2016) miRNA Deregulation in Cancer Cells and the Tumor Microenvironment. *Cancer Discov* http://cancerdiscovery.aacrjournals.org/content/early/2016/02/09/2159-8290.CD-15-0893. Accessed March 6, 2016.

Russell, A.B., Hood, R.D., Bui, N.K., LeRoux, M., Vollmer, W., and Mougous, J.D. (2011) Type VI secretion delivers bacteriolytic effectors to target cells. *Nature* **475**: 343–347.

Rutgeerts, P., D'Haens, G., Targan, S., Vasiliauskas, E., Hanauer, S.B., Present, D.H., *et al.* (1999) Efficacy and safety of retreatment with anti–tumor necrosis factor antibody (infliximab) to maintain remission in Crohn's disease. *Gastroenterology* **117**: 761–769.

Rutgeerts, P., Ghoos, Y., Vantrappen, G., and Fevery, J. (1986) Biliary lipid composition in patients with nonoperated Crohn's disease. *Dig Dis Sci* **31**: 27–32.

Saebo, A., Vik, E., Lange, O.J., and Matuszkiewicz, L. (2005) Inflammatory bowel disease associated with *Yersinia enterocolitica* O:3 infection. *Eur J Intern Med* **16**: 176–182.

Salacha, R., Kovacić, F., Brochier-Armanet, C., Wilhelm, S., Tommassen, J., Filloux, A., *et al.* (2010) The *Pseudomonas aeruginosa* patatin-like protein PlpD is the archetype of a novel Type V secretion system. *Environ Microbiol* **12**: 1498–1512.

Salim, S.Y., Silva, M.A., Keita, Å.V., Larsson, M., Andersson, P., Magnusson, K.-E., *et al.* (2009) CD83+CCR7− Dendritic Cells Accumulate in the Subepithelial Dome and Internalize Translocated *Escherichia coli* HB101 in the Peyer's Patches of Ileal Crohn's Disease. *Am J Pathol* **174**: 82–90.

Sanchez, B., Saad, N., Schmitter, J.-M., Bressollier, P., and Urdaci, M.C. (2010) Adhesive properties, extracellular protein production, and metabolism in the *Lactobacillus rhamnosus* GG strain when grown in the presence of mucin. *J Microbiol Biotechnol* **20**: 978–984.

Sanchez, L.M., Cheng, A.T., Warner, C.J.A., Townsley, L., Peach, K.C., Navarro, G., *et al.* (2016) Biofilm Formation and Detachment in Gram-Negative Pathogens Is Modulated by Select Bile Acids. *PLoS ONE* **11** http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4798295/. Accessed July 1, 2016.

Sandborn, W.J. (2007) Clinical perspectives in Crohn's disease. Moving forward with anti-TNF-alpha therapy: current needs and future treatments. *Rev Gastroenterol Disord* **7 Suppl 2**: S23-35.

Sandborn, W.J., and Hanauer, S.B. (2002) Infliximab in the treatment of Crohn's disease: a user's guide for clinicians. *Am J Gastroenterol* **97**: 2962–2972.

Sandkvist, M. (2001) Biology of type II secretion. *Mol Microbiol* **40**: 271–283.

Sartor, R.B. (1995) Current concepts of the etiology and pathogenesis of ulcerative colitis and Crohn's disease. *Gastroenterol Clin North Am* **24**: 475–507.

Sartor, R.B. (2005) Does *Mycobacterium avium* subspecies paratuberculosis cause Crohn's disease? *Gut* **54**: 896–898.

Sartor, R.B. (2006) Mechanisms of Disease: pathogenesis of Crohn's disease and ulcerative colitis. *Nat Clin Pract Gastroenterol Hepatol* **3**: 390–407.

Sartor, R.B. (2008) Microbial Influences in Inflammatory Bowel Diseases. *Gastroenterology* **134**: 577–594.

Sasaki, M., Sitaraman, S.V., Babbin, B.A., Gerner-Smidt, P., Ribot, E.M., Garrett, N., *et al.* (2007) Invasive *Escherichia coli* are a feature of Crohn's disease. *Lab Investig J Tech Methods Pathol* **87**: 1042–1054.

Sauri, A., Soprova, Z., Wickström, D., Gier, J.-W. de, Schors, R.C.V. der, Smit, A.B., *et al.* (2009) The Bam (Omp85) complex is involved in secretion of the autotransporter haemoglobin protease. *Microbiology* **155**: 3982–3991.

Saurí, A., Ten Hagen-Jongman, C.M., Ulsen, P. van, and Luirink, J. (2012) Estimating the size of the active translocation pore of an autotransporter. *J Mol Biol* **416**: 335–345.

Savage, D.C. (1977) Microbial Ecology of the Gastrointestinal Tract. *Annu Rev Microbiol* **31**: 107–133.

Scanlan, P.D., and Marchesi, J.R. (2008) Micro-eukaryotic diversity of the human distal gut microbiota: qualitative assessment using culture-dependent and -independent analysis of faeces. *ISME J* **2**: 1183–1193.

Schaffer, T., Müller, S., Flogerzi, B., Seibold-Schmid, B., Schoepfer, A.M., and Seibold, F. (2007) Anti-*Saccharomyces cerevisiae* mannan antibodies (ASCA) of Crohn's patients crossreact with mannan from other yeast strains, and murine ASCA IgM can be experimentally induced with *Candida albicans*. *Inflamm Bowel Dis* **13**: 1339–1346.

Schembri, M.A., Dalsgaard, D., and Klemm, P. (2004) Capsule Shields the Function of Short Bacterial Adhesins. *J Bacteriol* **186**: 1249–1257.

Schmitt, C.K., Ikeda, J.S., Darnell, S.C., Watson, P.R., Bispham, J., Wallis, T.S., *et al.* (2001) Absence of All Components of the Flagellar Export and Synthesis Machinery Differentially Alters Virulence of *Salmonella enterica* Serovar Typhimurium in Models of Typhoid Fever, Survival in Macrophages, Tissue Culture Invasiveness, and Calf Enterocolitis. *Infect Immun* **69**: 5619–5625.

Schnaitman, C.A., Parker, C.T., Klena, J.D., Pradel, E.L., Pearson, N.B., Sanderson, K.E., and MacClachlan, P.R. (1991) Physical maps of the rfa loci of *Escherichia coli* K-12 and *Salmonella typhimurium*. *J Bacteriol* **173**: 7410–7411.

Schooling, S.R., and Beveridge, T.J. (2006) Membrane Vesicles: an Overlooked Component of the Matrices of Biofilms. *J Bacteriol* **188**: 5945–5957.

Schreiber, S., Nikolaus, S., and Hampe, J. (1998) Activation of nuclear factor κB in inflammatory bowel disease. *Gut* **42**: 477–484.

Schultsz, C., Moussa, M., Ketel, R. van, Tytgat, G.N., and Dankert, J. (1997) Frequency of pathogenic and enteroadherent *Escherichia coli* in patients with inflammatory bowel disease and controls. *J Clin Pathol* **50**: 573–579.

Schwartz, D.A., and Herdman, C.R. (2004) The medical treatment of Crohn's perianal fistulas. *Aliment Pharmacol Ther* **19**: 953–967.

Schwarz, S., Hood, R.D., and Mougous, J.D. (2010) What is type VI secretion doing in all those bugs? *Trends Microbiol* **18**: 531–537.

Scribano, M.L., and Prantera, C. (2013) Use of antibiotics in the treatment of Crohn's disease. *World J Gastroenterol WJG* **19**: 648–653.

Sekirov, I., Russell, S.L., Antunes, L.C.M., and Finlay, B.B. (2010) Gut Microbiota in Health and Disease. *Physiol Rev* **90**: 859–904.

Seksik, P., Rigottier-Gois, L., Gramet, G., Sutren, M., Pochart, P., Marteau, P., *et al.* (2003) Alterations of the dominant faecal bacterial groups in patients with Crohn's disease of the colon. *Gut* **52**: 237–242.

Selkrig, J., Mosbahi, K., Webb, C.T., Belousoff, M.J., Perry, A.J., Wells, T.J., *et al.* (2012) Discovery of an archetypal protein transport system in bacterial outer membranes. *Nat Struct Mol Biol* **19**: 506–510.

Seminerio, J.L., Jr, E.V.L., Colombel, J.-F., Thapa, P., and Sandborn, W.J. (2012) Infliximab for Crohn's Disease: The First 500 Patients Followed Up Through 2009. *Dig Dis Sci* **58**: 797–806.

Sengupta, C., Ray, S., and Chowdhury, R. (2014) Fine tuning of virulence regulatory pathways in enteric bacteria in response to varying bile and oxygen concentrations in the gastrointestinal tract. *Gut Pathog* **6**: 38.

Serra, D.O., Richter, A.M., and Hengge, R. (2013) Cellulose as an Architectural Element in Spatially Structured *Escherichia coli* Biofilms. *J Bacteriol* **195**: 5540–5554.

Shabala, L., and Ross, T. (2008) Cyclopropane fatty acids improve *Escherichia coli* survival in acidified minimal media by reducing membrane permeability to H+ and enhanced ability to extrude H+. *Res Microbiol* **159**: 458–461.

Shere, K.D., Sallustio, S., Manessis, A., D'Aversa, T.G., and Goldberg, M.B. (1997) Disruption of IcsP, the major Shigella protease that cleaves IcsA, accelerates actin-based motility. *Mol Microbiol* **25**: 451–462.

Sherwood (2006) *Physiologie humaine*. de Boeck.,.

Shoda, R., Matsueda, K., Yamato, S., and Umeda, N. (1996) Epidemiologic analysis of Crohn disease in Japan: increased dietary intake of n-6 polyunsaturated fatty acids and animal protein relates to the increased incidence of Crohn disease in Japan. *Am J Clin Nutr* **63**: 741–745.

Silva, A.J., Pham, K., and Benitez, J.A. (2003) Haemagglutinin/protease expression and mucin gel penetration in El Tor biotype *Vibrio cholerae*. *Microbiol Read Engl* **149**: 1883–1891.

Silverman, J.M., Brunet, Y.R., Cascales, E., and Mougous, J.D. (2012) Structure and regulation of the type VI secretion system. *Annu Rev Microbiol* **66**: 453–472.

Simon, E.G., Ghosh, S., Iacucci, M., and Moran, G.W. (2016) Ustekinumab for the treatment of Crohn's disease: can it find its niche? *Ther Adv Gastroenterol* **9**: 26–36.

Singh, S.B., Davis, A.S., Taylor, G.A., and Deretic, V. (2006) Human IRGM Induces Autophagy to Eliminate Intracellular Mycobacteria. *Science* **313**: 1438–1441.

Sipka, S., and Bruckner, G. (2014) The Immunomodulatory Role of Bile Acids. *Int Arch Allergy Immunol* **165**: 1–8.

Slomiany, B.L., Sarosiek, J., and Slomiany, A. (1987) Gastric mucus and the mucosal barrier. *Dig Dis Basel Switz* **5**: 125–145.

Small, P., Blankenhorn, D., Welty, D., Zinser, E., and Slonczewski, J.L. (1994) Acid and base resistance in *Escherichia coli* and *Shigella flexneri*: role of rpoS and growth pH. *J Bacteriol* **176**: 1729–1737.

Snoeck, V., Goddeeris, B., and Cox, E. (2005) The role of enterocytes in the intestinal barrier function and antigen uptake. *Microbes Infect* **7**: 997–1004.

Sokol, H., Leducq, V., Aschard, H., Pham, H.-P., Jegou, S., Landman, C., *et al.* (2016) Fungal microbiota dysbiosis in IBD. *Gut* gutjnl-2015-310746.

Sokol, H., Pigneur, B., Watterlot, L., Lakhdari, O., Bermúdez-Humarán, L.G., Gratadoux, J.-J., *et al.* (2008) *Faecalibacterium prausnitzii* is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. *Proc Natl Acad Sci* **105**: 16731–16736.

Sokol, H., and Seksik, P. (2010) The intestinal microbiota in inflammatory bowel diseases: time to connect with the host. *Curr Opin Gastroenterol* **26**: 327–331.

Sokol, H., Seksik, P., Furet, J.P., Firmesse, O., Nion-Larmurier, I., Beaugerie, L., *et al.* (2009) Low counts of *Faecalibacterium prausnitzii* in colitis microbiota. *Inflamm Bowel Dis* **15**: 1183–1189.

Sokol, H., Seksik, P., Rigottier-Gois, L., Lay, C., Lepage, P., Podglajen, I., *et al.* (2006) Specificities of the fecal microbiota in inflammatory bowel disease. *Inflamm Bowel Dis* **12**: 106–111.

Sommer, F., and Bäckhed, F. (2013) The gut microbiota — masters of host development and physiology. *Nat Rev Microbiol* **11**: 227–238.

Sonnenburg, J.L., Xu, J., Leip, D.D., Chen, C.-H., Westover, B.P., Weatherford, J., *et al.* (2005) Glycan Foraging in Vivo by an Intestine-Adapted Bacterial Symbiont. *Science* **307**: 1955–1959.

St Geme, J.W., and Yeo, H.-J. (2009) A prototype two-partner secretion pathway: the *Haemophilus influenzae* HMW1 and HMW2 adhesin systems. *Trends Microbiol* **17**: 355–360.

Stancik, L.M., Stancik, D.M., Schmidt, B., Barnhart, D.M., Yoncheva, Y.N., and Slonczewski, J.L. (2002) pH-Dependent Expression of Periplasmic Proteins and Amino Acid Catabolism in *Escherichia coli*. *J Bacteriol* **184**: 4246–4258.

Standaert–Vitse, A., Jouault, T., Vandewalle, P., Mille, C., Seddik, M., Sendid, B., *et al.* (2006) *Candida albicans* Is an Immunogen for Anti–*Saccharomyces cerevisiae* Antibody Markers of Crohn's Disease. *Gastroenterology* **130**: 1764–1775.

Stappenbeck, T.S., Hooper, L.V., and Gordon, J.I. (2002) Developmental regulation of intestinal angiogenesis by indigenous microbes via Paneth cells. *Proc Natl Acad Sci U S A* **99**: 15451–15455.

Stathopoulos, C., Provence, D.L., and Curtiss, R. (1999) Characterization of the Avian Pathogenic *Escherichia coli* Hemagglutinin Tsh, a Member of the Immunoglobulin A Protease-Type Family of Autotransporters. *Infect Immun* **67**: 772–781.

Steinhart, A.H., Ewe, K., Griffiths, A.M., Modigliani, R., and Thomsen, O.O. (2003) Corticosteroids for maintenance of remission in Crohn's disease. In *The Cochrane Library*. John Wiley & Sons, Ltd, http://onlinelibrary.wiley.com.gate2.inist.fr/doi/10.1002/14651858.CD000301/abstract. Accessed February 19, 2016.

Stevenson, G., Andrianopoulos, K., Hobbs, M., and Reeves, P.R. (1996) Organization of the *Escherichia coli* K-12 gene cluster responsible for production of the extracellular polysaccharide colanic acid. *J Bacteriol* **178**: 4885–4893.

Stewart, D.B., and Hegarty, J.P. (2013) Correlation between virulence gene expression and proton pump inhibitors and ambient pH in *Clostridium difficile*: results of an in vitro study. *J Med Microbiol* **62**: 1517–1523.

Strober, W., Fuss, I., and Mannon, P. (2007) The fundamental basis of inflammatory bowel disease. *J Clin Invest* **117**: 514–521.

Subramanian, S., Roberts, C.L., Hart, C.A., Martin, H.M., Edwards, S.W., Rhodes, J.M., and Campbell, B.J. (2008) Replication of Colonic Crohn's Disease Mucosal *Escherichia coli* Isolates within Macrophages and Their Susceptibility to Antibiotics. *Antimicrob Agents Chemother* **52**: 427–434.

Sura, R., Colombel, J.-F., and Van Kruiningen, H.J. (2011) Lymphatics, tertiary lymphoid organs and the granulomas of Crohn's disease: an immunohistochemical study. *Aliment Pharmacol Ther* **33**: 930–939.

Sutton, C.L., Kim, J., Yamane, A., Dalwadi, H., Wei, B., Landers, C., *et al.* (2000) Identification of a novel bacterial sequence associated with Crohn's disease. *Gastroenterology* **119**: 23–31.

Swann, J.R., Want, E.J., Geier, F.M., Spagou, K., Wilson, I.D., Sidaway, J.E., *et al.* (2011) Systemic gut microbial modulation of bile acid metabolism in host tissue compartments. *Proc Natl Acad Sci U S A* **108**: 4523–4530.

Swidsinski, A., Ladhoff, A., Pernthaler, A., Swidsinski, S., Loening-Baucke, V., Ortner, M., *et al.* (2002) Mucosal flora in inflammatory bowel disease. *Gastroenterology* **122**: 44–54.

Swidsinski, A., Loening-Baucke, V., and Herber, A. (2009) Mucosal flora in Crohn's disease and ulcerative colitis - an overview. *J Physiol Pharmacol Off J Pol Physiol Soc* **60 Suppl 6**: 61–71.

Swidsinski, A., Weber, J., Loening-Baucke, V., Hale, L.P., and Lochs, H. (2005) Spatial Organization and Composition of the Mucosal Flora in Patients with Inflammatory Bowel Disease. *J Clin Microbiol* **43**: 3380–3389.

Tabaqchali, S., O'Donoghue, D.P., and Bettelheim, K.A. (1978) *Escherichia coli* antibodies in patients with inflammatory bowel disease. *Gut* **19**: 108–113.

Takesue, Y., Ohge, H., Uemura, K., Imamura, Y., Murakami, Y., Yokoyama, T., *et al.* (2002) Bacterial translocation in patients with Crohn's disease undergoing surgery. *Dis Colon Rectum* **45**: 1665–1671.

Tamboli, C.P., Neut, C., Desreumaux, P., and Colombel, J.F. (2004) Dysbiosis in inflammatory bowel disease. *Gut* **53**: 1–4.

Targan, S.R., Landers, C.J., Yang, H., Lodes, M.J., Cong, Y., Papadakis, K.A., *et al.* (2005) Antibodies to CBir1 Flagellin Define a Unique Response That Is Associated Independently With Complicated Crohn's Disease. *Gastroenterology* **128**: 2020–2028.

Tedelind, S., Westberg, F., Kjerrulf, M., and Vidal, A. (2007) Anti-inflammatory properties of the short-chain fatty acids acetate and propionate: a study with relevance to inflammatory bowel disease. *World J Gastroenterol WJG* **13**: 2826–2832.

Tenaillon, O., Skurnik, D., Picard, B., and Denamur, E. (2010) The population genetics of commensal *Escherichia coli*. *Nat Rev Microbiol* **8**: 207–217.

Tennant, S.M., Hartland, E.L., Phumoonna, T., Lyras, D., Rood, J.I., Robins-Browne, R.M., and Driel, I.R. van (2008) Influence of Gastric Acid on Susceptibility to Infection with Ingested Bacterial Pathogens. *Infect Immun* **76**: 639–645.

Thanassi, D.G., Cheng, L.W., and Nikaido, H. (1997) Active efflux of bile salts by *Escherichia coli*. *J Bacteriol* **179**: 2512–2518.

Theodoratou, E., Campbell, H., Ventham, N.T., Kolarich, D., Pučić-Baković, M., Zoldoš, V., *et al.* (2014) The role of glycosylation in IBD. *Nat Rev Gastroenterol Hepatol* **11**: 588–600.

Thiennimitr, P., Winter, S.E., Winter, M.G., Xavier, M.N., Tolstikov, V., Huseby, D.L., *et al.* (2011) Intestinal inflammation allows *Salmonella* to use ethanolamine to compete with the microbiota. *Proc Natl Acad Sci* **108**: 17480–17485.

Thim, L., Madsen, F., and Poulsen, S.S. (2002) Effect of trefoil factors on the viscoelastic properties of mucus gels. *Eur J Clin Invest* **32**: 519–527.

Thomson, N.R., Howard, S., Wren, B.W., Holden, M.T.G., Crossman, L., Challis, G.L., *et al.* (2006) The Complete Genome Sequence and Comparative Genome Analysis of the High Pathogenicity *Yersinia enterocolitica* Strain 8081. *PLoS Genet* **2** http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1698947/. Accessed July 3, 2016.

Török, H.-P., Glas, J., Tonenchi, L., Bruennler, G., Folwaczny, M., and Folwaczny, C. (2004) Crohn's disease is associated with a toll-like receptor-9 polymorphism. *Gastroenterology* **127**: 365–366.

Torres, A.G., Giron, J.A., Perna, N.T., Burland, V., Blattner, F.R., Avelino-Flores, F., and Kaper, J.B. (2002) Identification and Characterization of lpfABCC′DE, a Fimbrial Operon of Enterohemorrhagic *Escherichia coli* O157:H7. *Infect Immun* **70**: 5416–5427.

Toussirot, E. (2012) The IL23/Th17 pathway as a therapeutic target in chronic inflammatory diseases. *Inflamm Allergy Drug Targets* **11**: 159–168.

Travassos, L.H., Carneiro, L.A.M., Ramjeet, M., Hussey, S., Kim, Y.-G., Magalhães, J.G., *et al.* (2010) Nod1 and Nod2 direct autophagy by recruiting ATG16L1 to the plasma membrane at the site of bacterial entry. *Nat Immunol* **11**: 55–62.

Triantafillidis, J.K., Merikas, E., and Georgopoulos, F. (2011) Current and emerging drugs for the treatment of inflammatory bowel disease. *Drug Des Devel Ther* **5**: 185–210.

Tsai, J.C., Yen, M.-R., Castillo, R., Leyton, D.L., Henderson, I.R., and Saier, M.H. (2010) The bacterial intimins and invasins: a large and novel family of secreted proteins. *PloS One* **5**: e14403.

Tseng, T.-T., Tyler, B.M., and Setubal, J.C. (2009) Protein secretion systems in bacterial-host associations, and their description in the Gene Ontology. *BMC Microbiol* **9**: S2.

Tsukita, S., Furuse, M., and Itoh, M. (2001) Multifunctional strands in tight junctions. *Nat Rev Mol Cell Biol* **2**: 285–293.

Tu, Q.V., McGuckin, M.A., and Mendz, G.L. (2008) *Campylobacter jejuni* response to human mucin MUC2: modulation of colonization and pathogenicity determinants. *J Med Microbiol* **57**: 795–802.

Tucker, D.L., Tucker, N., and Conway, T. (2002) Gene Expression Profiling of the pH Response in *Escherichia coli*. *J Bacteriol* **184**: 6551–6558.

Turnbaugh, P.J., Hamady, M., Yatsunenko, T., Cantarel, B.L., Duncan, A., Ley, R.E., *et al.* (2009) A core gut microbiome in obese and lean twins. *Nature* **457**: 480–484.

Turner, J.R. (2009) Intestinal mucosal barrier function in health and disease. *Nat Rev Immunol* **9**: 799–809.

Tysk, C., Lindberg, E., Järnerot, G., and Flodérus-Myrhed, B. (1988) Ulcerative colitis and Crohn's disease in an unselected population of monozygotic and dizygotic twins. A study of heritability and the influence of smoking. *Gut* **29**: 990–996.

Ullman, T.A., and Itzkowitz, S.H. (2011) Intestinal Inflammation and Cancer. *Gastroenterology* **140**: 1807–1816.e1.

Ulsen, P. van, Alphen, L. van, Hove, J. ten, Fransen, F., Ley, P. van der, and Tommassen, J. (2003) A Neisserial autotransporter NalP modulating the processing of other autotransporters. *Mol Microbiol* **50**: 1017–1030.

Ursing, B., and Kamme, C. (1975) Metronidazole for Crohn's disease. *The Lancet* **305**: 775–777.

Vaishnava, S., Behrendt, C.L., Ismail, A.S., Eckmann, L., and Hooper, L.V. (2008) Paneth cells directly sense gut commensals and maintain homeostasis at the intestinal host-microbial interface. *Proc Natl Acad Sci U S A* **105**: 20858–20863.

Van Assche, G., Dignass, A., Panes, J., Beaugerie, L., Karagiannis, J., Allez, M., *et al.* (2010) The second European evidence-based Consensus on the diagnosis and management of Crohn's disease: Definitions and diagnosis. *J Crohns Colitis* **4**: 7–27.

Vandenbroucke, K., Hans, W., Van Huysse, J., Neirynck, S., Demetter, P., Remaut, E., *et al.* (2004) Active delivery of trefoil factors by genetically modified *Lactococcus lactis* prevents and heals acute colitis in mice. *Gastroenterology* **127**: 502–513.

Varet, H., Brillet-Guéguen, L., Coppée, J.-Y., and Dillies, M.-A. (2016) SARTools: A DESeq2- and EdgeR-Based R Pipeline for Comprehensive Differential Analysis of RNA-Seq Data. *PLOS ONE* **11**: e0157022.

Vazeille, E., Chassaing, B., Buisson, A., Dubois, A., Vallée, A. de, Billard, E., *et al.* (2016) GipA Factor Supports Colonization of Peyer's Patches by Crohn's Disease-associated *Escherichia coli*. *Inflamm Bowel Dis* **22**: 68–81.

Vejborg, R.M., Hancock, V., Petersen, A.M., Krogfelt, K.A., and Klemm, P. (2011) Comparative genomics of *Escherichia coli* isolated from patients with inflammatory bowel disease. *BMC Genomics* **12**: 316.

Velcich, A., Yang, W., Heyer, J., Fragale, A., Nicholas, C., Viani, S., *et al.* (2002) Colorectal Cancer in Mice Genetically Deficient in the Mucin Muc2. *Science* **295**: 1726–1729.

Verhulst, S.L., Vael, C., Beunckens, C., Nelen, V., Goossens, H., and Desager, K. (2008) A Longitudinal Analysis on the Association Between Antibiotic Use, Intestinal Microflora, and Wheezing During the First Year of Life. *J Asthma* **45**: 828–832.

Villani, A.-C., Lemire, M., Fortin, G., Louis, E., Silverberg, M.S., Collette, C., *et al.* (2009) Common variants in the NLRP3 region contribute to Crohn's disease susceptibility. *Nat Genet* **41**: 71–76.

Vinall, L.E., Hill, A.S., Pigny, P., Pratt, W.S., Toribara, N., Gum, J.R., *et al.* (1998) Variable number tandem repeat polymorphism of the mucin genes located in the complex on 11p15.5. *Hum Genet* **102**: 357–366.

Vinolo, M.A.R., Rodrigues, H.G., Nachbar, R.T., and Curi, R. (2011) Regulation of Inflammation by Short Chain Fatty Acids. *Nutrients* **3**: 858–876.

Voelkerding, K.V., Dames, S., and Durtschi, J.D. (2010) Next Generation Sequencing for Clinical Diagnostics-Principles and Application to Targeted Resequencing for Hypertrophic Cardiomyopathy. *J Mol Diagn JMD* **12**: 539–551.

Voulhoux, R., Ball, G., Ize, B., Vasil, M.L., Lazdunski, A., Wu, L.F., and Filloux, A. (2001) Involvement of the twin-arginine translocation system in protein secretion via the type II pathway. *EMBO J* **20**: 6735–6741.

Waar, K., Mei, H.C. van der, Harmsen, H.J.M., Degener, J.E., and Busscher, H.J. (2002) Adhesion to Bile Drain Materials and Physicochemical Surface Properties of *Enterococcus faecalis* Strains Grown in the Presence of Bile. *Appl Environ Microbiol* **68**: 3855–3858.

Wai, S.N., Lindmark, B., Söderblom, T., Takade, A., Westermark, M., Oscarsson, J., *et al.* (2003) Vesicle-Mediated Export and Assembly of Pore-Forming Oligomers of the Enterobacterial ClyA Cytotoxin. *Cell* **115**: 25–35.

Wakefield, A.J., Fox, J.D., Sawyerr, A.M., Taylor, J.E., Sweenie, C.H., Smith, M., *et al.* (1992) Detection of herpesvirus DNA in the large intestine of patients with ulcerative colitis and Crohn's disease using the nested polymerase chain reaction. *J Med Virol* **38**: 183–190.

Wakefield, A.J., Pittilo, R.M., Sim, R., Cosby, S.L., Stephenson, J.R., Dhillon, A.P., and Pounder, R.E. (1993) Evidence of persistent measles virus infection in Crohn's disease. *J Med Virol* **39**: 345–353.

Waksman, G. (2012) Bacterial secretion comes of age. *Philos Trans R Soc B Biol Sci* **367**: 1014–1015.

Walk, S.T., Alm, E.W., Calhoun, L.M., Mladonicky, J.M., and Whittam, T.S. (2007) Genetic diversity and population structure of *Escherichia coli* isolated from freshwater beaches. *Environ Microbiol* **9**: 2274–2288.

Wall, R., Ross, R.P., Ryan, C.A., Hussey, S., Murphy, B., Fitzgerald, G.F., and Stanton, C. (2009) Role of gut microbiota in early infant development. *Clin Med Pediatr* **3**: 45–54.

Walter, J., and Ley, R. (2011) The Human Gut Microbiome: Ecology and Recent Evolutionary Changes. *Annu Rev Microbiol* **65**: 411–429.

Wang, Y., Nordhues, B., Zhong, D., and De Guzman, R.N. (2010) NMR characterization of the interaction of the *Salmonella* type III secretion system protein SipD and bile salts. *Biochemistry (Mosc)* **49**: 4220–4226.

Wehkamp, J., Harder, J., Weichenthal, M., Schwab, M., Schäffeler, E., Schlee, M., *et al.* (2004) NOD2 (CARD15) mutations in Crohn's disease are associated with diminished mucosal alpha-defensin expression. *Gut* **53**: 1658–1664.

Wehkamp, J., Salzman, N.H., Porter, E., Nuding, S., Weichenthal, M., Petras, R.E., *et al.* (2005) Reduced Paneth cell alpha-defensins in ileal Crohn's disease. *Proc Natl Acad Sci U S A* **102**: 18129–18134.

Weinberg, I., Neuman, T., Margalit, M., Ayman, F., Wolf, D.G., and Ben-Yehuda, A. (2009) Epstein-Barr Virus-Related Diarrhea or Exacerbation of Inflammatory Bowel Disease: Diagnostic Dilemma. *J Clin Microbiol* **47**: 1588–1590.

Weingarden, A.R., Dosa, P.I., DeWinter, E., Steer, C.J., Shaughnessy, M.K., Johnson, J.R., *et al.* (2016) Changes in Colonic Bile Acid Composition following Fecal Microbiota Transplantation Are Sufficient to Control *Clostridium difficile* Germination and Growth. *PLoS ONE* **11** https://www-ncbi-nlm-nih-gov.gate2.inist.fr/pmc/articles/PMC4720481/. Accessed September 4, 2016.

Weiss, A., and Brockmeyer, J. (2012) Prevalence, Biogenesis, and Functionality of the Serine Protease Autotransporter EspP. *Toxins* **5**: 25–48.

Wellcome Trust Case Control Consortium (2007) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447**: 661–678.

White, B.A., Lipsky, R.L., Fricke, R.J., and Hylemon, P.B. (1980) Bile acid induction specificity of 7 alpha-dehydroxylase activity in an intestinal Eubacterium species. *Steroids* **35**: 103–109.

Whitfield, C. (2006) Biosynthesis and Assembly of Capsular Polysaccharides in *Escherichia coli*. *Annu Rev Biochem* **75**: 39–68.

Wibbenmeyer, J.A., Provenzano, D., Landry, C.F., Klose, K.E., and Delcour, A.H. (2002) *Vibrio cholerae* OmpU and OmpT Porins Are Differentially Affected by Bile. *Infect Immun* **70**: 121–126.

Wiles, T.J., Kulesus, R.R., and Mulvey, M.A. (2008) Origins and Virulence Mechanisms of Uropathogenic *Escherichia coli*. *Exp Mol Pathol* **85**: 11–19.

Willing, B., Halfvarson, J., Dicksved, J., Rosenquist, M., Järnerot, G., Engstrand, L., *et al.* (2009) Twin studies reveal specific imbalances in the mucosa-associated microbiota of patients with ileal Crohn's disease. *Inflamm Bowel Dis* **15**: 653–660.

Wimley, W.C., Selsted, M.E., and White, S.H. (1994) Interactions between human defensins and lipid bilayers: evidence for formation of multimeric pores. *Protein Sci Publ Protein Soc* **3**: 1362–1373.

Wright, E.K., De Cruz, P., Gearry, R., Day, A.S., and Kamm, M.A. (2014) Fecal biomarkers in the diagnosis and monitoring of Crohn's disease. *Inflamm Bowel Dis* **20**: 1668–1677.

Wu, J., and Kaufman, R.J. (2006) From acute ER stress to physiological roles of the Unfolded Protein Response. *Cell Death Differ* **13**: 374–384.

Xavier, R.J., and Podolsky, D.K. (2007) Unravelling the pathogenesis of inflammatory bowel disease. *Nature* **448**: 427–434.

Yang, S.K., Loftus, E.V., and Sandborn, W.J. (2001) Epidemiology of inflammatory bowel disease in Asia. *Inflamm Bowel Dis* **7**: 260–270.

Yang, Y., and Jobin, C. (2014) Microbial imbalance and intestinal pathologies: connections and contributions. *Dis Model Mech* **7**: 1131–1142.

Yao, R., Burr, D.H., Doig, P., Trust, T.J., Niu, H., and Guerry, P. (1994) Isolation of motile and non-motile insertional mutants of *Campylobacter jejuni*: the role of motility in adherence and invasion of eukaryotic cells. *Mol Microbiol* **14**: 883–893.

Yen, M.R., Peabody, C.R., Partovi, S.M., Zhai, Y., Tseng, Y.H., and Saier, M.H. (2002) Protein-translocating outer membrane porins of Gram-negative bacteria. *Biochim Biophys Acta* **1562**: 6–31.

Yin, L., Laevsky, G., and Giardina, C. (2001) Butyrate Suppression of Colonocyte NF-κB Activation and Cellular Proteasome Activity. *J Biol Chem* **276**: 44641–44646.

Yoder-Himes, D.R., Chain, P.S.G., Zhu, Y., Wurtzel, O., Rubin, E.M., Tiedje, J.M., and Sorek, R. (2009) Mapping the *Burkholderia cenocepacia* niche response via high-throughput sequencing. *Proc Natl Acad Sci U S A* **106**: 3976–3981.

Yohannes, E., Barnhart, D.M., and Slonczewski, J.L. (2004) pH-Dependent Catabolic Protein Expression during Anaerobic Growth of *Escherichia coli* K-12. *J Bacteriol* **186**: 192–199.

Younes, H., Coudray, C., Bellanger, J., Demigné, C., Rayssiguier, Y., and Rémésy, C. (2001) Effects of two fermentable carbohydrates (inulin and resistant starch) and their combination on calcium and magnesium balance in rats. *Br J Nutr* **86**: 479–485.

Young, G.M., Schmiel, D.H., and Miller, V.L. (1999) A new pathway for the secretion of virulence factors by bacteria: The flagellar export apparatus functions as a protein-secretion system. *Proc Natl Acad Sci U S A* **96**: 6456–6461.

Zaki, M.H., Lamkanfi, M., and Kanneganti, T.-D. (2011) The Nlrp3 inflammasome in IBD and colorectal tumorigenesis. *Trends Immunol* **32**: 171–179.

Zhang, Y., Hong, J.-Y., Rockwell, C.E., Copple, B.L., Jaeschke, H., and Klaassen, C.D. (2012) Effect of bile duct ligation on bile acid composition in mouse serum and liver. *Liver Int* **32**: 58–69.

Zhang, Y., Rowehl, L., Krumsiek, J.M., Orner, E.P., Shaikh, N., Tarr, P.I., *et al.* (2015) Identification of Candidate Adherent-Invasive *E. coli* Signature Transcripts by Genomic/Transcriptomic Analysis. *PLoS ONE* **10** http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4509574/. Accessed March 17, 2016.

Zhong, Y., Kinio, A., and Saleh, M. (2013) Functions of NOD-Like Receptors in Human Diseases. *Front Immunol* **4** http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3797414/. Accessed March 21, 2016.

Zou, Y., He, L., Wu, C.-H., Cao, H., Xie, Z.-H., Ouyang, Y., *et al.* (2007) PSF is an IbeA-binding protein contributing to meningitic *Escherichia coli* K1 invasion of human brain microvascular endothelial cells. *Med Microbiol Immunol (Berl)* **196**: 135–143.

Zucchelli, M., Torkvist, L., Bresso, F., Halfvarson, J., Hellquist, A., Anedda, F., *et al.* (2009) PepT1 oligopeptide transporter (SLC15A1) gene polymorphism in inflammatory bowel disease. *Inflamm Bowel Dis* **15**: 1562–1569.

# Annexes

## Annexe n°1

Small-molecule inhibitors prevent the genotoxic and protumoural effects induced by colibactin-producing bacteria

Cougnoux A, Delmas J, **Gibold L**, Faïs T, Romagnoli C, Robin F, Cuevas-Ramos G, Oswald E, Darfeuille-Michaud A, Prati F, Dalmasso G, Bonnet R.

Gut. 2016 Feb;65(2):278-85. doi: 10.1136/gutjnl-2014-307241. Epub 2015 Jan 14.

ORIGINAL ARTICLE

# Small-molecule inhibitors prevent the genotoxic and protumoural effects induced by colibactin-producing bacteria

Antony Cougnoux,[1] Julien Delmas,[1,2] Lucie Gibold,[1,2] Tiphanie Faïs,[1] Chiara Romagnoli,[3] Frederic Robin,[1,2] Gabriel Cuevas-Ramos,[4,5,6] Eric Oswald,[4,5,6,7,8] Arlette Darfeuille-Michaud,[1,9†] Fabio Prati[3] Guillaume Dalmasso,[1] Richard Bonnet[1,2]

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/gutjnl-2014-307241).

[1]Clermont Université, Université d'Auvergne; Inserm U1071; INRA USC2018, Clermont-Ferrand, France
[2]Centre Hospitalier Universitaire, Clermont-Ferrand, France
[3]Department of Life Sciences, University of Modena and Reggio Emilia, Modena, Italy
[4]INRA; USC 1360, Université de Toulouse, Toulouse, France
[5]Inserm; UMR1043, Université de Toulouse, Toulouse, France
[6]CNRS; UMR5282, Université de Toulouse, Toulouse, France
[7]UPS, Centre de Physiopathologie de Toulouse Purpan (CPTP), Toulouse, France
[8]CHU Toulouse; Hôpital Purpan; Service de bactériologie-Hygiène, Toulouse, France
[9]Institut Universitaire de Technologie, Université d'Auvergne, Aubière, France

**Correspondence to**
Dr Richard Bonnet, CHU, Laboratoire de Bacteriologie, 58 rue Montalembert, Clermont-Ferrand 63000, France; rbonnet@chu-clermontferrand.fr

AC and JD contributed equally.
†This work is dedicated to the memory of AD-M.

Received 17 March 2014
Revised 7 December 2014
Accepted 10 December 2014
Published Online First 14 January 2015

**To cite:** Cougnoux A, Delmas J, Gibold L, *et al*. *Gut* 2016;**65**:278–285.

## ABSTRACT

**Objective** Colorectal cancers (CRCs) are frequently colonised by colibactin toxin-producing *Escherichia coli* bacteria that induce DNA damage in host cells and exhibit protumoural activities. Our objective was to identify small molecules inhibiting the toxic effects induced by these colibactin-producing bacteria.

**Design** A structural approach was adopted for the identification of a putative ligand for the ClbP enzyme involved in the synthesis of colibactin. Intestinal epithelial cells and a CRC mouse model were used to assess the activity of the selected compounds in vitro and in vivo.

**Results** Docking experiments identified two boron-based compounds with computed ligand efficiency values (−0.8 and −0.9 kcal/mol/atom) consistent with data expected for medicinal chemistry leads. The crystalline structure of ClbP in complex with the compounds confirmed that the compounds were binding to the active site of ClbP. The two compounds (2 mM) suppressed the genotoxic activity of colibactin-producing *E coli* both in vitro and in vivo. The mean degree of suppression of DNA damage for the most efficient compound was 98±2% (95% CI). This compound also prevented cell proliferation and colibactin-producing *E coli*-induced tumourigenesis in mice. In a CRC murine model colonised by colibactin-producing *E coli,* the number of tumours decreased by 3.5-fold in animals receiving the compound in drinking water ($p<0.01$).

**Conclusions** These results demonstrate that targeting colibactin production controls the genotoxic and protumoural effects induced by this toxin.

### Significance of this study

**What is already known on this subject?**
- More than 50% of colorectal cancers are colonised by colibactin-producing *Escherichia coli*.
- Colibactin-producing *E coli* induce DNA damage and cellular senescence and promote colorectal tumourigenesis.
- Active serine peptidases are involved in colibactin synthesis.

**What are the new findings?**
- Two boron-based compounds were identified as ligands of the active serine peptidase ClbP.
- Incubation of the boron-based compounds with bacteria both in vitro and in vivo prevents the genotoxicity induced by colibactin-producing *E coli*.
- Boron-based compounds prevent the tumourigenesis induced by colibactin-producing *E coli*.

**How might it impact on clinical practice in the foreseeable future?**
- The results confirm the importance of colibactin-producing *E coli* in colon cancer tumourigenesis.
- Our data show the feasibility of targeting the activity of *pks+ E coli* with boron-based small molecules in preclinical models.

## INTRODUCTION

Mounting evidence suggests a role of microbiota in the development of colorectal cancer (CRC).[1–3] *Escherichia coli*, a common member of the human intestinal microbiota, usually coexists with its mammalian host in harmony. However, mucosa-associated *E coli* are more frequently identified in the colon tissue from patients with adenocarcinomas than in controls and can have procarcinogenic features.[4–6] Certain *E coli* strains harbour the pathogenic island *pks*, which encodes enzymes responsible for the synthesis of a genotoxin called colibactin. Colibactin induces double-strand DNA breaks and cell cycle arrest.[7] DNA damage is also associated with mutations and chromosomal instability,[8] thereby inducing lesions frequently observed in cancers. Interestingly, *pks*-harbouring *E coli* strains are observed in 50%–60% of human CRC specimens compared with 20% in controls[9 10] and have a carcinogenic effect that is unrelated to inflammation in susceptible mice.[9 11]

Genetic and functional analyses of the *pks* island indicate that it encodes the non-ribosomal peptide (NRP) synthetases and polyketide (PK) synthetases necessary for the production of one or more PK–NRP hybrid compounds, known as colibactin.[7] In addition, the *pks* island encodes the ClbP enzyme.[12] This active serine enzyme and its distant functional homologues form a subgroup of peptidases that activate NRP compounds.[12–14]

Owing to their unique structural features, boronic acid compounds are receiving increasing attention as potent inhibitors of active serine enzymes.[15] The empty *p*-orbital of boron atoms

can be occupied by a lone pair from other nucleophiles, thereby enabling boron to form a dative bond with biological nucleophiles such as reactive serine. In addition, the modulation of their substitutions enables boron-based compounds to improve their selectivity towards a desired target. Boronic acids can, therefore, act as competitive inhibitors that are specific for active serine enzymes[15] such as ClbP.[16]

In the present study, we identified boron-based molecules that bind the ClbP active site. These compounds suppress the genotoxic activity of *pks*-harbouring *E coli* both in vitro and in vivo and prevent the cell proliferation and protumourigenic effects induced by *pks*+ *E coli*.

## MATERIALS AND METHODS

### Strains, plasmids and compounds

For in vitro infections, we used the *E coli* strain DH10B hosting pBeloBAC bearing the pks island (*pks*+ *E coli*) and the *E coli* strain DH10B hosting the *pks* island devoid of the ClbP-encoding gene (*pks*− *E coli*).[7 11] The previously reported clinical *pks*+ *E coli* strain (11G5) and its isogenic mutant that was defective for the *pks* island were used for the in vivo experiments.[11] Compound 1 and the closely related compounds 1a to 1e were purchased from Sigma-Aldrich (Saint-Quentin-Fallavier, France). Compounds 2, 2a and 3 were synthesised as previously described.[17 18]

### In vitro cell assays

In vitro infections were performed at a multiplicity of infection of 500 using HeLa and HTC116 cells as previously described.[7 11] The comet assays (a single-cell gel electrophoresis assay) were performed at 24 h postinfection as previously described.[7] γH2Ax immunofluorescence staining was performed using antiphospho(Ser139)-H2Ax antibodies (Cell Signaling Technology, Leiden, The Netherlands). DNA was stained with Hoechst (Sigma-Aldrich), and images were acquired with a confocal microscope. Cytolysis was investigated by measuring lactate dehydrogenase (LDH) release using an LDH kit (Roche Diagnostics) according to the manufacturer's recommendations. Cellular proliferation was assessed using the Cell Proliferation Kit I (MTT; Roche Diagnostics, Meylan, France) as previously reported.[11]

### Senescence-associated β-galactosidase (SA-β-gal) activity

Senescent cells were visualised using the Senescence Cells Histochemical Staining Kit (Sigma-Aldrich) according to the manufacturer's instructions. The SA-β-gal activity was assessed in mouse tissues as previously described.[11] Enzymatic activity was normalised against protein concentration.

### Quantitative real-time reverse transcriptase PCR (qRT-PCR)

To quantify mRNA expression, the total RNA was isolated using TRIzol reagent (Invitrogen, Cergy Pontoise, France) and reverse transcribed using the first-strand cDNA synthesis kit (Euromedex, Souffelweyersheim, France) according to the manufacturer's instructions. The qRT-PCRs were performed with a SYBR Green qPCR Master Mix (Roche Diagnostics) on a MasterCycler RealPlex4 (Eppendorf, Montesson, France) using the primers reported in the online supplementary table S1. The fold inductions were calculated using the Ct method.

### Western blotting

The total protein was resolved on polyacrylamide gels and transferred to polyvinylidene difluoride membranes (Amersham Biosciences, München, Germany), which were blocked and then probed overnight at 4°C with antiphospho(Ser139)-H2Ax (Cell Signaling Technology) or anti-H2Ax (Cell Signaling Technology) antibodies. The immunoreactive bands were detected with the appropriate horseradish peroxidase-conjugated secondary antibodies (Cell Signaling Technology) using an enhanced chemiluminescence detection system according to the manufacturer's recommendations (Amersham Biosciences).

### Murine colon loop model

The colon loop assay was performed as previously described.[8] Briefly, BALB/cJ mice (4 weeks old, Charles Rivers) were anaesthetised, a midline abdominal incision was made and a ligature placed at the proximal and distal ends of the colon. Approximately $3\times10^9$ bacteria (300 µL) were then injected into the colonic lumen. The incision was sutured, and the animals were allowed to recover. The mice were euthanised 6 h later, the colons were collected immediately and γH2Ax immunostaining was performed as previously described.[8]

### Xenograft model

HCT116 cells were infected in vitro, as described above, with *pks*− *E coli*, *pks*+ *E coli* or *pks*+ *E coli* treated with the compounds. The *pks*+ *E coli*-induced senescence was checked 5 days postinfection using the SA-β-galactosidase assay as described above. The cells were trypsinised at 5 days postinfection, and $0.4\times10^6$ previously infected cells were mixed with $1.6\times10^6$ uninfected HCT116 cells. The resulting cellular mixes were embedded in Growth Factor Reduced Matrigel (Becton Dickinson Biosciences), subcutaneously injected into the dorsal flap of 5-week-old nude mice and tumour size was measured as previously described.[11]

### Azoxymethane (AOM)/dextran sodium sulfate (DSS) murine model

Colon cancer was induced in C57/B6 mice (6 weeks old, Charles Rivers) as previously described[19] with minor modifications (see online supplementary figure S1).[10] The animals were intraperitoneally injected with AOM (10 mg/kg body weight), and streptomycin (2 mg/mL) was added to the drinking water to facilitate *E coli* gut colonisation.[11 20] After 2 days, the antibiotic was discontinued, and 24 h later the mice received $10^9$ colony-forming units of a clinical *pks*+ *E coli* strain (11G5)[11] or its isogenic mutant (11G5Δ*pks*)[11] by gavage, after which they were maintained on a regular diet and water for 5 days. The mice were then subjected to two cycles of DSS treatment in which each cycle consisted of 2% DSS for 7 days followed by a 14-day recovery period with regular water. For the mice treated with boronic acid, the drug was added to the drinking water (2 mM) from gavage up to euthanasia, after which the macroscopic colonic tumours were counted. In parallel, bacterial colonisation was assessed by spreading faeces on selective agar medium.

### Ethics statement

All animal experiments were performed in accordance with the European Guidelines for the Care and Use of Animals for Research Purposes (86/609/CEE). The protocols were approved by the committee on the ethics of animal experiments at the Department of Auvergne following international directive (no CE16-09). All surgeries were performed under anaesthesia, and every effort was made to minimise animal suffering.

### Crystallisation and structure determination

ClbP was purified and crystallised as previously reported.[12] The crystals were soaked in the crystallisation buffer supplemented

with 50 mM boronic acid compounds. Data were collected at the European Synchrotron Radiation Facility (Grenoble, France) and analysed as previously reported.[12] The processing and crystallographic refinement statistics are listed in the online supplementary table S2. The structures have been deposited in the RCSB Protein Data Bank (PDB IDs: 4E6W and 4E6X).

### Molecular docking

A set of 450 boronic acid structures was provided by suppliers (Boron Molecular, Tianjin Boron PharmaTech and Sigma-Aldrich) and screened via computational simulation for their ability to bind ClbP active site. The screening scheme consisted of molecular docking with AutoDock Vina 1.1[21] and molecular dynamics simulation with Amber 10, followed by MM-PBSA binding free energy estimates as previously described.[22] The predicting binding modes were ranked according to ligand efficiency as described elsewhere.[23] The drug-like properties of the compounds (see online supplementary table S2) were computed with PaDEL-Descriptor as previously reported.[24 25]

### Statistical analyses

Statistical analyses were performed with the R software (V.2.13.2, http://www.r-project.org/). For experiments with multiple treatment groups, an analysis of variance was performed followed by pairwise comparisons using Bonferroni's multiple comparison tests. The statistical comparisons were performed using the Mann–Whitney non-parametric test for two-group comparisons. A p value $<0.05$ was considered statistically significant (*$p<0.05$, **$p<0.01$ and ***$p<0.001$).

## RESULTS

### Selection of putative boron-based ligands of ClbP

Because of the charge and the geometry of ClbP catalytic pockets,[12 14] we selected small boronic acids devoid of negatively charged groups (compounds 1 and 2; see online supplementary table S2) as putative ligands of the ClbP active site and a large negatively charged compound as a putative inactive compound (compound 3; see online supplementary table S3). The three compounds were added to a set of 450 boronic acids, which were docked in silico into the ClbP active site. Compounds 1 and 2 and closely related compounds (see online supplementary table S2) were accommodated in the ClbP active serine (compound 3 was not) and demonstrated a positioning compatible with the acylation of the active serine residue. The classification of compounds according to the ligand efficiency[23] ranked compounds 1 and 2 above the 16th position (see online supplementary table S2), suggesting that they are ligands of the ClbP active site. The ligand efficiency of inhibitors 1 and 2 (−0.9 and −0.8 kcal/mol/atom) suggested that the selected molecules were interesting leads as medicinal chemistry compounds.[23]

### Crystal structure of ClbP in complex with compounds

The structure of ClbP in complex with compounds 1 and 2 was determined at a resolution of 2.19 and 2.24 Å via X-ray diffraction to confirm the compounds' ability to bind the ClbP active site. The electron density map (see online supplementary figure S1) showed a well-defined density throughout the majority of the structure. The root-mean-square deviation of the Cα positions was 0.31 Å between the complex structures and the apoenzyme, indicating that the binding of compounds does not cause dramatic changes in ClbP geometry.

The positions of compounds 1 and 2 in the active site were unambiguously identified in the 2Fo-Fc omit density maps (see online supplementary figure S2). In contrast, no crystal structure was obtained from ClbP in complex with compound 3, confirming that compound 3 is unable to target ClbP. The electron density connected the Oγ of the catalytic residue Ser95 to the boron atoms of the compounds, confirming the acylation of the active serine by compounds 1 and 2. The other atoms of compounds 1 and 2 were accommodated near β-strand β11 (residues 324–332) between residues 327–330 and Tyr150 (figure 1). Intriguingly, most atoms of the compounds overlapped perfectly despite major differences in their chemical structure.

The moieties of compounds 1 and 2 established similar conventional interactions (Van der Waals and hydrogen bonds) with the ClbP active site. Non-canonical interactions were also observed (figure 2). For compound 1, two C atoms of the aromatic ring were at 2.8 Å and 3.0 Å of the Tyr188 hydroxyl group and the N atom of Gln330, respectively. The geometry of these close contacts is compatible with an NH–π interaction involving Gln330 and a CH–O interaction involving Tyr186.[26 27] The plane of the aromatic ring was parallel with the β-strand β11 and established a π-stacking-like interaction[28] with the peptide bond between residues Gly328 and Gly329. For compound 2, the chlorine atom established a Cl–π electrostatic interaction with the aromatic ring of Tyr188 (Chlorine-aromatic centroid distance: 3.6 Å), which exhibited a face-on geometry.[29] One oxygen atom of the sulfonamide was positioned in close contact with the C and N atoms of the Gly328–Gly329 peptide bond (3.1 and 3.6 Å, respectively). Due to the π electron resonance cloud of peptide bonds, this close interaction resembled a lone pair–π interaction.[30–32] In conclusion, the interactions differed between the molecules due to differences in their chemical structure but involved similar binding zones in the ClbP active site.

### Inhibitors prevent *pks+ E coli*-induced DNA damage both in vitro and in vivo

Colibactin-producing *E coli* (*pks+ E coli*) induce DNA damage both in vivo and in vitro.[7–9] We, therefore, investigated the ability of compounds 1 and 2 to prevent this genotoxic activity.

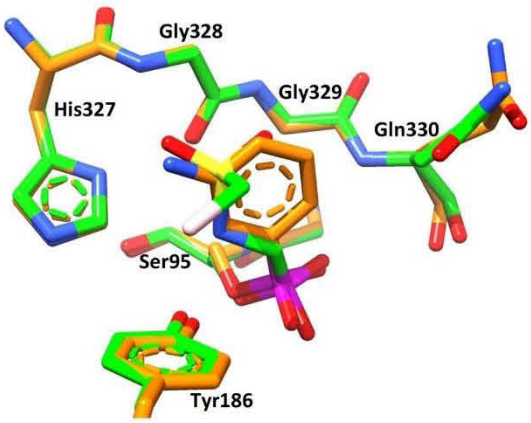



**Figure 1** Superimposition of the ClbP active site in complex with compounds 1 and 2. The atoms of compounds 1 and 2 overlap and are accommodated in the zone surrounded by residues 95, 186 and 327–330. Carbon atoms are shown in gold for ClbP in complex with compound 1 and in *green for ClbP in complex with compound* 2. Oxygen atoms are shown in red, nitrogen atoms in blue, boron atoms in magenta and sulfur atoms in yellow. The aromatic cycles are indicated by dashed circles.

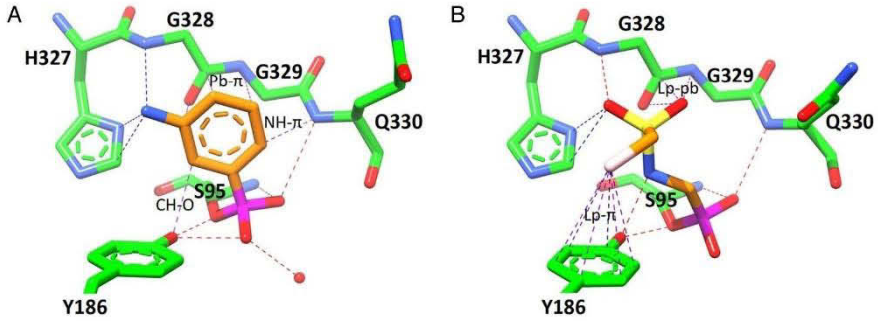



**Figure 2** Structure of the ClbP active site in complex with compounds 1 (A) and 2 (B). Carbon atoms are shown in *green for ClbP and in gold for the compounds*. Oxygen atoms are shown in red, nitrogen atoms in blue, boron in magenta, sulfur in yellow and chloride in white. The aromatic systems are indicated by dashed circles. The hydrogen bonds are shown as red dashed lines. Van der Waals contacts are indicated by blue dashed lines. Non-canonical interactions are indicated by purple dashed lines. CH–O, CH–O hydrogen bond; Lp–π, lone pair–π interaction; Lp–pb, lone pair–peptide bond interaction; NH–π, NH–π interaction; Pb–π, peptide bond–π stacking.

Epithelial cells were infected in vitro for 3 h with *pks+ E coli* pretreated overnight with or without increasing concentrations of compounds 1 and 2 (0.5–2 mM). DNA damage was assessed using comet assays (figure 3A). As expected, the percentage of cells exhibiting a comet shape was 20-fold higher in cells infected with *pks+ E coli* than in those infected with an isogenic mutant defective for the *pks* island (*pks– E coli*) or in uninfected cells ($p<0.01$). Compounds 1 and 2 significantly decreased the number of comets induced by *pks+ E coli* (2 mM compound; $p<0.01$ and $p=0.01$, respectively). The inhibition

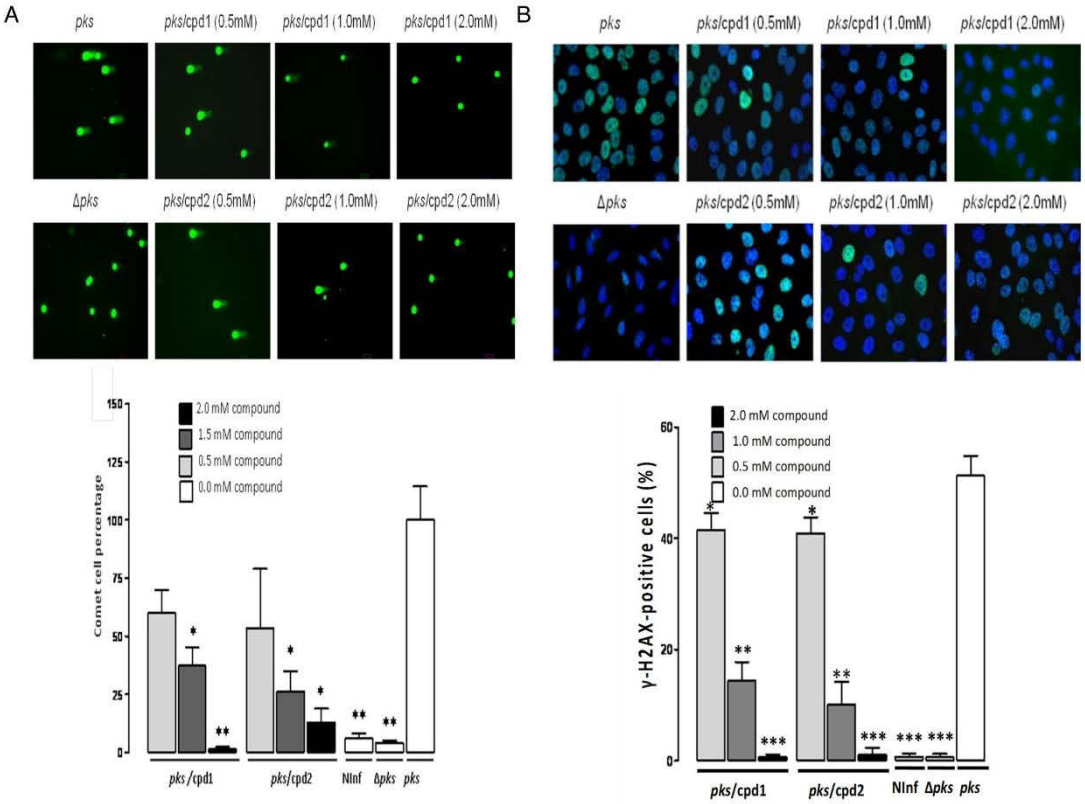



**Figure 3** Compounds 1 and 2 prevented DNA damage induced by *pks+ Escherichia coli*. (A and B) HeLa cells were infected for 3 h and DNA damage was revealed 24 h postinfection via comet assays (A) and γ-H2Ax immunofluorescence staining (B). (B) γ-H2Ax was labelled in green and DNA was stained in blue using 4′,6-diamidino-2-phenylindole. *pks*, cells infected with *pks+ E coli*; *Δpks*, cells infected with *pks– E coli*; Ninf, non-infected cells; *pks*/cpd1 and 2, cells infected with *pks+ E coli* treated with the compounds. The bar graphs on the bottom show the mean±SEM of positive cells obtained from three independent experiments (*p<0.05; **p<0.001; ***p<0.0001 vs *pks+* via analysis of variance).

of DNA damage was also observed with the molecules that were related to compounds 1 and 2 (1a to 1e and 2a, 2 mM). We next assessed the DNA damage by immunostaining for Ser139-phosphorylation of histone H2Ax (γ-H2Ax), a canonical and sensitive marker of DNA double-strand breaks. As shown in figure 3B, the γ-H2Ax staining induced by *pks+ E coli* was significantly decreased in a dose-dependent manner by the compounds. This inhibition was independent of the effect of antibiotics because the compounds themselves did not alter bacterial growth (see online supplementary figure S3). Furthermore, 24 h of cell treatment with the compounds (2 mM) did not affect cellular viability (see online supplementary figure S4), suggesting that the inhibitors act by directly targeting colibactin production. To substantiate these results, we investigated the ability of the compounds to prevent the DNA damage induced by *pks+ E coli* in a mouse colon loop model. Colon loops were infected for 6 h with *E coli* that were pretreated or untreated with compounds 1 and 2 (2 mM). γ-H2Ax was assessed in colonocytes via Western blotting and immunohistochemistry. These analyses revealed that compounds 1 and 2 prevented γ-H2Ax accumulation in intestinal loops infected with *pks+ E coli* (figure 4). Taken together, our results indicate that the compounds were able to block the DNA damage induced by *pks+ E coli* in vitro and in vivo with no evidence of toxicity against bacteria or eukaryotic cells.

### Inhibitors prevent *pks+ E coli*-induced cell proliferation

Colibactin-producing *E coli* induce the emergence of senescent cells, which stimulate the proliferation of uninfected cells via the secretion of hepatocyte growth factor (HGF).[11] We, therefore, investigated the ability of ClbP inhibitors to block cell proliferation because of *pks+ E coli*-induced senescence. As expected, conditioned medium derived from intestinal epithelial HCT116 cells infected with *pks+ E coli* stimulated the proliferation of uninfected HCT116 cells (unlike conditioned medium obtained from non-infected cells or from *pks− E coli*-infected cells; figure 5A). Interestingly, treatment of *pks+ E coli* with compounds 1 and 2 significantly inhibited the cell proliferation observed in conditioned media (figure 5A). In parallel, SA-β-gal activity was assessed in HCT116 cells infected with *pks+ E coli*, infected with *pks+ E coli* treated with compounds 1 and 2 or non-infected cells. As shown in figure 5B, compounds 1 and 2 inhibited *pks+ E coli*-induced cellular senescence. Accordingly, the compounds significantly decreased the level of HGF mRNA transcripts that were expressed by *pks+ E coli*-infected cells (figure 5C). These results suggest that compounds 1 and 2 prevent the senescence induced by *pks+ E coli* and, consequently, the secretion of HGF from senescent cells, which are not able to stimulate the proliferation of uninfected cells.

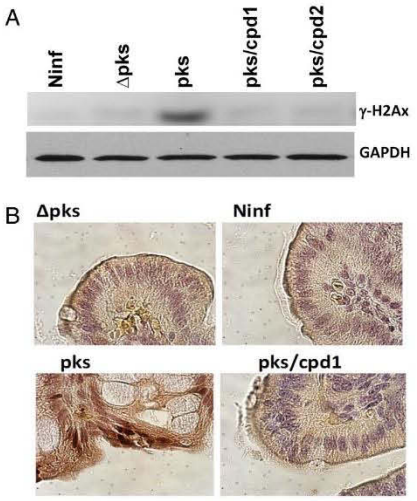



**Figure 4** Compounds 1 and 2 inhibited DNA damage induced by *pks+ Escherichia coli* in a mouse colon loop model. (A) Western blot analysis of γ-H2A$_X$ in colonocytes 6 h after the inoculation of colon loops with *pks− E coli* or *pks+ E coli* treated or untreated with compounds 1 and 2 (2 mM). (B) γ-H2AX immunohistochemical staining in colon loops. *pks*/cpd 1/2, colon loops infected with *pks+ E coli* treated with compounds 1 and 2 (2 mM); *Δpks*, colon loops infected with *pks− E coli*; Ninf, non-infected colon loops; *pks*, colon loops infected with *pks+ E coli*.

### Inhibitors prevent *pks+ E coli*-induced tumourigenesis in vivo

*pks+ E coli* promote the growth of xenograft tumours and intestinal carcinoma.[9 11] We, therefore, investigated the ability of compounds 1 and 2 to prevent the protumoural effect induced by *pks+ E coli* in a xenograft tumour model and in a mouse colon cancer model.

As previously reported,[11] unlike cells infected with *pks− E coli*, intestinal epithelial cells infected with *pks+ E coli* stimulated the growth of cotransplanted cells in immunodeficient mice (figure 5D). Interestingly, a significant decrease in tumour growth was observed from 26 days post-transplantation ($p<0.05$) when *pks+ E coli*-infected cells were treated with compounds 1 and 2 (figure 5D). The statistical significance of this inhibition of tumour growth increased between 26 days ($p<0.05$) and 34 days postengraftment ($p<0.001$). Likewise, the treatment of *pks+ E coli* with compounds 1 and 2 significantly decreased the level of the HGF mRNA transcripts in the tumours (figure 5E).

To substantiate these results, we next investigated the effect of compound 1 on the *pks+ E coli* 11G5 strain, which was isolated from a human CRC biopsy.[10] As shown in the online supplementary figure S5, compound 1 effectively prevented in vitro DNA damage, cellular senescence, HGF mRNA expression and conditioned medium-induced proliferation of uninfected cells, which are all mediated by the 11G5 strain. The impact of compound 1 on strain 11G5 was then assessed in vivo using the AOM/DSS mouse model. For this purpose, the AOM/DSS-treated mice were colonised with a clinical *pks+* strain 11G5 (n=10). To functionally link the presence of the *pks* island with CRC development, a group of mice (n=10) was colonised with an isogenic *pks*-deficient *E coli* 11G5 strain (11G5Δpks).[11] At 7 weeks postinfection, the absence of *pks* was associated with a significantly reduced number of macroscopic tumours (figure 6A), as previously observed.[9 11] The 11G5-colonised animals, which received compound 1 in their drinking water, had macroscopic tumour numbers similar to those of mice colonised with 11G5Δpks (figure 6A, B). This antitumoural effect was accompanied by a decrease in cell proliferation as assessed via proliferative cell nuclear antigen and cyclin D1 mRNA levels (figure 6C). Interestingly, compound 1 protected intestinal epithelial cells against colibactin-induced DNA damage (figure 6D). The same compound also decreased the SA-β-gal activity that was measured in tumours formed in 11G5-colonised mice treated with compound 1 compared with that in tumours formed in untreated 11G5-colonised mice (figure 6E). This decrease in SA-β-gal activity was associated

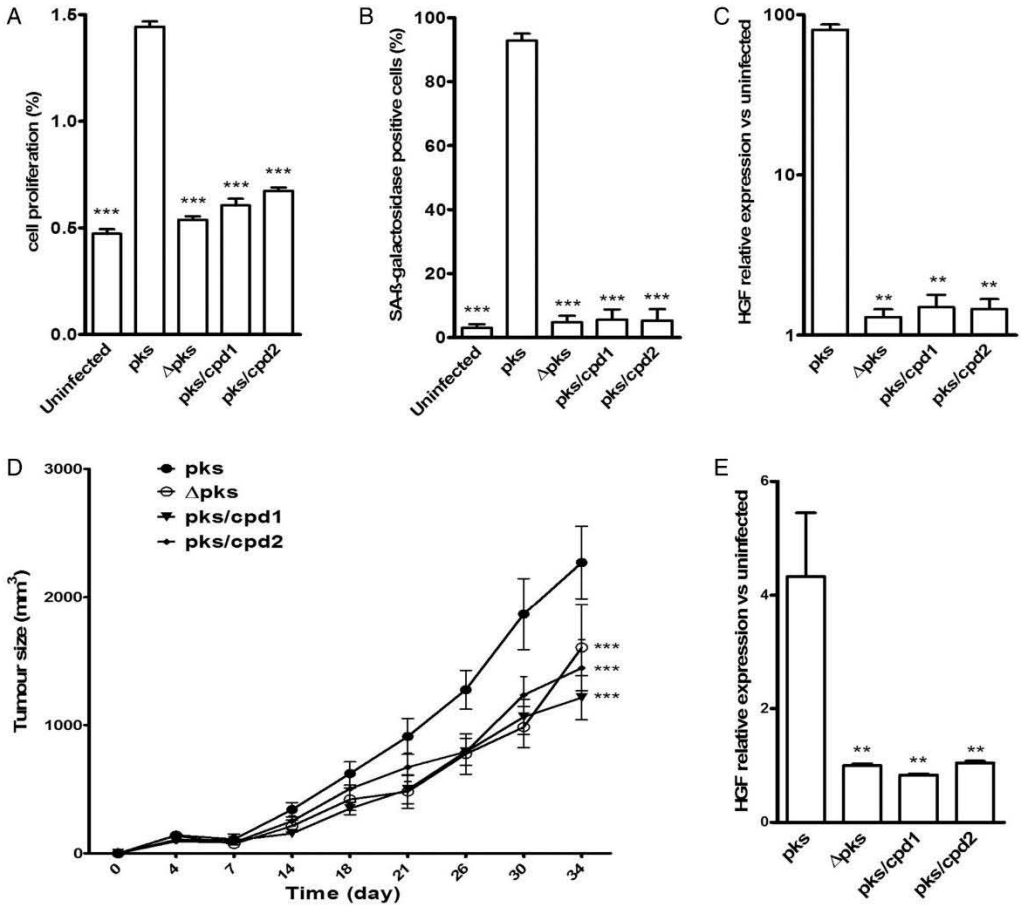



**Figure 5** Compounds 1 and 2 prevented cellular senescence, growth factor production, cell proliferation and tumour growth induced by *pks*+ *Escherichia coli*. (A) Conditioned media derived from uninfected HCT116 cells or HCT116 cells infected with *pks*+ *E coli* (*pks*), *pks*− *E coli* (Δpks) or *pks*+ *E coli* treated with compounds (*pks*/cpd) 1 and 2 (2 mM) collected at day 4 postinfection were used to stimulate the growth of uninfected HCT116 cells. Cellular growth was assessed using an MTT-based cell proliferation assay 24 h after the addition of conditioned media. Data are from triplicate experiments performed twice (n=6 per group; ***p<0.001 vs *pks* via analysis of variance (ANOVA)) and are presented as the mean±SEM. (B and C) HCT116 cells were infected with *pks*+ *E coli* (pks), *pks*− *E coli* (Δpks) or *pks*+ *E coli* treated with compounds (pks/cpd) 1 and 2 (2 mM). The percentage of senescence-associated β-galactosidase-positive cells (B) and hepatocyte growth factor (HGF) mRNA levels (C) were determined at 5 days postinfection (values represent the mean±SEM; n=3 independent experiments in duplicate; **p<0.001 vs *pks*, ***p<0.001 vs *pks* via ANOVA). (D and E) $0.4\times10^6$ HCT116 cells infected in vitro with *pks*− *E coli* (Δpks), *pks*+ *E coli* (pks) or *pks*+ *E coli* treated with compounds (pks/cpd) 1 and 2 (2 mM) were mixed with $1.6\times10^6$ non-infected HCT116 cells and subcutaneously injected into nude mice. (D) Tumour size was monitored during the indicated period (values represent the mean±SEM; n=5 mice per group; ***p<0.001 vs *pks* via ANOVA). (E) HGF mRNA levels in mouse tumours were assessed via quantitative real-time reverse transcriptase PCR at 34 days postengraftment (values represent the mean±SEM; n=5 mice per group; **p<0.01 vs *pks* via ANOVA).

with an inhibition of HGF mRNA synthesis (figure 6C). The production of cytokines/chemokines was also significantly decreased in the tumours from compound 1-treated mice (figure 6C). Neither treatment with compound 1 nor the absence of *pks* impaired intestinal colonisation ability, suggesting that the protective effect of compound 1 was due to the inhibition of colibactin production rather than the modification of 11G5 intestinal implantation (figure 6F). Altogether, our results show that oral administration of compound 1 prevents the emergence of CRC induced by *pks*+ bacteria.

## DISCUSSION

NRPs and NRP–PK hybrids are highly bioactive molecules that are frequently toxic and thus of potential interest in the drug development.[33] Their biosynthesis occurs on multimodular enzymatic assembly lines called megasynthases.[34] The NRP–PK compound colibactin is processed to maturity by the ClbP NRP-associated serine peptidases.[12 13] In this work, we identified two compounds that bind the ClbP active site and prevent the DNA damage and proliferation/tumourigenesis induced by colibactin-producing *E coli* both in vitro and in vivo. These compounds are devoid of apparent toxicity against eukaryotic and prokaryotic cells. Our findings agree with our earlier observation that ClbP is necessary for the production of colibactin[12 14] and emphasise the importance of colibactin-producing *E coli* in colon cancer tumourigenesis.[9 11]

*E coli* is primarily a commensal inhabitant of the mammalian digestive tract, and the *pks* island is hosted by commensal *E coli*

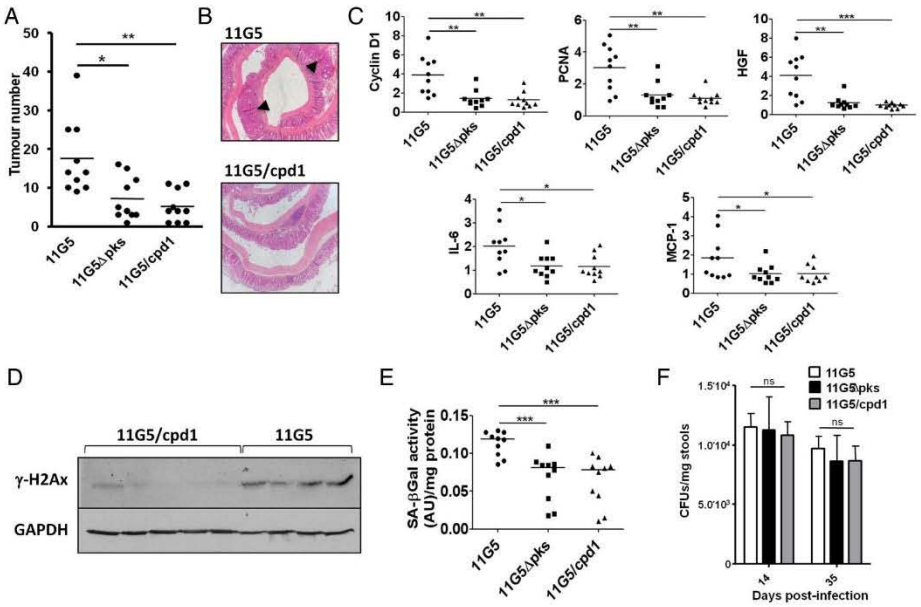



**Figure 6** Compound 1 prevented the emergence of colon tumours in azoxymethane (AOM)/dextran sodium sulfate (DSS)-treated mice colonised by *pks* + *Escherichia coli*. (A) The number of tumours in the colon per mouse at the end of the AOM/DSS protocol (*p<0.05, **p<0.01 via analysis of variance (ANOVA)). (B) Representative H&E staining of a mouse colon at the end of the AOM/DSS experiment. Black arrowheads indicate tumours. (C) Colonic mRNA level of proliferative cell nuclear antigen (PCNA), cyclin D1, hepatocyte growth factor (HGF) and cytokines/chemokines (interleukin (IL)-6 and monocyte chemoattractant protein-1 (MCP-1)) analysed via quantitative real-time reverse transcriptase PCR in AOM/DSS-treated mice. The mRNA levels are expressed as relative values (**p<0.01 via ANOVA). (D) Western blot analysis of γ-H2A$_X$ in tumours from AOM/DSS-treated mice. (E) Senescence-associated β-galactosidase (SA-β-gal) activity was assessed in colon tumours from AOM/DSS-treated mice (*p<0.05, **p<0.01, ***p<0.001 via ANOVA). (F) Quantification of 11G5 and 11G5Δpks in the faeces of mice at 14 and 35 days postgavage. Data are from triplicate experiments performed twice (mean±SEM, ANOVA). CFU, colony-forming units; GADPH, glyceraldehyde-3-phosphate dehydrogenase; ns, not significant.

strains belonging to the B2 phylogroup, which are commonly found and persistent in human microbiota.[7 35–39] Recent studies have shown that *pks*-harbouring bacteria are over-represented in human CRC specimens and promote tumour progression.[9 11] The *pks* island is also found in 72% of the *E coli* strains that cause prostatitis, a disease associated with a significantly increased risk of prostate cancer.[40] The DNA damage induced by *pks*+ *E coli* strains is estimated to be similar to that of a 0.5-Gy γ-ray whole-body irradiation and triggers mammalian cell transformation, gene mutations and chromosomal instability,[8] which are central events in tumourigenesis.[41 42] The *pks*+ *E coli* strains, therefore, represent a potential aetiological component of colorectal and prostate cancers, the third most frequently diagnosed malignancy in both sexes and the most frequent cancer in men, respectively.[43] In the present study, we report that the treatment of *pks*+ *E coli* with boronic acid-based drug-like compounds prevents the genotoxic and tumourigenic activity induced by the *pks* island both in vitro and in vivo. The usual dietary intake of boron in humans is estimated to be between 0.3 and 4.2 mg/day, and the background concentration of boron in mouse plasma is ~200 ng/mL, suggesting that the body is familiar with boron.[15] There have been no documented reports of toxicity associated with boronic acid isosteres. One boron-containing therapeutic agent was recently approved by the Food and Drug Administration for the treatment of multiple myeloma, and several more are in various stages of clinical trials in diverse medical fields,[16] suggesting that this class of compound holds promise for therapeutic use. However, our data only show the feasibility of targeting *pks*+ *E coli* activity with boron-based small molecules in preclinical models. Further investigations are required to draw conclusions on toxicity and efficacy in a clinical setting.

The compounds used in this study harbour a boronic acid function bound to a clearly different moiety, which suggests that two distinct families of inhibitors could be derived from these compounds and could be used to improve current treatments. Although they exhibit differing chemical structures, these two inhibitors were accommodated in the same area of the ClbP binding site (towards the N-terminal extremity of the β11 strand between the reactive serine Ser64 and residues Tyr188 and Gly328). The N-terminal extremity of the β11 strand forms a broad groove that is accessible to solvent; mutagenesis experiments have shown that, unlike the C-terminal extremity, the N-terminal contains residues that are essential to ClbP activity.[12 14] The N-terminal extremity of the β11 strand may, therefore, serve as a primary site of colibactin accommodation.

The binding of the inhibitors to the ClbP active site involves hydrogen bonds and Van der Waals contacts with residues 95, 186, 327–328 and 330. Non-canonical interactions are known to play an important role in protein stability and ligand binding. The binding of inhibitors in the ClbP active site revealed an atypically high density of such interactions, which may be key determinants of their activity. The non-conventional interactions established by the inhibitors differ between molecules due to differences in chemical structure. However, the structural elements of the ClbP active site are also involved in the interactions; binding, therefore, uses a similar logic. Consequently, the structures indicate different strategies for triggering the same

effect and have an interesting inventory of unusual interactions that can be built into the designs of other inhibitors. Crystallographic data showed that the compounds are able to bind the ClbP active site. However, we cannot exclude the possibility that the compounds inhibit other enzymes because the biochemical synthetic pathway and the structure of colibactin are unknown.

In conclusion, we identified boronic acid-based compounds that suppress DNA damage and the proliferation/tumourigenesis induced by *pks*-harbouring bacteria. This family of inhibitors may be useful in reducing the carcinogenesis induced by these bacteria.

**Acknowledgements** We are grateful to our local contact at the European Synchrotron Radiation Facility (ESRF, Grenoble, France) for providing assistance in using beamline ID23-1. We also thank the ICCF platform of the Université d'Auvergne for their technical assistance with histology.

**Contributors** Study design: AC, JD, GD and RB. Acquisition and interpretation of data: AC, JD, GD, LG, TF, CR, FR, GC-R and EO. Drafting of the manuscript: RB, JD, GD and AC. Critical revision of the manuscript: RB and AD-M.

**Funding** This work was supported by the Ministère de la Recherche et de la Technologie, the Institut National de la Santé et de la Recherche Médicale (UMR 1071), the Institut National de la Recherche Agronomique (USC-2018) and the Centre Hospitalier Régional Universitaire de Clermont-Ferrand, France.

**Competing interests** None.

**Provenance and peer review** Not commissioned; externally peer reviewed.

## REFERENCES

1 Dove WF, Clipson L, Gould KA, *et al*. Intestinal neoplasia in the ApcMin mouse: independence from the microbial and natural killer (beige locus) status. *Cancer Res* 1997;57:812–14.
2 Uronis JM, Mühlbauer M, Herfarth HH, *et al*. Modulation of the intestinal microbiota alters colitis-associated colorectal cancer susceptibility. *PLoS ONE* 2009;4:e6026.
3 Tjalsma H, Boleij A, Marchesi JR, *et al*. A bacterial driver-passenger model for colorectal cancer: beyond the usual suspects. *Nat Rev Microbiol* 2012;10:575–82.
4 Swidsinski A, Khilkin M, Kerjaschki D, *et al*. Association between intraepithelial *Escherichia coli* and colorectal cancer. *Gastroenterology* 1998;115:281–6.
5 Martin HM, Campbell BJ, Hart CA, *et al*. Enhanced *Escherichia coli* adherence and invasion in Crohn's disease and colon cancer. *Gastroenterology* 2004;127:80–93.
6 Prorok-Hamon M, Friswell MK, Alswied A, *et al*. Colonic mucosa-associated diffusely adherent afaC+ *Escherichia coli* expressing lpfA and pks are increased in inflammatory bowel disease and colon cancer. *Gut* 2014;63:761–70.
7 Nougayrède J-P, Homburg S, Taieb F, *et al*. *Escherichia coli* induces DNA double-strand breaks in eukaryotic cells. *Science* 2006;313:848–51.
8 Cuevas-Ramos G, Petit CR, Marcq I, *et al*. *Escherichia coli* induces DNA damage in vivo and triggers genomic instability in mammalian cells. *Proc Natl Acad Sci USA* 2010;107:11537–42.
9 Arthur JC, Perez-Chanona E, Mühlbauer M, *et al*. Intestinal Inflammation Targets Cancer-Inducing Activity of the Microbiota. *Science* 2012;338:120–3.
10 Buc E, Dubois D, Sauvanet P, *et al*. High prevalence of mucosa-associated *E. coli* producing cyclomodulin and genotoxin in colon cancer. *PLoS ONE* 2013;8:e56964.
11 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.
12 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* 2010;286:35562–70.
13 Reimer D, Pos KM, Thines M, *et al*. A natural prodrug activation mechanism in nonribosomal peptide synthesis. *Nat Chem Biol* 2011;7:888–90.
14 Cougnoux A, Gibold L, Robin F, *et al*. Analysis of structure-function relationships in the colibactin-maturating enzyme ClbP. *J Mol Biol* 2012;424:203–14.
15 Baker SJ, Ding CZ, Akama T, *et al*. Therapeutic potential of boron-containing compounds. *Future Med Chem* 2009;1:1275–88.
16 Trippier PC, McGuigan C. Boronic acids in medicinal chemistry: anticancer, antibacterial and antiviral applications. *Med Chem Commun* 2010;1:183–98.
17 Eidam O, Romagnoli C, Caselli E, *et al*. Design, synthesis, crystal structures, and antimicrobial activity of sulfonamide boronic acids as beta-lactamase inhibitors. *J Med Chem* 2010;53:7852–63.
18 Morandi F, Caselli E, Morandi S, *et al*. Nanomolar inhibitors of AmpC beta-lactamase. *J Am Chem Soc* 2003;125:685–95.
19 Tanaka T, Kohno H, Suzuki R, *et al*. A novel inflammation-related mouse colon carcinogenesis model induced by azoxymethane and dextran sodium sulfate. *Cancer Sci* 2003;94:965–73.
20 Stecher B, Robbiani R, Walker AW, *et al*. *Salmonella enterica* serovar typhimurium exploits inflammation to compete with the intestinal microbiota. *PLoS Biol* 2007;5:2177–89.
21 Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* 2010;31:455–61.
22 Rastelli G, Del Rio A, Degliesposti G, *et al*. Fast and accurate predictions of binding free energies using MM-PBSA and MM-GBSA. *J Comput Chem* 2010;31:797–810.
23 Hopkins AL, Groom CR, Alex A. Ligand efficiency: a useful metric for lead selection. *Drug Discov Today* 2004;9:430–1.
24 Yap CW. PaDEL-descriptor: an open source software to calculate molecular descriptors and fingerprints. *J Comput Chem* 2011;32:1466–74.
25 Lipinski CA. Lead- and drug-like compounds: the rule-of-five revolution. *Drug Discov Today Technol* 2004;1:337–41.
26 Cotesta S, Stahl M. The environment of amide groups in protein-ligand complexes: H-bonds and beyond. *J Mol Model* 2006;12:436–44.
27 Malone JF, Murray CM, Charlton MH, *et al*. X-H...pi(phenyl) interactions Theoretical and crystallographic observations. *J Chem Soc Faraday Trans* 1997;93:3429–36.
28 Bendova L, Jurecka P, Hobza P, *et al*. Model of peptide bond-aromatic ring interaction: correlated ab initio quantum chemical study. *J Phys Chem B* 2007;111:9975–9.
29 Imai YN, Inoue Y, Nakanishi I, *et al*. Cl-pi interactions in protein-ligand complexes. *Protein Sci* 2008;17:1129–37.
30 Egli M, Sarkhel S. Lone pair-aromatic interactions: to stabilize or not to stabilize. *Acc Chem Res* 2007;40:197–205.
31 Jain A, Ramanathan V, Sankararamakrishnan R. Lone pair ... pi interactions between water oxygens and aromatic residues: Quantum chemical studies based on high-resolution protein structures and model compounds. *Protein Sci* 2009;18:595–605.
32 Mooibroek TJ, Gamez P, Reedijk J. Lone pair—π interactions: a new supramolecular bond? *Cryst Eng Comm* 2008;10:1501–15.
33 Li JW-H, Vederas JC. Drug Discovery and Natural Products: End of an Era or an Endless frontier? *Science* 2009;325:161–5.
34 Sattely ES, Fischbach MA, Walsh CT. Total biosynthesis: in vitro reconstitution of polyketide and nonribosomal peptide pathways. *Nat Prod Rep* 2008;25:757–93.
35 Escobar-Paramo P, Grenet K, Le Menac'h A, *et al*. Large-scale population structure of human commensal *Escherichia coli* isolates. *Appl Environ Microbiol* 2004;70:5698–700.
36 Johnson JR, Johnston B, Kuskowski MA, *et al*. Molecular epidemiology and phylogenetic distribution of the *Escherichia coli* pks genomic island. *J Clin Microbiol* 2008;46:3906–11.
37 Kurokawa K, Itoh T, Kuwahara T, *et al*. Comparative metagenomics revealed commonly enriched gene sets in human gut microbiomes. *DNA Res* 2007;14:169–81.
38 Putze J, Hennequin C, Nougayrede JP, *et al*. Genetic structure and distribution of the colibactin genomic island among members of the family Enterobacteriaceae. *Infect Immun* 2009;77:4696–703.
39 Zhang L, Foxman B, Marrs C. Both urinary and rectal *Escherichia coli* isolates are dominated by strains of phylogenetic group B2. *J Clin Microbiol* 2002;40:3951–5.
40 De Marzo AM, Platz EA, Sutcliffe S, *et al*. Inflammation in prostate carcinogenesis. *Nat Rev Cancer* 2007;7:256–69.
41 Grady WM, Carethers JM. Genomic and epigenetic instability in colorectal cancer pathogenesis. *Gastroenterology* 2008;135:1079–99.
42 Lengauer C, Kinzler KW, Vogelstein B. DNA methylation and genetic instability in colorectal cancer cells. *Proc Natl Acad Sci USA* 1997;94:2545–50.
43 Siegel R, Ward E, Brawley O, *et al*. Cancer statistics, 2011: the impact of eliminating socioeconomic and racial disparities on premature cancer deaths. *CA Cancer J Clin* 2011;61:212–36.

## Annexe n°2

Bacterial genotoxin colibactin promotes colon tumour growth by inducing a senescence-associated secretory phenotype.

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.

Gut. 2014 Dec;63(12):1932-42. doi: 10.1136/gutjnl-2013-305257. Epub 2014 Mar 21.

ORIGINAL ARTICLE

# Bacterial genotoxin colibactin promotes colon tumour growth by inducing a senescence-associated secretory phenotype

Antony Cougnoux,[1,2] Guillaume Dalmasso,[1,2] Ruben Martinez,[3] Emmanuel Buc,[1,2,4] Julien Delmas,[1,2,5] Lucie Gibold,[1,2,5] Pierre Sauvanet,[1,2,4] Claude Darcha,[6] Pierre Déchelotte,[6] Mathilde Bonnet,[1,2] Denis Pezet,[1,2,4] Harald Wodrich,[3] Arlette Darfeuille-Michaud,[1,2] Richard Bonnet[1,2,5]

▸ Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/gutjnl-2013-305257).

[1]Clermont Université, UMR 1071 Inserm/Université d'Auvergne, Clermont-Ferrand, France
[2]INRA, USC 2018, Clermont-Ferrand, France
[3]Microbiologie Fondamentale et Pathogénicité, CNRS UMR 5234, Université Bordeaux Segalen, Bordeaux, France
[4]Service de Chirurgie Digestive, Centre Hospitalier Universitaire, Clermont-Ferrand, France
[5]Service de Bactériologie, Centre Hospitalier Universitaire, Clermont-Ferrand, France
[6]Servie d'anatomo-pathologie, CHU de Clermont-Ferrand, Clermont-Ferrand, France

**Correspondence to**
Richard Bonnet, CHU, Laboratoire de Bacteriologie, 58 rue Montalembert, Clermont-Ferrand 63000, France; rbonnet@chu-clermontferrand.fr

AC and GD contributed equally.

Received 13 May 2013
Revised 21 January 2014
Accepted 28 January 2014
Published Online First 21 March 2014

Linked

▸ http://dx.doi.org/10.1136/gutjnl-2014-306890

CrossMark

**To cite:** Cougnoux A, Dalmasso G, Martinez R, *et al*. *Gut* 2014;**63**: 1932–1942.

## ABSTRACT

**Background** *Escherichia coli* strains harbouring the *pks* island (*pks*+ *E. coli*) are often seen in human colorectal tumours and have a carcinogenic effect independent of inflammation in an AOM/IL-$10^{-/-}$ (azoxymethane/interleukin) mouse model.

**Objective** To investigate the mechanism sustaining *pks*+ *E. coli*-induced carcinogenesis.

**Method** Underlying cell processes were investigated *in vitro* and *in vivo* (xenograft model) using intestinal epithelial cells infected by *pks*+ *E. coli* or by an isogenic mutant defective for *pks* (*pks*− *E. coli*). The results were supported by data obtained from an AOM/DSS (azoxymethane/dextran sodium sulphate) colon cancer mouse model and from human colon cancer biopsy specimens colonised by *pks*+ *E. coli* or *pks*− *E. coli*.

**Results** Colibactin-producing *E. coli* enhanced tumour growth in both xenograft and AOM/DSS models. Growth was sustained by cellular senescence (a direct consequence of small ubiquitin-like modifier (SUMO)-conjugated p53 accumulation), which was accompanied by the production of hepatocyte growth factor (HGF). The underlying mechanisms involve microRNA-20a-5p, which targets SENP1, a key protein regulating p53 deSUMOylation. These results are consistent with the expression of SENP1, microRNA-20a-5p, HGF and phosphorylation of HGF receptor found in human and mouse colon cancers colonised by *pks*+ *E. coli*.

**Conclusion** These data reveal a new paradigm for carcinogenesis, in which colibactin-induced senescence has an important role.

## Significance of this study

**What is already known about this subject?**

- Colibactin-producing *E. coli* have been found in 50–60% of human colorectal tumours compared with 20% in controls.
- Colibactin-producing *E. coli* induce DNA damage, cell cycle arrest, mutations and chromosomal instability.
- Colibactin-producing *E. coli* promote colorectal cancer in a murine AOM/IL-$10^{-/-}$ (azoxymethane/interleukin) mouse model.

**What are the new findings?**

- Colibactin-producing *E. coli* indirectly enhance tumour growth by inducing the emergence of senescent cells secreting growth factors.
- The hepatocyte growth factor (HGF) is the main mechanistic link between *pks*+ *E. coli*-induced senescence and tumour growth.
- The mechanism underlying colibactin-induced senescence involves microRNA-20a-5p, which targets SENP1, a key protein in the regulation of p53 SUMOylation.
- The expression of microRNA-20a-5p, SENP1, HGF and activated HGF receptor in human colon cancers is influenced by the presence of *pks*+ *E. coli*.

**How might it impact on clinical practice in the foreseeable future?**

- The data support a new paradigm for colorectal carcinogenesis in which colibactin-producing bacteria favour the emergence of senescent cells secreting growth factors. The colonisation of colon tumours by *pks*+ bacteria may affect the prognosis of colorectal cancer and therefore the management of curative and/or preventive treatments. Targeting colibactin production may be a strategy to restrain the production of pro-tumourigenic factors.

## INTRODUCTION

Cancer is defined as uncontrolled cell proliferation caused by accumulated genetic and epigenetic mutations.[1–3] The origin of these mutations can be multifactorial and 20% of cancers worldwide are estimated to be attributable to infections.[4 5] Many cancers occur in tissues with a high exposure to microbiota, especially colorectal cancers (CRCs), the fourth most common cancer with more than one million cases a year.[6] The high density of microorganisms seen in the colon ($10^{13}$ commensal bacteria colonising the colon vs from $10^{2}$ for the duodenum to $10^{8}$ for the distal ileum) is associated with a greater risk of cancer than from microorganisms in the small intestine.[5] The role of microbiota in colon cancer has been shown by studies in which mice genetically susceptible to CRC, or mice with chemically induced chronic inflammation, developed

significantly fewer tumours under germ-free conditions than when harbouring conventional microbiota.[7 8] The role of microbiota in the development of CRC might be related to the procarcinogenic features of the intestinal bacteria, which may initiate cancer development and/or have a competitive advantage in the tumour microenvironment that allows them to play a role in disease progression.[9 10] However, few putative procarcinogenic bacteria have been identified. The *Enterococcus faecalis* species produces superoxide, which can induce DNA damage and chromosomal instability in colonic epithelial cells.[11] Enterotoxigenic *Bacteroides fragilis* produce the fragilysin toxin, which is genotoxic,[10] and causes cell proliferation by cleavage of the tumour suppression factor E-cadherin,[12] and increases tumourigenesis by triggering production of interleukin-17.[13] Finally, colonic malignancies allow *Streptococcus gallolyticus* subsp *gallolyticus* to colonise established colorectal tumours, leading to tumour development by induction of the proinflammatory cyclo-oxygenase 2 pathway.[14]

*Escherichia coli* are the predominant aero-anaerobic Gram-negative bacteria in the normal intestinal microbiota. As a commensal, *E. coli* coexist harmoniously with their mammalian host, promote normal intestinal homeostasis and rarely cause disease. However, some virulent *E. coli* that have acquired pathogenicity islands can colonise the human gastrointestinal tract and induce disease. Mucosa-associated *E. coli*, which are more frequently identified in colon tissue from patients with adenocarcinomas than in controls, can have procarcinogenic features[15 16] that promote downregulation of DNA mismatch repair proteins.[17] In addition, various toxins produced by pathogenic *E. coli* modulate the cell cycle of host cells such as cytotoxic necrotising factor, cycle-inhibiting factor, cytolethal distending toxins and colibactin.[9 18] Colibactin is a genotoxin produced by *E. coli* harbouring the *pks* island, which encodes enzymes responsible for its synthesis.[19] Colibactin induces double-strand DNA breaks, cell cycle arrest and megalocytosis.[19] Colibactin-induced DNA damage is associated with mutations and chromosomal instability,[20] lesions often seen in cancers. Interestingly, *pks*-harbouring *E. coli* (*pks*+ *E. coli*) strains were recently reported in 50–60% of human colorectal tumours versus 20% in controls[21 22] and, as shown with an AOM/IL-$10^{-/-}$ (azoxymethane/interleukin) mouse model, *pks*+ *E. coli* have a carcinogenic effect independent of inflammation.[21]

In this study, we investigated the mechanisms underlying *pks*+ *E. coli*-induced carcinogenesis. We report that *pks*+ *E. coli* infection induces cellular senescence, characterised by a high production of growth factors that promote proliferation of uninfected cells and, subsequently, tumour growth.

## METHODS

### Bacterial strains

The laboratory *E. coli* strain DH10B hosting the pBeloBAC 11 bearing the *pks* island (*pks*+ *E. coli*), or hosting an empty pBeloBAC 11 (*pks*− *E. coli*) were used.[19] The clinical *pks*+ *E. coli* strain (CCR20) was isolated from a human colon tumour.[22] The isogenic mutant of *E. coli* CCR20 designated CCR20Δ*pks* was generated as indicated in the online supplementary section. Bacteria were grown at 37°C in Luria-Bertani (LB) medium. Bacterial inoculums were assessed at optical density 620 nm using a UV-1800 spectrophotometer (Shimadzu) and controlled by bacterial spreading on LB agar plates.

### Cell culture and infection

Intestinal epithelial cells were grown following ATCC guidelines and infected at a multiplicity of infection (MOI) of 500, as previously described.[19] After 3 h of infection, cells were washed three times with phosphate-buffered saline, and culture medium supplemented with 200 μg/mL gentamicin was added to kill remaining bacteria. The culture medium was then changed every 2 days. To prepare the conditioned medium (CM), $1.5\times10^5$ cells were infected as described above and at 1, 2 or 4 days after infection, the medium was removed, washed three times with phosphate-buffered saline and replaced by serum-free medium. Twenty-four hours later, the medium was collected and used as CM to culture $5.0\times10^4$ uninfected cells. For growth factor inhibition, conditioned media were supplemented with 4 nM hepatocyte growth factor (HGF) inhibitor (JNJ-38877605, Selleckchem), 4 nM fibroblast growth factor (FGF) inhibitor (PD173074, Selleckchem) or 0.1 mg/mL of anti-HGF, anti-granulocyte macrophage colony-stimulating factor (anti-GM-CSF), anti-FGF neutralising antibody (Sigma-Aldrich) or rat IgG isotype control (Biolegend). For all the xenograft experiments, cellular viability was assessed using trypan blue dye.

### *P53* status

HCT116, Int-407 and IEC-6 harbour a wild-type (WT) *p53* gene, whereas HCT8, Caco2 and HT-29 harbour a mutated *p53*.

### Tumourigenesis assay using *in vivo* infected cells

$2\times10^6$ HCT116 cells mixed with *pks*+ *E. coli* or *pks*− *E. coli* (MOI=20) were embedded in growth factor-reduced Matrigel (Becton Dickinson) to obtain a tumour easily measurable with a calliper and to decrease the bias due to growth factors present in Matrigel. The cells were then immediately subcutaneously injected (200 μL) into the dorsal flap of 5-week-old nude mice (Charles River). Each mouse received one xenograft. Three hours after injection, the mice received intraperitoneally 150 mg/kg of broad-spectrum antibiotic imipenem (MSD), and absence of bacteria was confirmed by plating tumour homogenates and blood on LB plates 24 h after treatment.

### Tumourigenesis assay using *in vitro* infected cells

HCT116 cells were infected *in vitro* as described above with *pks* + or *pks*− *E. coli*. *pks*+ *E. coli*-induced senescence was checked 5 days after infection using a senescence-associated β-galactosidase (SA-β-gal) assay as described below. Cells were trypsinised 5 days after infection, and $0.4\times10^6$ previously infected cells were mixed with $1.6\times10^6$ uninfected HCT116 cells. The resulting cellular mixes were embedded in growth factor-reduced Matrigel (Becton Dickinson), subcutaneously injected into the dorsal flap of 5-week-old nude mice and the tumour size was measured as previously described.[37] Each mouse received one xenograft.

### AOM/DSS murine model

Colon cancer was induced in C57/B6 mice (6 weeks old, Charles Rivers), as previously described,[23] with some modifications. Mice were intraperitoneally injected with AOM (10 mg/kg body weight) and treated with streptomycin in drinking water (2 mg/mL) for 2 days to facilitate bacterial colonisation.[24] One day after the end of the antibiotic treatment, they received by gavage $10^9$ colony forming units of a clinical *pks*+ *E. coli* strain (CCR20) or its isogenic mutant (CCR20Δ*pks*), and were maintained on regular diet and water for 5 days. They were subjected to two cycles of 2% dextran sodium sulfate (DSS) treatment[23] and then killed. Colonic tumours were counted and measured with a dissecting microscope, then Swiss-rolled and fixed for histology. Two macroscopic tumours from each mouse were collected and pooled for western blot and quantitative

real-time PCR. In parallel, bacterial colonisation was assessed by counting CCR20 and CCR20Δ*pks* in faeces, healthy colonic tissues and malignant colon tumours by spreading on selective agar (bioMérieux).

### Human biopsy specimens

Biopsy specimens of colon adenocarcinomas were collected in a previously published study,[22] from colon resections, which were required for the treatment of patients. Ethical approval for the study was granted by the Clermont-Ferrand research ethics committee. Verbal informed consent to participate in the research was obtained from all patients included in the study in accordance with the French bioethics law (Act No 2004-800 of 6 August 2004). Samples were taken on resected colon at the site of malignant tumours. Pathological analysis confirmed the neoplastic features of the samples. Age and sex of patients, tumour node metastases stage, neoplastic grade and inflammatory score are given in online supplementary table S1. We included randomly selected biopsy specimens colonised by *pks*+ *E. coli* (n=8), for which the functionality of *pks* island was confirmed *in vitro*.[19] The biopsy specimens colonised by *pks*− *E. coli* (n=9) were also randomly selected.

### Senescence-associated β-galactosidase staining

Senescent cells were visualised using the Senescence Cells Histochemical Staining Kit (Sigma-Aldrich) according to the manufacturer's instructions. *In vivo* β-galactosidase staining was performed as previously described.[25]

### Senescence-associated β-galactosidase activity

Human biopsy specimens and mouse tissues (~10 mg) were lysed using an ultra-turrax homogeniser (IKA) in a reaction buffer (200 µL containing 10 mM potassium ferricyanide, 300 mM NaCl, 4 mM $MgCl_2$ and 0.2 M phosphate buffer pH6.0). After centrifugation, 25 µL of supernatants was supplemented with an equal volume of 2 mg/mL 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (Sigma-Aldrich) and hydrolysis was monitored at 420 nm. Enzymatic activity was normalised against protein concentration. SA-β-gal activity was correlated ($r^2$=0.898) with the number of SA-β-gal-positive cells counted with the Senescence Cells Histochemical Staining Kit (Sigma-Aldrich).

### Statistical analysis

Statistical analyses were performed with R software (V.2.13.2, http://www.r-project.org/). A p value < 0.05 was considered statistically significant. Redundancy analyses (RDAs) were performed with the Vegan package (V.2.0-2, http://cran.r-project.org/web/packages/vegan/index.html). Significant differences were investigated by 1000 permutation tests.

## RESULTS

### Colibactin-producing *E. coli* infection increases tumour growth

Intestinal epithelial cells HCT116 mixed with *E. coli* DH10B strain producing colibactin (*pks*+) or not (*pks*−) were subcutaneously injected into nude mice. Three hours later, a broad-spectrum antibiotic (imipenem) was administered to kill bacteria. This single and transient exposure to *pks*+ *E. coli* significantly increased both tumour growth (figure 1A) and the number of Ki67-positive cells (a cellular proliferation marker) (figure 1B). Nude mice were then subcutaneously injected with HCT116 cells infected *in vitro* in combination with uninfected HCT116 cells (20% and 80%, respectively) and collected 5 days after infection. Xenografts containing cells infected by *pks*+ *E. coli* *in vitro* grew more rapidly and harboured a higher number of Ki67-positive cells than those containing *pks*− *E. coli*-infected cells (Figure 1C,D). Tumours composed of *pks*− *E. coli*-infected cells exhibited a similar development to that seen with tumours composed of uninfected cells (figure 1A,C), suggesting that *pks*− *E. coli* DH10B have no significant effect on tumour growth.

HCT116 cells were infected *in vitro* with *pks*+ or *pks*− *E. coli* for 3 h and cellular proliferation was assessed 24 h after infection. We observed a ~20% drop in cell proliferation of *pks*+ *E. coli*-infected cells compared with *pks*− *E. coli*-infected cells (see online supplementary figure S1). This decrease was probably due to colibactin-induced cell cycle arrest.[19] The drop in cell proliferation seen *in vitro* and the pro-proliferative effect seen in the xenograft mouse model in response to *pks*+ *E. coli* infection suggested an indirect effect of bacteria on tumour growth *in vivo*. We therefore hypothesised that cells infected with *pks*+ *E. coli* secrete factors that induce proliferation of uninfected cells. We assessed the ability of conditioned medium (CM) derived from HCT116 cells infected with *E. coli* for 3 h to promote cell proliferation. CMs were obtained by collecting serum-free supernatant of HCT116 cells at days 1, 3 or 5 after infection. Cells incubated with CM derived from *pks*+ *E. coli*-infected cells at 5 days after infection proliferated as fast as cells incubated with a medium containing 10% fetal bovine serum, whereas CMs derived from *pks*− *E. coli*-infected cells did not affect cell proliferation (see online supplementary figure S2). Similar results were obtained using various human intestinal epithelial cells (Int-407, HCT8, Caco2, HT-29; see online supplementary figure S3) and IEC-6 non-transformed cells (see online supplementary figure S3). These results confirm an indirect effect of colibactin-producing *E. coli* on tumour growth.

### Colibactin-producing *E. coli* induce senescence of intestinal epithelial cells

Colibactin is known to induce megalocytosis and cell cycle arrest,[19] which are features of cellular senescence.[26] SA-β-gal activity was therefore assessed in infected cells. *pks*+ *E. coli* infection induced an increase in the number of SA-β-bal positive megalocytes 5 days after infection (figure 2A,B and see online supplementary figure S4A). Similar results were obtained with various human intestinal epithelial cells and non-transformed cells (see online supplementary figure S4B). In contrast, the number of SA-β-gal positive cells induced by *pks*− *E. coli* was similar to that of non-infected cells (figure 2A,B), indicating that the induction of senescence was due to colibactin synthesis. Of note, we did not find a significant increase in either apoptosis or necrosis of *pks*+ *E. coli*-infected cells compared with *pks*− *E. coli*-infected cells, as assessed by lactate dehydrogenase release and caspase 3 activation (data not shown).

At day 5 after infection *pks*+ *E. coli*-infected cells exhibited an accumulation of phospho-p53, $p21^{Cip}$ and Rb, and a decrease in E2F-1 expression (figure 2C), a protein expression profile characteristic of senescent cells,[27] confirming that *pks*+ *E. coli* induce senescence of human intestinal epithelial cells. To investigate whether *pks*+ *E. coli* promote cellular senescence *in vivo* at early time points in tumour development, nude mice were subcutaneously injected with HCT116 cells mixed with *E. coli* as described in figure 1A and tumours were analysed at day 10 after implantation. Tumour sizes were comparable at this early point in animals injected with cells infected by *pks*+ *E. coli* and by *pks*− *E. coli* (figure 1A). However, tumours derived from *pks*+ *E. coli*-infected HCT116 cells showed a significant increase in SA-β-gal activity compared with tumours derived from *pks*− *E. coli*-infected cells (figure 2D).

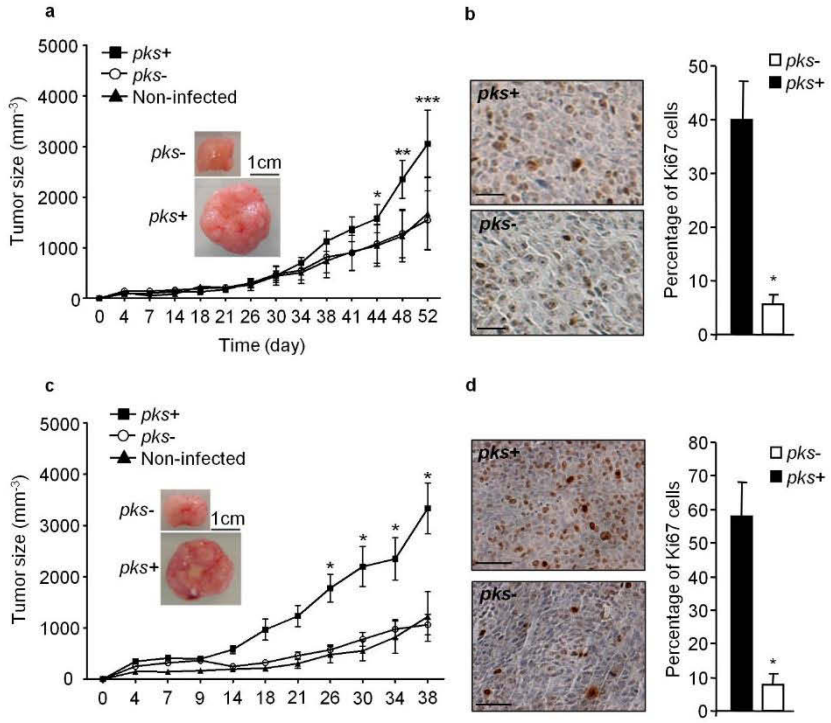



**Figure 1** Infection with *pks+ E. coli* promotes tumour growth and cell proliferation. (A) $2\times10^6$ HCT116 cells mixed with *pks+ E. coli* or *pks− E. coli* (MOI=20) were embedded in Matrigel, and immediately subcutaneously injected into the dorsal flap of 5-week-old nude mice. Three hours after the injection, the mice received intraperitoneally 150 mg/kg of the broad-spectrum antibiotic imipenem to kill bacteria. Tumour size was monitored during the indicated period (n=5/group; **p<0.01; ***p<0.001, by analysis of variance (ANOVA)). Representative photographs of tumours are presented (scale bar: 1 cm). (B) Representative Ki67 immunohistochemical staining in xenograft 52 days after infection. Scale bars: 100 μm. The bar graph on the right represents measurement of Ki67 positive cells. Data are presented as means±SEM (*p<0.05, by Mann–Whitney test). (C) $0.4\times10^6$ HCT116 cells, infected *in vitro* by *pks+* or *pks− E. coli* and collected 5 days after infection, were mixed with $1.6\times10^6$ non-infected HCT116 cells and subcutaneously injected into nude mice. Tumour size was monitored during the indicated period (n=10/group; *p<0.05, by ANOVA). Representative photographs of tumours are presented (scale bar: 1 cm). (D) Representative Ki67 immunohistochemical staining in xenograft 38 days after infection. Scale bars: 100 μm. The bar graph on the right represents measurement of Ki67 positive cells. Data are presented as means ±SEM (*p<0.05, by Mann–Whitney test).

Senescent cells are metabolically active and usually acquire a senescence-associated secretory phenotype (SASP).[26 28] SASP was therefore investigated at day 5 after infection. Interestingly, *pks+ E. coli*-infected cells expressed significantly higher levels of the growth factors HGF, FGF, GM-CSF, bone morphogenetic protein 4 (BMP4) and vascular endothelial growth factor (VEGF) than *pks− E. coli*-infected cells (see online supplementary figure S5A). Similar results were obtained with five human intestinal epithelial cell lines and non-transformed cells (see online supplementary figure S5B). Of note, expression levels of FGF, GM-CSF and HGF only were also increased in xenografts 10 days after *pks+ E. coli* infection compared with xenografts infected with *pks− E. coli* (figure 2E).

To identify the relevant secreted mediators involved in cell proliferation, HCT116 cells were cultured with conditioned media supplemented with growth factor inhibitors. HGF pathway inhibitor, in contrast to vehicle or FGF inhibitor, abrogated *in vitro* the pro-proliferative activity of conditioned media derived from *pks+ E. coli*-infected cells (figure 2F). Similar results were obtained using neutralising antibodies; HGF antibodies, in contrast to GM-CSF or FGF antibodies, blocked the pro-proliferative effect (see online supplementary figure S6). As seen *in vitro*, HGF inhibitor significantly blocked the growth of xenograft obtained from cells infected with *pks+ E. coli* (figure 2G). Together, these data show that the pro-proliferative effect of *pks+ E. coli*-induced senescent cells is mainly mediated by HGF.

### SENP1 downregulation and p53 SUMOylation are key features of *pks+ E. coli*-induced senescence

The SUMOylation of p53 has emerged as a key determinant of cellular senescence.[29] Using HA-tagged small ubiquitin-like modifier 1 (SUMO1)-expressing cells infected with *E. coli*, we observed that *pks+ E. coli* increased the SUMOylation of p53 (figure 3A) and that this modification was associated with a downregulation of SENP1 (figure 3B), a negative regulator of p53 SUMOylation.[30] To confirm the role of SENP1 in *pks+ E. coli*-induced senescence, HCT116 cells were transfected with a vector encoding SENP1.[31] Overexpression of SENP1 significantly decreased the number of senescent cells induced by *pks+ E. coli* infection to a level similar to that found with *pks− E. coli*-infected cells (figure 3C). This effect was dependent on SENP1 enzymatic activity since the expression of an inactive mutant of SENP1 (C602S) did not protect against *pks+*

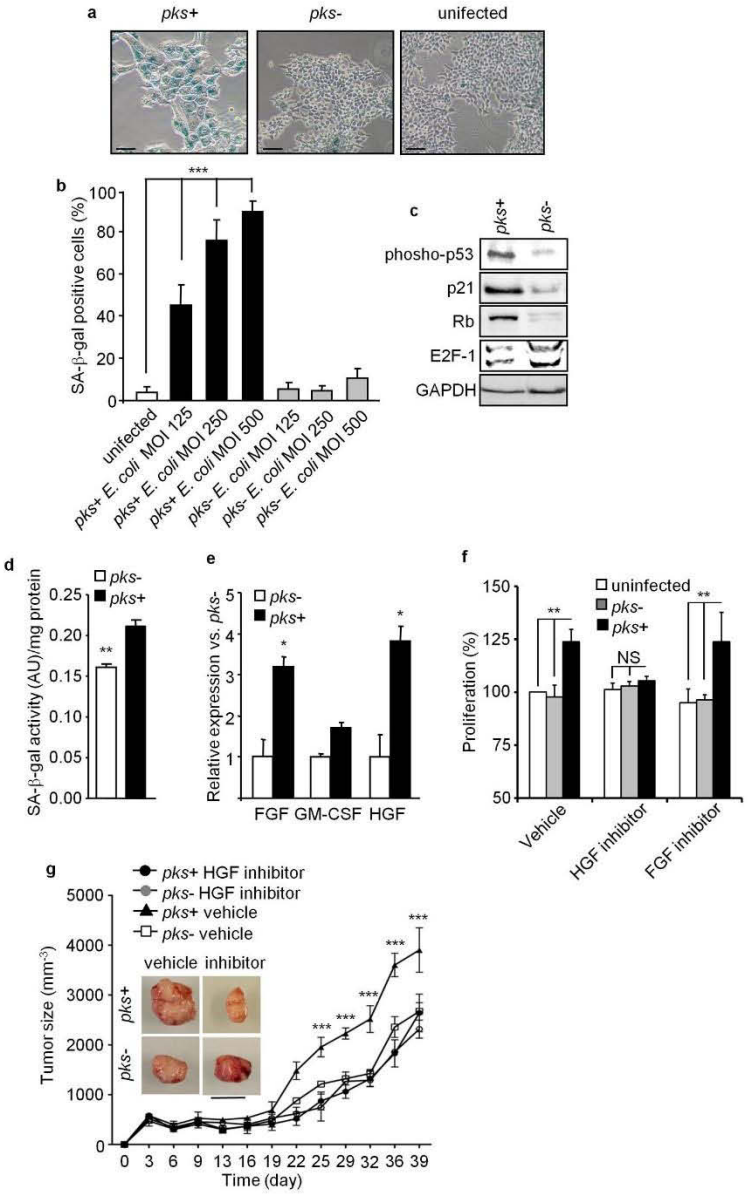



**Figure 2** Colibactin-producing *E. coli* induce cellular senescence. (A) HCT116 cells were infected with *pks*+ *E. coli* or *pks*− *E. coli*, and SA-β-gal positive cells were assessed 5 days after infection. The three figures are at the same magnification (scale bars: 100 μm). (B) Percentage of SA-β-gal positive cells determined 5 days after infection. Data are from three experiments performed twice and are presented as means±SEM (n=6/group; ***p<0.001 vs uninfected cells, by Mann–Whitney test). (C) Whole-cell lysates were analysed by immunoblotting experiments 5 days after infection. (D) $2\times10^6$ HCT116 cells mixed with *pks*+ *E. coli* or *pks*− *E. coli* (MOI=20) were embedded in Matrigel, and immediately subcutaneously injected (200 μL) into the dorsal flap of 5-week-old nude mice. Three hours after injection, the mice received intraperitoneally 150 mg/kg of broad-spectrum antibiotic imipenem. SA-β-gal activity was assessed in tumours at 10 days after infection (n=5/group; **p<0.05, by Mann–Whitney test). (E) Expression levels of growth factors mRNA were assessed by qRT-PCR in mice tumours 10 days after infection. Data are presented as means±SEM (n=5/group,*p<0.05; by Mann–Whitney test). (F) HCT116 cells were cultured in presence of conditioned media derived from *pks*+ *E. coli* or *pks*− *E. coli*-infected HCT116 cells and supplemented with growth factor inhibitors. Cellular growth was assessed by a MTT assay after 24 h incubation (n=8/group; **p<0.01; NS, not significant vs *pks*−, by Mann–Whitney test). Data are from three experiments performed twice and are presented as means±SEM. (G) $0.4\times10^6$ HCT116 cells collected 5 days after *pks*+ or *pks*− *E. coli* infection, were mixed with $1.6\times10^6$ non-infected HCT116 cells and subcutaneously injected into nude mice. The mice received 3.4 μg/kg hepatocyte growth factor (HGF) inhibitor or vehicle twice a week. Tumour size was monitored during the indicated period (n=5 mice/group: ***p<0.001 vs other groups, by two-way analysis of variance). Representative photographs of tumours are presented (scale bar: 1 cm).

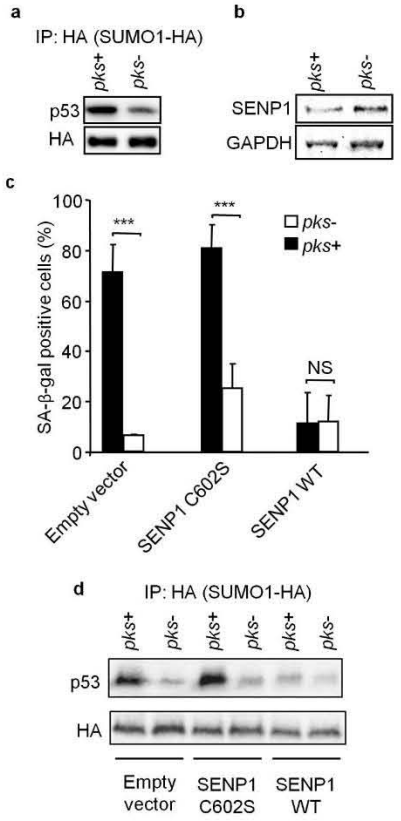



**Figure 3** SUMOylation controls senescence upon infection with *pks+ E. coli*. (A) Small ubiquitin-like modifier (SUMO)-conjugated p53 was assessed in HCT116 intestinal epithelial cells in response to infection by *pks+ E. coli* or *pks− E. coli*. HCT116 cells transiently transfected with a vector encoding SUMO1 HA-tagged protein, were lysed 5 days after infection and then immunoprecipitated using an antibody specific for HA. The amount of co-immunoprecipitated p53 was examined by western blot. (B) Immunoblot analysis of SENP1 expression in HCT116 cells 5 days after infection. (C, D) HCT116 cells transiently transfected with an empty vector or a vector encoding SENP1 (SENP1 WT) or an inactive mutant (SENP1 C602S) were infected with *pks+ E. coli* or *pks− E. coli*. (C) The bar graph represents the percentage of SA-β-gal positive cells 5 days after infection. SUMO-conjugated p53 was assessed by co-immunoprecipitation and western blot analysis as indicated in panel A. Data are from three experiments repeated twice and are presented as means±SEM (n=6/group; NS, not significant; ***p<0.001, by Mann–Whitney test). (D) SUMO1-conjugated protein patterns were assessed 5 days after infection.

*E. coli*-induced senescence (figure 3C). In addition, overexpression of wild-type SENP1, unlike overexpression of the inactive SENP1, abolished the *pks+ E. coli*-induced accumulation in SUMO1-conjugated p53 (figure 3D). During these experiments, the genotoxic effect of *pks+ E. coli* was not affected (see online supplementary figure S7) and conditioned media derived from *pks+ E. coli*-infected cells overexpressing functional SENP1 did not promote HCT116 cell proliferation (see online supplementary figure S8). These data show that SENP1 downregulation and subsequent p53 SUMOylation are key features in *pks+ E. coli*-induced cell senescence.

### Colibactin-producing *E. coli* decrease SENP1 expression via miR-20a-5p and c-Myc

Among the microRNAs previously reported to be deregulated during senescence,[32] *in silico* predictions showed that miR-20a-5p potentially targets SENP1 (see online supplementary figure S9). Interestingly, miR-20a-5p expression was significantly upregulated in *pks+ E. coli* infected HCT116 cells compared with *pks− E. coli* infected cells (figure 4A). Furthermore, transfection of cells with mature miR-20a-5p decreased SENP1 expression at both mRNA and protein levels (Figure 4B,C). The *SENP1* mRNA 3′-UTR was cloned downstream of the luciferase gene and this construct was transfected into cells in the presence of miR-20a-5p or miR-control. MiR-20a-5p significantly repressed luciferase activity (figure 4D), suggesting the binding of miR-20a-5p to the *SENP1* mRNA 3′-UTR. These results indicate that *pks+ E. coli* infection induces miR-20a-5p expression, which in turn downregulates SENP1 expression by targeting *SENP1* mRNA 3′-UTR.

We next investigated the role of miR-20a-5p in senescence and cell proliferation. HCT116 cells were transfected with mature miR-20a-5p or a miR-control and SA-β-gal positive cells were counted 5 days after transfection. In cells transfected with miR-20a-5p, a significant increase in the number of senescent cells was seen compared with cells transfected with miR-control (figure 4E). In addition, the transfection of an anti-miR-20a-5p in HCT116 cells significantly decreased the percentage of senescent cells and the level of SUMO-conjugated p53 upon *pks+ E. coli* infection, in contrast to the transfection of an anti-miR control (figure 4F and online supplementary figure S10A, respectively). As observed after *pks+ E. coli* infection, cells transfected with miR-20a-5p produced a significant level of HGF and their conditioned media induced cell proliferation (see online supplementary figures S11A,B). In contrast, anti-miR-20a-5p abolished *pks+ E. coli*-induced HGF expression and cell proliferation induced by conditioned media obtained from *pks+ E. coli*-infected cells (see online supplementary figures S11C,D). These results show that *pks+ E. coli* induce senescence via miR-20a-5p and, subsequently, can stimulate cell proliferation.

MiR-20a-5p is known to be regulated by the transcription factor c-Myc,[33] which is involved in DNA damage response.[34] We observed that *pks+ E. coli* infection induced c-Myc expression (figure 4G) and c-Myc binding to the miR-20a-5p promoter (see online supplementary figure S12). Transfection of *pks+ E. coli*-infected cells with c-Myc siRNA reduced the level of SUMO-conjugated p53 (see online supplementary figure S10B), strongly reduced the expression of miR20a-5p, increased the expression of SENP1, decreased the number of *pks+ E. coli*-induced senescent cells and consequently, abolished the pro-proliferative effect of conditioned media derived from *pks+ E. coli*-infected cells (figure 4H–K). These findings suggest that c-Myc plays an upstream role in the cellular signalling cascade induced by *pks+ E. coli*. Finally, we observed that *pks+ E. coli* induced c-Myc expression 3 days after infection (see online supplementary figure S13A), followed 2 days later by modifications of miR-20a-5p (see online supplementary figure S13B) and SENP1 expression levels (see online supplementary figure S13A). Altogether, our data demonstrate that *pks+* bacteria regulate miR-20a-5p expression and cellular senescence via c-Myc.

### Modified expression of miR-20a-5p, SENP1 and growth factors in mouse and human colon tumours colonised by *pks+ E. coli*

To confirm the reliability of our findings, we explored the impact of a clinical *E. coli* strain isolated from a human CRC biopsy

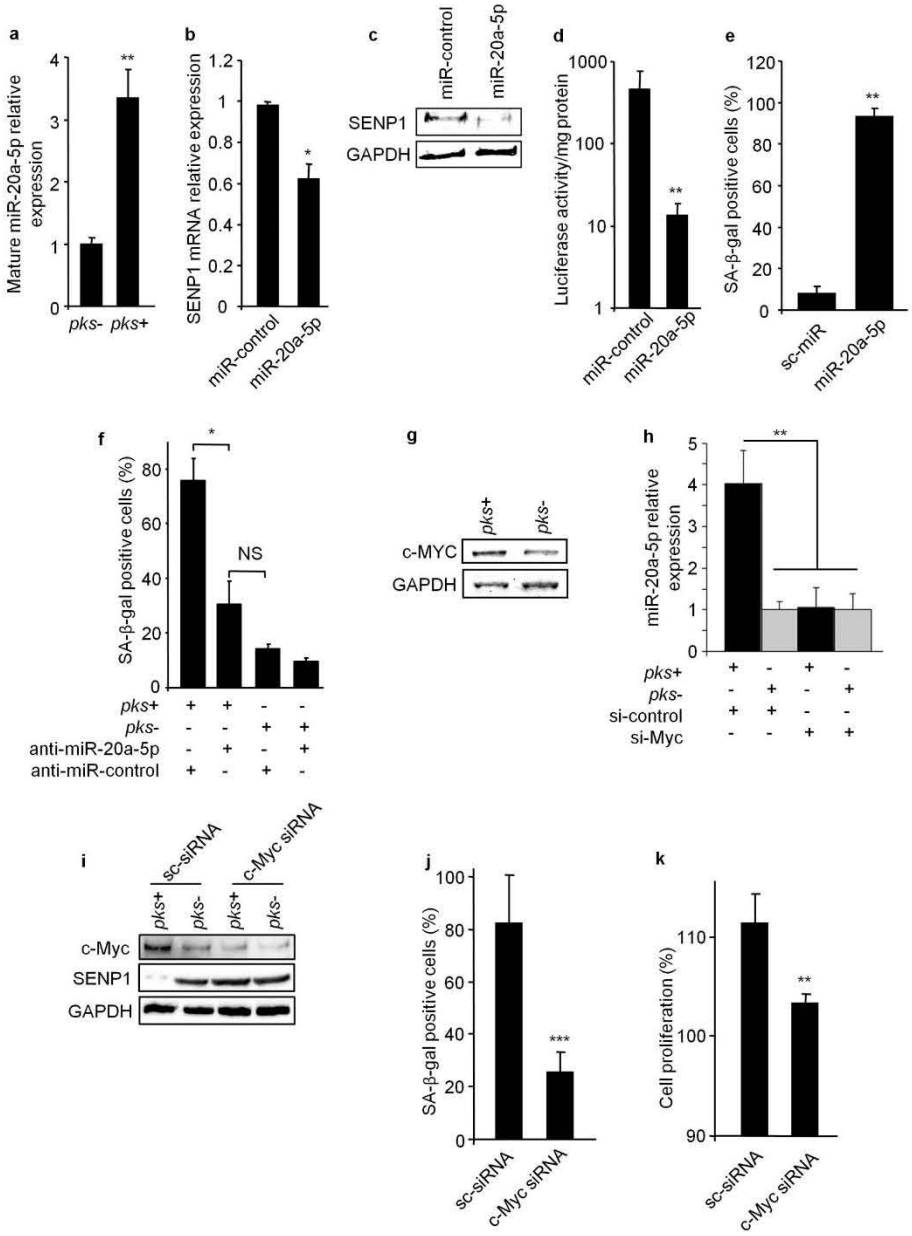



**Figure 4** MiR-20a-5p represses SENP1 expression and controls the senescence induced by *pks+ E. coli*. (A) Mature forms of miR-20a-5p were quantified by qRT-PCR in HCT116 cells 1, 3 and 5 days after infection. (B, C) HCT116 cells were transfected with precursors of miR-20a-5p or miR-control, and SENP1 expression was assessed by (B) qRT-PCR and (C) western blot. (D) HCT116 cells were transfected with a luciferase vector encoding the *SENP1* mRNA 3′-UTR in the presence of miR-20a-5p or miR-control. Luciferase activity was determined. (E) HCT116 cells were transfected with the precursor of miR-20a-5p or miR-control and senescence was assessed at day 5 by detecting SA-β-gal positive cells. (F) HCT116 cells were transfected with anti-miR-20a-5p or anti-miR-control and infected with *pks+ E. coli* or *pks− E. coli*. Five days after infection senescent cells were assessed by detecting SA-β-gal positive cells. (G) HCT116 cells were infected with *pks+* or *pks− E. coli* and c-Myc expression was analysed by western blot at day 5 after infection. (H) HCT116 cells were transfected with c-Myc siRNA or scramble (sc) siRNA and miR-20a-5p expression level was assessed by qRT-PCR. (I) HCT116 cells transfected with c-Myc siRNA or sc-siRNA were infected with *pks+* or *pks− E. coli* and c-Myc and SENP1 expression levels were assessed by western blot. (J) HCT116 cells transfected with c-Myc siRNA or sc-siRNA were infected with *pks+ E. coli*, and SA-β-gal positive cells were determined. (K) Conditioned media were derived from HCT116 cells transfected with c-Myc siRNA or sc-siRNA and infected with *pks+ E. coli*. Conditioned media were used to culture HCT116 cells, and cellular growth was assessed by a MTT assay. All data are presented as means±SEM (n=6/group; NS, not significant; *p<0.05; **p<0.01; ***p<0.001, by Mann–Whitney test (A, B, D, E, J and K) or by analysis of variance (F)).

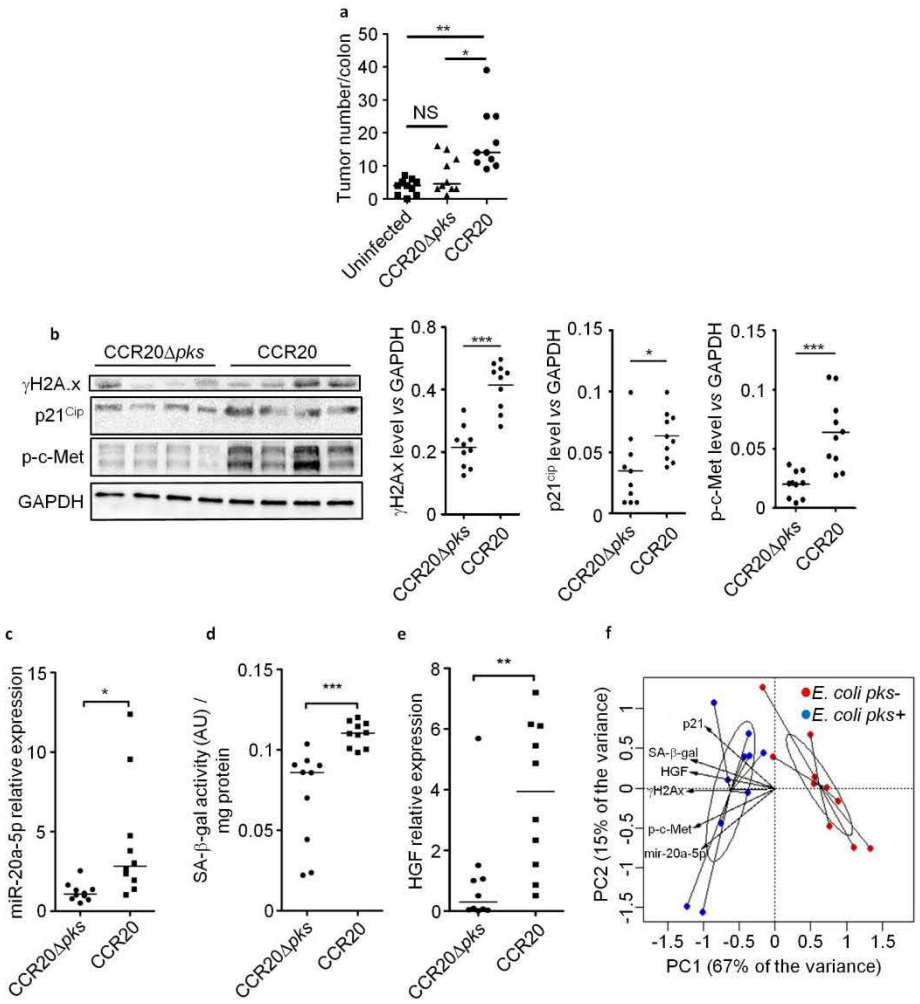



**Figure 5** Clinical *pks+ E. coli* (CCR20) induced the miR-20a-5p/HGF pathway and consequently promoted tumourigenesis in an AOM/DSS mouse model. (A) Number of tumours per mouse at the end of the AOM/DSS protocol (n=10/group; *p<0.05; **p<0.01; NS, not significant by analysis of variance). (B) γH2A.x, p21$^{cip}$, and phospho-c-Met (p-c-Met) protein levels in colon tumours were assessed by western blot. The picture presents representative results obtained from four samples in each group (colonised with *pks+ E. coli* or *pks− E. coli*). Bar graphs on the right represent the densitometric quantification in the entire tissue collection (n=10/group; *p<0.05; ***p<0.001, by Mann–Whitney test). (C) Mature miR-20a-5p levels in colon tumours from AOM/DSS-treated mice were assessed by qRT-PCR (n=10/group; *p<0.05, by Mann–Whitney test). (D) SA-β-gal activity was assessed in colon tumours from AOM/DSS-treated mice (n=10/group; ***p<0.001, by Mann–Whitney test). (E) HGF mRNA levels in colon tumours from AOM/DSS-treated mice were assessed by qRT-PCR (n=10/group; **p<0.01, by Mann–Whitney test). (F) Redundancy analysis of SA-β-gal activity and expression levels of miR-20a-5p, HGF mRNA, γH2A.x, p21$^{cip}$ and p-c-MET in colon tumours from *pks+ E. coli* or *pks− E. coli* colonised mice after AOM/DSS treatment (n=10/group).

specimen and harbouring the *pks* island (CCR20) or its isogenic mutant (CCR20Δ*pks*) in an AOM/DSS colon tumour mouse model (see online supplementary figure S14A). Bacterial colonisation of the intestinal tract (faeces and intestine tissues) was persistent and similar in CCR20- and CCR20Δ*pks*-infected mice (see online supplementary figures S14B,C). CCR20 and CCR20Δ*pks* were predominantly found in colon tumours (see online supplementary figure S14C), as previously reported in human CRC biopsy specimens.[15 16] CCR20 colonisation significantly increased the number of tumours in comparison with CCR20Δ*pks* colonisation (figure 5A), without affecting the inflammatory score, neoplastic grade or the size of the tumours (see online supplementary figures S14D–F). Furthermore, no significant difference in tumour numbers was seen between CCR20Δ*pks*-infected mice and uninfected mice (figure 5A). In addition, DNA damage (γH2A.x), miR-20a-5p expression level and senescence markers (SA-β-gal activity and p21$^{Cip}$ expression) were significantly higher in tumours isolated from mice colonised by CCR20 than in those isolated from mice colonised by CCR20Δ*pks* (figure 5B–D). Finally, CCR20 induced high levels in tumours of both HGF mRNA and activated HGF receptors (phosphorylated c-Met) (figures 5B,E, respectively). Interestingly, redundancy analysis (RDA) showed a correlation between these factors and discriminated between tumours according to their *pks E. coli* status (figure 5F).

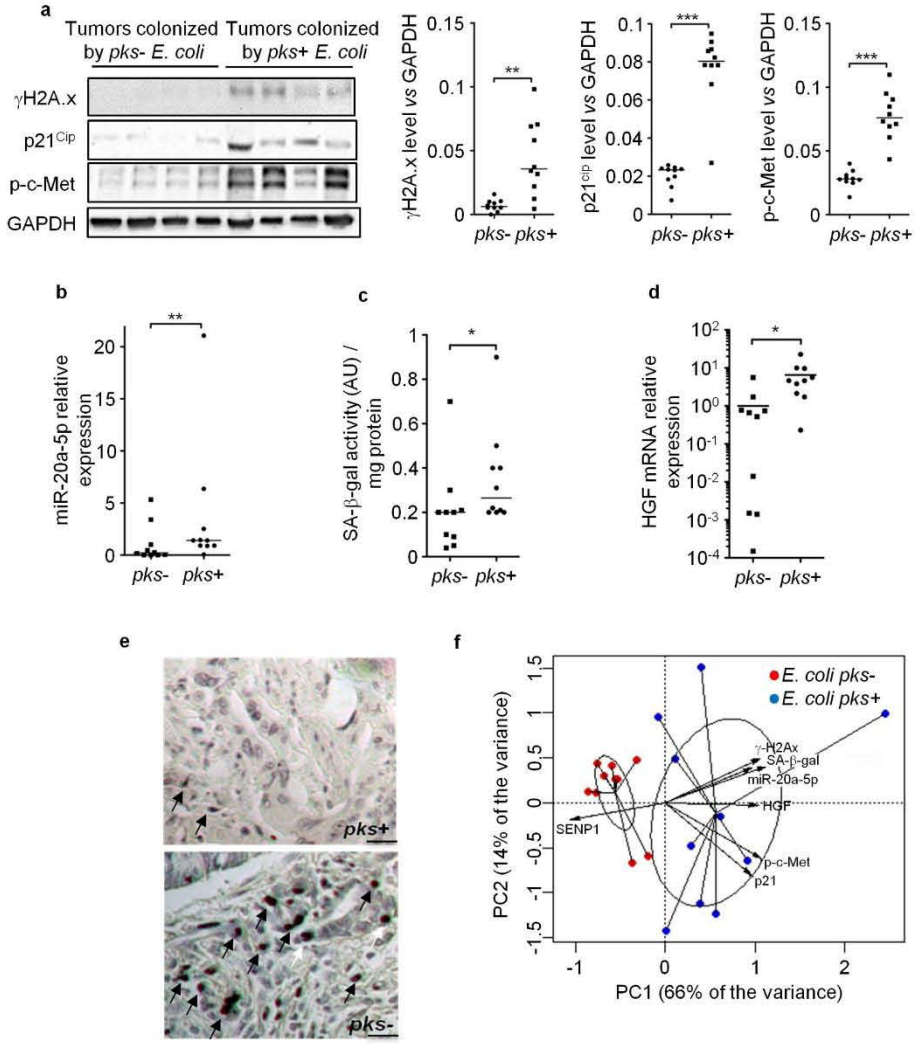



**Figure 6** Modified expression pattern of miR-20a-5p, SENP1 and HGF in human *pks+ E. coli*-colonised colon tumours. (A) γH2A.x, p21[cip], and phospho-c-Met (p-c-Met) protein were assessed by western blot in biopsy samples of human colon cancers. Representative results obtained from four samples in each group (colonised with *pks+ E. coli* or *pks− E. coli*) are shown. Bar graphs on the right represent the densitometric quantification in the entire tissue collection (**p<0.01; ***p<0.001 by Mann–Whitney test). (B) Mature miR-20a-5p levels were assessed by qRT-PCR (**p<0.01, by Mann–Whitney test). (C) SA-β-gal activity was assessed in human biopsy samples (n=10/group; *p<0.05, by Mann–Whitney test). (D) HGF mRNA expression levels were assessed by qRT-PCR (*p<0.05, by Mann–Whitney test). (E) SENP1 immunohistochemical staining in a human biopsy samples. Scale bars: 50 μm. (F) Redundancy analysis of SA-β-gal activity, number of SENP1 positive epithelial cells, and expression levels of miR-20a-5p, HGF mRNA, γH2A.x, p21[cip] and p-c-Met in human biopsy samples colonised with pks+ *E. coli* or pks− *E. coli*.

To substantiate these results, we investigated human colon adenocarcinomas colonised by *pks+ E. coli* or *pks− E. coli* (eight and nine patients/group, respectively). There was no significant difference in tumour node metastases stage, neoplastic grade, inflammatory score, or bacterial colonisation between tumours from the two groups (see online supplementary table S1). However, human biopsy specimens colonised by *pks+ E. coli* harboured high expression levels of γH2A.x (DNA damage marker), miR-20a-5p, senescence markers (SA-β-gal, p21[Cip]), HGF mRNA and activated HGF receptors, as found in the AOM/DSS mouse model, and a decrease in SENP1-expressing cells (figure 6A–E). These biological stigmas significantly differentiated human colon adenocarcinomas colonised by *pks+ E. coli* from those colonised by *pks− E. coli* using RDA (figure 6F). Overall, these data indicate that the miR-20a-5p/SENP1/senescence/HGF pathway is activated by *pks+ E. coli* in human and murine colon adenocarcinomas, as seen *in vitro*, and therefore might contribute to the development of CRC.

## DISCUSSION

*Pks+ E. coli* were recently shown to be over-represented in human colorectal tumours[21 22] and to have a carcinogenic

effect.[21] In addition, it has been reported that genes present in the *pks* island are significantly induced during development of inflammation and CRC.[35] In this study, using mouse models and human CRC biopsy specimens, we show that colibactin-producing *E. coli* contribute to the emergence of senescent cells, which enhance tumour promotion via growth factor secretion (see online supplementary figure S15). A short time of contact between tumour cells and colibactin-producing *E. coli* is sufficient to stimulate tumour growth in the xenograft model, suggesting that even a transient colonisation could affect tumour fate.

The toxic effect of *pks* requires direct bacteria–host cell contact[19] and thus an environment in which bacteria can more readily access the epithelium. Consequently, *pks*+ *E. coli* might not induce adenocarcinomas in healthy colon epithelium,[21] which is protected by a mucus layer and the production of antibacterial peptides. Arthur *et al*[21] developed a model in which inflammation mediated by *IL-10* gene knockout alters the natural barrier function, thereby facilitating adhesion of *pks*+ bacteria to colonic epithelial cells. In the murine AOM/DSS model, the non-genotoxic compound DSS is known to promote inflammation and may therefore favour contact between bacteria and intestinal epithelial cells. In addition, tumoural tissues are known to have altered mucosa barrier function,[36–38] suggesting that colon cancer also creates an environment suitable for the adhesion of bacteria to the colonic epithelial cells.

Initiation of cancer requires genomic changes within cells such as point mutations or chromosomal rearrangements, which activate oncogenes or inactivate tumour suppressors.[3] While genetic changes are responsible for many aspects of cancer development, they cannot alone, promote tumour growth and progression. Inflammation is known to play a major role in tumour promotion especially in CRC.[39] However, like Arthur *et al*[21] we did not find any significant increase in inflammatory score in the colon of *pks*+ *E. coli*-infected mice, which suggests that these bacteria may enhance tumour growth by mechanisms other than inflammation.

We observed that *pks*+ *E. coli* could sustain tumour development by inducing cellular senescence. Colibactin-producing *E. coli*-induced senescent cells acquire a SASP characterised by the production of growth factors, as seen for tumour-associated fibroblast.[40] We found that HGF produced by senescent intestinal epithelial cells induced proliferation of uninfected cells both *in vitro* and *in vivo*. Interestingly, the production of HGF, which is a key determinant of colon cancer progression, a marker of poor prognosis and a target for treatment,[41 42] was significantly increased in CRC biopsy specimens colonised by *pks*+ *E. coli* compared with those harbouring *pks*− *E. coli*. This was correlated with an increase in SA-β-gal activity in mouse and human tumours colonised by *pks*+ *E. coli* compared with that seen in tumours colonised by *pks*− *E. coli*. Most senescent cells undergo cell cycle arrest; however, experiments using intestinal epithelial cells treated with doxorubicin showed that a small fraction of senescent cells escaped from senescence and underwent improper divisions.[43] We cannot rule out the possibility that these cells participated in the tumour growth seen in our study.

The mechanism underlying *pks*+ *E. coli*-induced senescence involves c-Myc, miR-20a-5p and SENP1 (see online supplementary figure S15). Colibactin-producing *E. coli* increased miR-20a-5p expression in infected cells (via c-Myc transcription factor), which bound to *SENP1* mRNA 3'UTR, resulting in its translational silencing. Interestingly, we observed a similar deregulation of SENP1 and miR-20a-5p expression in human CRC biopsy specimens colonised by *pks*+ *E. coli*. SENP1 is a negative regulator of p53 SUMOylation.[30 44] Accordingly, SENP1 downregulation induced by colibactin-producing *E. coli* or by transfection of miR-20a-5p triggered an accumulation of SUMO-conjugated p53, which is consistent with the predominant enzymatic activity of SENP1[44] and known to be a key step in the regulating framework that triggers cellular senescence.[29 30] However, *pks*+ *E. coli*-induced senescence was also found in $p53^{-/-}$ cells suggesting that other pathways, in addition to p53, may be involved in the senescent process.

In conclusion, our data support a new paradigm for colorectal carcinogenesis, in which colibactin-induced senescence has a critical role. Colibactin-producing bacteria induce the emergence of senescent cells, which promote tumour growth by the secretion of growth factors, notably, HGF. Targeting colibactin production may therefore be an efficient strategy to restrain the production of pro-tumourigenic factors.

**Acknowledgements** We thank Professor Andrew D Sharrocks (Faculty of Life Sciences, University of Manchester, UK) for kindly providing us with vector-encoding wild-type or mutated SENP1. We are grateful to Marlène Jan for her excellent technical assistance and the ICCF platform of the Université d'Auvergne for technical assistance with histology.

**Contributors** Study design: AC, GD, RB; acquisition of data and interpretation: AC, GD, RM, EB, LG, MB, CD, PS, PD, JD, DP, HW and RB; drafting of the manuscript: RB, GD and AC; critical revision of the manuscript: RB and AD-M.

**Funding** This work was supported by the Ministère de la Recherche et de la Technologie, the Institut National de la Santé et de la Recherche Médicale (UMR Inserm U1071), the Institut National de la Recherche Agronomique (USC-2018), the Ligue Contre le Cancer and the Centre Hospitalier Régional Universitaire de Clermont-Ferrand, France.

**Competing interests** None.

**Ethics approval** Research ethics committee of University of Clermont-Ferrand, France.

**Provenance and peer review** Not commissioned; externally peer reviewed.

## REFERENCES

1 Esteller M. Cancer epigenomics: DNA methylomes and histone-modification maps. *Nat Rev Genet* 2007;8:286–98.
2 Stratton MR, Campbell PJ, Futreal PA. The cancer genome. *Nature* 2009;458:719–24.
3 Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. *Cell* 1990;61:759–67.
4 De Martel C, Ferlay J, Franceschi S, *et al*. Global burden of cancers attributable to infections in 2008: a review and synthetic analysis. *Lancet Oncol* 2012;13:607–15.
5 Tjalsma H, Boleij A, Marchesi JR, *et al*. A bacterial driver-passenger model for colorectal cancer: beyond the usual suspects. *Nat Rev Microbiol* 2012;10:575–82.
6 Ferlay J, Shin H-R, Bray F, *et al*. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer* 2010;127:2893–917.
7 Dove WF, Clipson L, Gould KA, *et al*. Intestinal neoplasia in the ApcMin mouse: independence from the microbial and natural killer (beige locus) status. *Cancer Res* 1997;57:812–14.
8 Uronis JM, Mühlbauer M, Herfarth HH, *et al*. Modulation of the intestinal microbiota alters colitis-associated colorectal cancer susceptibility. *PLoS ONE* 2009;4:e6026.
9 Lax AJ. Opinion: bacterial toxins and cancer—a case to answer? *Nat Rev Microbiol* 2005;3:343–9.
10 Goodwin AC, Destefano Shields CE, Wu S, *et al*. Polyamine catabolism contributes to enterotoxigenic Bacteroides fragilis-induced colon tumorigenesis. *Proc Natl Acad Sci USA* 2011;108:15354–9.
11 Wang X, Huycke MM. Extracellular superoxide production by Enterococcus faecalis promotes chromosomal instability in mammalian cells. *Gastroenterology* 2007;132:551–61.
12 Wu S, Rhee K-J, Zhang M, *et al*. Bacteroides fragilis toxin stimulates intestinal epithelial cell shedding and gamma-secretase-dependent E-cadherin cleavage. *J Cell Sci* 2007;120:1944–52.
13 Wu S, Rhee K-J, Albesiano E, *et al*. A human colonic commensal promotes colon tumorigenesis via activation of T helper type 17 T cell responses. *Nat Med* 2009;15:1016–22.
14 Abdulamir AS, Hafidh RR, Bakar FA. Molecular detection, quantification, and isolation of Streptococcus gallolyticus bacteria colonizing colorectal tumors:

inflammation-driven potential of carcinogenesis via IL-1, COX-2, and IL-8. *Mol Cancer* 2010;9:249.
15 Swidsinski A, Khilkin M, Kerjaschki D, *et al*. Association between intraepithelial Escherichia coli and colorectal cancer. *Gastroenterology* 1998;115:281–6.
16 Martin HM, Campbell BJ, Hart CA, *et al*. Enhanced Escherichia coli adherence and invasion in Crohn's disease and colon cancer. *Gastroenterology* 2004;127:80–93.
17 Maddocks ODK, Short AJ, Donnenberg MS, *et al*. Attaching and effacing Escherichia coli downregulate DNA mismatch repair protein *in vitro* and are associated with colorectal adenocarcinomas in humans. *PLoS ONE* 2009;4:e5517.
18 Nougayrède J-P, Taieb F, De Rycke J, *et al*. Cyclomodulins: bacterial effectors that modulate the eukaryotic cell cycle. *Trends Microbiol* 2005;13:103–10.
19 Nougayrède J-P, Homburg S, Taieb F, *et al*. Escherichia coli induces DNA double-strand breaks in eukaryotic cells. *Science* 2006;313:848–51.
20 Cuevas-Ramos G, Petit CR, Marcq I, *et al*. Escherichia coli induces DNA damage *in vivo* and triggers genomic instability in mammalian cells. *Proc Natl Acad Sci USA* 2010;107:11537–42.
21 Arthur JC, Perez-Chanona E, Mühlbauer M, *et al*. Intestinal inflammation targets cancer-inducing activity of the microbiota. *Science* 2012;338:120–3.
22 Buc E, Dubois D, Sauvanet P, *et al*. High prevalence of mucosa-associated *E. coli* producing cyclomodulin and genotoxin in colon cancer. *PLoS ONE* 2013;8:e56964.
23 Tanaka T, Kohno H, Suzuki R, *et al*. A novel inflammation-related mouse colon carcinogenesis model induced by azoxymethane and dextran sodium sulfate. *Cancer Sci* 2003;94:965–73.
24 Dreux N, Denizot J, Martinez-Medina M, *et al*. Point mutations in FimH adhesin of Crohn's disease-associated adherent-invasive Escherichia coli enhance intestinal inflammatory response. *PLoS Pathog* 2013;9:e1003141.
25 Debacq-Chainiaux F, Erusalimsky JD, Campisi J, *et al*. Protocols to detect senescence-associated beta-galactosidase (SA-betagal) activity, a biomarker of senescent cells in culture and *in vivo*. *Nat Protoc* 2009;4:1798–806.
26 Rodier F, Campisi J. Four faces of cellular senescence. *J Cell Biol* 2011;192:547–56.
27 Ben-Porath I, Weinberg RA. The signals and pathways activating cellular senescence. *Int J Biochem Cell Biol* 2005;37:961–76.
28 Kuilman T, Peeper DS. Senescence-messaging secretome: SMS-ing cellular stress. *Nat Rev Cancer* 2009;9:81–94.
29 Bischof O, Dejean A. SUMO is growing senescent. *Cell Cycle* 2007;6:677–81.
30 Yates KE, Korbel GA, Shtutman M, *et al*. Repression of the SUMO-specific protease Senp1 induces p53-dependent premature senescence in normal human fibroblasts. *Aging Cell* 2008;7:609–21.
31 Witty J, Aguilar-Martinez E, Sharrocks AD. SENP1 participates in the dynamic regulation of Elk-1 SUMOylation. *Biochem J* 2010;428:247–54.
32 Bonifacio LN, Jarstfer MB. MiRNA profile associated with replicative senescence, extended cell culture, and ectopic telomerase expression in human foreskin fibroblasts. *PLoS ONE* 2010;5:pii: e12519.
33 O'Donnell KA, Wentzel EA, Zeller KI, *et al*. c-Myc-regulated microRNAs modulate E2F1 expression. *Nature* 2005;435:839–43.
34 Guerra L, Albihn A, Tronnersjö S, *et al*. Myc is required for activation of the ATM-dependent checkpoints in response to DNA damage. *PLoS ONE* 2010;5:e8924.
35 Muehlbauer M, Gharaibeh R, Arthur JC, *et al*. Intestinal inflammation targets *E. coli* NC101 transcriptome response and promotes development of colorectal cancer (CRC). *Gastroenterology* 2013;144:S144–5.
36 Campbell BJ, Finnie IA, Hounsell EF, *et al*. Direct demonstration of increased expression of Thomsen-Friedenreich (TF) antigen in colonic adenocarcinoma and ulcerative colitis mucin and its concealment in normal mucin. *J Clin Invest* 1995;95:571–6.
37 Matsuo K, Ota H, Akamatsu T, *et al*. Histochemistry of the surface mucous gel layer of the human colon. *Gut* 1997;40:782–9.
38 Boland CR, Deshmukh GD. The carbohydrate composition of mucin in colonic cancer. *Gastroenterology* 1990;98:1170–7.
39 Greten FR, Eckmann L, Greten TF, *et al*. IKKbeta links inflammation and tumorigenesis in a mouse model of colitis-associated cancer. *Cell* 2004;118:285–96.
40 Krtolica A, Parrinello S, Lockett S, *et al*. Senescent fibroblasts promote epithelial cell growth and tumorigenesis: a link between cancer and aging. *Proc Natl Acad Sci USA* 2001;98:12072–7.
41 Stein U, Walther W, Arlt F, *et al*. MACC1, a newly identified key regulator of HGF-MET signaling, predicts colon cancer metastasis. *Nat Med* 2009;15:59–67.
42 Matsui S, Osada S, Tomita H, *et al*. Clinical significance of aggressive hepatectomy for colorectal liver metastasis, evaluated from the HGF/c-Met pathway. *Int J Oncol* 2010;37:289–97.
43 Sliwinska MA, Mosieniak G, Wolanin K, *et al*. Induction of senescence with doxorubicin leads to increase genomic instability of HCT166 cells. *Mech Ageing Dev* 2009;130:24–32.
44 Mikolajczyk J, Drag M, Békés M, *et al*. Small ubiquitin-related modifier (SUMO)-specific proteases: profiling the specificities and activities of human SENPs. *J Biol Chem* 2007;282:26217–24.

# Annexe n°3

Analysis of structure-function relationships in the colibactin-maturating enzyme ClbP.

Cougnoux A, **Gibold L**, Robin F, Dubois D, Pradel N, Darfeuille-Michaud A, Dalmasso G, Delmas J, Bonnet R.

J Mol Biol. 2012 Dec 7;424(3-4):203-14. doi: 10.1016/j.jmb.2012.09.017. Epub 2012 Oct 2

jmb Article

# Analysis of Structure–Function Relationships in the Colibactin-Maturating Enzyme ClbP

Antony Cougnoux[1], Lucie Gibold[1,2], Frederic Robin[1,2], Damien Dubois[3,4], Nathalie Pradel[5], Arlette Darfeuille-Michaud[1], Guillaume Dalmasso[1], Julien Delmas[1,2] and Richard Bonnet[1,2]

*1 - **M2iSH,** INSERM U1071, INRA USC2018, Clermont Université, Université d'Auvergne, 63000 Clermont-Ferrand, France*
*2 - **Centre Hospitalier Universitaire,** 63000 Clermont-Ferrand, France*
*3 - **Université de Toulouse,** ENVT, UMR 1225, Toulouse F-31076, France*
*4 - **INSERM,** UMR 1043, Toulouse F-31300, France*
*5 - **IRD,** UMR 180, IFR-BAIM, Universités de la Méditerranée, 13288 Marseille, France*

***Correspondence to Richard Bonnet:** Faculte de Medecine, Laboratoire de Bacteriologie, 28 Place Henri Dunant, 63000 Clermont-Ferrand, France. richard.bonnet@u-clermont1.fr*
http://dx.doi.org/10.1016/j.jmb.2012.09.017
***Edited by R. Huber***

## Abstract

*pks* genomic island of *Escherichia coli* is involved in the synthesis of the non-ribosomal peptide-type genotoxin colibactin, which has been suggesting as affecting the host immune response and having an impact on cancer development. The *pks*-encoded enzyme ClbP is an atypical peptidase that contributes to the synthesis of colibactin. In this work, we identified key features of ClbP. Bacterial fractionation and Western-blot analysis revealed the docking of ClbP to the bacterial inner membrane *via* a C-terminal domain harboring three predicted transmembrane helices. Whereas only one helix was necessary for the location in the inner membrane, the complete sequence of the C-terminal domain was necessary for ClbP bioactivity. In addition, the N-terminal sequence of ClbP allowed the SRP/Sec/YidC- and MreB-dependent translocation of the enzymatic domain in the periplasmic compartment, a feature also essential for ClbP bioactivity. Finally, the comparison of ClbP structure with that of the paralogs FmtA-like and AmpC revealed at an extremity of the catalytic groove a negative electrostatic potential surface characteristic of ClbP. Site-directed mutagenesis experiments identified in this zone two aspartic residues that were important for ClbP bioactivity. Overall, these results suggest a model for precolibactin activation by ClbP and pave a way for the design of inhibitors targeting colibactin production.

© 2012 Elsevier Ltd. All rights reserved.

## Introduction

Nougayrède *et al.* recently reported a genomic island designated *pks* that is involved in the production of a genotoxin named colibactin.[1] The *pks* island is observed in both *Escherichia coli*, a versatile commensal bacterial inhabitant of human large intestine, and other *Enterobacteriaceae*.[2] *pks*-harboring bacteria (*pks*+) induce DNA double-strand breaks in eukaryotic cells both *in vitro* and *in vivo*.[3] These lesions activate the ATM-CHK2 signaling pathway. However, the cells can present signs of incomplete DNA repair, which are marked by megalocytosis, G2/M cell cycle arrest, gene mutations, chromosomal instability and anchorage-independent colony formation, demonstrating the carcinogenic potential of colibactin.[1,3]

Colibactin has never been purified despite several attempts, and its precise structure remains unknown. However, genetic and functional analyses of the *pks* island indicate that it encodes non-ribosomal peptide synthetases (NRPSs) and polyketide synthetases (PKSs), which are required for the production of one or more polyketide (PK)–non-ribosomal peptide (NRP) hybrid compounds.[4] In addition to NRPS and PKS, the *pks* island encodes nine accessory, tailoring and editing enzymes, among which ClbP is crucial in colibactin maturation.[5] We recently determined the crystal structure of the ClbP major domain and identified

functional homologous proteins sharing about 30% identity with ClbP.[5] These proteins, such as ZmaM and XcnG, which are involved in the synthesis of the PK–NRP compounds zwittermicin and xenocoumacin, respectively,[6,7] belong to the S12 enzyme family of the MEROPS classification,[8] share a similar putative topology and are able to transcomplement *pks* island defective for *clbP*. They form a new subgroup of enzymes associated with NRP assembly line and are required for the maturation and activation of NRP compounds.[5,7] These enzymes, designated NRP-associated peptidases, could be an interesting target for molecular probes or molecules affecting the production of NRP bioactive molecules.

The protein FmtA-like, a protein of unknown function,[9] and AmpC, a β-lactamase,[10,11] also belong to the S12 enzyme family of the MEROPS classification and share 30% identity with ClbP. FmtA-like, in addition, exhibits a predicted topology similar to that of NRP-associated peptidases. However, these proteins are not able to transcomplement *pks* island defective for ClbP, indicating that they are paralogs of ClbP. By analogy with the well-known enzyme AmpC, Dubois *et al.* identified the catalytic triad required for ClbP activity.[5] However, the structural elements, which explain the difference in activity among FmtA-like, AmpC and ClbP, remain unknown. In addition, the role of the N- and C-terminal sequences of ClbP has never been investigated.

In the present study, we investigated the role of ClbP C- and N-terminal sequences in the location and bioactivity of ClbP. By comparison with AmpC and FmtA-like crystal structures, we also identified key determinants of ClbP activity far from the catalytic site, suggesting an atypically large binding site.

## Results

### ClbP is an inner membrane protein

ClbP (504 amino acids) was predicted *in silico* to be an inner membrane protein harboring three C-terminal transmembrane helices (residues 390–412, 433–455 and 465–485). A C-terminal RGS6His-tagged ClbP protein was expressed in *E. coli*, and Western-blot analysis was undertaken from whole cell lysates and cellular fractions to establish the location of ClbP in bacteria. A reactive band with a molecular mass compatible with RGS6His-tagged ClbP mass was observed in the whole cell lysate of *E. coli* (Fig. 1a). In addition, RGS6His-tagged ClbP was predominantly found in the inner membrane fraction and the spheroplast fraction (a cell fraction containing the inner membrane), but not in the periplasmic or cytoplasmic fractions (Fig. 1a). These results show that ClbP is located in the bacterial inner membrane.

HA-tagged ClbP harboring three, two, one or no C-terminal helices were expressed in *E. coli*. Anti-HA Western-blot analysis was performed on cellular fractionation experiments. It revealed that only one helix was required to associate ClbP with the inner membrane (Fig. 1b). To investigate the role of helices in ClbP activity, we expressed constructs in *E. coli* containing a *pks* island defective for *clbP* (pBAC*pks*Δ*clbP*) and we infected eukaryotic cells with the bacteria. ClbP derivatives harboring one, two or no helices did not induce megalocytosis and cell cycle arrest in response to the infection (Fig. 1c and d). In contrast, expression of the construct harboring the three putative helices induced megalocytosis and cell cycle arrest, which show restoration of a full functional *pks* island (Fig. 1c and d). In conclusion, the location of ClbP at the inner membrane is necessary but not sufficient to assume colibactin production and the complete ClbP C-terminal domain is required for ClbP bioactivity and colibactin production.

### ClbP possesses an N-terminal cleavable signal sequence, which is SRP/Sec/YidC and MreB dependent

The N-terminal residues of ClbP were predicted *in silico* to be a transmembrane helix (residues 12–30) or a cleavable Sec-dependent signal sequence (ALA-QE 30-31). This putative signal sequence of ClbP (ClbPss) plus the first six amino acids of the mature part were fused with the GFP (*g*reen *f*luorescent *p*rotein) reporter to form the fusion protein ClbPss-GFP. Western-blot analysis using a GFP-specific antibody was performed against *E. coli* cells expressing ClbPss-GFP. It revealed the precursor (ClbPss-GFP) and mature (GFP) forms of the construct (Fig. 1e). Cellular fractionation of *E. coli* expressing ClbPss-GFP revealed only little ClbPss-GFP in cytoplasmic fraction and an accumulation of both the complete construct and its processed form in periplasmic fraction (Fig. 1d). In addition, we analyzed the bioactivity of ClbP harboring its N-terminal sequence or not as described above. The results showed that the N-terminal sequence of ClbP was required to transcomplement *E. coli* harboring a *pks* island defective for *clbP* (Fig. 1c). All together, these results confirm the presence of a cleavable signal sequence in ClbP N-terminal sequence, which mediate a translocation to the periplasm that was required for ClbP bioactivity.

### ClbP signal sequence is SRP/Sec/YidC and MreB dependent

SRP, SecY and YidC are major components of the translocation/insertion pathway used by inner membrane proteins. This translocation system recognizes in the protein target an N-terminal signal

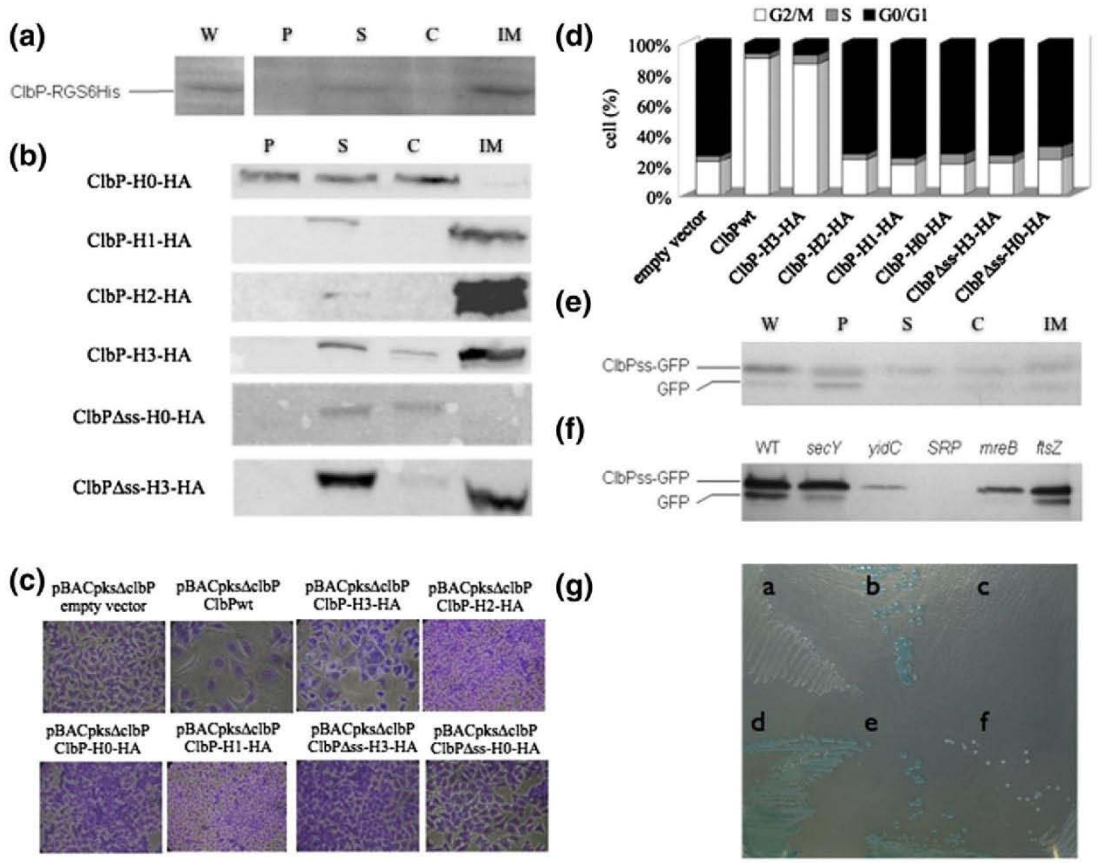



**Fig. 1.** The role of C- and N-terminal sequences in the location and bioactivity of ClbP. Each construction used was represented in Fig. S1. (a) Western-blot analysis of RGS6His-tagged ClbP location in *E. coli* cellular fractions. (b) Western-blot analysis of HA-tagged ClbP derivatives harboring one, two or three transmembrane helices in *E. coli* cellular fractions. (c and d) The genotoxicity of *E. coli* DH10B pBAC*pks*Δ*clbP* expressing ClbP derivatives. The genotoxicity was assessed by megalocytosis assays (c) and cell cycle analysis, which is reported as percentage of cells in phase G0/G1, S and G2/M (d). (e) Cellular location of ClbPss-GFP and its processed mature counterpart GFP in *E. coli* cellular fractions. (f) Cellular location of ClbPss-GFP and its processed mature counterpart GFP in wild-type *E. coli* (WT) and mutants for *secY*, *yidC*, *mreB*, *SRP* and *ftsZ*. (g) Alkaline phosphate (PhoA) export monitored on Mueller Hinton agar containing the chromogenic substrate BCIP. (g-a) *E. coli* defective for PhoA; (g-b) *E. coli* defective for PhoA transcomplemented by a plasmid encoding PhoA; (g-c) no bacteria; (g-d, g-e and g-f) *E. coli* defective for PhoA transcomplemented by a plasmid encoding PhoA fused to ClbP-signal peptide (ClbPss-PhoA), (g-d) ClbP-signal peptide and the three transmembrane helices of ClbP (ClbPss-PhoA-ClbP3TM) (g-e) or the three transmembrane helices of ClbP (PhoA-ClbP3TM) (g-f). Blue staining occurs on *E. coli* expressing PhoA or PhoA-fusion proteins in the bacterial periplasm. W, whole cells; P, periplasm; S, spheroplast; C, cytoplasm; SM, solubilized inner membrane fraction. ClbP-H0-HA, ClbP derivative without transmembrane helix; ClbP-H1-HA, ClbP derivative harboring one transmembrane helix; ClbP-H2-HA, ClbP derivative harboring two helices; ClbP-H3-HA, ClbP harboring three helices; ClbPΔss-H0-HA, ClbP derivative devoid of both signal sequence and transmembrane helix; ClbPΔss-H3-HA, ClbP derivative devoid of signal sequence; ClbPss-GFP, ClbP signal sequence fused to GFP.

sequence, which was cleaved at the time of translocation into periplasm and/or insertion to the inner membrane.[12] MreB and FtsZ are proteins of bacterial cytoskeleton, and it has been suggested that MreB is involved in cytoplasmic trafficking of proteins toward the Sec translocation system.[13] To further analyze the translocation system involved in ClbP maturation, we analyzed the cleavage of ClbPss-GFP fusion protein by Western blot using *SRP*, *secY*, *yidC*, *mreB* and *ftsZ E. coli* knockout or conditional isogenic mutants. The processing of ClbPss-GFP was impaired in *secY*, *yidC* and *mreB* mutants in comparison with the wild-type strain, but not in *ftsZ* mutant (Fig. 1e). No signal was observed with the *SRP* mutant, suggesting an absence of translocation and a degradation of the nascent peptide in the cytoplasm. These data suggest that actin-like cytoskeleton MreB and the SRP/Sec/YidC

pathway perform the translocation of ClbP to the inner membrane by targeting ClbP signal sequence.

**The N-terminal domain of ClbP is periplasmic**

ClbP is associated with the inner membrane because of its C-terminal sequence and harbors a cleavable N-terminal signal sequence allowing the translocation to periplasm, and hence, we decided to investigate whether the central enzymatic domain associated with ClbP C-terminal sequence is located at the periplasmic side of the bacterial inner membrane. Alkaline phosphatase enzymatic domain of PhoA, which is only active in periplasm, was used as a reporter of the periplasmic location as previously reported.[14,15] The PhoA domain was therefore fused to ClbP signal sequences, ClbP C-terminal sequence (380–504) or to both, and the constructs were expressed in *E. coli* DH5αF′. The bacteria expressing PhoA fused to ClbPss and fused to ClbPss plus ClbP C-terminal sequence harbored in the presence of PhoA chromogene substrate a blue staining, which revealed the presence of a PhoA active domain. This activity indicated a location in the periplasmic compartment (Fig. 1g-d and g-e). In contrast, *E. coli* DH5αF′ expressing PhoA fused to ClbP C-terminal sequence alone remained white (Fig. 1g-f), although the construct was expressed in the bacteria, as assessed by Western-blot analysis (data not shown). The absence of signal sequence prevents the location of enzymatic domain at the periplasmic side (Fig. 1g-f) without affecting the inner membrane association (Fig. 1b). Overall, these data suggest that an enzymatic domain docked to the inner membrane but facing inward prevents ClbP bioactivity (Fig. 1c and d). In addition, these results show that the N-terminal cleavable signal sequence of ClbP is sufficient and necessary to translocate the partner enzymatic domain into periplasm and confirmed therefore the location of the ClbP enzymatic domain on the periplasmic side of the bacterial inner membrane.

*FmtA-like has peptidase activity*

FmtA-like is the most closely related paralog of ClbP.[1,5] We therefore speculated that comparison of these enzymes could help in understanding the specificity of ClbP activity. We first compared their enzymatic activities against peptidase substrates formed by amino acids conjugated to *p*-nitroanilines. The catalytic efficacy $k_{cat}/K_m$ of FmtA-like against Gly and L-Ala-*p*-nitroanilines was 154.4 and 78.9 $\mu M^{-1} min^{-1}$, respectively, which is more than 10,000-fold higher than that of ClbP (0.005 and 0.020 $\mu M^{-1} min^{-1}$). However, FmtA-like had no detectable activity against D-Ala-D-Ala- and D-Ala-*p*-nitroanilines ($k_{cat} < 10^{-4} min^{-1}$). These results suggest that FmtA-like is not a DD-peptidase and confirm the differences in enzymatic activity between ClbP and FmtA-like.

*Comparison of FmtA-like, AmpC and ClbP structures*

FmtA-like is the most closely related paralog of ClbP. However, its structure is unknown. We therefore determined the crystal structure of FmtA-like enzymatic domain (33–369) by X-ray diffraction to compare its structure with those of ClbP and AmpC, another paralog of ClbP.[10,11] The crystal structure was solved at 3.2 Å resolution by molecular replacement using AmpC and ClbP as search models. The stereochemical parameters of the model are reported in Table S3. The Ramachandran plot[16] showed 97.6% in favored region and 2.4% in allowed region of ϕ/ψ space. The final model included 4 FmtA-like enzymatic domain monomers and 35 water molecules.

The overall structure of FmtA-like enzymatic domain is composed of two regions: one α/β region and one all-helical region (Fig. S2). The organization of the corresponding secondary structure elements was conserved in ClbP and FmtA-like except two additional small helices (H5 and H10) in ClbP, but differed from that of AmpC (Fig. S3). The α/β domain consisted of a main 10-stranded antiparallel β-sheet flanked by seven helices, and the all-helical region consisted of four helices. Between these two structural domains was a groove, which spreads over the β3 strand and contains the catalytic triad, as observed for ClbP, AmpC and other S12 enzymes. The first motif of the catalytic triad (consensus: S-X-X-K) is strictly conserved in S12 MEROPS enzymes and consists of serine residue, acting as catalytic nucleophile during the acylation step of the catalytic reaction, plus a lysine residue one helix turn downstream. The second motif (consensus: Y-X-N/S) consists of Y190, which is usually responsible for the deacetylation step, and N192, which is frequently involved in substrate binding in S12 enzymes. The third motif, which is less conserved (consensus: K/R/H-T/S/G-G) and usually also involved in substrate binding, consists of the HSG sequence in FmtA-like, KTG in AmpC and HGG in ClbP. The positioning of these catalytic motifs was strikingly conserved in the active sites of FmtA-like and ClbP [PDB (*P*rotein *D*ata *B*ank) ID: 3O3V] and AmpC (PDB ID: 1FSY) (Fig. S4), suggesting a common catalytic mechanism of acylation–deacylation.

The structural elements surrounding the catalytic motifs are known to determine the specificity of S12 MEROPS enzymes.[17] We therefore investigated the positioning of these elements in ClbP and in the two paralogs FmtA-like and AmpC. The major differences observed between these enzymes were the positioning and the size of the loop, which connects helix H9 and the β5 β-strand of FmtA-like. In ClbP

and FmtA-like, this loop was composed of 21 and 19 residues, respectively. The corresponding structural element comprised a loop (279–310) and the helices H13 (280–287) and H14 (289–292) in AmpC. In ClbP structure, it is positioned toward the solvent and, consequently, allows a large extension of the catalytic groove toward the N-terminal extremity of the β3 strand. In AmpC and, at lesser extent, in FmtA-like, the loop shifted toward the catalytic groove in close contact with the helices H16 and H11 in AmpC and FmtA-like, respectively, and covered the N-terminal extremity of the β3 strand (Fig. 2). These elements interacted in FmtA-like *via* one close contact between the residue 307 of the loop and the residue 363 of helix H11 (distance between the $O^{\delta 1}$ atom of Asn307 and the $O^{\varepsilon 2}$ atom of Glu363, 4.4 Å; Fig. 2a), despite a 2.3- to 2.7-Å shift of helix H11 toward the solvent compared to AmpC and ClbP. Likewise, 26 close contacts (distances comprised between 3.2 and 4.5 Å; Fig. 2b) were observed in AmpC in this area, since the distances between the corresponding residues are higher than 10 Å in ClbP (Fig. 2c). Consequently, the long catalytic groove observed in ClbP is severely shortened in FmtA-like and particularly in AmpC at the N-terminal extremity of the β3 strand.

*The surface electrostatic potentials of ClbP, FmtA-like and AmpC*

Electrostatic fields around the surface of proteins are known to play an important role in molecular recognition, owing to the attractive or repulsive effect of the coulombic potential at long-range distance.[18] We therefore calculated the electrostatic potential molecular surface of ClbP, FmtA-like and AmpC enzymes (Fig. 3a–c) using APBS software.[19] The results revealed very different electrostatic potentials in their catalytic grooves. On the side harboring the catalytic groove, we observed in ClbP a strong negative electrostatic potential surface restricted to the catalytic groove, especially at the two extremities (Fig. 3a). A diffuse negative electrostatic potential was observed in FmtA-like with strong charged environment at the C-terminal extremity of the β3 strand, as in ClbP, but a weak or neutral charged environment at the N-terminal extremity of the β3 strand (Fig. 3c). In contrast, the electrostatic potential surface of AmpC catalytic pocket was mainly positive (Fig. 3b) as observed with other closely related S12 enzymes such as carboxypeptidases or transpeptidases (data not shown). Overall, ClbP differed therefore from AmpC and FmtA-like

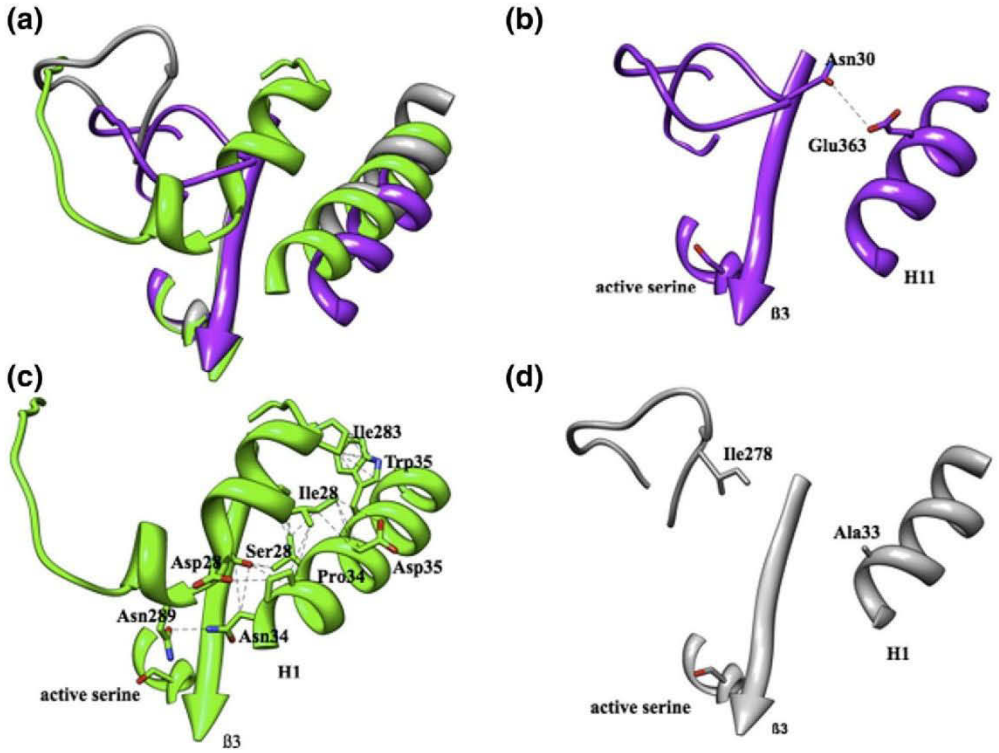



**Fig. 2.** Superimposed structures of ClbP, FmtA-like and AmpC (a) and close contact analysis in FmtA-like (b), AmpC (c) and ClbP (d) and structures at the vicinity of β3 strand. FmtA-like is purple, AmpC is green and ClbP is gray. Close contacts between residues are indicated by gray broken lines. Carbon atoms are purple for FmtA-like, green for AmpC and gray for ClbP; oxygen atoms are red and nitrogen is blue.

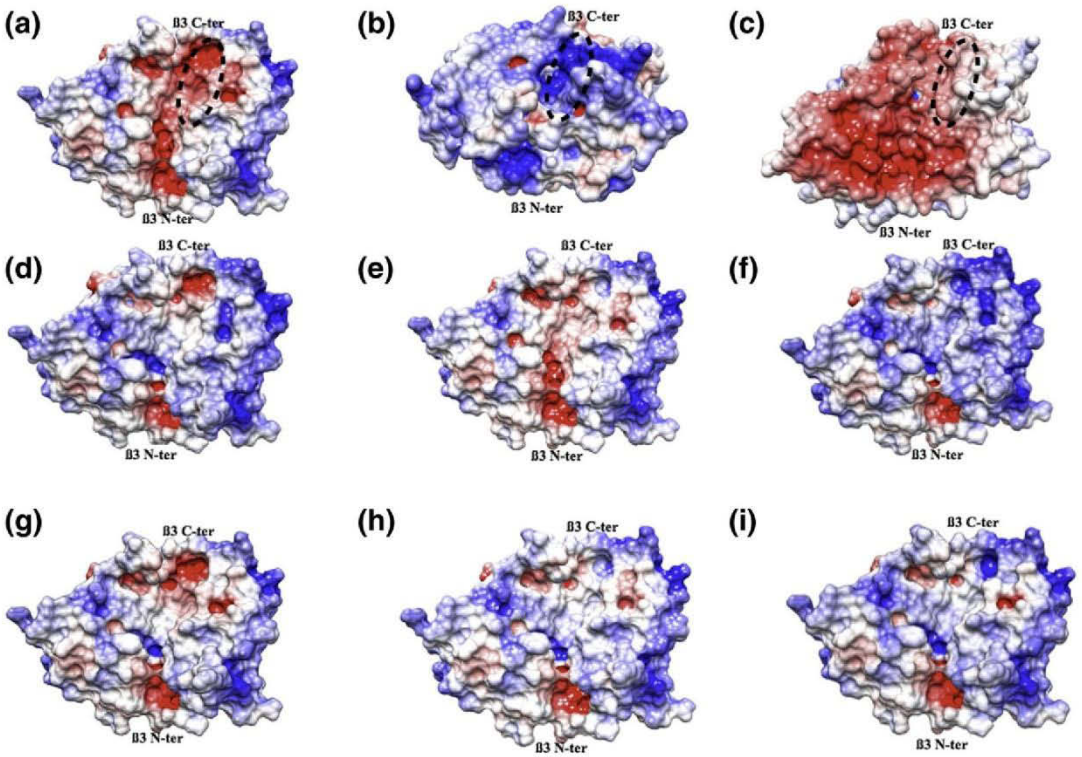



**Fig. 3.** Molecular surfaces strained according to the electrostatic potential for ClbP (a), AmpC (b), FmtA-like (c) and ClbP mutants harboring the substitutions D334Q (d), D336L (e), D334Q-D336L (f), D334E (g), D336N (h) and D3343-D336N (i). The surfaces were colored red for negative charge, white for neutral and blue for positive. Helix H13 from ClbP and the corresponding helices in AmpC (H16) and FmtA-like (H11) were circled in black.

by a strong negative electrostatic potential at the N-terminal extremity of the β3 strand in an area edged by the C-terminal part of helix H11.

The distribution of charges on protein surface was compared to that of known substrates or putative ligands. Benzylpenicillin and cephalothin, which are the substrates of AmpC, exhibited an overall negative electrostatic potential (Fig. S5). This can favor their binding into the positive-charged catalytic pocket of AmpC. Interestingly, the final bioactive PK–NRP compound zwittermicin and its putative preprocessing counterpart pre-zwittermicin, which Kevany *et al.* suggest is cleaved by the NRP-associated peptidase ZmaM,[6] exhibited overall a positive electrostatic potential surface (Fig. S5). The positive potential surface of NRP compounds can therefore match with the negative potential surface of ClbP, which may be a key feature of this enzyme.

The residues, which may be responsible for the negative electrostatic potential surface in ClbP catalytic groove, were investigated using molecular modeling approaches. The negative-charged residues were replaced *in silico* by the corresponding residues in AmpC or FmtA-like, and the electrostatic potential was calculated to evaluate their contribution to the electrostatic properties of ClbP. We found at the C-terminal part of helix H11 two positions harboring negative-charged residues, D334 and D336. The substitutions D334Q and D336L, which introduce residues observed in AmpC, abrogated and decreased the negative electrostatic potential (Fig. 3d–f), respectively. In combination, this substitution conferred a positive electrostatic potential, as observed in AmpC. The substitutions D334E and D336N, which introduce residues observed in FmtA-like, slightly decreased and abolished the negative electrostatic potential observed at the N-terminal extremity of the β3 strand (Fig. 3g–i). Overall, these results suggest that residues D334 and D336 may be key elements in ClbP bioactivity.

### *The positions 334 and 336 of ClbP are key residues for ClbP bioactivity*

To evaluate the contribution of residues 334 and 336 to the bioactivity of ClbP, we introduced substitutions D334E (FmtA-like residue), D334Q (AmpC residue), D336N (FmtA-like residue) and D336L (AmpC residue) alone and in combination into RGS6His-tagged ClbP by site-directed mutagenesis.

These constructs and wide-type RGS6His-tagged ClbP were expressed in *E. coli* containing a *pks* island defective for ClbP. The constructs were expressed at a similar level, as assessed by Western-blot analysis (data not shown). The genotoxicity of bacteria was then investigated using comet assays, which show at early-time postinfection (24 h) the DNA damage induced by colibactin-producing *E. coli*. Megalocytosis assays and cell cycle analysis were also performed to reveal the presence of patent lesions at 72 h. RGS6His-tagged ClbP mutants at position 334 or 336 induced a significant decrease in mean comet tail moment in comparison with wide-type RGS6His-tagged ClbP (RGS6His-tagged ClbP *versus* RGS6His-tagged ClbP-D334E/Q or RGS6His-tagged ClbP D336L/N, $p<0.05$; Fig. 4a). This decrease revealed an alteration in the genotoxic activity of bacteria expressing ClbP mutants. However, the genotoxicity was not enough impaired to affect the cell cycle arrest and megalocytosis induced by the bacteria at 72 h (Fig. 4b and c). In contrast, the expression of the mutants harboring the combinations of substitutions at positions 334 and 336 strongly impaired the genotoxicity of bacteria; DNA damage was almost abrogated at 24 h (95% and 92.8%, $p<0.01$; Fig. 4a), cell cycle arrest was significantly affected ($p<0.001$; Fig. 4b) and megalocytosis decreased at 72 h (Fig. 4c). All together, these results demonstrate that positions 334 and 336, which are located far from the catalytic triad (distance between $C^{\alpha}$ atoms of active serine Ser64 and residues 334/336, 19.5/19.0 Å, respectively), are still important for ClbP bioactivity.

## Discussion

The bacterial genomic island *pks* is involved in the production of enzymes, which synthesize a PK–NRP-type genotoxin named colibactin.[1] This type of toxin can have a major impact on host–microbe relationships.[20,21] It has been suggested that colibactin affects the host immune response and has an impact on colorectal cancer development.[1,3] The *pks*-encoded enzyme ClbP has been identified as a prototype of an atypical peptidase subgroup of the S12 MEROPS family. These enzymes, designated NRP-associated peptidases, contribute to the maturation of NRP compounds such as colibactin.[5] In this work, we identified key features of ClbP bioactivity: (i) the association of ClbP with the bacterial inner membrane, (ii) the exposition of its enzymatic domain in the bacterial periplasm and (iii) key residues of the enzymatic domain involved in the negative electrostatic potential of the catalytic groove. These findings enlighten our understanding of colibactin synthesis and pave the way for the design of inhibitors targeting colibactin production.

We show that the N-terminal sequence is a cleavable signal sequence mediating a translocation

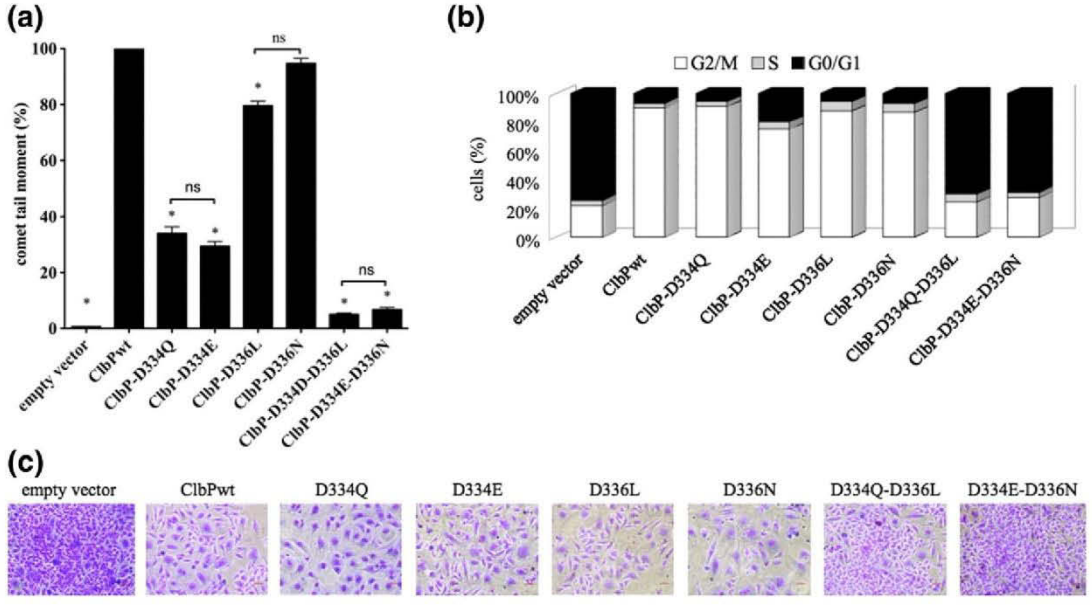



**Fig. 4.** The genotoxicity of *E. coli* pBAC*pks*Δ*clbP* expressing ClbP or ClbP mutants at position 334 and/or 336 was investigated with HeLa cells. (a) DNA damage at 24 h postinfection assessed by comet assays. The results are expressed in percentage of comet tail moment (100% was attributed to bacteria expressing wide-type ClbP). (b) Cell cycle analysis at 72 h postinfection reported as percentage of cells in phase G0/G1, S and G2/M. (c) Cytopathic effect (megalocytosis) at 72 h postinfection.

of the passenger domain to periplasm *via* the actin-like cytoskeleton MreB and the SRP/Sec/YidC pathway. A similar putative signal sequence is observed in other NRP-associated peptidases,[6,7] suggesting that the enzymatic domain of NRP-associated peptidases such as ClbP is located outside the cytoplasm. In contrast, the other colibactin-producing enzymes, essentially NRPS and PKS, are devoid of putative signal sequence and are known to be in the cytoplasm. ClbP acts therefore outside the cytoplasm, after assembly line chains of NRP biosynthesis, and thus is involved in the final steps of colibactin maturation. It allows maturation of the genotoxin colibactin outside the cytoplasm at a distance from the DNA of the producing bacteria probably in order to protect its genome. This location also makes ClbP and the other NRP-associated peptidases more accessible than the other enzymes of NRP assembly lines and therefore renders them key targets for the control of the bioactivity of PK–NRP-producing gene clusters.

We observed that ClbP is not free in the periplasm but associated with the inner membrane because of the C-terminal domain. Other S12 enzymes of MEROPS classification such as transpeptidases are associated with the inner membrane with their enzymatic domain in the periplasm.[22,23] They are docked to the membrane *via* a single N-terminal transmembrane helix,[12] while the docking of ClbP to the membrane is dependent on the C-terminal domain, which contains three predicted transmembrane helices. Only one helix was sufficient to assure the docking of ClbP to the membrane. However, the complete C-terminal domain was required for the bioactivity of ClbP. The presence of three transmembrane helices is compatible with a periplasmic enzymatic domain in the periplasm and a C-terminal extremity in the cytoplasm. Interestingly, several NRP-associated peptidases are fused with ABC exporters belonging to the family of cyclic peptides.[5–7] The role of ClbP C-terminal domain may also be important therefore for interactions with the cytoplasmic assembly line and/or efflux pumps carrying out the export of precolibactin from the cytoplasm to the periplasm.

We analyzed the phylogeny of the condensation domain in the NRPS enzymes encoded by *pks* island as previously reported[24] and found typical starter C-type domain in ClbN. This type of domain acylates the first amino acid of NRP compounds with typically β-hydroxyl fatty acids to form a lipopeptide.[25] Reimer *et al.* recently reported the maturation of the lipopeptides pre-xenocoumacin in an NRP compound devoid of fatty acid (xenocoumacin) by the NRP-associated peptidase XcnG.[7] Kevany *et al.* previously proposed a similar biosynthesis pathway for zwittermicin A.[6] Accordingly, precolibactin may therefore be a lipopeptide inserted in the bacterial inner membrane and actively released from this membrane by the inner-membrane-located enzyme ClbP. This model (Fig. 5) can explain the importance of ClbP location at the inner membrane and why *pks*-harboring bacteria killed by heat or gentamicin are not genotoxic.[1]

The enzymatic domain of ClbP was compared to those of the two paralogs FmtA-like and AmpC. AmpC is a class C β-lactamase, which is well known to confer a resistance to β-lactam antibiotics and to be widespread in Gram-negative bacteria. FmtA-like is encoded by the genome of *Staphylococcus* isolates. We found that this protein harbors a peptidase activity different from that of ClbP, but its biological function remains unknown. Phylogenetic analysis has shown that FmtA-like is the most closely related paralog of ClbP, while AmpC is more distantly related.[5] The determination of FmtA-like crystal structure shows that this phylogenetic lineage is also apparent in the three-dimensional structure of these enzymes and helps to identify interesting characteristic of the ClbP catalytic groove. From AmpC to ClbP, we observed a progressive opening of the catalytic groove at the vicinity of the N- and C-terminal extremities of the β3 strand and helix H11 (or their structural equivalent H16 and H13), respectively (Fig. 2). This opening is associated with an inversion of the electrostatic potential at the surface of the catalytic groove. Whereas the AmpC electrostatic potential surface of AmpC active site is overall positive, we observed a strong negative electrostatic potential focused in the catalytic groove of ClbP. These negative charges may be conserved around the NRP-associated peptidases and fit the positive charges of NRP compounds produced by assembly lines related to the *pks*-encoded assembly line.

In FmtA-like, the negative electrostatic potential is more diffuse than in ClbP and the charge is less marked especially in the vicinity of the N- and C-terminal extremities of the β3 strand and helix

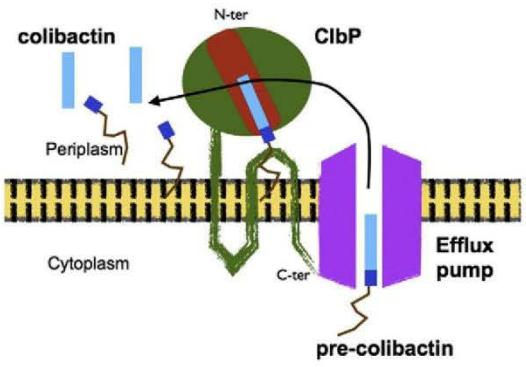



**Fig. 5.** ClbP role in the maturation of colibactin. Negatively charge active site of ClbP was colored red. The putative positively charged colibactin compound was colored blue.

H11. These data suggest that this latter zone is crucial for the specificity of ClbP bioactivity. Accordingly, we found two acidic residues important for ClbP bioactivity in this zone in helix H13 (positions 334 and 336). Their substitutions with the corresponding residues of FmtA-like and AmpC affected the bioactivity of ClbP in a synergistic manner and, consequently, eliminated the genotoxicity mediated by *pks*-harboring *E. coli*. This inhibition of ClbP bioactivity is probably due to a modification of the electrostatic potential, but we cannot exclude that ClbP activity is impaired by steric hindrance. Several residues have been mutated in the vicinity of the catalytic triad and the C-terminal part of the β3 strand without affecting ClbP bioactivity,[5] suggesting that the non-catalytical residues in this zone are less important for ClbP bioactivity than the extremity of the catalytic groove around the N-terminal part of the β3 strand. These results also reveal that the ClbP catalytic groove is atypically long and is designed to accommodate a filiform compound.

In conclusion, the NRP-associated peptidase ClbP is a periplasmic enzyme that is docked to the bacterial inner membrane by its C-terminal domain. This topology allows maturation of the genotoxin colibactin outside the cytoplasm at a distance from the genomic material of the producing bacteria. It also makes ClbP an interesting drug target for control of the genotoxic effect of colibactin-producing bacteria. We also identified at a distance from the catalytic triad an electronegative area as a key determinant of ClbP bioactivity that may be targeted by drugs.

## Materials and Methods

### Strains, plasmids and media

Strains and plasmids used in this study are listed in Table S1, and the primers are in Table S2. *E. coli* strains and *Staphylococcus aureus* mu50 were grown at 37 °C in Mueller Hinton medium or on Luria–Bertani (LB) plates. Amoxicillin (32 μg/ml) (Sigma), kanamycin (50 μg/ml) (Sigma), chloramphenicol (25 μg/ml), glucose (0.2%), arabinose (0.2%), lactose (0.2%), isopropyl β-D-1-thiogalactopyranosid (1 mM) (Euromedex) or tetracyclin (200 ng/ml) (Sigma) was added as required to the medium.

### Cell induction assays of translocation pathway and fractionation

Cell induction assays of translocation pathway were performed as previously described.[26–28] We used the cold-sensitive *secY-cs* and *yidC* conditional mutants CU164A and MYC-cs, in which protein export is partially slowed down at 37 °C and is severely affected upon exposure to 23 °C. The following strains harboring p61ClbPssGFP were incubated in LB–amoxicillin–glucose (0.2%) medium at permissive temperature (37 °C for cold-sensitive strains) with shaking until early exponential phase: MC4100A, MC1060 and MC1000 wild type; conditional translocation mutant CU164A (Δ*secY*) and MYC-cs (Δ*yidC*); HPT406A (SRP mutant), YLS-3 (Δ*mreB*) and YLS-24 (Δ*ftsZ*) strains. The cultures were incubated on ice for 3 min and then kept at 23 °C (nonpermissive temperature) for 1 h for conditional mutant induction. The cells were harvested by centrifugation, washed once with LB and resuspended in 0.2% LB–amoxicillin–arabinose medium and incubated for a further 3 h at 23 °C with shaking for induction of ClbPss-GFP expression.

Cellular fractionation experiments were performed as previously described.[26,27] Crude extracts were prepared by centrifugation of 1 ml culture at $OD_{600\,nm}$=0.5. The pellets were suspended in 0.1 ml Tris–HCl buffer (40 mM, pH 8.0) and 0.02 ml of 1% protamine sulfate. Periplasm and spheroplasts were prepared by lysozyme/ethylenediaminetetraacetic acid/cold osmoshock method and separated by centrifugation. Membranes and cytoplasmic soluble fractions were prepared from spheroplasts by breaking the cells by sonication and separated by ultracentrifugation at 120,000*g* for 90 min. Protease inhibitor Complete (Roche Diagnostics, Mannheim, Germany) was added at the maximum recommended concentration in all buffers used.

### Western-blot analysis

Western blots were performed with polyclonal rabbit anti-GFP (Invitrogen) at a dilution of 1:10,000 and monoclonal anti-RGS6His according to the manufacturer's instructions (Qiagen). Western blots performed with monoclonal rabbit anti-HA (Sigma) at a dilution of 1:1000, mice anti-alkaline phosphatase (Millipore) at a dilution of 1:5000 and horseradish-peroxidase-tagged goat anti-rabbit or goat anti-mouse secondary antibodies (Cell Signaling) at a dilution of 1:10,000 according to the manufacturer's instructions.

### Monitoring of alkaline phosphatase activity

ClbP fused to alkaline phosphatase were obtained by recombination using primers listed in Table S2. Cloning experiments were performed with In-Fusion HD cloning kit (Clontech) using the primers described in Table S2. *E. coli* cells were grown on LB plates with amoxicillin (32 μg/ml) and 50 μg/ml of 5-bromo-4-chloro-3-indolyl phosphate (the chromogenic substrate of alkaline phosphatase purchased from Sigma) at 30 °C for 2 days. The blue-stained colony-forming units reveal the presence of the PhoA alkaline phosphatase domain in the periplasm as previously reported.[15]

### Cloning and site-directed mutagenesis

*fmtA*-like gene was amplified from *S. aureus* strain mu50. After a double digestion with NdeI and NotI (New England Biolabs), PCR products were cloned in pET-M13 (Bioscience) expression vectors. ClbP mutants were obtained by site-directed mutagenesis using primers listed in Table S2. Site-directed mutagenesis experiments were performed with QuikChange™ Site-Directed Mutagenesis Kit II (Stratagene, La Jolla, CA) using the

mutagenic primers described in Table S2. The substitutions D334E, D336N, D334E-D336N, D334Q, D336L and D334Q-D336L were individually introduced in ClbP cloned into p33ClbP plasmid. The changes were confirmed by double-stranded DNA sequencing. The final p33clbP derivative vectors were used to transform *E. coli* DH10B pBAC*pks*Δ*clbP* strain, which is defective for ClbP production.

### Eukaryotic cell culture, bacterial infections, comet assay and cell cycle analysis

Cell assays were performed as previously described, with minor modifications.[1,5] For bacterial infections, overnight LB–amoxicillin cultures of bacteria were diluted in interaction medium [Dulbecco's modified Eagle's medium, 10% decomplemented fetal calf serum (Lonza), 25 mM Hepes (Sigma) and 32 μg/ml amoxicillin]. Tetracyclin was added to overnight cultures and interacting medium to induce the expression of ClbP and derivatives. Approximately 50% confluent HeLa cell cultures (American Type Culture Collection CCL2) were then infected with a multiplicity of infection (number of bacteria per eukaryotic cell) of 500. Cells were washed twice in phosphate-buffered saline (pH 7.4) 3 h after inoculation and incubated in Dulbecco's modified Eagle's medium, 10% fetal calf serum and 200 μg/ml gentamicin (Sigma) to stop the infection and were kept until analysis. Cell morphology was examined by Giemsa staining at 72 h postinfection. For cell cycle analysis, cellular suspensions were fixed and permeabilized with 70% ethanol at 48 h postinfection, then treated with 50 μg/ml propidium iodide and 250 μg/ml RNase A (Roche). DNA-content data were acquired with an SLR II or FACScalibur flow cytometer (Becton Dickinson) and analyzed with FlowJo software (Tree Star). DNA breaks were analyzed by comet assay using SYBR Green as fluorescent DNA staining agent as described previously.[29,30] Comets were scored with the image analysis software CometScore™ (TriTek Corporation) using the mean comet tail moment (product of tail length and fraction of DNA in the tail).

### FmtA-like production and purification

pFmtA-like plasmid encodes the peptidase domain of FmtA-like without the first 87 nucleotides and the last 384 nucleotides of the *fmtA*-like gene. These sequences encode the signal peptide and three putative hydrophobic transmembrane helices that were removed to improve cytoplasmic expression and protein purification. *E. coli* BL21(DE3)Δ*ampC* transformed with pFmtA-like plasmid was cultured in Terrific broth medium supplemented with 50 μg/ml kanamycin, 0.4 M D-sorbitol (MP) and 2.5 mM β-betaine (Sigma). The cell culture was grown at 37 °C with shaking up to reach an $OD_{620nm}$=0.7. Overexpression of the *fmtA*-like encoding gene was induced with 0.2 mM isopropyl β-D-thiogalactopyranoside overnight at 25 °C with shaking. Cells were harvested by centrifugation (10,000***g*** for 10 min at 4 °C) and suspended in 20 mM Tris–HCl buffer (pH 7.5). After addition of 5 μg lysozyme (2 h at 4 °C), the cells were disrupted by ultrasonic treatment (four times for 30 s, each time at 20 W). The extract was clarified by centrifugation at 14,000***g*** for 60 min at 4 °C. After addition of 2 mg DNase I (Roche) for 15 min at room temperature, the supernatant (10,000***g*** for 60 min at 4 °C) was dialyzed overnight against 20 mM Tris–HCl buffer (pH 7.5) and filtered (0.22-μm filter) (Millipore). Purification was carried out by ion-exchange chromatography on a 10-ml HiTrap™ Q Sepharose™ high-performance column (GE Healthcare) equilibrated with 20 mM Tris–HCl (pH 7.5) and eluted with a linear NaCl gradient (0–1 M). The FmtA-like-containing elution peak (160 mM NaCl) was then dialyzed 4 h against 50 mM of 4-morpholineethanesulfonic acid buffer (pH 6.0) and purified by a second ion-exchange chromatography on a HiTrap™ Q Sepharose™ high-performance column equilibrated with 50 mM of 4-morpholineethanesulfonic acid buffer (pH 6.0) and eluted with a linear NaCl gradient (0–1 M). The FmtA-like-containing elution peak (50 mM NaCl) was extensively dialyzed overnight against 5 mM Tris–HCl and 50 mM NaCl buffer (pH 7) and concentrated by ultrafiltration to 10 mg/ml for crystallization. The concentration was carried out using YM-10000 centricon (Millipore). The enzyme was more than 95% homogeneous as determined by Coomassie blue staining after 10% SDS-PAGE analysis.

### Enzymatic activity

Standard photometric assays using *p*-nitroaniline-conjugated amino acids X-pNA (X: Gly, Gly-Gly, L-Ala, D-Ala, D-Ala-D-Ala, Bz-Asn or L-Lys) as substrates were performed at 25 °C in 5 mM Tris–HCl and 50 mM NaCl, pH 7.0, containing 4 mM of the substrate. Data were collected on the UVICON930 system. Absorption coefficients at 405 nm of 9450 $M^{-1}\,cm^{-1}$ for X-pNA were used to calculate the enzymatic activity. $K_m$ and $k_{cat}$ were determined from kinetic data using ENCORA software (Kluyver laboratory).

### Crystallization and structure determination

*fmtA*-like crystals were grown in hanging drops over a well solution of 0.2 M ammonium acetate (Sigma), 0.1 M sodium acetate trihydrate (pH 4.6) (Sigma) and 30% (w/v) polyethylene glycol 4000 (Sigma). Crystals reached a size of 40 μm×40 μm×70 μm within 4 months. Before data collection, crystals were immersed for about 30 s in a cryoprotectant solution [30% sucrose, 0.2 M ammonium acetate, 0.1 M sodium acetate trihydrate (pH 4.6) and 30% (w/v) polyethylene glycol 4000] and were flashcooled in liquid nitrogen. Data were collected using a Q315r ADSC-CCD detector on European Synchrotron Radiation Facility (ESRF) beamline ID-23-1 at the ESRF (Grenoble, France). Reflections were indexed, integrated and scaled using CCP4 package.[31–33] The initial model was obtained by molecular replacement with the program Phaser,[34] and the ClbP X-ray structures (PDB ID: 3O3V) were used as search models. The structure was automatically and manually refined with the REFMAC5 and Coot programs, respectively.[35,36] Cross-validation was used throughout, and 5% of the data were used for the $R_{free}$ calculation. The stereochemical quality of the models was monitored with the PROCHECK program.[37] Ramachandran plots were calculated by RAMPAGE.[16] Processing and crystallographic refinement statistics for FmtA-like crystal structure are listed in Table S3 (PDB ID: 4GDN).

### Bioinformatic analysis and molecular modeling

ClbP primary sequence (accession number CAJ76284.1) was analyzed using InterProScan.[38] Signal peptide analysis was performed using SignalP 4.0 server[39] and lipoP 1.0 server.[40] Transmembrane helix location was predicted using TMHMM server† and TMBASE.[41]

Substitutions at position 334 and/or 336 were introduced in ClbP structure using Chimera version 1.5.3‡ visualization software and Dunbrack rotamer library.[42] The models were then protonated with PROPKA using the PDB2PQR server,[43] and the protons were subjected to energy minimization (100 steps steepest descent) using the MMTK routine under Chimera[44] with AMBER parameters. The quality and stereochemistry of the models were evaluated using the program PROCHECK program.[37] The APBS program[19] was used to calculate the electrostatic surface potential map of protein models, which were visualized with the Chimera program.

## Acknowledgements

We thank Christelle Blavignac, Claire Szczepaniak and Lorraine Gameiro from the Centre Imagerie Cellulaire Santé (Université d'Auvergne) for their technical assistance in photonic microscopy and in flow cytometry. We thank Marlène Jan and Roland Perroux for their excellent technical assistance. This work was supported by the Ministère de la Recherche et de la Technologie, l'Institut National de la Santé et de la Recherche Médicale (INSERM U1071) and l'Institut National de la Recherche Agronomique (USC INRA 2018). These experiments were performed on the ID-23-1 beamline at the ESRF, Grenoble, France. We are grateful to Local Contact at ESRF for providing assistance in using beamline ID-23-1.

## Supplementary Data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jmb.2012.09.017

*Received 17 August 2012;*
*Received in revised form 20 September 2012;*
*Accepted 21 September 2012*
Available online 2 October 2012

***Keywords:***
serine peptidases;
colibactin;
non-ribosomal peptide;
electrostatic surface charge;
*Escherichia coli*

†software@cbs.dtu.dk
‡http://plato.cgl.ucsf.edu/chimera/

***Abbreviations used:***
ESRF, European Synchrotron Radiation Facility; PK, polyketide; PKS, polyketide synthetase; NRP, non-ribosomal peptide; NRPS, non-ribosomal peptide synthetase.

## References

1. Nougayrède, J.-P., Homburg, S., Taieb, F., Boury, M., Brzuszkiewicz, E., Gottschalk, G. *et al.* (2006). *Escherichia coli* induces DNA double-strand breaks in eukaryotic cells. *Science*, **313**, 848–851.
2. Putze, J., Hennequin, C., Nougayrède, J.-P., Zhang, W., Homburg, S., Karch, H. *et al.* (2009). Genetic structure and distribution of the colibactin genomic island among members of the family *Enterobacteriaceae*. *Infect. Immun.* **77**, 4696–4703.
3. Cuevas-Ramos, G., Petit, C. R., 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. USA*, **107**, 11537–11542.
4. 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.
5. Dubois, D., Baron, O., Cougnoux, A., Delmas, J., Pradel, N., Boury, M. *et al.* (2011). ClbP is the prototype of a peptidase subgroup involved in the biosynthesis of nonribosomal peptides. *J. Biol. Chem.* http://dx.doi.org/10.1074/jbc.M111.221960.
6. Kevany, B. M., Rasko, D. A. & Thomas, M. G. (2009). Characterization of the complete zwittermicin A biosynthesis gene cluster from *Bacillus cereus*. *Appl. Environ. Microbiol.* **75**, 1144–1155.
7. Reimer, D., Pos, K. M., Thines, M., Grün, P. & Bode, H. B. (2011). A natural prodrug activation mechanism in nonribosomal peptide synthesis. *Nat. Chem. Biol.* **7**, 888–890.
8. Rawlings, N. D., Morton, F. R., Kok, C. Y., Kong, J. & Barrett, A. J. (2008). MEROPS: the peptidase database. *Nucleic Acids Res.* **36**, D320–D325.
9. Komatsuzawa, H., Choi, G. H., Fujiwara, T., Huang, Y., Ohta, K., Sugai, M. & Suginaka, H. (2000). Identification of a fmtA-like gene that has similarity to other PBPs and β-lactamases in *Staphylococcus aureus*. *FEMS Microbiol. Lett.* **188**, 35–39.
10. Usher, K. C., Blaszczak, L. C., Weston, G. S., Shoichet, B. K. & Remington, S. J. (1998). Three-dimensional structure of AmpC β-lactamase from *Escherichia coli* bound to a transition-state analogue: possible implications for the oxyanion hypothesis and for inhibitor design. *Biochemistry*, **37**, 16082–16092.
11. Jacoby, G. A. (2009). AmpC β-lactamases. *Clin. Microbiol. Rev.* **22**, 161–182; Table of Contents.
12. Izard, J. W. & Kendall, D. A. (1994). Signal peptides: exquisitely designed transport promoters. *Mol. Microbiol.* **13**, 765–773.
13. Shih, Y.-L., Kawagishi, I. & Rothfield, L. (2005). The MreB and Min cytoskeletal-like systems play independent roles in prokaryotic polar differentiation. *Mol. Microbiol.* **58**, 917–928.

14. Marrichi, M., Camacho, L., Russell, D. G. & DeLisa, M. P. (2008). Genetic toggling of alkaline phosphatase folding reveals signal peptides for all major modes of transport across the inner membrane of bacteria. *J. Biol. Chem.* **283**, 35223–35235.
15. Drew, D., Sjöstrand, D., Nilsson, J., Urbig, T., Chin, C., de Gier, J.-W. & von Heijne, G. (2002). Rapid topology mapping of *Escherichia coli* inner-membrane proteins by prediction and PhoA/GFP fusion analysis. *Proc. Natl Acad. Sci. USA*, **99**, 2690–2695.
16. Laskar, A., Chatterjee, S., Roy, A., Kumar Dey, S. & Mandal, C. (2011). Molecular modeling and structural analysis of five SE clan (S12 family) serine proteases. *Asian J. Biotechnol.* **3**, 435–448.
17. Sharp, K. A. & Honig, B. (1990). Electrostatic interactions in macromolecules: theory and applications. *Annu. Rev. Biophys. Biophys. Chem.* **19**, 301–332.
18. Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. (2001). Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl Acad. Sci. USA*, **98**, 10037–10041.
19. Demangel, C., Stinear, T. P. & Cole, S. T. (2009). Buruli ulcer: reductive evolution enhances pathogenicity of *Mycobacterium ulcerans*. *Nat. Rev., Microbiol.* **7**, 50–60.
20. Fischbach, M. A. & Walsh, C. T. (2006). Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: logic, machinery, and mechanisms. *Chem. Rev.* **106**, 3468–3496.
21. Delfosse, V., Girard, E., Birck, C., Delmarcelle, M., Delarue, M., Poch, O. *et al.* (2009). Structure of the archaeal pab87 peptidase reveals a novel self-compartmentalizing protease family. *PLoS One*, **4**, e4712.
22. Antalis, T. M., Buzza, M. S., Hodge, K. M., Hooper, J. D. & Netzel-Arnett, S. (2010). The cutting edge: membrane-anchored serine protease activities in the pericellular microenvironment. *Biochem. J.* **428**, 325–346.
23. Rausch, C., Hoof, I., Weber, T., Wohlleben, W. & Huson, D. H. (2007). Phylogenetic analysis of condensation domains in NRPS sheds light on their functional evolution. *BMC Evol. Biol.* **7**, 78.
24. Konz, D., Doekel, S. & Marahiel, M. A. (1999). Molecular and biochemical characterization of the protein template controlling biosynthesis of the lipopeptide lichenysin. *J. Bacteriol.* **181**, 133–140.
25. Pradel, N., Delmas, J., Wu, L. F., Santini, C. L. & Bonnet, R. (2009). Sec- and Tat-dependent translocation of β-lactamases across the *Escherichia coli* inner membrane. *Antimicrob. Agents Chemother.* **53**, 242–248.
26. Santini, C. L., Ize, B., Chanal, A., Müller, M., Giordano, G. & Wu, L. F. (1998). A novel Sec-independent periplasmic protein translocation pathway in *Escherichia coli*. *EMBO J.* **17**, 101–112.
27. Chen, M., Xie, K., Jiang, F., Yi, L. & Dalbey, R. E. (2002). YidC, a newly defined evolutionarily conserved protein, mediates membrane protein assembly in bacteria. *Biol. Chem.* **383**, 1565–1572.
28. Olive, P. L. & Banáth, J. P. (2006). The comet assay: a method to measure DNA damage in individual cells. *Nat. Protoc.* **1**, 23–29.
29. Collins, A. R. (2004). The comet assay for DNA damage and repair: principles, applications, and limitations. *Mol. Biotechnol.* **26**, 249–261.
30. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R. *et al.* (2011). Overview of the CCP4 suite and current developments. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **67**, 235–242.
31. Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **67**, 271–281.
32. Evans, P. (2006). Scaling and assessment of data quality. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **62**, 72–82.
33. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674.
34. Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A. *et al.* (2011). REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **67**, 355–367.
35. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **66**, 486–501.
36. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291.
37. Lovell, S. C., Davis, I. W., Arendall, W. B., 3rd, de Bakker, P. I. W., Word, J. M., Prisant, M. G. *et al.* (2003). Structure validation by Calpha geometry: phi, psi and Cbeta deviation. *Proteins*, **50**, 437–450.
38. Syed, A. & Upton, C. (2006). Java GUI for InterProScan (JIPS): a tool to help process multiple InterProScans and perform ortholog analysis. *BMC Bioinformatics*, **7**, 462.
39. Petersen, T. N., Brunak, S., von Heijne, G. & Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods*, **8**, 785–786.
40. Rahman, O., Cummings, S. P., Harrington, D. J. & Sutcliffe, I. C. (2008). Methods for the bioinformatic identification of bacterial lipoproteins encoded in the genomes of Gram-positive bacteria. *World J. Microbiol. Biotechnol.* **24**, 2377–2382.
41. Hofmann, K. & Stoffel, W. (1993). TMBASE—a database of membrane spanning protein segments. *Biol. Chem. Hoppe-Seyler*, **374**, 166.
42. Dunbrack, R. L., Jr. (2002). Rotamer libraries in the 21st century. *Curr. Opin. Struct. Biol.* **12**, 431–440.
43. Dolinsky, T. J., Czodrowski, P., Li, H., Nielsen, J. E., Jensen, J. H., Klebe, G. & Baker, N. A. (2007). PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucleic Acids Res.* **35**, W522–W525.
44. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004). UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612.

## Annexe n°4

Covalent docking of large libraries for the discovery of chemical probes.

London N, Miller RM, Krishnan S, Uchida K, Irwin JJ, Eidam O, **Gibold L**, Cimermančič P, Bonnet R, Shoichet BK, Taunton J.

Nat Chem Biol. 2014 Dec;10(12):1066-72. doi: 10.1038/nchembio.1666. Epub 2014 Oct 26.

Erratum in: Nat Chem Biol. 2015 Mar;11(3):235.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH Public Access
Author Manuscript
*Nat Chem Biol*. Author manuscript; available in PMC 2015 June 01.

Published in final edited form as:
*Nat Chem Biol*. 2014 December ; 10(12): 1066–1072. doi:10.1038/nchembio.1666.

# Covalent Docking of Large Libraries for the Discovery of Chemical Probes

**Nir London**[1,*], **Rand M. Miller**[2,*], **Shyam Krishnan**[3,*], **Kenji Uchida**[3], **John J. Irwin**[4], **Oliv Eidam**[1], **Lucie Gibold**[5,6,7], **Peter Cimermančič**[3], **Richard Bonnet**[5,6,7], **Brian K. Shoichet**[1,4,§], and **Jack Taunton**[3,§]

[1]Department of Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, CA 94158, USA

[2]Chemistry and Chemical Biology Graduate Program, University of California, San Francisco, San Francisco, CA 94158, USA

[3]Department of Cellular and Molecular Pharmacology, Howard Hughes Medical Institute, University of California, San Francisco, San Francisco, CA 94158, USA

[4]Faculty of Pharmacy & Ontario Institute for Cancer Research, University of Toronto, Toronto, Canada

[5]Clermont Université, UMR 1071 Inserm/Université d'Auvergne, 63000 Clermont-Ferrand, France

[6]INRA, USC 2018, 63000 Clermont-Ferrand, France

[7]Service de Bactériologie, Centre Hospitalier Universitaire, 63000 Clermont-Ferrand, France

## Abstract

Chemical probes that form a covalent bond with a protein target often show enhanced selectivity, potency, and utility for biological studies. Despite these advantages, protein-reactive compounds are usually avoided in high-throughput screening campaigns. Here we describe a general method (DOCKovalent) for screening large virtual libraries of electrophilic small molecules. We apply this method prospectively to discover reversible covalent fragments that target distinct protein nucleophiles, including the catalytic serine of AmpC β-lactamase and noncatalytic cysteines in RSK2, MSK1, and JAK3 kinases. We identify submicromolar to low-nanomolar hits with high ligand efficiency, cellular activity and selectivity, including the first reported reversible covalent

[§]Corresponding authors: bshoichet@gmail.com (for correspondence relating to the docking method and to covalent inhibition of β-lactamase); jack.taunton@ucsf.edu (for correspondence relating to covalent inhibition of kinases).
[*]Equal contribution

**AUTHOR CONTRIBUTIONS**

BKS and JT directed the project. NL designed the algorithms, wrote the covalent docking code, and executed the docking simulations. NL performed β-lactamase assays and crystallography with help from OE. JJI designed and implemented the DOCKovalent web tool. PC performed the proteasome experiments. LG and RB performed bacterial cell culture experiments. RMM executed synthetic chemistry, kinase assays, crystallography, and cell-based assays for RSK2/MSK1. SK and KU performed synthetic chemistry for JAK3, SK performed JAK1/3 kinase assays. NL, RMM, BKS, and JT wrote the paper. All authors contributed to the manuscript in its final form.

**Competing Financial Interests**

J.T., R.M.M. and S.K. have filed patent applications on cyanoacrylamide kinase inhibitors (licensed to Principia Biopharma, of which J.T. is a co-founder).

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript



inhibitors of JAK3. Crystal structures of inhibitor complexes with AmpC and RSK2 confirm the docking predictions and guide further optimization. As covalent virtual screening may have broad utility for the rapid discovery of chemical probes, we have made the method freely available through an automated web server (http://covalent.docking.org).

**Keywords**

covalent virtual screening; beta-lactamase; RSK2 kinase; MSK1 kinase; JAK3 kinase; boronic acid; cyanoacrylamide; covalent electrophile; covalent inhibition; covalent fragments; reversible covalent

## Introduction

Small-molecule modulators of protein activity are useful as tools for investigating biology and as leads for drug discovery. Development of genuinely useful probes typically involves iterative rounds of medicinal chemistry to optimize the potency and selectivity of initial active molecules. An effective strategy for enhancing both properties is via covalent bond formation with a nucleophilic residue that is specific to a target of interest and ideally absent from off-targets. Such covalent-acting chemical probes have increasingly been used in proteome-wide target identification [1], imaging [2], and for finding inhibitors with high specificity among related enzymes and enzyme isoforms [3,4]. Covalent drugs [5] and natural products [6] are also well known. A challenge in developing covalent probes is identifying reactive functional groups ("warheads") that do not make the molecule so reactive as to be promiscuous. Less recognized is the challenge of screening a wide variety of scaffolds for optimal presentation of such reactive functionality.

The most widely used technique for novel ligand discovery is high-throughput screening (HTS), and one could potentially screen extant libraries for new small molecules that react covalently. However, protein-reactive compounds are rarely screened [7] and are typically avoided in HTS [8] or flagged as artifacts due to concerns about promiscuous activity [9]. Whereas this is sensible for drug discovery, it removes potential starting points for covalent chemical probes [10].

Covalent ligands can target either catalytically essential nucleophiles such as those in serine and cysteine hydrolases, or noncatalytic nucleophiles, usually cysteine, found in small-molecule binding sites on proteins, including proteins without enzymatic activity (e.g., GPCRs and nuclear receptors) [11]. Depending on the electrophile and nucleophile, they can bind reversibly or irreversibly. In all cases, specific noncovalent interactions contributed by the scaffold are critical for orienting the electrophile relative to the protein nucleophile, thereby increasing the rate (and stability, in the case of reversible covalent ligands) and selectivity of covalent bond formation. A key unsolved problem in the discovery of covalent probes is how to identify a protein-binding scaffold that optimally orients the electrophile, while minimizing the number of compounds that must be synthesized and tested.

In principle, structure-based docking screens [12,13] can address the gap left by HTS and its libraries. Given the structure of a protein target, docking programs computationally screen

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript



large compound libraries for molecules predicted to bind favorably within a defined binding site. The technique has been widely used for the discovery of reversible, non-covalent ligands [14,15]. To date, there have been few docking screens for covalent ligands. Key obstacles include combining classical non-covalent scoring with covalent restraints and bond energies, and developing compound libraries suited to covalent modification of proteins. For the compound libraries, one would prefer not only commercially available electrophiles, but also the ability to design new, readily synthesizable molecules bearing a particular electrophile. Recently, there has been encouraging progress in developing covalent docking methods, but these have been restricted to retrospective recapitulation of covalent complexes [16–18], or to screens of a few hundred compounds [19,20]; we are unaware of prospective, large-scale covalent docking screens to find novel inhibitors, or of any such screens against targets for which covalent ligands are unprecedented.

Here, we adapt the non-covalent docking program DOCK3.6 to large-scale, covalent virtual screening of electrophilic small molecules, including low-molecular weight electrophilic fragments (Fig. 1). Nine libraries of ligands bearing different electrophiles, amounting to over 650,000 commercially available or synthetically accessible small molecules, are developed for use with the method. We used the method (DOCKovalent) to prospectively screen compound libraries against three targets of therapeutic interest: AmpC β-lactamase, RSK2 kinase and JAK3 kinase. Multiple potent, reversible covalent inhibitors were found against all three targets. X-ray crystal structures of predicted ligands, and the occasional false negative, illuminated not only the method's ability to prospectively identify ligands and to predict their structures, but also its limitations. Several of the new covalent ligands were tested in cell culture experiments that established biological efficacy and target engagement. To ensure that the method may be used by a broad community, it has been made available on an easy to use web server (http://covalent.docking.org).

## Results

### Overview of the method

We begin by constructing large virtual libraries of electrophiles, either commercially available or synthetically accessible in 1–2 steps. We created libraries of well-studied electrophiles including: α,β-unsaturated carbonyls, aldehydes, boronic acids, α-cyanoacrylamides, alkyl halides, carbamates, α-ketoamides, and epoxides (Supplementary Results, Supplementary Fig. 1, Supplementary Table 1); other electrophilic chemotypes, such as vinyl sulfones, sulfonyl fluorides, 2-chloropyridines, and cyanopyrimidines, are also imaginable. All stereoisomers, protonation states and conformations of the covalent adduct are pre-generated for each ligand, enabling rapid docking of the library to any target (Fig. 1).

For each ligand, DOCKovalent exhaustively samples all poses and ligand conformations with respect to the covalent bond to the target nucleophile, constrained by ideal bond lengths and angles (Fig. 1 and Supplementary Fig. 2). The nucleophile is immobile during the sampling, and a separate screen should be run for each likely rotamer of the nucleophile. Each sampled conformation is scored using the physics-based scoring function in DOCK3.6 [21], which evaluates the ligand's van der Waals and electrostatic interactions, and corrects for its desolvation. Using this scoring function, the entire library is ranked from

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript



most to least favorable. The top 1–3% of the ranked list is inspected for mis-docked ligands, which are common in large-scale docking screen, and molecules with incorrect ionization states, tautomers, or strained conformations are removed. The remaining molecules are prioritized for experimental testing based on their availability or synthetic accessibility, the presence of unprecedented chemotypes, and diversity of chemical structure.

### Retrospective assessment of covalent docking

We first tested method's ability to find known covalent ligands in five retrospective screens against four targets, and to recapitulate geometries for a previously published benchmark of covalent ligand complexes [16]. DOCKovalent performed well in pose recapitulation (Supplementary Table 2,3) and in four of the five retrospective screens (Supplementary Note 1 and Supplementary Table 4). We thus turned to prosecuting *prospective* screens for new reversible covalent ligands for three enzymes.

### New boronic acid inhibitors of AmpC β-lactamase

AmpC β-lactamase is the leading cause of resistance to cephalosporin antibiotics in clinical settings [22], and several new β-lactamase inhibitors are in clinical trials [23]. Boronic acids inhibit AmpC by forming a reversible covalent adduct with its active-site nucleophilic serine (Ser64). We first assessed the ability of our covalent docking method to recapitulate known boronic acid complexes with AmpC. In 15 of 23 cases, the ligand pose was accurately recovered to less than 2 Å RMSD (Supplementary Table 5 and Supplementary Fig. 3). Surprisingly, a relatively simple compound, *m*-aminophenylboronic acid (MAPB), failed this retrospective test. The boronic acid in the published MAPB/AmpC structure (PDB: 3BLS; Fig. 2a) adopted a pose that differed from other boronic acid complexes with the enzyme, and from our other docking poses. To investigate this discrepancy, we re-determined the crystal structure of AmpC in complex with MAPB. In this 1.65 Å resolution x-ray structure we observed unambiguous density for the ligand that corresponds to the docking prediction (0.46 Å RMSD, Fig. 2a).

Encouraged by these results, we used the method to covalently dock a library of 23,000 commercially available boronic acids against AmpC. Among the top-ranked 4.5% of the library, we sought boronic acids with scaffolds that had not been tested previously against AmpC. Five such compounds ranked between 11 and 646 of the docked library (top 0.04 to 2%) and were purchased and tested (compounds **1**–**5;** Fig. 2b). An additional, lower ranking compound (**6**) was purchased as a proxy for structurally related, high-ranking predictions that were commercially unavailable. Five of the six compounds inhibited AmpC with $K_i$ values ranging from 40 nM to 3.55 μM (Table 1), and three of these exhibited submicromolar potency (ligand efficiency, LE 0.38–0.66, Table 1; see docking predictions in Fig. 2c–d and Supplementary Fig. 4). None of the five inhibitors resemble a known boronic acid inhibitor of AmpC (Tanimoto coefficients < 0.3 to AmpC boronic acids in ChEMBL, using ECFP4-based fingerprints).

A 1.74 Å crystal structure of compound **3**, the most potent inhibitor from our initial set of six compounds ($K_i$ = 40 nM), confirmed the docking pose prediction (1.38 Å RMSD, Fig. 2c). The boronic acid occupies the oxyanion hole formed by the backbone amides of Ala318

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

compound **7** was the most potent at reversing antibiotic resistance, lowering the MIC for six strains to ≤2 μg mL, the empirical threshold for resistance defined by the Clinical and Laboratory Standards Institute [25]. None of the compounds had significant antibiotic activity in the absence of cefotaxime (Supplementary Table 8).

### New cyanoacrylamide inhibitors for RSK2 and MSK1 kinases

The C-terminal kinase domains (CTD) of p90 ribosomal protein S6 kinase-2 (RSK2) and the closely related paralog, mitogen- and stress-activated kinase-1 (MSK1), contain a noncatalytic active-site cysteine shared by only 11 of the 518 human protein kinases. Starting with an established kinase inhibitor scaffold, we previously designed irreversible [3] and reversible [26,27] covalent inhibitors that target this cysteine (Cys436 in RSK2). To achieve reversible covalent inhibition, we exploited the atypical reactivity of cyanoacrylamide Michael acceptors, which react rapidly and reversibly with cysteine thiols at physiological pH. Both RSK2 and MSK1 are attractive therapeutic targets implicated in tumor metastasis [28,29], neurodegeneration [30], and atherosclerosis [31], among other pathological conditions. We sought novel RSK2/MSK1 inhibitors by covalent docking screens of thousands of cyanoacrylamide fragments.

As an initial blind test, we used the method to predict the poses of two cyanoacrylamide fragments bound to RSK2, prior to determining the crystal structures. The predicted binding modes anticipated the experimental structures to 1.93 Å and 1.56 Å RMSD (Fig. 3a and 3b). Retrospective docking of two larger cyanoacrylamides also recapitulated their crystal structures (0.66 Å and 1.52 Å RMSD; Supplementary Fig. 8a–b). In each prediction, the scaffold portion of the molecule, which forms critical non-covalent interactions with RSK2, closely matched the x-ray structures (0.91–1.36 Å RMSD).

Encouraged by these results, we used covalent docking to screen for novel cyanoacrylamide inhibitors. Cyanoacrylamide fragments are rare in commercial collections (602 out of 474,770 of the "fragments in-stock" in ZINC [32]). However, β-substituted cyanoacrylamides can be synthesized in one step by condensing aldehydes with cyanoacetamide (Supplementary Fig. 9). We therefore assembled ~12,000 aldehyde fragments from ZINC [32] (molecular weight < 250 Da). These were converted *in silico* to generate a virtual library of cyanoacrylamide fragments.

We docked this library against Cys436 of RSK2. After manually inspecting the top-ranked compounds for novelty, diversity, and accessibility, we pursued eight virtual cyanoacrylamide fragments ranked between 96 and 391 (top 3%; Compounds **19**–**26;** Fig. 3c). The corresponding aldehydes were purchased and converted to the cyanoacrylamides, which were tested against wild-type RSK2 and the T493M gatekeeper mutant (Table 2). We have previously used this mutant as a biochemical surrogate for MSK1, as MSK1 CTD kinase activity has yet to be reconstituted *in vitro* [27]. Five of the eight high-ranking cyanoacrylamides inhibited RSK2 with $IC_{50}$ values less than 10 μM in the presence of competing 0.1 mM ATP and 10 mM reduced glutathione. Pyridine **21** and pyrazole **24** (Fig. 3d–e) were the most potent against wild-type RSK2 and inhibited the T493M mutant with submicromolar potency ($IC_{50}$ 430 nM and 370 nM, respectively).

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

We determined the co-crystal structure of compound **24** bound to T493M RSK2 at 3.0 Å resolution. Even at this modest resolution, the electron density allowed unambiguous modeling of the phenylpyrazole fragment and the covalent bond to Cys436 (Supplementary Fig. 10). The crystallographic structure superposed well with the docking prediction, with an RMSD of 0.86 Å over the phenylpyrazole fragment (Fig. 3d). The pyrazole forms two hydrogen bonds with the hinge region (N1 to the carbonyl of Glu494 and N2 to the amide NH of Met496). This ring also packs edgewise against the methionine gatekeeper, and the additional van der Waals contacts may explain the enhanced potency for the T493M mutant.

We tested **21** and **24** for activity in mammalian cells stimulated with phorbol myristate acetate (PMA), which activates kinase cascades upstream of RSK2 and MSK1. Both compounds inhibited the activating autophosphorylation of wild-type MSK1 ($EC_{50}$ < 5μM; Fig. 3f). While these compounds inhibit WT RSK2 less potently in cells, quantitation of the normalized pS380 signal reveals dose-dependent inhibition ($EC_{50}$ ~ 20 uM). A cysteine to valine substitution in the C-terminal kinase domain of MSK1 (C458V) conferred complete resistance to both inhibitors, consistent with on-target efficacy. Hence, the unoptimized cyanoacrylamide fragments identified by covalent docking inactivate the target kinases in cells without affecting upstream kinases such as RAF, MEK, and ERK.

The docking pose of **21** (Fig. 3e) suggested the possibility of improving its potency by fusing a pyrrole ring, an H-bond donor, to either a pyridine H-bond acceptor, as in **21**, or a pyrimidine ring. Exemplified by pyrrolopyrimidine **27** (Fig. 3g), this would bury more hydrophobic surface area and form an additional hydrogen bond with the hinge region, as suggested by covalent docking to RSK2 (Fig. 3h). Compound **27** was not included in the original screen, as the corresponding aldehyde is not commercially available. On synthesis and testing, compound **27** inhibited WT RSK2 kinase *in vitro* with an $IC_{50}$ of 42 nM, over 25-fold better than **21** (Fig. 3g). Correspondingly, **27** was substantially more potent than **21** in cells, blocking MSK1 autophosphorylation with an $EC_{50}$ < 1 μM (Fig. 3i).

### Selective, reversible covalent inhibitors of JAK3 kinase

Members of the Janus kinase family, comprised of JAK1, JAK2, JAK3, and TYK2, are essential for signaling downstream of many cytokine receptors [33]. JAK3 is expressed predominantly in immune cells and is a potential therapeutic target for autoimmune diseases like rheumatoid arthritis (RA) [34]. A pan-JAK inhibitor, tofacitinib [35], was recently approved for RA, but it suffers from adverse effects such as elevated liver enzymes and LDL cholesterol [36]. Selective JAK3 inhibitors may avoid such toxicities, and moreover, could help illuminate JAK3-specific roles in cytokine signaling. To date, development of selective JAK3 inhibitors has been hampered by the high sequence identity among JAK-family kinases [37]. JAK3 contains a solvent-exposed cysteine residue just outside the ATP binding site (Cys909), which is not found in JAK1, JAK2, or TYK2, and is present in only nine other human kinases. We used DOCKovalent in an effort to find the first reversible covalent inhibitors of JAK3, which might be expected to have specificity over closely related JAK kinases that lack Cys909.

The vector from Cys909 to the hinge differs greatly from the previously targeted Cys436 of RSK2. A preliminary screen of the virtual cyanoacrylamide fragment library developed

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

initially for RSK2 suggested that greater diversity and perhaps larger fragments would be required to engage both Cys909 and the hinge of JAK3. Inspired by the simple two-step synthesis of **27**, we designed a combinatorial virtual library based on two synthetic transformations: a Suzuki-Miyaura cross-coupling reaction between an aryl or heteroaryl bromide and an aldehyde-containing boronic acid, followed by a Knoevenagel condensation of the aldehyde with cyanoacetamide. We selected 50 commercially available boronic acids and 4,400 aryl bromides, which were converted to their corresponding products *in silico*. This approach afforded a diverse virtual library of 220,000 hetero-biaryl cyanoacrylamides (Supplementary Table 1, Supplementary Fig. 11), which was screened against eight JAK3 crystal structures.

We purchased eight arylbromides and three boronic acids common among the top 0.2% of the docked library (Supplementary Table 9). From these building blocks, we synthesized 15 inhibitors (Compounds **28**–**42**; Fig. 4a). Compound **42** was prepared from a commercially available aldehyde. Each compound was initially tested against JAK3 at two concentrations. Nine of the 15 compounds inhibited the kinase by > 50% at 5 μM, and five maintained > 50% inhibition at 1 μM (Fig. 4b).

We focused on the two most potent compounds **31** and **33** (for docking poses, see Supplementary Fig. 12). To assess their selectivity, we measured full dose-response curves against JAK3 and against the other three JAK-family kinases lacking Cys909. Cyanoacrylamides **31** and **33** inhibited JAK3 with $IC_{50}$ values of 49 nM and 93 nM, respectively, but neither inhibitor affected any other JAK kinase at concentrations up to 10 μM (Fig. 4c). This marked selectivity for JAK3 may be largely attributed to covalent bond formation with the target cysteine. By virtue of this covalent targeting of a noncatalytic cysteine, **31** and **33** are among the most selective JAK3 inhibitors reported to date [37]. We further assessed the selectivity of **31** against nine additional human kinases that have a cysteine at the equivalent position to JAK3. **31** potently inhibited three of the nine kinases ($IC_{50}$= 22 nM, 44 nM and 221 nM for BLK, ERB-B4 and ITK, respectively) but had at least 30-fold specificity for JAK3 over the remaining six kinases ($IC_{50}$ > 1 μM; Supplementary Fig. 13).

## Discussion

Covalent probes play a crucial role in chemical biology [1–4], yet electrophilic molecules that might serve as initial hits for developing such probes have largely been expunged from empirical screening libraries. Accordingly, we sought to enable large-scale screening for covalent probe molecules via structure-based docking. Although the method has limitations, its utility is supported by the discovery of new chemotypes in three prospective docking screens. For AmpC, the new inhibitors bear little topological resemblance with previously known inhibitors — with ECFP4-based Tanimoto coefficients < 0.3 to the most similar previous inhibitor — despite intense study of boronic acids against this enzyme over the last 30 years. For RSK2/MSK1, docking identified the active phenylpyridine **21**, even though isomeric phenylpyridines were previously found to be inactive [27]. The predicted pose of **21** supported the design of the unique pyrrolopyrimidine variant **27**. Finally, the 4-phenylindazole scaffold of the newly discovered JAK3 inhibitor **31** has little precedent

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

among kinase inhibitors; a related scaffold found in linifanib binds in a completely different orientation to VEGFR2 (PDB: 1YWN).

For all targets, the hit rates were high — five of six molecules predicted and tested for β-lactamase (Table 1), five of eight for RSK2 kinase (Table 2) and nine of fifteen for JAK3 kinase (Fig. 4). These hit rates did not reflect broad promiscuity on the part of the electrophiles. The boronic acids active against β-lactamase were typically inactive against related serine proteases (Supplementary Table 6) and against the proteasome (Supplementary Fig. 6), and most poorly ranked boronic acids were inactive against the enzyme (Supplementary Table 7). The one exception to this was shown by crystallography to reflect an unexpected and previously unobserved conformational change. Similarly, MSK1 inhibitors were selective in cells, being inactive against the Cys458→Val substitution of MSK1, as well as the upstream kinases, MEK and ERK, which also have noncatalytic cysteines in their active sites. Finally, the most potent JAK3 inhibitors were virtually inactive against other JAK kinases, underscoring the advantages of covalently targeting poorly conserved noncatalytic cysteines. Moreover, compound **31** showed selectivity over six of the nine human kinases bearing a homologous cysteine. Given its low molecular weight (288 Da) and simple architecture, accessible in two steps from inexpensive building blocks, the selectivity of **31** among these cysteine-containing kinases is remarkably high, comparing favorably with the approved covalent drug, ibrutinib [38], which targets the same cysteine in BTK.

Whereas most of the initial hits do not have the potency typically associated with optimized chemical probes, the best examples from the AmpC and RSK/MSK screens were sufficiently potent and selective to be active in cell-based assays. Hits from all screens were remarkably ligand efficient, with LE values as high as 0.66. Moreover, the initial hits could be improved, as shown in two cases. Boronic acid **7** has four-fold higher potency relative to the initial docking hit, **3**. Indeed, this inhibitor appears to have the highest ligand efficiency of any serine β-lactamase inhibitor [24]. As for RSK2/MSK1, compound **27** was improved 25-fold versus its parent screening hit, had sub-micromolar cellular activity, and is one of the most ligand-efficient RSK2 inhibitors known. The origins of the selectivity and affinities of these ligands are captured by their docking poses, which superpose well with the subsequently determined x-ray structures. The high fidelity of the docking to the x-ray structures suggests that the covalently docked poses alone may guide ligand optimization in cases where obtaining experimental structures is challenging, as demonstrated by **27**.

An innovation introduced in this study, one with few precedents in docking as a field, is the development and experimental testing of truly virtual, readily synthesizable, molecular libraries. The few commercially available cyanoacrylamides necessitated the design of larger and more diverse libraries, which is likely the case for other classes of reactive molecules as well.

One of our concerns at the outset was that, in mixing covalent and non-covalent scoring terms, the former would overwhelm the latter [39]. To avoid this, we did not explicitly score the covalent bond energy of the adducts, but rather imposed distance and angle constraints on bond formation and then scored the docked poses by our standard physics-based scoring

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

function, ignoring the covalent bond. An advantage is that the scoring function is not dominated by the covalent bond to the electrophile, and instead favors structural complementarity and specificity provided by noncovalent interactions with the scaffold. The approach is especially well suited to the reversible, and hence thermodynamically driven, inhibitors discovered here (our best compounds were explicitly shown to be rapidly reversible by dilution; Supplementary Fig. 14). Still, our docking method can be extended to model high-energy intermediate states of the covalent bond-forming reaction [40], as reflected in the high enrichment in retrospective docking screens against FAAH and AChE (Supplementary Table 4), where the carbamates were modeled as high-energy intermediates (Supplementary Fig. 1). The method also may be suited to irreversible Michael acceptors, such as acrylamide-based inhibitors of the EGFR kinase, where retrospective enrichment was also high (Supplementary Table 4).

The failure to explicitly model the energy of covalent bond formation is one of several methodological gaps that merit mention. Without considering the covalent bond energy, we cannot directly compare molecules bearing different electrophiles. Thus, currently each library of different electrophiles is docked and ranked separately. Incorporating quantum-level approaches to predict the reactivity of specific scaffold/electrophile combinations may be a direction forward [41]. Even within the current constraint-based approach, sampling can be improved by relaxing around ligand dihedral angles, and perhaps by accounting for bonded energies. Lastly, as demonstrated by compound **14** with AmpC, allowing for receptor flexibility can be essential for accurate modeling of receptor/ligand interactions. This remains an area of active research [12,42].

These caveats should not obscure the potential benefits offered by the method. We hope it will begin to address what has been a substantial gap in our ability to screen widely for useful covalent chemical probes, and we have made it and its associated libraries freely available through a general use, public access web-portal (http://covalent.docking.org).

## Online Methods

### Ligand Generation

Ligand flexibility is sampled by generating ligand conformations prior to docking. Given a SMILES string of a ligand with a specific electrophile, we use the OEChem library (OEChem TK 1.7.4; Openeye Scientific Software: Santa Fe, NM. http://www.eyesopen.com) to convert the ligand to its final, reacted form (Supplementary Fig. 1). The receptor's nucleophilic atom involved in the covalent bond is represented by a dummy atom (silicon, for technical reasons). Following the generation of the 'reacted' electrophile, the ligand's 3D structures and stereoisomers are built by Corina [43] (Molecular Networks, Erlangen, Germany), and then protonated and tautomerized by EPIK (Schrodinger software, Catsville NY). Partial atomic charges and solvation energies are calculated for each of these structures with AMSOL [44]. The electrophile serves as a starting rigid body fragment, and conformations are generated using Omega [45] (Omega parameters: EnergyWindow=30.0; MaxConfs=10,000; RMSThreshold=0.5). The collection of pre-generated ligand conformations in the reacted state is saved to a DOCK-readable flexibase format file.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

*Sampling* of ligand poses within the protein binding-site is restricted to exhaustive ligand placement with respect to the covalent bond (Supplementary Fig. 2). The covalent attachment point is sampled in steps of 20° around the terminal dihedral of the nucleophilic side chain. Based on the electrophile geometry determined during ligand generation, and user provided parameters, the vectors of the covalent bond from the ligand and receptor sides are aligned and the ligand is rotated around this vector in 20° steps. For each placement, all of the pre-generated ligand conformations are scored and the score for the best pose is saved. This process is repeated for different values of the covalent bond length and angles, centered on ideal values (Supplementary Fig. 2). The magnitude of deviation from the ideal values, as well as the step sizes, are user specified.

*Scoring* is performed as previously described, using pre-calculated van der-Waals, electrostatic, and ligand solvent-excluded desolvation grids, correcting for ligand desolvation [21]. Receptor structures were prepared using an automated procedure as described in [46] using DELPHI [47] for electrostatics. The ligand's electrophilic atom participating in the bond is omitted from the overall ligand score.

## Availability

As noted, the method is accessible through a public web-server (http://covalent.docking.org), and for download as part of the next DOCK3.x release (http://dock.compbio.ucsf.edu/).

## Virtual ligand libraries

For the curation of ligand libraries, the electrophiles were represented by SMARTS regular expressions (Supplementary Table 1). The Full, Lead-like, and Fragment-like subsets of the ZINC database of commercially available molecules were filtered using these patterns to identify electrophile-bearing molecules [32]. Known inhibitors for the five retrospective virtual screens were collected from ChEMBL14 [48] for AChE, FAAH (carbamates with <1uM activity), and NS3 (α-ketoamides with <100nM activity). For EGFR ligands, we used the same ligand set as in [16], and FAAH boronic acid inhibitors ($K_i$ < 10 nM) were collected from [49]. The best receptor structure was selected based on its ability to enrich known ligands.

Known AmpC binders were defined as having < 1uM activity in ChEMBL16 [48]. Compound similarity was calculated using ECFP4-based Tanimoto coefficients [50] as implemented in Pipeline-Pilot version 6.1 (SciTegic Inc. San Diego, CA). Boron is not parameterized in several of our ligand generation programs, to overcome this in the construction of our boronic acids library, we searched ZINC's commercial catalogs for boronic acids, replacing the boron with a carbon during ligand generation. This doesn't affect ranking, as the energy of the ligand's bonding atom is omitted from the docking score.

The one step cyanoacrylamide based library is based on filtering ZINC's fragment-like subset for fragments containing an aldehyde. The two step, Suzuki-Miyaura based cyanoacrylamide library is made according to (Supplementary Fig. 11) by combining any of 50 commercially available aldehyde containing boronic acids with 4,397 aryl bromide

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

fragments from ZINC's fragments-in-stock subset, containing one of the following motifs: 'ncN', [NX3][CX3](=[OX1])[#6], or a pyridine, pyrimidine, pyrrole, pyrazine, pyrazole, or triazole.

### Covalent Virtual Screening

For the AmpC retrospective pose recapitulation benchmark (Supplementary Table 5) and prospective screen, PDB: 4E3N was used as the receptor. The flexible side chain of Gln120, was truncated at Cβ. The backbone of residues Ala318 and Ser64 and side chains of Asn152 and Tyr150 were polarized to emphasize the electrostatic interactions with the boronic acid. Covalent bond sampling parameters were set to d=1.5±0.1Å a=116.0°±5° and b=109.5°±5°. (Step size of 0.05Å and 1°; Supplementary Fig. 2). Compounds **1**, **3**, **4**, and **14–18** were selected for testing based on this screen. Following the determination of the crystal structure of AmpC in complex with **3,** a second screen was performed using the new structure (PDB: 4LV1; in this screen, the bond length was fixed to d=1.6Å, and Gln120 was not truncated). Compounds **2, 5** and **6** were selected for testing. Run time for the entire screen was 227 CPU hours (ran in parallel on an 800-CPU computer cluster; Wall time < 1 hour). The RSK2 screen used PDB: 4D9T for the receptor structure. Given that nearly all kinase inhibitors form one or more hydrogen bonds with the hinge region, the backbone amides of the kinase hinge residues (Glu494 and Met496) were polarized to emphasize hinge-binding hydrogen bonds. A technique we have long used. Covalent bond sampling parameters were set to d=2.0Å a=109.5°±10° and b=109.5°±10° (Step size of 2.5°; Supplementary Fig. 2). This bond length was chosen based on the two available RSK2 cyanoacrylamide complexes (PDBs: 4D9U,4D9T). While longer than a typical thioether by about 0.2Å, this helps ensure that the scores are dominated by the non-covalent terms, minimizing hard van der Waals repulsion in the region of the new bond. We have not attempted to optimize this term. Calculation time for the docking screen was 103 CPU hours, elapsed wall time was less than 1 hour owing to use of a cluster. JAK3 docking was performed using eight available crystal structures PDB: 1YVJ, 3LXK, 3LXL, 3PJC, 4HVD, 4HVG, 4HVH, and 4HVI. For each structure we used both the native cysteine rotamer as well as an alternative rotamer (χ1= −60°). Docking to the native rotamer produced few plausible results, likely due to steric clashes of the electrophile with nearby residues, and only one candidate was chosen based on this rotamer (Supplementary Table 9). The backbone amides of the hinge residues (Glu903 and Leu905) were polarized similarly to RSK2. Covalent bond sampling parameters were set to d=1.8Å a=109.5°±10° and b=109.5°±10° (Step size of 2.5°; Supplementary Fig. 2). Calculation time for the entire docking screen was 11,740 CPU hours, elapsed wall time ~14 hours.

### Selection criteria

As noted, following docking, the top 500 (kinases) −1000 (AmpC) molecules were manually inspected for exclusion criteria based on considerations that are orthogonal to the docking scoring function such as structural diversity, commercial availability and or synthetic accessibility, perceived ability to derivatize and improve hits, correct representation of the molecule, and internal strain (ligand internal energy is not part of the scoring function). We remove redundant, highly similar molecules and compounds with potentially reactive or unstable functional groups. Additionally, for AmpC we mostly selected poses in which the

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

boronic acid was predicted to occupy the oxyanion hole. For the kinase inhibitors we only selected poses predicted to form at least one hydrogen bond with the hinge. For JAK3 inhibitors we usually restricted the $C\beta_{cysteine}$-$S\gamma_{cysteine}$-$C\beta_{acrylamide}$-$C\alpha_{acrylamide}$ dihedral angle to be >90° to avoid internal strain.

### RMSD calculations

RMSD values were calculated using the Hungarian matching algorithm as implemented in DOCK6 [51]. For comparison purposes, RMSDs for the β-lactam benchmark were calculated using software generously provided by Xuchang Ouyang [16]. For AmpC, receptors were superimposed based on chain A before calculations. For RSK2, receptors were superimposed based on C436, M496, and C560.

### Enzymology

AmpC Enzyme inhibition was measured from initial rates using curve fitting in the native Agilent software. Compounds were initially dissolved in DMSO at 100 mM and individually diluted from such stocks. AmpC activity on CENTA ($K_m$ =15 μM) was monitored by the change of absorbance at 405 nm [24]. AmpC was expressed and purified as described [24], CENTA was purchased from Tydock Pharma (Modena, Italy). $IC_{50}$ values were obtained by fitting percent inhibition to a sigmoidal dose-response equation using GraphPad Prism (GraphPad Software Inc., Supplementary Fig. 16). $K_i$ values (average of two biological experiments) were determined using the Cheng-Prusoff equation assuming competitive inhibition. Reactions were performed at room temperature in 50 mM sodium-cacodylate, pH 6.5 in the presence of 0.01% Triton-X-100 in 1 mL cuvettes with 50–100 μM CENTA and initiated by addition of AmpC.

### Reversibility

Compounds **3** (1 μM) and **7** (276 nM) were assayed for AmpC (1nM) inhibition with no incubation as described above or after 5 minutes incubation with 10nM AmpC. In the latter, the reaction was started by 10x dilution of AmpC and the incubated inhibitor into a reaction buffer with substrate. Reversibility experiments were performed in 50 mM potassium phosphate buffer, pH 7.0

### Selectivity

Compounds **2**, **3**, **5**, and **7** were tested against porcine pancreas elastase (Sigma E-0258), porcine pancreas trypsin (Sigma T-0134), and bovine pancreas α-chymotrypsin (Sigma C-7762). The following substrates were used (Bachem Biosciences): Suc-Ala-Ala-Pro-Ala-pNA (L1775), Suc-Ala-Ala-Pro-Arg-pNA (L1720), and Suc-Ala-Ala-Pro-Phe-pNA (L1400) respectively. Serine protease activities were assayed at a concentration of 0.01 mg/mL enzyme in 50 mM Tris buffer, pH 7.0, with 0.01% v/v Triton X-100. The reaction was initiated by the addition of 200 μM substrate and monitored at 405 nm. $IC_{50}$ values were calculated from single inhibitor concentration measurements (usually 100/1000 μM), and $K_i$ values were estimated for each inhibitor and protein pair using reported $K_m$ values: 190 μM for elastase [52], 37 μM for trypsin [53] and 50 μM for α-chymotrypsin [54].

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript



### In vitro proteasome activity assay

Each reaction contained the substrate Suc-LLVY-AMC (R&D Systems) at a final concentration of 150 μM; assay buffer (25 mM HEPES pH 7.4, 100 mM KCl, 20 mM MgCl2, and 10% glycerol); 1% DMSO, 100 μM compound, or 5 nM PA26; and 1 nM of the yeast 20S proteasome (except for the control reaction). Following the addition of the proteasome, fluorescence was read every 20 seconds, for 20 minutes. Experiments were conducted in duplicates at 23°C. Activity rates were calculated based on the last 10 minutes of a reaction. The yeast 20S proteasome and 26S proteasome activator complex were generous gifts from the laboratory of Philip Coffino. 100 μl reactions were performed in 96-well plates using a SpectraMax M5 Microplate Reader.

### Microbiology

Susceptibility testing was performed and interpreted following the guidelines of the Clinical and Laboratory Standards Institute [25]. The compounds were dissolved in DMSO, and dilutions were made into Muller Hinton medium, keeping DMSO < 5%. Inhibitors were tested for synergy with the third-generation β-lactam cefotaxime against clinical bacteria. The ratio of β-lactam to inhibitor was 1:4. Each value reported reflects the average of three independent experiments. The bacteria exhibited high levels of resistance to cefotaxime because of the expression of class C β-lactamases (AmpC) or class A extended-spectrum β-lactamases (ESBLs). Three Escherichia coli strains and one strain each of Citrobacter freundii, Enterobacter aerogenes and Enterobacter cloacae showed an AmpC-overproduction phenotype. Finally, two of the E. coli strains produced the plasmid-mediated class A ESBLs, TEM-3 and CTX-M-14 (Table 1).

### Crystallography

Co-crystals of AmpC in complex with the inhibitors MAPB, **3**, **7**, and **14** (PDB codes: 4LV0, 4LV1, 4LV2, 4LV3) were grown by the hanging drop vapor diffusion method equilibrated over 1.7 M potassium phosphate buffer (pH 8.6–8.8). Protein (4 mg/mL) was mixed with 1 mM inhibitors and incubated for 20–30 minutes. Drops were set up by mixing 2 μL of protein-inhibitor with 2 μL of well solution; 1 μL of microseeding solution was added to promote crystal growth. Crystals appeared after 2–7 days of equilibration at 20°C. Before data collection, crystals were immersed in a cryoprotectant solution composed of 25% sucrose, 1.7 M potassium phosphate, pH 8.7, for about 30s and were flash-cooled in liquid nitrogen. The cryoprotectant solution also contained the respective inhibitor at 1 mM concentration.

Diffraction was measured at beamline 8.3.1 of the Advance Light Source (ALS, Lawrence Berkeley National Laboratory). Reflections were indexed, integrated and scaled using the XDS package [55] in the space group C2 with two molecules in the asymmetric unit. Structure refinement was carried out using Phenix [56]. Coot [57] was used for model building and eLBOW [58] was used to generate coordinates and ligand restraints; The Ser64 Oxygen-Boron bond length was set to 1.45 Å.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript



The initial phasing model was based on apo-AmpC (PDB: 1KE4), with water molecules and ions removed. B-factors were refined isotropically and the protein was subjected to TLS refinement; three TLS groups were determined for each chain using the TLSMD server [59].

Similarly, co-crystals of RSK2 T493M in complex with compound **24** (PDB: 4M8T) were grown by hanging drop diffusion as described previously [27], and diffraction data were collected at the ALS on beamline 8.2.2. Diffraction images were indexed and integrated using XDS. Molecular replacement was performed using apo-RSK2 CTD as a starting model (PDB: 2QR8) using Phaser, and TLS refinement was carried out using Phenix. See Supplementary Table 10 for crystallographic statistics.

### RSK2 kinase assays

Wild-type and T493M RSK2 kinase activity were assayed as reported previously [26]. Briefly, ERK2-activated RSK2 CTD (5 nM) was incubated with varying concentrations of each inhibitor for 30 minutes in the presence of 100 μM ATP and 10 mM GSH. Each reaction was initiated by the addition of 167 μM substrate peptide (RRQLFRGFSFVAK) and 0.3μCi/μL $\gamma$-$^{32}$P-ATP in a final volume of 25 μL for an additional 30 minutes. Reactions were spotted on phosphocellulose membranes, washed once with 10% AcOH, twice with 0.1% $H_3PO_4$, and once with MeOH prior to drying. Blots were exposed to a phosphor storage plate (Amersham Biosciences), imaged with a Typhoon scanner (GE Healthcare), and quantitated using the SPOT program [60]. $IC_{50}$ values were calculated using a sigmoidal dose response fitting in the Prism program (Graphpad) and are reported as the average of two biological experiments. Full $IC_{50}$ curves are presented in Supplementary Fig. 17.

### JAK3 kinase assays

JAK3 kinase activity was assayed using recombinant JAK3 (Invitrogen, catalog # PV5774). JAK3 (3.1 nM) was incubated with varying concentrations of each inhibitor for 30 minutes in the presence of 11.5 μM ATP. Each reaction was initiated by the addition of 17.9 μM substrate peptide and 0.3 μCi/μL $\gamma$-$^{32}$P-ATP in a final volume of 25 μL for an additional 60 minutes. Reactions were spotted on phosphocellulose membranes, washed once with 10% AcOH, twice with 0.1% $H_3PO_4$, and once with MeOH prior to drying. Blots were exposed to a phosphor storage plate (Amersham Biosciences), imaged with a Typhoon scanner (GE Healthcare), and quantitated using the SPOT program[60].

### JAK kinase selectivity

Kinase selectivity dose-response experiments for **31** and **33** were performed by Nanosyn (Santa-Clara, CA). Test compounds were diluted in 100% DMSO using 3-fold dilution steps. Final compound concentration in assay ranged from 10μM to 0.056 nM. Compounds were tested in a single well for each dilution, and the final concentration of DMSO in all assays was kept at 1%. All assays were performed with a substrate concentration of 1 μM and $K_m$ ATP concentration. Enzyme concentrations ranged 0.1–2 nM, and incubation times were 2–4 hours. See Supplementary Table 11 for exact concentrations. $IC_{50}$ values were calculated using a sigmoidal dose response fitting in the Prism program (Graphpad) and are reported as the average of two biological experiments.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript



### Assay demonstrating dissociation of covalent inhibitors from RSK2 CTD and JAK3

Compound **24** (5 μM) and WT RSK2 CTD (100 nM) were incubated in the presence of 10 mM GSH and 100 μM ATP for 1 hour in RSK kinase buffer (20 mM HEPES buffer, pH 8.0 with 10 mM $MgCl_2$, 0.2 mg/mL BSA and 2 mM Tris(carboxyethyl)phosphine (TCEP)). The solution was then diluted 20-fold in duplicate into the RSK kinase buffer or RSK kinase buffer containing compound **24** (final concentration 5 μM). After 1 hour, kinase activity was then assayed as previously described.

Compound **31** (2 μM) and JAK3 (62.6 nM) were incubated in Invitrogen's recommended JAK3 kinase buffer (50 mM HEPES, pH 7.5, 0.01% BRIJ-35, 10 mM $MgCl_2$ and 1mM EGTA) in the presence of 11.5 μM ATP, 2.5 mM DTT and 0.2 mg/ml BSA for 1 hour. The solution was then diluted 20-fold in duplicate into the JAK3 kinase buffer or JAK3 kinase buffer containing compound **31** (final concentration 2 μM) in the presence of 100 μM ATP, 2.5 mM DTT and 0.2 mg/ml BSA for 1 hour. After an additional hour, kinase activity was assayed as previously described.

### Cell-based assay and Western blotting

Confluent COS-7 cells were transfected overnight with HA-tagged full-length RSK2, MSK1, and MSK1 C458V, as reported previously [27]. Transfected cells were seeded into 6-well plates at 500,000 cells/well in 2 mL DMEM supplemented with 10% FBS (Axenia), 100 units/mL penicillin and 100 μg/mL streptomycin (Gibco) and allowed to adhere for 4 hours. Cells were serum starved for 18 hours, then treated with inhibitor or DMSO for an additional 2 hours. Cells were stimulated with PMA (100 ng/mL) for 30 minutes, washed with cold PBS, and frozen. Cells were thawed into 60 μL of lysis buffer (50 mM HEPES, pH 7.4, 150 mM NaCl, 0.1% Triton X100 supplemented with Roche phosphatase and protease inhibitor cocktails). Lysates were clarified by centrifugation at 14,000 rpm, normalized by Bradford assay, denatured in SDS and separated by 7.5% acrylamide SDS-PAGE. Gels were transferred to nitrocellulose, blocked with Odyssey LiCOR blocking buffer for one hour, and probed with 1:1000 HA, 1:1000 pS380 RSK2 or 1:1000 pS376 MSK1 antibody dilutions. After thorough washing with TBST, blots were incubated with 1:10,000 dilutions of fluorescent secondary antibodies (Odyssey) for one hour, washed with TBST, and scanned on an Odyssey LiCOR instrument. Raw gel images can be found in Supplementary Fig. 15. phospho-S380 RSK2 (Cat #9335S), phospho-S376 MSK1 (Cat #9591S) antibodies were purchased from Cell Signaling. The HA antibody was 12CA5 from Roche, Cat #11 666 606 001. COS-7 cells were obtained from the ATCC.

### Compound sources

Compounds were sourced from the following vendors: Combi-Blocks: **1–11**, **13–15,**, Matrix Scientific: **12,** Alfa-Aesar: **16**, TCI: **17**, PepTech: **18**. These compounds were sourced at 95% or greater purity as described by the vendors. Compounds **19–42** were synthesized in-house, see Supplementary Note 2 for synthetic chemistry details and characterization.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript



### Structural Data Deposition

All crystal structures reported here were deposited in the Protein Data Bank (PDB) with accession codes 4LV0, 4LV1, 4LV2, 4LV3 and 4M8T.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

Computational methods supported by US NIH grant GM59957, web portal supported by NIH GM71896. Also supported by the Ministère de la Recherche et de la Technologie, the Institut national de la santé et de la recherche médicale (UMR Inserm U1071), the Institut National de la Recherche Agronomique (USC-2018), and the Centre Hospitalier Régional Universitaire de Clermont-Ferrand, France (to RB). We thank Dr. M. Fischer and Dr. D Shaya for help with x-ray data collection, Dr. A. O'Donoghue for protease substrates, Dr. P. Coffino and Dr. S. Menant for the proteasome and the PA26 complex sample and Dr. S. Barelier for reading of this manuscript. NL was supported by an EMBO long term fellowship (ALTF 1121-2011) and the UCSF Program for Breakthrough Biomedical Research, which is funded in part by the Sandler Foundation. SK was supported by a fellowship from California TRDRP (#19FT-0091). PC was supported by Howard Hughes Medical Institute Predoctoral Fellowhip.

## References

1. Weerapana E, Simon GM, Cravatt BF. Disparate proteome reactivity profiles of carbon electrophiles. Nat Chem Biol. 2008; 4:405–7. [PubMed: 18488014]
2. Blum G, von Degenfeld G, Merchant MJ, Blau HM, Bogyo M. Noninvasive optical imaging of cysteine protease activity using fluorescently quenched activity-based probes. Nat Chem Biol. 2007; 3:668–77. [PubMed: 17828252]
3. Cohen MS, Zhang C, Shokat KM, Taunton J. Structural bioinformatics-based design of selective, irreversible kinase inhibitors. Science. 2005; 308:1318–21. [PubMed: 15919995]
4. Chang JW, Nomura DK, Cravatt BF. A potent and selective inhibitor of KIAA1363/AADACL1 that impairs prostate cancer pathogenesis. Chem Biol. 2011; 18:476–84. [PubMed: 21513884]
5. Robertson JG. Mechanistic basis of enzyme-targeted drugs. Biochemistry. 2005; 44:5561–71. [PubMed: 15823014]
6. Drahl C, Cravatt BF, Sorensen EJ. Protein-reactive natural products. Angew Chem Int Ed Engl. 2005; 44:5788–809. [PubMed: 16149114]
7. Kathman SG, Xu Z, Statsyuk AV. A fragment-based method to discover irreversible covalent inhibitors of cysteine proteases. J Med Chem. 2014; 57:4969–74. [PubMed: 24870364]
8. Sirois S, Hatzakis G, Wei D, Du Q, Chou KC. Assessment of chemical libraries for their druggability. Comput Biol Chem. 2005; 29:55–67. [PubMed: 15680586]
9. Baell JB, Holloway GA. New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. J Med Chem. 2010; 53:2719–40. [PubMed: 20131845]
10. Potashman MH, Duggan ME. Covalent modifiers: an orthogonal approach to drug design. J Med Chem. 2009; 52:1231–46. [PubMed: 19203292]
11. Weerapana E, et al. Quantitative reactivity profiling predicts functional cysteines in proteomes. Nature. 2010; 468:790–5. [PubMed: 21085121]
12. Totrov M, Abagyan R. Flexible ligand docking to multiple receptor conformations: a practical alternative. Curr Opin Struct Biol. 2008; 18:178–84. [PubMed: 18302984]
13. Shoichet BK. Virtual screening of chemical libraries. Nature. 2004; 432:862–5. [PubMed: 15602552]
14. Kitchen DB, Decornez H, Furr JR, Bajorath J. Docking and scoring in virtual screening for drug discovery: methods and applications. Nat Rev Drug Discov. 2004; 3:935–49. [PubMed: 15520816]
15. Glick M, Jacoby E. The role of computational methods in the identification of bioactive compounds. Curr Opin Chem Biol. 2011; 15:540–6. [PubMed: 21411361]

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

16. Ouyang X, et al. CovalentDock: automated covalent docking with parameterized covalent linkage energy estimation and molecular geometry constraints. J Comput Chem. 2013; 34:326–36. [PubMed: 23034731]
17. Toledo Warshaviak D, Golan G, Borrelli KW, Zhu K, Kalid O. Structure-based virtual screening approach for discovery of covalently bound ligands. J Chem Inf Model. 2014; 54:1941–50. [PubMed: 24932913]
18. Zhu K, et al. Docking covalent inhibitors: a parameter free approach to pose prediction and scoring. J Chem Inf Model. 2014; 54:1932–40. [PubMed: 24916536]
19. Schroder J, et al. Docking-based virtual screening of covalently binding ligands: an orthogonal lead discovery approach. J Med Chem. 2013; 56:1478–90. [PubMed: 23350811]
20. De Cesco S, et al. Virtual screening and computational optimization for the discovery of covalent prolyl oligopeptidase inhibitors with activity in human cells. J Med Chem. 2012; 55:6306–15. [PubMed: 22765237]
21. Mysinger MM, Shoichet BK. Rapid context-dependent ligand desolvation in molecular docking. J Chem Inf Model. 2010; 50:1561–73. [PubMed: 20735049]
22. Jacoby GA. AmpC beta-lactamases. Clin Microbiol Rev. 2009; 22:161–82. Table of Contents. [PubMed: 19136439]
23. Livermore DM, Mushtaq S. Activity of biapenem (RPX2003) combined with the boronate beta-lactamase inhibitor RPX7009 against carbapenem-resistant Enterobacteriaceae. J Antimicrob Chemother. 2013; 68:1825–31. [PubMed: 23580564]
24. Eidam O, et al. Fragment-guided design of subnanomolar beta-lactamase inhibitors active in vivo. Proc Natl Acad Sci U S A. 2012; 109:17448–53. [PubMed: 23043117]
25. Wikler, MA. Performance Standards for Antimicrobial Susceptibility Testing: Twentieth Informational Supplement. Clinical and Laboratory Standards Institute; 2010.
26. Serafimova IM, et al. Reversible targeting of noncatalytic cysteines with chemically tuned electrophiles. Nat Chem Biol. 2012; 8:471–6. [PubMed: 22466421]
27. Miller RM, Paavilainen VO, Krishnan S, Serafimova IM, Taunton J. Electrophilic fragment-based design of reversible covalent kinase inhibitors. J Am Chem Soc. 2013; 135:5298–301. [PubMed: 23540679]
28. Doehn U, et al. RSK is a principal effector of the RAS-ERK pathway for eliciting a coordinate promotile/invasive gene program and phenotype in epithelial cells. Mol Cell. 2009; 35:511–22. [PubMed: 19716794]
29. Kang S, et al. p90 ribosomal S6 kinase 2 promotes invasion and metastasis of human head and neck squamous cell carcinoma cells. J Clin Invest. 2010; 120:1165–77. [PubMed: 20234090]
30. Park J, et al. RAS-MAPK-MSK1 pathway modulates ataxin 1 protein levels and toxicity in SCA1. Nature. 2013; 498:325–31. [PubMed: 23719381]
31. Le NT, et al. A crucial role for p90RSK-mediated reduction of ERK5 transcriptional activity in endothelial dysfunction and atherosclerosis. Circulation. 2013; 127:486–99. [PubMed: 23243209]
32. Irwin JJ, Sterling T, Mysinger MM, Bolstad ES, Coleman RG. ZINC: a free tool to discover chemistry for biology. J Chem Inf Model. 2012; 52:1757–68. [PubMed: 22587354]
33. Yamaoka K, et al. The Janus kinases (Jaks). Genome Biol. 2004; 5:253. [PubMed: 15575979]
34. Kremer JM, et al. A phase IIb dose-ranging study of the oral JAK inhibitor tofacitinib (CP-690,550) versus placebo in combination with background methotrexate in patients with active rheumatoid arthritis and an inadequate response to methotrexate alone. Arthritis Rheum. 2012; 64:970–81. [PubMed: 22006202]
35. Jiang JK, et al. Examining the chirality, conformation and selective kinase inhibition of 3-((3R, 4R)-4-methyl-3-(methyl(7H-pyrrolo[2,3-d]pyrimidin-4-yl)amino)piperidin-1-y l)-3-oxopropanenitrile (CP-690,550). J Med Chem. 2008; 51:8012–8. [PubMed: 19053756]
36. Fleischmann R, et al. Placebo-controlled trial of tofacitinib monotherapy in rheumatoid arthritis. N Engl J Med. 2012; 367:495–507. [PubMed: 22873530]
37. Clark JD, Flanagan ME, Telliez JB. Discovery and development of Janus kinase (JAK) inhibitors for inflammatory diseases. J Med Chem. 2014; 57:5023–38. [PubMed: 24417533]

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript


38. Honigberg LA, et al. The Bruton tyrosine kinase inhibitor PCI-32765 blocks B-cell activation and is efficacious in models of autoimmune disease and B-cell malignancy. Proc Natl Acad Sci U S A. 2010; 107:13075–80. [PubMed: 20615965]
39. Smith AJ, Zhang X, Leach AG, Houk KN. Beyond picomolar affinities: quantitative aspects of noncovalent and covalent binding of drugs to proteins. J Med Chem. 2009; 52:225–33. [PubMed: 19053779]
40. Hermann JC, et al. Structure-based activity prediction for an enzyme of unknown function. Nature. 2007; 448:775–9. [PubMed: 17603473]
41. Schwobel JA, et al. Prediction of michael-type acceptor reactivity toward glutathione. Chem Res Toxicol. 2010; 23:1576–85. [PubMed: 20882991]
42. Fischer M, Coleman RG, Fraser JS, Shoichet BK. Incorporation of protein flexibility and conformational energy penalties in docking screens to improve ligand discovery. Nat Chem. 2014; 6:575–83. [PubMed: 24950326]
43. Gasteiger J, Rudolph C, Sadowski J. Automatic generation of 3D-atomic coordinates for organic molecules. Tetrahedron Computer Methodology. 1990; 3:537–547.
44. Li J, et al. Extension of the platform of applicability of the SM5. 42R universal solvation model. Theoretical Chemistry Accounts. 1999; 103:9–63.
45. Hawkins PC, Skillman AG, Warren GL, Ellingson BA, Stahl MT. Conformer generation with OMEGA: algorithm and validation using high quality structures from the Protein Databank and Cambridge Structural Database. J Chem Inf Model. 2010; 50:572–84. [PubMed: 20235588]
46. Mysinger MM, Carchia M, Irwin JJ, Shoichet BK. Directory of useful decoys, enhanced (DUD-E): better ligands and decoys for better benchmarking. J Med Chem. 2012; 55:6582–94. [PubMed: 22716043]
47. Gilson MK, Sharp KA, Honig BH. Calculating the electrostatic potential of molecules in solution: method and error assessment. Journal of computational chemistry. 1988; 9:327–335.
48. Gaulton A, et al. ChEMBL: a large-scale bioactivity database for drug discovery. Nucleic Acids Res. 2012; 40:D1100–7. [PubMed: 21948594]
49. ADAMS, J., et al. Boronic acids and esters as inhibitors of fatty acid amide hydrolase. WO Patent. 2,008,063,300. 2008.
50. Rogers D, Hahn M. Extended-connectivity fingerprints. J Chem Inf Model. 2010; 50:742–54. [PubMed: 20426451]
51. Brozell SR, et al. Evaluation of DOCK 6 as a pose generation and database enrichment tool. J Comput Aided Mol Des. 2012; 26:749–73. [PubMed: 22569593]
52. Del Mar EG, Largman C, Brodrick JW, Fassett M, Geokas MC. Substrate specificity of human pancreatic elastase 2. Biochemistry. 1980; 19:468–72. [PubMed: 6898442]
53. Pouvreau L, et al. Effect of pea and bovine trypsin inhibitors on wild-type and modified trypsins. FEBS Lett. 1998; 423:167–72. [PubMed: 9512351]
54. Rodriguez-Martinez JA, Rivera-Rivera I, Sola RJ, Griebenow K. Enzymatic activity and thermal stability of PEG-alpha-chymotrypsin conjugates. Biotechnol Lett. 2009; 31:883–7. [PubMed: 19224136]
55. Kabsch W. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. Journal of applied crystallography. 1993; 26:795–800.
56. Adams PD, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr. 2010; 66:213–21. [PubMed: 20124702]
57. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr. 2004; 60:2126–32. [PubMed: 15572765]
58. Moriarty NW, Grosse-Kunstleve RW, Adams PD. electronic Ligand Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint generation. Acta Crystallogr D Biol Crystallogr. 2009; 65:1074–80. [PubMed: 19770504]
59. Painter J, Merritt EA. TLSMD web server for the generation of multi-group TLS models. Journal of Applied Crystallography. 2006; 39:109–111.
60. Knight ZA, Feldman ME, Balla A, Balla T, Shokat KM. A membrane capture assay for lipid kinase activity. Nat Protoc. 2007; 2:2459–66. [PubMed: 17947987]

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript



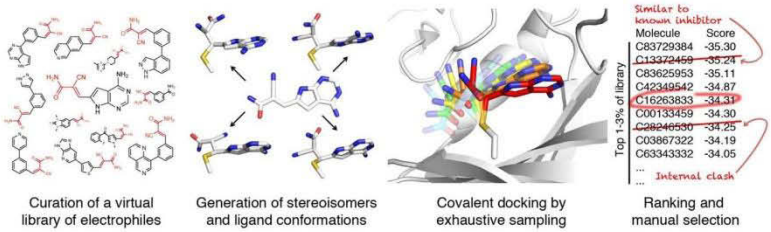



**Figure 1. Overview of the DOCKovalent methodology**

A library of commercially available or easily synthesized small molecules containing a specific electrophile is constructed virtually. In this example, the cyanoacrylamide electrophile is shown in red. All stereoisomers, protonation states and conformations of each ligand are pre-generated. Conformational space is exhaustively sampled around the covalent bond for each pre-generated ligand state, and each pose is scored using a physics-based energy function. Each molecule is represented by its best scoring pose, and high-ranking candidates are manually selected for experimental validation.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript



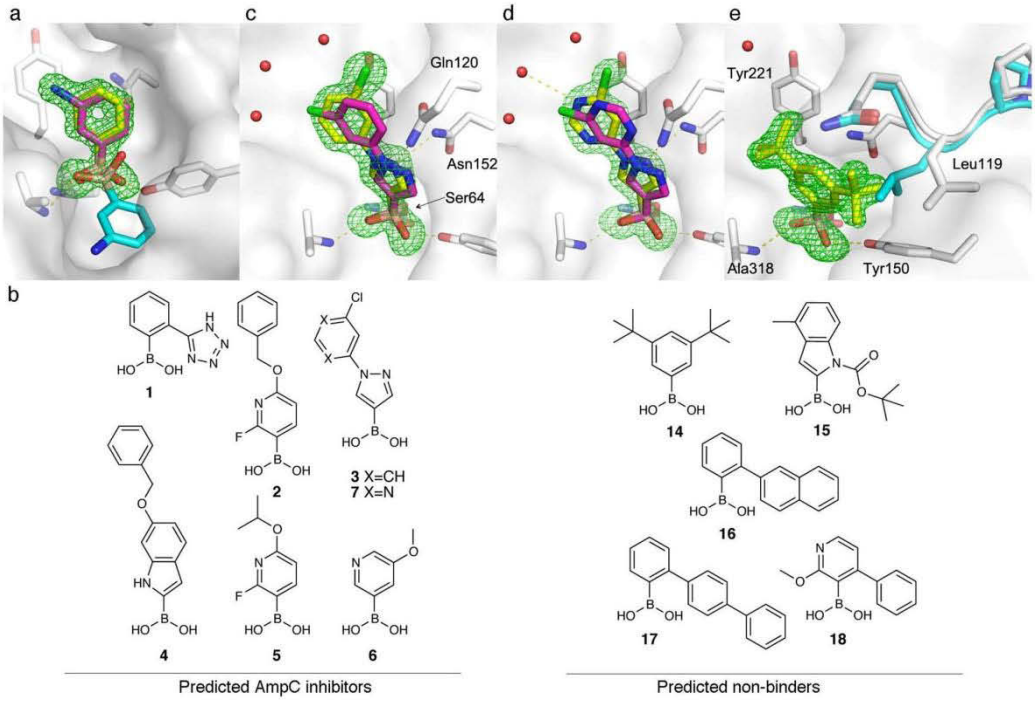



**Figure 2. Boronic acid inhibitors of AmpC predicted by virtual screening**

Crystal structures of boronic acids (yellow) covalently attached to AmpC are overlaid with their respective docking predictions (magenta). The omit $F_o$-$F_c$ electron map is shown in green. **a.** Crystal structure of MAPB superposed on the docking prediction and the published structure (cyan, PDB: 3BLS) **b.** Chemical structures of predicted binders **1**–**6** and non-binders **14**–**18. c** and **d.** X-ray structures of **3** and **7** superposed on their docking predictions. **e.** Compound **14** induces an unanticipated rotamer change in Leu119 and a rearrangement of loop 117–120 relative to the published structure of apo-AmpC (cyan, PDB: 1KE4).

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript



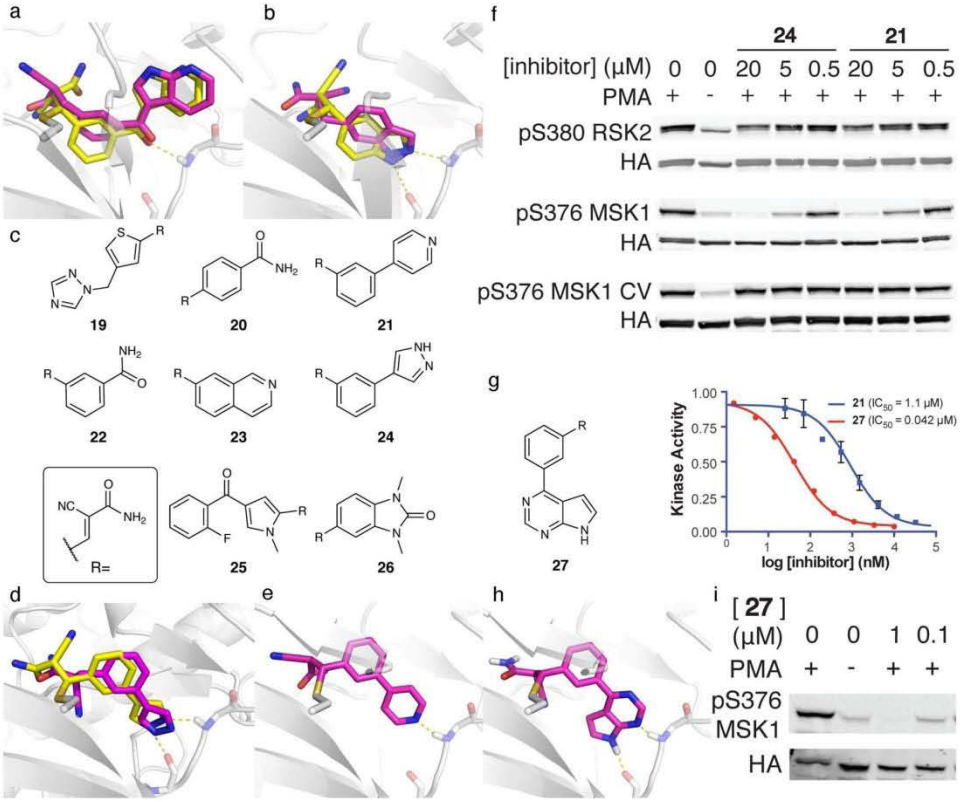



**Figure 3. Cyanoacrylamide inhibitors of RSK2 and MSK1 predicted by covalent docking**

**a** and **b.** Blind docking predictions of two cyanoacrylamide fragments covalently bound to RSK2 (magenta) recapitulate their crystallographic poses (yellow, PDB: 4JG7,4JG6). **c.** Chemical structures of cyanoacrylamide fragments selected for synthesis and testing. **d.** Docking prediction for the most potent fragment **24** corresponds well to the experimental structure. **e.** Docking prediction of the binding mode of compound **21. f.** Compounds **24** and **21** inhibit autophosphorylation of RSK2 and MSK1 in PMA-stimulated cells. Neither compound inhibits the cysteine to valine mutant of MSK1 at concentrations up to 20 μM. Western blots are representative of duplicate biological measurements. **g.** Dose-response curves comparing pyrrolopyrimidine **27** and **21** vs. WT RSK2. **27** was designed based on the docked structure of **21** (See **e.**). Data are plotted as the mean of duplicate measurements ± the range. **h.** Docked pose of **27. i.** Compound **27** inhibits MSK1 autophosphorylation in PMA-stimulated cells. All western blots are representative of duplicate experiments. Full gel images can be found in Supplementary Fig. 15.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript



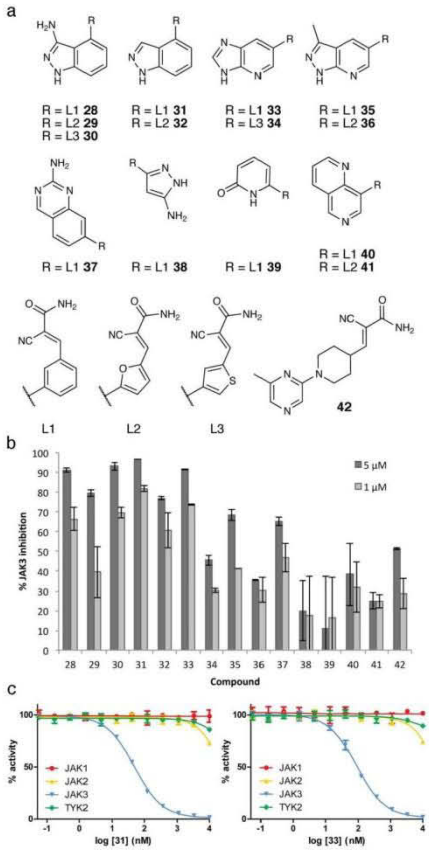



**Figure 4. Reversible covalent JAK3 inhibitors discovered via docking**

**a.** First- and second-generation virtual libraries of cyanoacrylamide fragments were screened by DOCKovalent vs. JAK3. Compounds **28–42** were selected and synthesized as described in the Supplementary Information. **b.** JAK3 inhibition at 1 μM and 5 μM. **c.** Cyanoacrylamides **31** and **33** are selective for JAK3 over JAK1, JAK2, and TYK2. JAK3 $IC_{50}$ = 49 nM and 93 nM**,** respectively. Data represent mean values of two independent experiments ± s.d.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript



**Table 1**

Docking rank, *in vitro* $K_i$ values and minimum inhibitory concentrations of boronic acids against AmpC

| | | | | Minimum inhibitory concentrations (MICs; μg/mL)[a] | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | AmpC overproducer | | | | | | ESBL[c] producers | |
| Compound | Dock Rank | $K_i$ [μM] | Ligand Eff.[b] | Enterobacter cloacae | Enterobacter aerogenes | Citrobacter freundii | Escherichia coli | Escherichia coli | Escherichia coli | Escherichia coli (TEM-3) | Escherichia coli (CTX-M-14) |
| Cefotaxime alone:[d] | | | | 64 | 32 | 16 | 16 | 8 | 4 | 8 | 256 |
| **1** | 11 | N/A[e] | N/A | | | | | | | | |
| **2** | 63 | 0.17[f] | 0.50 | 8 | 8 | 4 | 8 | 4 | 4 | 4 | 64 |
| **3** | 95 | 0.04[f] | 0.66 | 4 | 8 | 4 | 4 | 4 | 2 | 2 | 32 |
| **4** | 420 | 2.37[e] | 0.38 | | | | | | | | |
| **5** | 646 | 0.48[f] | 0.61 | 8 | 8 | 8 | 8 | 4 | 8 | 4 | 256 |
| **6** | 5240[h] | 3.55[e] | 0.66 | | | | | | | | |
| **7** | 280 | 0.01[f] | 0.71 | 1 | 4 | 2 | 2 | 1 | 1 | 1 | 32 |

[a] Compounds were dosed at a cefotaxime:inhibitor ratio of 1:4.

[b] Ligand efficiency based on the calculated $K_i$

[c] Extended-spectrum β-lactamase producers.

[d] MICs for cefotaxime alone.

[e] N/A: < 10% inhibition at 10 μM

[f] $IC_{50}$ was calculated based on a full dose response curve (Supplementary Fig. 16)

[g] $IC_{50}$ was calculated based on a single point measurement

[h] The docking hit list was dominated by larger analogs of this compound that were unavailable for purchase. Compound **6** did not rank well, but was purchased as a proxy for structurally related, high-ranking predictions.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript



**Table 2**

Docking rank and *in vitro* $IC_{50}$ values for cyanoacrylamides 19 – 26 against RSK2 WT and T493M mutant C-terminal kinase domain.

| Compound | DOCK rank | $IC_{50}$ (μM) | |
|---|---|---|---|
| | | RSK2 WT | RSK2 T493M |
| **19** | 66 | 50.4 | 27.9 |
| **20** | 96 | 7 | 5.2 |
| **21** | 122 | 1.1 | 0.43 |
| **22** | 132 | 3.3 | 6.8 |
| **23** | 142 | 12.7 | 6.4 |
| **24** | 200 | 1.2 | 0.37 |
| **25** | 368 | >100 | >100 |
| **26** | 391 | 6 | 7.1 |

ERRATUM

## Rick Morimoto

Catherine Goodman

*Nat. Chem. Biol.* **10**, 875 (2014); published online 17 October 2014; corrected after print 11 December 2014

In the version of this article initially published, Mark Wilson was inadvertently described as 'the' founding member of the PDRC, instead of 'a' founding member. The error has been corrected in the HTML and PDF versions of the article.

ERRATUM

## Covalent docking of large libraries for the discovery of chemical probes

Nir London, Rand M Miller, Shyam Krishnan, Kenji Uchida, John J Irwin, Oliv Eidam, Lucie Gibold, Peter Cimermančič, Richard Bonnet, Brian K Shoichet & Jack Taunton

***Nat. Chem. Biol.* 10, 1066–1072 (2014); published online 26 October 2014; corrected after print 22 December 2014**

In the version of this article initially published, the scaffold structure shown for compounds 33 and 34 was incorrect in Figure 4a, with a nitrogen atom misplaced within the five-membered ring moiety. The error has been corrected in the HTML and PDF versions of the article.

ERRATUM

## Subcellular metal imaging identifies dynamic sites of Cu accumulation in *Chlamydomonas*

Anne Hong-Hermesdorf, Marcus Miethke, Sean D Gallaher, Janette Kropat, Sheel C Dodani, Jefferson Chan, Dulmini Barupala, Dylan W Domaille, Dyna I Shirasaki, Joseph A Loo, Peter K Weber, Jennifer Pett-Ridge, Timothy L Stemmler, Christopher J Chang & Sabeeha S Merchant

***Nat. Chem. Biol.* 10, 1034–1042 (2014); published online 26 October 2014; corrected after print 6 January 2015**

In the version of this article initially published, the labels in Figure 4b were incorrect. The error has been corrected in the HTML and PDF versions of the article.

CORRIGENDUM

## Hydrolysis of 2′3′-cGAMP by ENPP1 and design of nonhydrolyzable analogs

Lingyin Li, Qian Yin, Pia Kuss, Zoltan Maliga, José L Millán, Hao Wu & Timothy J Mitchison

***Nat. Chem. Biol.* 10, 1043–1048 (2014); published online 26 October 2014; corrected after print 29 January 2015**

In the version of this article initially published, the chemical structures of the two cyclic dinucleotides in Figure 4a were incorrect at the triphosphate group, with each of the phosphates shown as containing two extra hydrogen atoms. The error has been corrected in the HTML and PDF versions of the article.

© 2015 Nature America, Inc. All rights reserved.

## Annexe n°5

Evaluation of the efficiency of cefoxitin/cefepime combination against Enterobacteriaceae resistant to expanded-spectrum cephalosporins.

Serres A, **Gibold L**, Dalmasso G, Robin F, Bonnet R, Delmas J.

Int J Antimicrob Agents. 2015 Jan;45(1):86-7. doi: 10.1016/j.ijantimicag.2014.10.001. Epub 2014 Oct 22.

## Evaluation of the efficiency of cefoxitin/cefepime combination against Enterobacteriaceae resistant to expanded-spectrum cephalosporins

*Keywords:*
Antibiotic resistance
Gram-negative
ESBL
AmpC β-lactamase
Cefoxitin
Cefepime

*Sir,*

Production of extended-spectrum β-lactamases (ESBLs) and overproduction of AmpC β-lactamases are the most frequent mechanisms by which antibiotic multiresistance occurs. Cephamycins such as cefoxitin have consistent activity in vitro against Enterobacteriaceae and exhibit good stability against hydrolysis by ESBLs [1]. Cefepime has intrinsic activity against AmpC-overproducing strains and has been used for the treatment of invasive infections due to AmpC β-lactamase-producing organisms [2]. Theoretically, the combination of cefoxitin with cefepime should cover most Enterobacteriaceae resistant to expanded-spectrum cephalosporins.

Minimum inhibitory concentrations (MICs) of clinical strains, chosen to be representative of the dominant resistance groupings found in clinical samples (Table 1), were determined using a microdilution method and were interpreted according to the recommendations of the European Committee on Antimicrobial Susceptibility Testing (EUCAST). Unlike ESBL-producing strains (MICs of 2–128 mg/L), strains overproducing AmpC β-lactamases were susceptible to cefepime (0.125–1 mg/L). Conversely, no strain overproducing chromosomal or plasmid AmpC β-lactamases was susceptible to cefoxitin (MICs of 16 mg/L to ≥64 mg/L), whilst cefoxitin MICs were in the susceptible range for ESBL-producing strains, except for the CTX-M-9/TEM-3-producing *Escherichia coli* strain, the CTX-M-15-producing *Klebsiella pneumoniae* strain and the TEM-24-producing *Enterobacter aerogenes* strain (MICs of 16 mg/L to >64 mg/L).

The in vitro efficacy of the cefoxitin/cefepime combination was assessed by the chequerboard method. With regard to ESBL-producing strains, the combination showed an additive effect against CTX-M-producing strains [fractional inhibitory concentration index (FICI) range 0.62–0.75] and a synergistic effect against TEM-producing strains (FICI = 0.31). In combination with 8 mg/L cefoxitin (susceptibility breakpoint), cefepime MICs, except for CTX-M-15 producing *K. pneumoniae*, decreased substantially, typically by 32-fold or more, and dropped into the susceptible range (MICs ≤ 0.5 mg/L). With regard to AmpC-producing strains, the combination revealed no antagonistic effect between cefoxitin and cefepime for AmpC enzymes (FICI range 0.75–2), suggesting that cefepime conserved activity against AmpC-producing strains.

The efficacy of the cefoxitin/cefepime combination against a clinical CTX-M-14-producing *E. coli* and a clinical AmpC-overproducing *Enterobacter cloacae* was investigated in a mouse sepsis model. These strains are representative of those found in clinical laboratories in terms of epidemiology and MICs. Following infection of mice (OF1) by intraperitoneal administration, antibiotics were intraperitoneally administered at previously reported concentrations (50 mg/kg body weight dose of cefepime [3], 50 mg/kg dose of imipenem [4], 200 mg/kg dose of cefoxitin [1] and 50/200 mg/kg dose of cefoxitin/cefepime) combination at 0.5, 3 and 6 h post infection [4]. Survival of mice (eight animals per group) was monitored for 24 h post infection. Upon infection with CTX-M-14-producing *E. coli*, the survival rate of mice treated with the cefoxitin/cefepime combination (88%) was similar to that of the imipenem-treated group (88%), significantly higher than in mice receiving cefepime alone (25%) ($P = 0.037$) and different to that of the group with cefoxitin treatment (75%) ($P > 0.05$). Upon infection with AmpC-overproducing *E. cloacae*, only 50% of mice treated with cefoxitin or cefepime alone survived at 24 h post

**Table 1**
Fractional inhibitory concentration indexes (FICIs) and minimum inhibitory concentration (MICs) of cefoxitin (FOX) and cefepime (CEF) alone and in combination against clinical strains producing extended-spectrum β-lactamases and AmpC β-lactamases, and frequency of mutants resistant to FOX.

| Strain and β-lactamase(s) | FICI[a] | MIC (mg/L) | | | | | | Frequency of mutants resistant to FOX ($\times 10^{-8}$) | |
|---|---|---|---|---|---|---|---|---|---|
| | | FOX[b] | FOX with CEF[c] (1 mg/L) | FOX with CEF[d] (4 mg/L) | CEF[e] | CEF with FOX[f] (1 mg/L) | CEF with FOX[g] (8 mg/L) | FOX 8 mg/L | FOX 8 mg/L with CEF 1 mg/L |
| *Escherichia coli* CTX-M-1 | 0.62 | 2 | 1 | 0.50 | 4 | 1 | ≤0.015 | 5.8 (±3.6) | <0.01 |
| *E. coli* CTX-M-2 | 0.62 | 8 | 8 | 2 | 16 | 8 | 0.5 | 60 (±17) | 0.02 (±0.05) |
| *E. coli* CTX-M-14 | 0.75 | 2 | 0.5 | ≤0.015 | 2 | 0.25 | ≤0.015 | 2.4 (±1.3) | <0.01 |
| *E. coli* CTX-M-9/TEM-3 | 0.31 | 16 | 8 | 0.25 | 16 | 8 | ≤0.015 | NR[h] | 0.18 (±2.2) |
| *E. coli* CTX-M-15 | 0.62 | 2 | 0.5 | ≤0.015 | 2 | 0.25 | ≤0.015 | 1.5 ± 1.0 | <0.01 |
| *Enterobacter aerogenes* TEM-24 | 0.31 | 128 | ≤0.015 | ≤0.015 | 2 | 0.5 | 0.5 | NR | NR |
| *Klebsiella pneumoniae* CTX-M-15 | 0.62 | 16 | 16 | 16 | 128 | 32 | 16 | NR | NR |
| *E. coli* AmpC | 0.75 | 64 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | NR | NR |
| *E. coli* AmpC | 0.75 | 64 | 0.50 | 0.50 | 0.25 | 0.25 | 0.25 | NR | NR |
| *E. coli* AmpC DHA-1 | 1.5 | 16 | 0.50 | ≤0.015 | 0.25 | 0.5 | 0.25 | NR | NR |
| *E. coli* AmpC CMY-2 | 1.1 | 64 | 0.125 | ≤0.015 | 0.25 | 1 | 1 | NR | NR |
| *Enterobacter cloacae* AmpC | 2 | 64 | 64 | 0.25 | 1 | 2 | 2 | NR | NR |
| *E. aerogenes* AmpC | 2 | 64 | 0.25 | ≤0.015 | 0.25 | 0.5 | 0.5 | NR | NR |
| *Citrobacter freundii* AmpC | 1.2 | 256 | 64 | 0.25 | 1 | 2 | 2 | NR | NR |
| *K. pneumoniae* AmpC DHA-1 | 2 | 64 | 0.125 | ≤0.015 | 0.25 | 0.5 | 0.5 | NR | NR |

S, susceptible; R, resistant.
[a] The interaction was defined as synergistic if the FICI was ≤0.5, additive if the FICI was >0.5 to ≤1.0, indifferent if the FICI was >1.0 to ≤2 and antagonistic if the FICI was >2.
[b] Breakpoints of FOX: S ≤ 8 mg/L; R > 32 mg/L.
[c] MIC of FOX in combination with CEF at fixed concentration of 1 mg/L.
[d] MIC of FOX in combination with CEF at fixed concentration of 4 mg/L.
[e] Breakpoints of CEF: S ≤ 1 mg/L; R > 4 mg/L.
[f] MIC of CEF in combination with FOX at fixed concentration of 1 mg/L.
[g] MIC of CEF in combination with FOX at fixed concentration of 8 mg/L.
[h] NR, non-realisable (strain MIC was greater than FOX concentration used).

infection. In contrast, combined cefoxitin/cefepime and imipenem rescued all animals. To complete these data, bacterial loads in the blood, kidney, liver and spleen of treated mice (five to eight mice per group) were determined at 0.5, 3, 6 and 24 h post infection (Supplementary Fig. S1). The cefoxitin/cefepime combination was significantly more effective against the two clinical strains than cefoxitin or cefepime alone in the blood, spleen and liver ($P = 0.004–0.008$ vs. cefoxitin and $P \leq 0.008$ vs. cefepime). However, the efficacy of the cefoxitin/cefepime combination in eliminating CTX-M-14-producing *E. coli* from the kidneys ($P = 0.092$) was comparable with that of cefoxitin, probably because of the high urinary concentrations observed with this antibiotic. The effectiveness of the cefoxitin/cefepime combination was confirmed by a rapid and significant decrease in bacterial blood counts (≥3 log in 6 h).

Supplementary material related to this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ijantimicag.2014.10.001.

The emergence of clinical isolates resistant to cefoxitin by depletion or loss of porins has been observed in *E. coli* [5]. The frequency of mutations conferring cefoxitin resistance to ESBL-producing strains was investigated in vitro at a cefoxitin concentration of 8 mg/L (susceptibility breakpoint) and varied from $1.5 \times 10^{-8}$ to $6.0 \times 10^{-7}$. When 1 mg/L cefepime (susceptibility breakpoint) was added, this frequency decreased substantially (>100 fold) (Table 1). The cefoxitin/cefepime combination therefore prevented the appearance of cefoxitin-resistant mutants in CTX-M-producing *E. coli* strains at clinically relevant concentrations.

In conclusion, cefoxitin and cefepime in combination improve antibacterial activity against *E. coli* and *E. aerogenes* strains producing ESBLs and is effective against AmpC-overproducing strains but not against CTX-M-15-producing *Klebsiella*. This combination also prevented the appearance of mutants resistant to cefoxitin in CTX-M-producing *E. coli* strains and may therefore be used to treat *E. coli* and *Enterobacter* infections as an alternative to carbapenems, which are considered as last-line agents. These results are promising but need to be backed up by additional studies, in particular with other strains and other ESBLs.

**Acknowledgments**

The authors thank Marlene Jan for technical assistance and would like to thank H. N'Guyen for reading and commenting on the manuscript.

*Funding*: This work was supported by the Ministère Français de l'Éducation Nationale, de la Recherche et de la Technologie, l'Institut national de la santé et de la recherche médicale (INSERM U1071), and l'Institut National de la Recherche Agronomique (USC-2018). This work was also supported by the Centre Hospitalier Régional Universitaire de Clermont-Ferrand (Clermont-Ferrand, France).

*Competing interests*: None declared.

*Ethical approval*: The animal protocol was approved by the Committee for Research and Ethical Issues of the Department of Auvergne (CEMEA Auvergne) following international directive 86/609/CEE [n° CE16-09].

**References**

[1] Lepeule R, Ruppé E, Le P, Massias L, Chau F, Nucci A, et al. Cefoxitin as an alternative to carbapenems in a murine model of urinary tract infection due to *Escherichia coli* harboring CTX-M-15-type extended-spectrum β-lactamase. Antimicrob Agents Chemother 2012;56:1376–81.
[2] Tamma PD, Girdwood SCT, Gopaul R, Tekle T, Roberts AA, Harris AD, et al. The use of cefepime for treating AmpC β-lactamase-producing Enterobacteriaceae. Clin Infect Dis 2013;57:781–8.
[3] Pechère J-C, Vladoianu IR. Development of resistance during ceftazidime and cefepime therapy in a murine peritonitis model. J Antimicrob Chemother 1992;29:563–73.
[4] Eidam O, Romagnoli C, Dalmasso G, Barelier S, Caselli E, Bonnet R, et al. Fragment-guided design of subnanomolar β-lactamase inhibitors active in vivo. Proc Natl Acad Sci U S A 2012;109:17448–53.
[5] Martínez-Martínez L. Extended-spectrum β-lactamases and the permeability barrier. Clin Microbiol Infect 2008;14(Suppl. 1):82–9. Erratum in: Clin Microbiol Infect 2008;14(Suppl. 5):21–4.

Arnaud Serres
*Laboratoire de Bactériologie, Centre Hospitalo-Universitaire Clermont-Ferrand, Clermont-Ferrand, France*

Lucie Gibold [a,b]
[a] *Laboratoire de Bactériologie, Centre Hospitalo-Universitaire Clermont-Ferrand, Clermont-Ferrand, France*
[b] *Microbes, Intestins, Inflammation et Susceptibilité de l'Hôte, INSERM U1071, INRA USC2018, Université d'Auvergne, Clermont-Ferrand, France*

Guillaume Dalmasso
*Microbes, Intestins, Inflammation et Susceptibilité de l'Hôte, INSERM U1071, INRA USC2018, Université d'Auvergne, Clermont-Ferrand, France*

Frédéric Robin [a,b]
Richard Bonnet [a,b]
Julien Delmas [a,b,*]
[a] *Laboratoire de Bactériologie, Centre Hospitalo-Universitaire Clermont-Ferrand, Clermont-Ferrand, France*
[b] *Microbes, Intestins, Inflammation et Susceptibilité de l'Hôte, INSERM U1071, INRA USC2018, Université d'Auvergne, Clermont-Ferrand, France*

* Corresponding author at: Laboratoire de Bactériologie, CHU Clermont-Ferrand, 58 rue Montalembert, 63003 Clermont-Ferrand, France. Tel.: +33 4 7375 4920; fax: +33 4 7375 4922.
*E-mail address:* jdelmas@chu-clermontferrand.fr (J. Delmas)

22 September 2014
http://dx.doi.org/10.1016/j.ijantimicag.2014.10.001

## Optimal loading regimen and achievement of trough concentrations for teicoplanin using Japanese population parameters

*Sir,*

Teicoplanin (TEIC) is an effective glycopeptide antibiotic against infections with Gram-positive cocci. As TEIC has a long half-life, large volume of distribution and low clearance, the blood concentration requires time to achieve the therapeutic range at steady state. A loading dose (LD) is therefore administered on Day 1 for a rapid therapeutic response [1]. However, as the LD of TEIC is decided by the physician, the probability of achieving an optimal LD regimen remains unclear.

We investigated the optimal LD and maintenance dose (MD) as well as the probability of achieving the therapeutic range at steady state using a Monte Carlo simulation. The pseudorandom number of the pharmacokinetic (PK) parameters for 30 000 virtual patients according to a normal distribution was estimated using R software v.3.0.2 (R Core Team, Vienna, Austria; http://www.R-project.org/) with the rnorm function, which was based on and consistent with previously reported population PK parameters in Japan [2]. The trough concentration ($C_{trough}$) of TEIC was calculated using a two-compartment model with the PK parameters obtained in the Monte Carlo simulation.

Of the 30 000 virtual patients, 14 959 were male and 15 041 were female, with a mean ± standard deviation age of 75.5 ± 11.6