HAL
open science

Caractérisation du rôle de l'Antigène 43 dans le processus de colonisation d'Escherichia coli O157 H7

Valentin Ageorges

To cite this version:

Valentin Ageorges. Caractérisation du rôle de l'Antigène 43 dans le processus de colonisation d'Escherichia coli O157 H7. Microbiologie et Parasitologie. Université Clermont Auvergne [2017-2020], 2019. Français. NNT : 2019CLFAC019 . tel-02910354

HAL Id: tel-02910354

https://theses.hal.science/tel-02910354v1

Submitted on 2 Aug 2020

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.

UNIVERSITE CLERMONT AUVERGNE
CLERMONT-FERRAND
Année 2019

ECOLE DOCTORALE
DES SCIENCES DE LA VIE, SANTE, AGRONOMIE ET ENVIRONNEMENT
N° d'ordre :

Thèse

Présentée à l'Université Clermont-Auvergne pour l'obtention du grade de

DOCTEUR D'UNIVERSITÉ
Spécialité : Microbiologie

Soutenue le 26 juin 2019

VALENTIN AGEORGES

Caractérisation du rôle de l'Antigène 43 dans le processus de colonisation d'*Escherichia coli* O157:H7

Directeur de thèse :
Mickaël Desvaux (DR, INRA)
Co-encadrants :
Frédérique Chaucheyras-Durand (IR, Lallemand/INRA)
Christian Provot (DS, Biofilm Control)

Membres du Jury

Présidente :
Valérie Livrelli (PU, Université Clermont-Auvergne, Clermont-Ferrand)
Rapporteurs :
Muriel Cocaign-Bousquet (DR, INSA, Toulouse)
Romé Voulhoux, (DR, CNRS, Marseille)
Christophe Beloin, (DR, Institut Pasteur, Paris)

Unité Microbiologie Environnement Digestif et Santé, UMR454, Université Clermont-Auvergne, INRA Auvergne Rhône-Alpes, Site de Theix

A ma grand mère ...

REMERCIEMENTS

Ma thèse a débuté le 1er février 2016 dans l'Unité de Microbiologie de l'INRA de Theix dirigée par Régine Talon que je remercie de m'avoir accueilli au sein de son laboratoire et de m'avoir soutenu durant ces 3 années. Lorsque ma thèse s'est achevée le 31 janvier 2019, l'Unité était devenue UMR MEDiS sous la direction de Pierre Peyret et d'Annick Bernalier-Donadille.

Tous mes remerciements vont à mon directeur de thèse Mickaël Desvaux pour son encadrement et la confiance qu'il m'a accordée durant ce travail. Merci pour ta grande disponibilité, ton soutien inconditionnel et la liberté que tu m'as offerte pour que je puisse réaliser ce travail de thèse dans les meilleures conditions possibles. Je n'oublierai pas tes grandes qualités humaines, ton humour, toutes les discussions que nous avons pu avoir, tous les conseils, scientifiques ou non, que tu as pu me délivrer pendant ces années et les bons moments partagés en congrès. Je te serai toujours reconnaissant pour ce que nos échanges ont pu m'apporté sur le plan aussi bien scientifique que personnel. Merci pour tout.

Je tiens également à remercier les partenaires industriels Lallemand et Biofilm Control et plus particulièrement mes co-encadrants Frédérique Chaucheyras-Durand et Christian Provot pour leur disponibilité, leurs encouragements et les précieux conseils qu'ils m'ont apportés tout au long de ce projet. Un grand merci à Frédérique et la société Lallemand de m'avoir permis de terminer la rédaction de mon manuscrit dans de bonnes conditions.

Un grand merci à Muriel Cocaign-Bousquet, Christophe Beloin et Romé Voulhoux de l'intérêt qu'ils ont porté à mon travail en acceptant d'être mes rapporteurs. Je remercie également Valérie Livrelli de m'avoir fait l'honneur d'accepter de présider mon jury de thèse.

Je remercie Nicolas Barnich et Damien Balestrino pour les conseils prodigués lors de mes comités de suivi de thèse.

Je tiens évidemment à adresser un merci tout particulier à Nelly Caccia. Merci pour tous tes conseils techniques, mon "GPS" du labo. Merci pour les (très) longues heures que tu as accordé à mes travaux et qui ont grandement contribué au bon déroulement de ma thèse. Merci pour ta disponibilité et ta bonne humeur quotidienne malgré tes courtes nuits. Merci également pour ton soutien dans les moments plus difficiles. Travailler avec toi toutes ces années aura été un véritable plaisir et je souhaite à tous les thésards d'avoir la chance de travailler quotidiennement avec quelqu'un comme toi. Merci Nelly.

Merci également à Ingrid pour ses nombreux conseils en protéomique. Merci pour ta disponibilité, ta gentillesse et ta bonne humeur. Je n'oublierai pas les moments qu'on a passé à Florence. Et aussi… merci pour la boite de nuit !

Merci à Michel Hébraud pour son expertise en protéomique, à Grégory Jubelin et Chloé Habouzit pour m'avoir formé à la cytométrie en flux, à Brigitte Gaillard-Martinie pour la microscopie à fluorescence, à Etienne Dague et Marion Schiavonne pour leur initiation à la microscopie en force atomique ainsi qu'à Eric Peyretaillade et Philippe Ruiz pour leurs grandes compétences en bioinformatique. Mes remerciements à Sabine Leroy pour son expertise en biologie moléculaire, ses conseils et son aide précieuse lors des derniers mois de ma thèse.

Je tiens à remercier chaleureusement l'ensemble du personnel de l'UMR MEDiS pour leur accueil et leur gentillesse. Merci à Alexandra pour son sourire quotidien, merci à Alexandre pour son humour, à Ricardo, Tiago et Geoffrey pour avoir facilité mon intégration. Un merci particulier à Carine Andant pour sa gentillesse, sa bonne humeur et ses encouragements, ce fut un plaisir de partager mon bureau avec toi pendant ces 3 années. Sans oublier Aurore, Jean-Paul, Diane, Roumeila, Stéphanie, Sylvie, Cédric, Justine, Aline, Bertrand … Merci à tous pour les bons moments passés pendant toutes ces années.

Un grand merci à mes amis qui ont toujours été présents et qui m'ont toujours soutenu sans jamais vraiment comprendre ce que je faisais. Merci à Renaud, Maxime, Julie, Kiki, Dub, Thib, Flo et Lise.

Je remercie mes parents et mon frère du fond du cœur. Merci d'avoir cru en moi et de m'avoir encouragé depuis toujours. Sans vous je n'en serai pas là aujourd'hui.

Enfin, je terminerai ces quelques lignes en remerciant plus particulièrement Alicia. Merci pour ta présence, ton soutien et l'amour que tu me portes quotidiennement depuis la moitié de ma thèse. Merci infiniment de tes encouragements et du réconfort que tu m'as apporté dans les moments plus difficiles. Tu as su rendre les choses beaucoup plus faciles, un immense merci.

TABLE DES MATIERES

REMERCIEMENTS **5**

TABLE DES MATIERES **9**

LISTE DES ABREVIATIONS **13**

INTRODUCTION **17**

CHAPITRE I : LES *ESCHERICHIA COLI* ENTEROHEMORRAGIQUES **19**

I. 1. CLASSIFICATION DES *ESCHERICHIA COLI* DIARRHEIQUES (DEC) **20**

I. 2. PATHOGENICITE DES EHEC **22**

I. 2. A. ASPECTS CLINIQUES **22**

I. 2. B. FACTEURS DE VIRULENCE DES EHEC **24**

Les Shiga-toxines (Stx) 24

Le locus d'effacement des entérocytes (LEE) 27

Autres facteurs de virulence des EHEC 29

I. 3. EHEC ET STEC **30**

I. 4. EPIDEMIOLOGIE DES INFECTIONS A EHEC **32**

I. 5. RESERVOIRS ET MODES DE TRANSMISSION DES EHEC **34**

CHAPITRE II : DETERMINANTS MOLECULAIRES DE SURFACE IMPLIQUES DANS LA COLONISATION DES EHEC **37**

II. 1. LES EXOPOLYSACHARIDES (EPS) **38**

II. 2. L'ADN EXTRACELLULAIRE **39**

II. 3. LE SYSTEME DE SECRETION DE TYPE 1 (T1SS) **39**

II. 4. LE SYSTEME DE SECRETION DE TYPE II (T2SS) **40**

II. 5. LE SYSTEME DE SECRETION DE TYPE III (T3SS) **41**

II. 6. LE SYSTEME DE SECRETION DE TYPE IV (T4SS) 41

II. 7. LE SYSTEME DE SECRETION DE TYPE V (T5SS)..... 42

II. 7. A. LES AT (AUTOTRANSPORTERS) 42

Les ATA (adhesins autotransporters) 43

Les SAAT (self-associating autotransporters) 43

Les SPATE (serine-protease autotransporters of Enterobacteriaceae) 44

II. 7. B. LES TAA (TRIMERIC AUTOTRANSPORTERS) 45

II. 7. C. LES IAT (INVERTED AUTOTRANSPORTERS) 46

II. 8. LE SYSTEME DE SECRETION DE TYPE VI (T6SS) 46

II. 9. LE SYSTEME DE SECRETION DE TYPE VII (T7SS)..... 46

II. 10. LE SYSTEME DE SECRETION DE TYPE VIII (T8SS) 48

II. 11. LE SYSTEME BAM (B -BARREL ASSEMBLY MACHINERY)..... 49

II. 12. AUTRES MECANISMES DE TRANSPORT DES PROTEINES..... 50

CHAPITRE III : L'AG43 D'*E. COLI* 52

III. 1. STRUCTURE DE LA PROTEINE AG43 53

III. 2. PHYLOGENIE ET DIVERSITE ALLELIQUE DU GENE *AGN43*..... 56

III. 3. REGULATION DE L'EXPRESSION DU GENE *AGN43* 59

III. 4. ROLES DE L'AG43..... 63

III. 4. A. L'AUTOAGREGATION..... 64

III. 4. B. LA FORMATION DE BIOFILMS ET L'ADHESION..... 67

III. 4. C. IMPLICATION DANS LA VIRULENCE..... 68

CHAPITRE IV : OBJECTIFS DE RECHERCHE 69

RÉSULTATS 73

CHAPITRE I : CARACTERISATION DE L'AG43 DANS L'ADHESION AUX PROTEINES DE L'ECM, L'AUTOAGREGATION ET SA REGULATION CHEZ *E. COLI* O157:H7 75

ARTICLE N°1 77

CHAPITRE II : ETUDE DE LA DIVERSITE PHYLOGENETIQUE, STRUCTURALE ET FONCTIONNELLE DE L'AG43 105

ARTICLE N°2 107

ARTICLE N°3 129

DISCUSSION ET PERSPECTIVES 161

IDENTIFICATIONS DE DETERMINANTS MOLECULAIRES IMPLIQUES DANS LA COLONISATION DE L'ECM PAR *E. COLI* O157:H7 163

DIVERSITE ET PREVALENCE DE L'AG43 CHEZ LES BACTERIES 164

CARACTERISATION DES CAPACITES D'INTERACTION HOMOTYPIQUES ET HETEROTYPIQUES DES DIFFERENTES CLASSES D'AG43 165

ROLE ET REGULATION DE L'AG43 D'*E. COLI* O157:H7 166

RÉFÉRENCES BIBLIOGRAPHIQUES 169

LISTE DES COMMUNICATIONS 197

COMMUNICATION ORALES : 199

POSTERS : 199

ANNEXES 201

LISTE DES ABREVIATIONS

AAF : Aggregative adherence fimbriae

Aap : Anti-aggregation protein

aat : Aggregative ABC transporter

AC : Autochaperon

ADN : Acide désoxyribonucléique

A/E : Attachement/effacement

AEEC : Attaching and effacing *E. coli*

Ag43 : Antigen 43

Ag43 $^{C1 - C6}$: Ag43 appartenant aux classes C1 à C6

AI-2 : Auto-inducteur 2

AIDA-I : Adhesin involved in diffuse adherence phenotype

AIEC : *E. coli* adhérents invasifs

ANSES : Agence nationale de la sécurité sanitaire de l'alimentation, de l'environnement et du travail

ARN : Acide ribonucléique

AT : Autotransporteur

ATA : Adhesins autotransporters

BaDiAl : Adaptation des Bactéries aux environnements Digestif et Alimentaire

BAM : β-barrel assembly machinery

BHI : Brain Heart Infusion

BL : Bottom of the letter L

BrkA : Bordetella resistance to killing

BSA : Bovine serum albumin

cah : calcium binding antigen 43 homologue

CDC : Center for Disease Control and Prevention

CLA : Chain-like adhesion

CoMBa : Colonisation de la Matrice par les Bactéries

DAEC : *E. coli* adhérents diffusifs

Dam : DNA adenine methylase

DEC : *E. coli* diarrhéiques

DMEM : Dulbecco's modified eagle medium

eae : *E. coli* attaching and effacing

EAEC : *E. coli* entéroagrégatifs

ECDC : European Center for Disease Control and Prevention

ECM : Matrice extracellulaire

ECP : *E. coli* common pilus

Efa-1 : EHEC factor for adherence 1

EFSA : European Food Safety Authority

Eha : EHEC autotransporters

EHEC : *E. coli* entérohémorragiques

Eib : *E. coli* immunoglobuin-binding proteins

EIEC : *E. coli* entéroinvasifs

EJ : Elbow joingin

ELF : *E. coli* laminin binding fimbriae

EPEC : *E. coli* entéropathogènes

EPS : Exopolysaccharides

Esp : *E. coli* secreted proteins

EspP : Extracellular serine protease plasmid encoded

ESPR : Extended signal peptide region

ETEC : *E. coli* entérotoxinogènes

ExPEC : *E. coli* pathogènes extraintestinaux

FDA : Food an Drug Administration

FdeC : Factor adherence of *E. coli*

flu : fluffing

FoodNet : Foodborne Diseases Active Surveillance Network

FWD : Food and Waterborne Diseases and Zoonoses

GAPDH : Glyceraldehyde 3-phosphate dehydrogenase

Gb3 : Globotriaosylceramide

Hes : Hemagglutinin from STEC

IAT : Inverted autotransporters

IcsA : Intra-cellular spread protein A

IhA : IrgA homologue adhesin

InPEC : *E. coli* pathogènes intestinaux

INRA : Institut National de la Recherche Agronomique

IrgA : Iron-regulated protein A

LB : Lysogeny broth

LEAT : AT lipase/esterase

LEE : Locus d'effacement des entérocytes

Ler : LEE-encoded regulator

LifA : lymphostatin

LPF : Long polar fimbriae

LPS : Lipopolysaccharide

Mat : Meningitidis-associated temperature dependent pilus

MEDiS : Microbiologie Environnement Digestif et Santé

MLST : Multilocus sequence typing

MSCRAMM : Microbial surface components recognising adhesive matrix molecule

Nle : Non LEE-encoded factor

NMEC : *E. coli* responsables de méningites néonatales

NTEC : *E. coli* nécrotoxinogènes

OI : O-Island

OM Outer membrane

OMP : β-barrel outer membrane protein

OmpA : outer membrane protein A

Pet : Pasmid encoded toxin

PHAT : AT phosphatase/hydrolase

Pic : Protein involved in colonisation

Pili HCP : Haemorrhagic coli pili

PNAG : Poly-β-1,6-N-acetyl-D-glucosamine

PssA : Protease secreted by STEC

PTT : Purpura thrombotique thrombocytopénique

RGD : Arg-Gly-Asp

Saa : STEC autoagglutination adhesin

SAAT : Self-associating autotransporters

Sab : STEC-autotransporter mediating biofilm formation

SDP : Secretion-dependant pathway

Sep : Secretion of EHEC proteins

SEPEC : *E. coli* associés à une septicémie

SHU : Syndrome hémolytique et urémique

SF-I : Sous-famille I

SF-II : Sous-famille II

Sfp : Sorbitol fermenting protein

SL : Stem of the letter L

SP : Peptide signal

SPATE : Serine-protease autotransporters of Enterobacteriaceae

SQ : Senseur du quorum

SslE : Secreted and surface-associated lipoprotein from *E. coli*

STEC : *E. coli* codant des Stx ou *E. coli* producteurs de Stx

Stx : Shiga-toxine

T1P : Pili de type 1

T1SS : Système de sécrétion de type I

T2SS : Système de sécrétion de type II

T2aSS : système de sécrétion de type II, sous-type a

T2bSS : système de sécrétion de type II, sous-type b

T2cSS : système de sécrétion de type II, sous-type c

T3SS : Système de sécrétion de type III

T3aSS : système de sécrétion de type III, sous-type a

T3bSS : système de sécrétion de type III, sous-type b

T4P : Pili de type 4

T4SS : Système de sécrétion de type IV

T5SS : Système de sécrétion de type V

T5aSS : Système de sécrétion de type V, sous-type a

T5bSS : Système de sécrétion de type V, sous-type b

T5cSS : Système de sécrétion de type V, sous-type c

T5dSS : Système de sécrétion de type V, sous-type d

T5eSS : Système de sécrétion de type V, sous-type e

T6SS : Système de sécrétion de type VI

T7SS : Système de sécrétion de type VII

T8SS : Système de sécrétion de type VIII

TAA : Trimeric autotransporters

Tat : Twin-arginine translocation pathway

TD : Tube digestif

TibA : Toxigenic invasion locus b

Tir : Translocated intimin receptor

tra : Transfert

TU : Translocation unit

Tubes HCP : Haemolysin-coregulated protein

UPEC : *E. coli* uropathogènes

USDA : United States department of Agriculture

VAT : AT associés à la vacuolisation

INTRODUCTION

Chapitre I : Les *Escherichia coli* entérohémorragiques

Le genre *Escherichia,* appartenant à la famille des *Enterobacteriaceae*, a été isolé pour la première fois en 1885 par le pédiatre Theodor Escherich et décrit comme commensal de l'intestin de l'enfant. Le nom *Escherichia coli* n'est cependant apparu qu'en 1919 lorsque Castellani et Chalmers décidèrent de reclasser une espèce préalablement connue sous les noms *Bacterium coli* commune ou *Bacillus coli* (Escherich 1988). L'espèce *Escherichia coli* a longtemps été la seule représentante du genre avant la découverte de 6 autres espèces : *E. adercarboxylata* (Leclerc 1962), *E. blattae* (Burgess *et al.* 1973), *E. hermanii* (Brenner *et al.* 1982), *E. vulneris* (Brenner *et al.* 1982), *E. fergusonii* (Farmer *et al.* 1985) et *E. albertii* (Huys *et al.* 2003). Cependant, des analyses phylogénétiques plus récentes ont révélé que le genre *Escherichia* devrait plutôt être divisé en 8 groupes contenant les 3 espèces *E. coli*, *E. fergusonii* et *E. albertii* ainsi que 5 clades numérotés de I à V isolés d'échantillons environnementaux, animaux ou humains (Lawrence *et al.* 1991, Walk *et al.* 2009). En effet, *E. adercaboxylata* a été reclassée en *Leclercia adercarboxylata* (Tamura *et al.* 1986), *E. blattae* en *Shimwellia blattae* (Priest *et al.* 2010)*, E. hermannii* en *Atlantibacter hermannii* (Hata *et al.* 2016) et *E. vulneri*s en *Pseudescherichia vulneris* (Alnajar *et al.* 2017). Les clades eux, ne peuvent pour l'instant pas être affiliés à une espèce en particulier à cause du manque de différences phénotypiques claires avec les autres espèces du genre (Walk *et al.* 2009, Clermont *et al.* 2011). Même si le genre *Escherichia* et l'espèce *E. coli* sont fortement étudiés, il reste probablement d'autres espèces à identifier et décrire, comme le suggère *E. senegalensis* qui attend encore d'être publiée (GenBank : AY217654). De plus, malgré un nom de genre différent, les *Shigella spp.* sont aujourd'hui reconnues comme appartenant à l'espèce *E. coli* (Brenner *et al.* 1972, Chaudhuri *et al.* 2012). En avril 2019, la base de données GenBank recense plus de 700 génomes complets d'*E. coli* séquencés (incluant *Shigella spp.*) contre 15 pour *E. albertii* and 2 *pour E. fergusonii*.

E. coli est une bactérie à Gram négatif ou plus précisément une bactérie diderme lipopolysaccharidique (Desvaux *et al.* 2009). Cette bactérie utilisée comme organisme modèle en microbiologie est avant tout un microorganisme commensal du tube digestif (TD) des animaux à sang chaud, dont l'être humain, (Mason *et al.* 1981). *E. coli* colonise très tôt le TD des jeunes enfants et représente en général moins de 1% du microbiote intestinal humain (Eckburg *et al.* 2005). Cependant, certaines souches d'*E. coli* peuvent provoquer des infections. D'une part les *E. coli* pathogènes extraintestinaux (ExPEC), comprenant les *E. coli*

uropathogènes (UPEC), les *E. coli* responsables de méningites néonatales (NMEC), les *E. coli* nécrotoxinogènes (NTEC) et les *E. coli* associés à une septicémie (SEPEC), d'autre part, les *E. coli* pathogènes intestinaux (InPEC) regroupant les *E. coli* entérotoxinogènes (ETEC), les *E. coli* entéroagrégatifs (EAEC), les *E. coli* entéropathogènes (EPEC), les *E. coli* entérohémorragiques (EHEC), les *E. coli* entéroinvasifs (EIEC), les *E. coli* adhérents diffusifs (DAEC) et *E. coli* adhérents invasifs (AIEC) (Boudeau *et al.* 1999, Kaper *et al.* 2004, Croxen *et al.* 2010, Mann *et al.* 2012). Ces 6 premiers pathotypes d'InPEC sont clairement associés à des diarrhées infectieuses et correspondent à des *E. coli* diarrhéiques (DEC) pouvant être retrouvés à différents niveaux de la chaine alimentaire (Croxen *et al.* 2013).

I. 1. Classification des *Escherichia coli* diarrhéiques (DEC)

Les *E. coli* peuvent être répartis dans 7 groupes phylogénétiques distincts (A, B1, B2, C, D , E et F) (Ochman *et al.* 1984, Jaureguy *et al.* 2008, Walk *et al.* 2009, Tenaillon *et al.* 2010, Clermont *et al.* 2013, Beghain *et al.* 2018). La classification générale d'*E. coli* se fait sur la base de la combinaison du sérotype O:H de l'antigène somatique O du lipopolysaccharide (LPS) et de l'antigène flagellaire H. Si la souche ne possède pas de flagelle, elle est identifiée O:NM (non motile). L'identification de ces antigènes permet la corrélation du sérotype d'une souche à son pouvoir pathogène (Bryan *et al.* 2015).

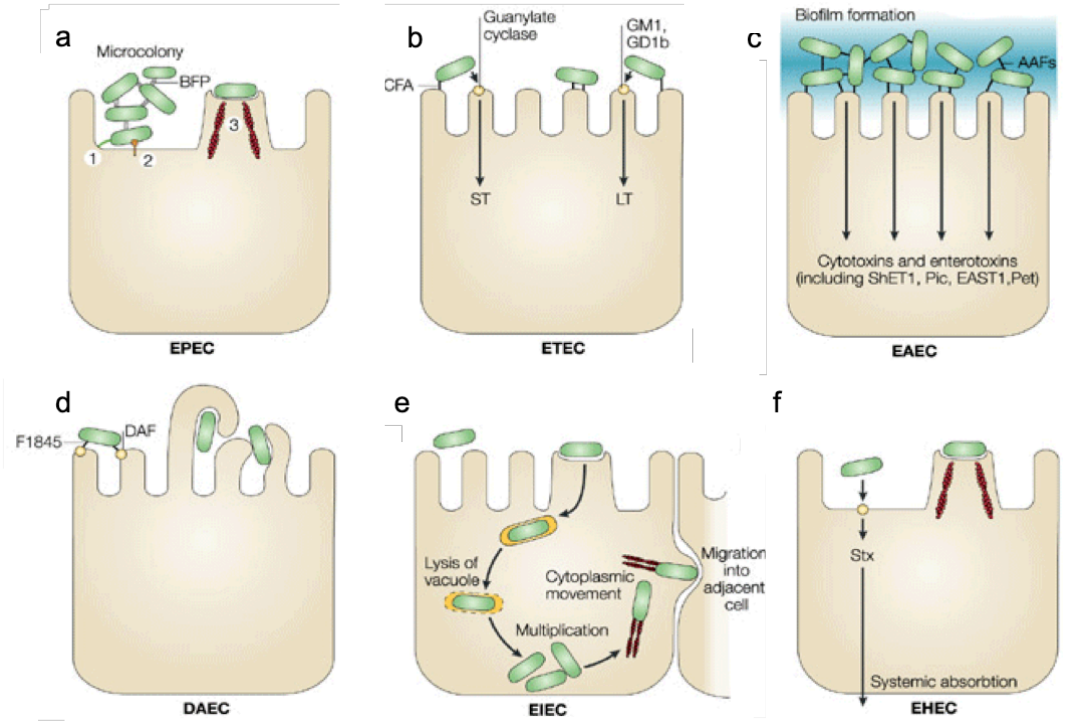

Figure 1 : Mécanismes de colonisation et de virulence des DEC. Les différents pathotypes de DEC possèdent des mécanismes de colonisation et de virulence qui leurs sont propres. (a) : EPEC; (b) : ETEC; (c) : EAEC; (d) : DAEC; (e) : EIEC; (f) : EHEC. AAF, aggregative adherence fimbriae; BFP, bundle-forming pilus; CFA, colonization factor antigen; DAF, decay-accelerating factor; EAST1, enteroaggregative *E. coli* ST1; LT, heat-labile enterotoxin; ShET1, *Shigella* enterotoxin 1; ST, heat-stable enterotoxin. (D'après (Kaper *et al.* 2004))

Les différents DEC sont catégorisés sur la base de paramètres cliniques, histologiques et de facteurs de virulence (Kaper *et al.* 2004, Croxen *et al.* 2013). Les DEC présentent différents mécanismes de colonisation et de virulence (Figure 1). Les EPEC colonisent l'intestin grêle en provoquant des lésions caractéristiques d'attachement/effacement (A/E) des microvillosités intestinales *via* l'expression du locus d'effacement des entérocytes (LEE). Les ETEC adhèrent aux entérocytes de l'intestin grêle *via* de nombreux facteurs de colonisation et libèrent des entérotoxines stables ou labiles. Les EAEC colonisent la muqueuse intestinale par établissement d'épais biofilms et adhèrent entre eux à la surface avec une configuration d'amas de briques. Ils libèrent de plus des cytotoxines et des entérotoxines. Les DAEC induisent un signal de transduction particulier au niveau des entérocytes de l'intestin grêle et se dispersent à l'apex de la cellule selon un motif d'adhésion diffuse. Les EIEC, très proches des *Shigella spp.*, sont des pathogènes intracellulaires qui pénètrent l'épithélium du côlon par trancystose via les cellules M, induisent la lyse du phagosome et se déplacent dans la cellule par nucléation des microfilaments d'actine. Les bactéries peuvent ensuite se multiplier dans les cellules et se

disperser dans les cellules adjacentes. Enfin, Les EHEC établissent également des lésions A/E des microvillosités intestinales grâce au LEE en libérant de plus des Shiga-toxines (Stx) pouvant provoquer des complications potentiellement mortelles pour les sujets à risque (Kaper *et al.* 2004). Sur la base de la présence de divers facteurs de virulence, un continuum des différents pathotypes d'*E. coli* a été récemment proposé (EFSA 2015).

Alors que les *E. coli* commensaux sont majoritairement associés au phylogroupe A, ces différents pathotypes appartiennent à des groupes différents. Par exemple, les EPEC sont répartis dans les groupes B2 et E, les ETEC principalement associés aux groupes A et B1, les EAEC retrouvés dans le groupe B1 et les EHEC dans les groupes A, B1 et D (Clermont *et al.* 2000, Croxen *et al.* 2013, Rossi *et al.* 2018). Ces différents groupements suggèrent une évolution parallèle des différents pathotypes avec l'intervention possible d'éléments mobiles permettant l'acquisition de combinaisons spécifiques de facteurs de virulence (Chaudhuri *et al.* 2012, Croxen *et al.* 2013).

I. 2. Pathogénicité des EHEC

Une classification des agents biologiques en 4 classes selon leurs effets pathogènes a été introduite dans le code du travail (Art. R.4421-3) (Légifrance 2008). Ainsi, sont considérés comme pathogènes les agents biologiques des groupes 2, 3 et 4. Parmi les DEC, les EHEC sont les seuls pathogènes de classe 3 pouvant de ce fait provoquer une maladie grave chez l'Homme, avec une propagation possible dans la communauté, sans traitement efficace à ce jour. De plus, ils constituent un danger sérieux pour les travailleurs exposés à ces pathogènes, notamment dans les industries agro-alimentaires et les laboratoires de recherche et d'enseignement. Les EHEC, dont le sérotype le plus virulent est O157:H7, sont caractérisés par leur capacité à provoquer chez l'Homme des diarrhées aqueuses et hémorragiques, dues à l'expression du LEE et la libération de Stx, pouvant évoluer vers des complications potentiellement mortelles pour les sujets à risques (O'Loughlin *et al.* 2001).

I. 2. A. Aspects cliniques

La dose infectieuse nécessaire pour déclencher les symptômes chez l'Homme est très faible (10 à 100 unités formant colonies) (Tilden *et al.* 1996, Tuttle *et al.* 1999). Différents symptômes caractérisent ces infections en apparaissant séquentiellement dans les 15 jours suivant l'ingestion de l'aliment contaminé. Ces symptômes, ainsi que l'issue de l'infection,

peuvent varier selon l'âge et l'état du système immunitaire de l'hôte. En général, les premiers symptômes apparaissent entre 1 et 9 jours après ingestion sous forme de diarrhées aqueuses modérées, douleurs abdominales sévères et déshydratation modérée. Ces symptômes coïncident avec l'adhésion des bactéries à la surface des cellules épithéliales intestinales causant des lésions A/E des microvillosités. Dans 90% des cas, les diarrhées deviennent sanglantes et caractéristiques d'une colite hémorragique. Les lésions sont localisées dans l'iléon et le côlon et correspondent à des œdèmes, hémorragies et nécroses de la muqueuse intestinale. Ces symptômes peuvent persister pendant 2 à 10 jours et peuvent être accompagnés de vomissements. Des complications potentiellement mortelles pour l'hôte peuvent apparaître dans 10% des cas, une semaine après le début des symptômes. Ces complications sont dues à la production de Stx par les EHEC dans l'intestin qui peuvent passer dans la circulation sanguine et atteindre des organes secondaires portant le récepteur globotriaosylceramide (Gb3) et entrainer des micro-angiopathies thrombotiques. Les toxines agissent ainsi au niveau de la sous-unité 28S du ribosome afin de bloquer la synthèse protéique et entrainer *de facto* la mort cellulaire. Si le tissu vasculaire des reins est touché, on parle de syndrome hémolytique et urémique (SHU). Il en résulte une anémie micro-angiopathique hémolytique, une thrombopénie ainsi qu'une insuffisance rénale aigue. Il est la première cause d'insuffisance rénale aigue en France et frappe essentiellement les enfants de moins de 5 ans. Des troubles du système nerveux central (léthargie, convulsions, encéphalopathie) ont également été observés (Tesh *et al.* 1993). Si c'est le cerveau qui est ciblé, on parle de purpura thrombotique thrombocytopénique (PTT) chez des patients adultes (Andreoli *et al.* 2002). Les caractéristiques physiopathologiques du PTT sont les mêmes que pour le SHU accompagnées de fièvre et symptômes neurologiques plus marqués. Ces manifestations cliniques sévères font des EHEC un problème majeur de santé publique.

A ce jour les seuls traitements contre les infections à EHEC sont essentiellement des traitements symptomatiques de la maladie : réhydratation, nutrition parentérale, transfusion et dialyse pour le SHU et exsanguino-transfusion ou plasmaphérèse associée à une perfusion de plasma frais pour le PTT (Andreoli *et al.* 2002). Contrairement à la majorité des infections microbiennes, l'antibiothérapie est ici très controversée. En effet, les Stx sont codées par des bactériophages de type λ en phase lysogénique et sont libérées lors du cycle lytique des phages (voir ci-après). Certains antibiotiques peuvent donc favoriser la libération de ces toxines très cytotoxiques pour l'endothélium micro-vasculaire rénal en activant le cycle lytique du phage et donc augmenter le risque de SHU. Devant ces difficultés, des traitements alternatifs sont en

cours de développement. Si l'utilisation d'anticorps anti-Stx chez des modèles animaux (Paton *et al.* 2000, Yamagami *et al.* 2001, Sheoran *et al.* 2005) et l'administration par voie orale de composés glucidiques (Takeda 1998) ou peptidiques (Nishikawa *et al.* 2006) ont montré des effets intéressants, les résultats les plus prometteurs sont du côté de l'Eculizumab, un anticorps qui inhibe l'activation du complément (qui est à l'origine des symptômes les plus graves de l'infection) par les Stx (Lapeyraque *et al.* 2011) (Delmas *et al.* 2014).

I. 2. B. Facteurs de virulence des EHEC

Les Shiga-toxines (Stx)

Les symptômes cliniques des infections à EHEC, notamment les colites hémorragiques et le SHU, sont dus à la production de Stx par ces bactéries. Ainsi les Stx, également connues sous le nom de Vérotoxines de par leur cytotoxicité sur les lignées cellulaires Vero, constituent le facteur de virulence majeur des EHEC. Les principales cibles de ces endotoxines sont les cellules endothéliales du côlon, des reins et du cerveau. Sur la base de leur différence de toxicité *in vitro* et *in vivo* et de leurs séquences, elles ont été réparties en deux catégories, Stx1 et Stx2 (Kruger *et al.* 2015). Chacune de ces catégories contient plusieurs variants distingués par leur différence de réactivité immunologique, d'activité biologique, de récepteurs au niveau des cellules cibles et de leur capacité à être activés par l'élastase du mucus intestinal (Scheutz *et al.* 2014). Ainsi, le groupe Stx1 est le plus homogène et contient les variants Stx1a, Stx1c et Stx1d alors que le groupe Stx2 comprend les variants Stx2a, Stx2b, Stx2c, Stx2d, Stx2e, Stx2f et Stx2g (Scheutz *et al.* 2012). Le type de variant retrouvé reflète l'origine des souches et leur pouvoir pathogène. Par exemple, les variants Stx1 et Stx2e, Stx2f et Stx2g sont rarement impliqués dans les infections humaines alors que Stx2a, Stx2c et Stx2d sont fréquemment associés à des cas de SHU (Kawano *et al.* 2008). En plus d'*E. coli*, d'autres espèces peuvent produire ces Stx, dont notamment *Shigella dysenteriae* type I, *Acinetobacter haemolyticus*, *Citrobacter freundii* ou *Enterobacter cloacae*.

Les gènes *stx* sont codés dans le génome de phages lambdoïdes, appelés phages *stx*. Ces phages sont des éléments génétiques mobiles permettant la transmission des gènes *stx* au cours du transit intestinal, dans l'eau, au sein de biofilms ou *via* les fécès (Imamovic *et al.* 2009, Solheim *et al.* 2013). Ces bactériophages s'intègrent dans le génome bactérien et s'y maintiennent en utilisant le même mécanisme que le bactériophage λ avec un cycle lytique/lysogénique (Scotland *et al.* 1983, O'Brien *et al.* 1984). Les gènes codant Stx1 ou Stx2

sont portés par des phages différents mais ayant un mécanisme d'induction similaire (Mellies *et al.* 2014). Lors du cycle lysogénique, le phage est intégré au chromosome bactérien au niveau d'un site préférentiel unique ou d'autres sites secondaires (Serra-Moreno *et al.* 2007) et reste stable en inhibant la synthèse de la plupart de ses gènes, dont les *stx*. En revanche, le prophage *stx* peut moduler l'expression de certains gènes de la bactérie avec par exemple une inhibition de l'expression du système de sécrétion de type III (T3SS) ou une augmentation de l'expression de gènes impliqués dans la motilité (Su 2010, Tozzoli 2014). De plus, il peut réguler l'interaction des bactéries avec l'épithélium intestinal en augmentant entre autres l'expression d'un des récepteurs de l'intimine : la nucléoline (Robinson 2006). Lorsque le prophage est induit et entre en cycle lytique, l'expression de ses gènes est déclenchée et les Stx sont ainsi produites puis libérées suite à la lyse de la bactérie. Plusieurs facteurs peuvent influencer l'expression des gènes *stx* : le stress oxydatif lors d'une infection chez l'Homme (Los *et al.* 2010), le traitement des infections par antibiotiques ciblant la synthèse d'acide désoxyribonucléique (ADN) (Zhang *et al.* 2000), la carence en fer (Mellies *et al.* 2014) ou la présence de diverses molécules telles l'éthanolamine (Kendall *et al.* 2012), l'adrénaline, la noradrénaline ou l'auto-inducteur 3 (AI-3) (Reading *et al.* 2009, Moreira *et al.* 2016).

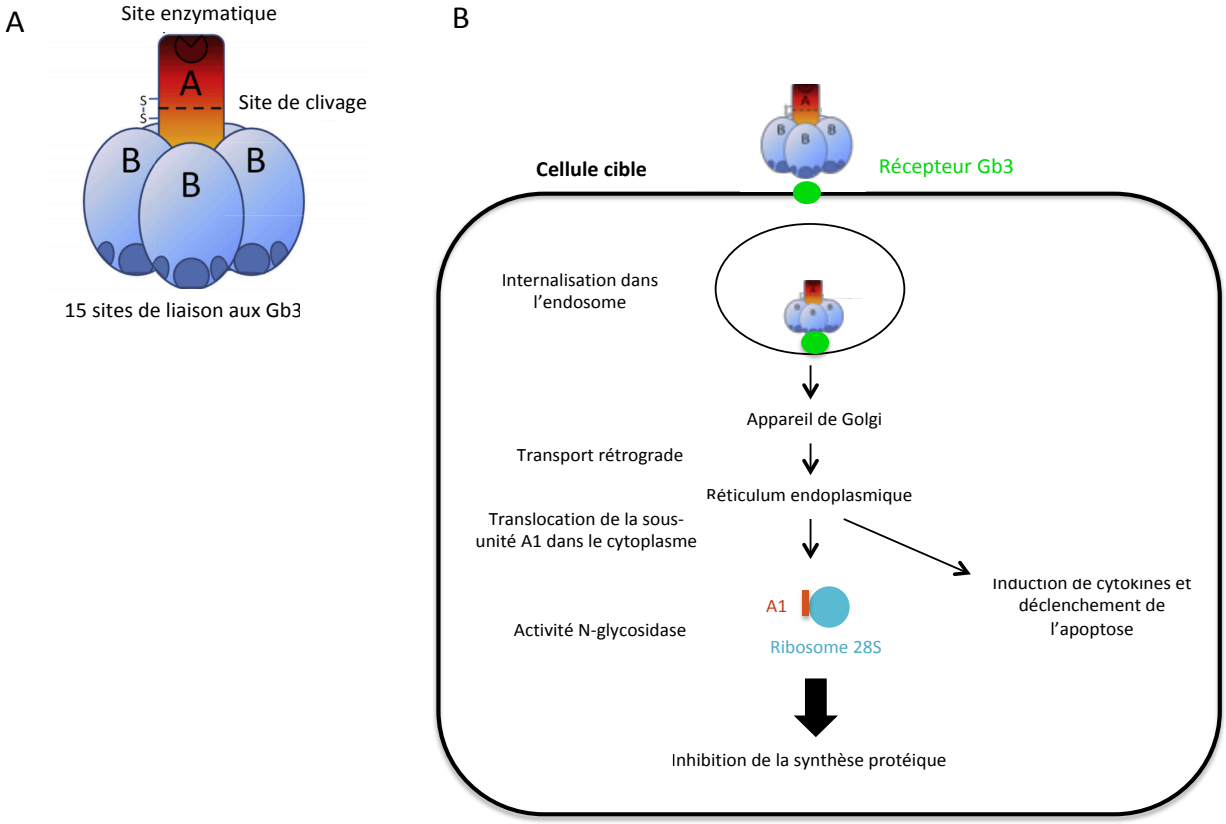

Figure 2 : Structure et mécanisme d'action des Stx. (A) Structure des Stx. (B) Mécanisme d'action des Stx. (D'après (Mariani-Kurkdjian *et al.* 2012) et (Bryan *et al.* 2015))

Sur le plan structurel, les Stx appartiennent à la famille des toxines AB_5. Elles sont donc composées d'une sous-unité A catalytique (StxA), divisée en 2 parties A1 et A2 reliées entre elles par un pont disulfure, liée de façon non covalente à 5 sous unités B identiques formant un pentamère (StxB) (Figure 2). Le pentamère de sous-unités B permet l'interaction avec le récepteur Gb3 situé au niveau des cellules endothéliales intestinales, rénales et cérébrales chez l'Homme (Bauwens *et al.* 2013). La sous-unité A possède quant à elle l'activité enzymatique N-glycosidase et peut donc altérer la synthèse protéique une fois internalisée dans la cellule cible. Si les toxines Stx2 et Stx1 ne présentent que 60% de similarité, elles possèdent néanmoins le même mode d'action (Figure 2) (Sandvig *et al.* 2010, Melton-Celsa 2014). Les bactéries traversent l'épithélium intestinal à travers les cellules M puis entrent et survivent à l'intérieur des macrophages. Après réplication des bactéries, la production de Stx induit la mort de la cellule hôte et les toxines libérées peuvent passer dans la circulation sanguine (Etienne-Mesmin *et al.* 2011). Les Stx peuvent ainsi atteindre et se fixer sur la membrane cytoplasmique des cellules cibles riches en Gb3 grâce à la sous-unité B (O'Loughlin *et al.* 2001). Le complexe Stx/Gb3 est ensuite internalisé par endocytose et fusionne avec l'endosome avant de migrer par un transport rétrograde de l'appareil de Golgi vers le réticulum endoplasmique. Pendant cette

étape, la sous-unité A est clivée et seule la forme A1 est libérée dans le cytosol pour atteindre l'enveloppe nucléaire et les ribosomes. L'activité N-glycosidase de A1 retire le résidu adénosine 2348 de l'ARN ribosomique 28S du ribosome 60S, induisant l'arrêt de la synthèse protéique et ainsi la mort cellulaire (Spooner *et al.* 2012). Les Stx pourraient également induire l'apoptose de différents types cellulaires (Tesh 2010).

Le locus d'effacement des entérocytes (LEE)

Les gènes responsables des lésions A/E induites par les EHEC sont portés par le LEE. Chez les EHEC, le LEE mesure 43 Kb et contient 41 gènes répartis entre 5 opérons *LEE1* à *LEE5* (Figure 3).

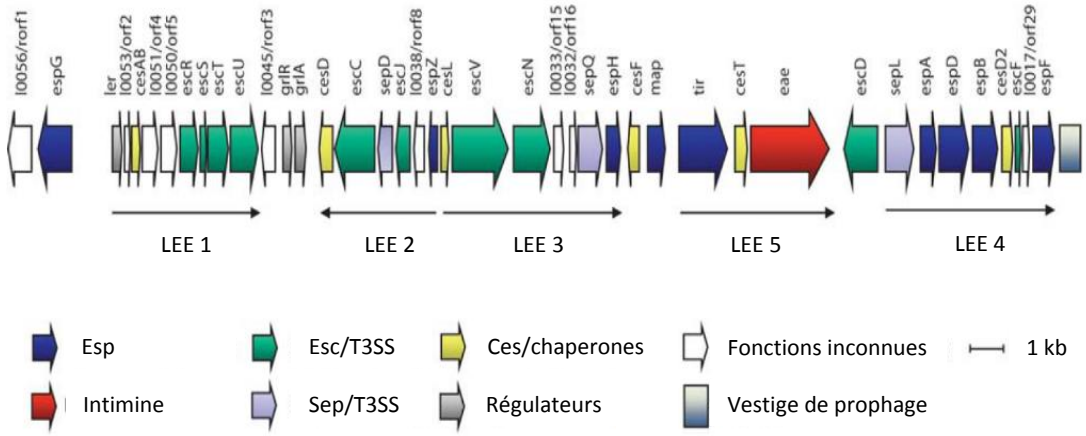

Figure 3 : Organisation génétique du LEE d'*E. coli* O157:H7. (D'après (Stevens *et al.* 2014))

Le LEE code notamment un système de sécrétion de type III, sous-type a (T3aSS) et des protéines effectrices dont l'intimine codée par le gène *eae* (*E. coli* attaching and effacing) et son récepteur Tir (*tir :* translocated intimin receptor) (Desvaux *et al.* 2006, Stevens *et al.* 2014). Le T3aSS est également connu sous le nom d'injectisome du fait de sa structure macromoléculaire permettant l'injection d'une cinquantaine d'effecteurs bactériens directement à l'intérieur de la cellule cible jouant ainsi le rôle de seringue moléculaire (Tobe *et al.* 2006, Gaytan *et al.* 2016) (Figure 4).

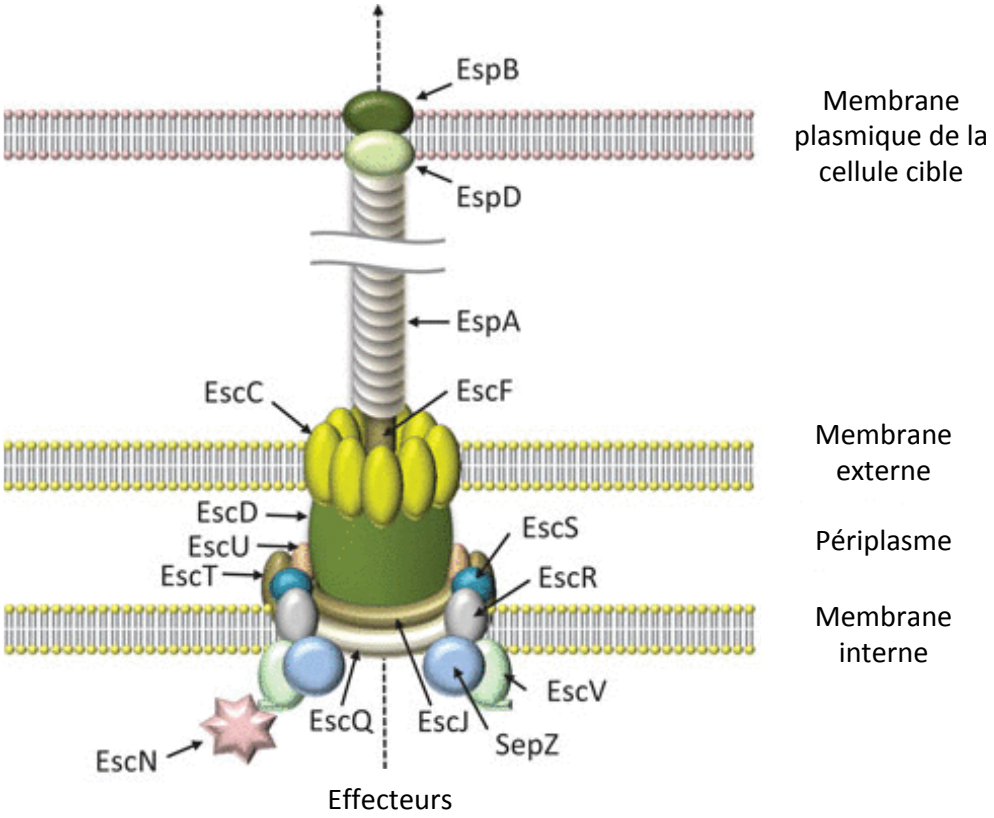

Figure 4 : Structure de l'injectisome d'*E. coli* O157:H7. Le T3aSS est divisé en 3 composantes principales : le corps basal, l'ATPase et les appendices extracellulaires. Le corps basal est composé de 3 anneaux étendus de la membrane interne à la membrane externe de la bactérie et connectés par une tige périplasmique. Le corps basal est associé à une ATPase cytoplasmique, codée par le gène *escN*, qui fournit l'énergie nécessaire pour assurer la sécrétion de protéines effectrices Esp. Enfin, les appendices extracellulaires comprennent un polymère formant une aiguille, associée à une structure filamenteuse creuse (filament EspA) qui permet la sécrétion de protéines formant le translocon (EspBD) dans la membrane plasmique de la cellule cible. (D'après (Stevens *et al.* 2014))

Les effecteurs injectés perturbent ainsi les processus cellulaires de la cellule hôte au profit de la bactérie (Stevens *et al.* 2014). Parmi ces effecteurs, Tir s'insère dans la membrane plasmique où il pourra être reconnu par l'intimine localisée à la surface de la bactérie, on parle alors d'adhésion intime. Les signaux déclenchés par cette interaction induisent une désorganisation du cytosquelette d'actine de l'entérocyte aboutissant à la formation d'un piédestal et à la destruction locale des microvillosités intestinales (Mellies *et al.* 2014) (Figure 5). Cependant, d'autres effecteurs non codés par le LEE mais sécrétés par le T3aSS sont nécessaires aux EHEC pour établir ces lésions A/E. C'est le cas de l'effecteur EspFU qui interagit avec Tir dans la cellule cible pour stimuler la polymérisation de l'actine et favoriser la formation du piédestal (Campellone *et al.* 2004).

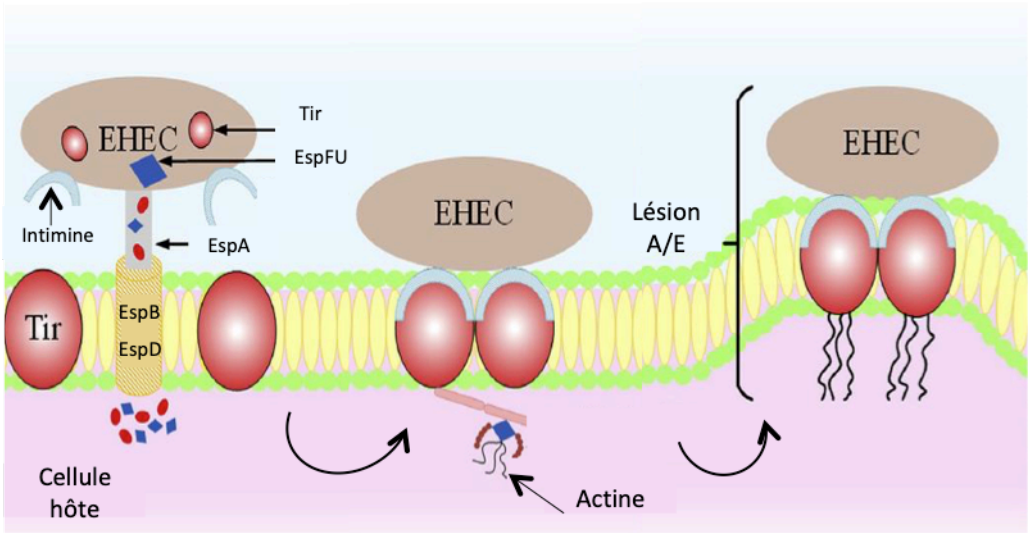

Figure 5 : Mécanisme de formation des lésions A/E par les EHEC sur les cellules épithéliales intestinales. (D'après (Saeedi *et al.* 2017))

La régulation complexe du LEE fait intervenir de nombreux acteurs et se fait aux niveaux transcriptionnels, traductionnels et post-traductionnels. Le régulateur transcriptionnel Ler (LEE-encoded regulator), dont le gène est codé sur l'opéron *LEE1*, est un acteur principal qui permet l'expression des 4 autres opérons en réponse à des signaux tels que la carence nutritionnelle ou le senseur du quorum (SQ) (Elliott *et al.* 2000, Mellies *et al.* 2014). De plus, l'assemblage du T3aSS et l'injection des effecteurs sont contrôlés spatio-temporellement de manière à échapper au système immunitaire de l'hôte. Cette régulation fait intervenir les protéines SepD et SepL (Secretion of EHEC proteins D and L) codées respectivement dans les opérons *LEE2* et *LEE4*. Ces protéines permettent la transition entre la sécrétion des translocateurs EspABD (*E. coli* secreted proteins A, B and D), nécessaires à l'assemblage du T3aSS, et la translocation des effecteurs dans la cellule hôte (Deng *et al.* 2005, Berger *et al.* 2012).

Autres facteurs de virulence des EHEC

En plus des deux facteurs de virulence principaux que sont les Stx et le LEE, les EHEC possèdent d'autres caractéristiques participant à leur pathogénicité. Tout d'abord d'un point de vue génomique, les souches EHEC O157:H7 possèdent un plasmide de virulence nommé pO157. Ce plasmide, dont la taille peut varier entre 92 et 104 Kb, est un plasmide de type F non conjugatif qui contient de nombreux éléments mobiles comme des transposons, des prophages, des séquences d'insertion ou des fragments d'autres plasmides bactériens (Perna *et al.* 2001). Une centaine de gènes a été identifiée sur ce plasmide dont une trentaine pourrait être impliquée

dans la virulence avec entre autres une hémolysine (Ehx), une sérine protéase (EspP), un système de sécrétion de type II (T2SS), une adhésine (ToxB), une catalase-péroxydase (KatP) ou une métalloprotéase (StcE) (Levine 1987, Lim *et al.* 2010). De plus, au sein même de leur chromosome, les EHEC possèdent des ilôts de pathogénicité appelés O-Islands (OI) qui les différencient des souches d'*E. coli* non pathogènes. La majorité de ces séquences semblent provenir de transferts horizontaux par l'intermédiaire de prophages (Herold *et al.* 2005). Parmi les 177 OI identifiés chez EHEC O157:H7 (Perna *et al.* 2001), certains regroupent des gènes codant pour des effecteurs du T3aSS nommés Nle (non LEE-encoded factor) (Tobe *et al.* 2006, Coombes *et al.* 2008).

Après ingestion, les EHEC transitent par l'estomac où ils doivent faire face au premier système de défense de l'hôte, à savoir l'acidité gastrique (pH compris entre 2 et 6). L'estomac constitue un milieu peu favorable au développement bactérien. Le suc gastrique est composé entre autres d'acide chlorhydrique responsable de la dénaturation des protéines en cassant les ponts disulfures et les liaisons hydrogènes ainsi que de lipases gastriques impliquées dans la digestion des lipides. Pour faire face à ce système de défense de l'hôte, les EHEC possèdent cinq systèmes de résistance à l'acidité gastrique : un système oxydatif, glutamate dépendant, arginine dépendant, lysine dépendant et ornithine dépendant (Price *et al.* 2004, Zhao *et al.* 2010). De plus, le SQ joue également un rôle important dans le contrôle de la tolérance à l'acidité des EHEC (Mellies *et al.* 2014).

I. 3. EHEC et STEC

Les EHEC dérivent d'EPEC qui auraient acquis les gènes codant les Stx responsables d'hémorragie et appartiennent donc à la catégorie des *E. coli* codant des Stx ou *E. coli* producteurs de Stx (STEC) (O'Loughlin *et al.* 2001, Monteiro *et al.* 2016). Ainsi, les EHEC sont des STEC responsables d'infections et donc isolés de cas cliniques. En revanche, si tous les EHEC sont des STEC, d'autres STEC peuvent être isolés de divers environnements et ne sont donc pas systématiquement des EHEC (Levine 1987). Des souches hybrides de STEC possèdent les facteurs de virulence caractéristiques d'autre pathotypes d'*E. coli* tels que les EAEC (Brzuszkiewicz *et al.* 2011) et les ETEC (Tozzoli *et al.* 2014, Nyholm *et al.* 2015, Nyholm *et al.* 2015). De plus, les souches possédant le LEE et donc capables d'établir des lésions A/E au niveau de l'iléon ou du côlon de l'Homme sont appelées AEEC (attaching and effacing *E. coli*). De ce fait, les souches EHEC qualifiées de typiques possèdent le LEE et

peuvent établir ce type de lésions alors que les EHEC atypiques ne possèdent pas le LEE (Nataro *et al.* 1998) (Figure 6).

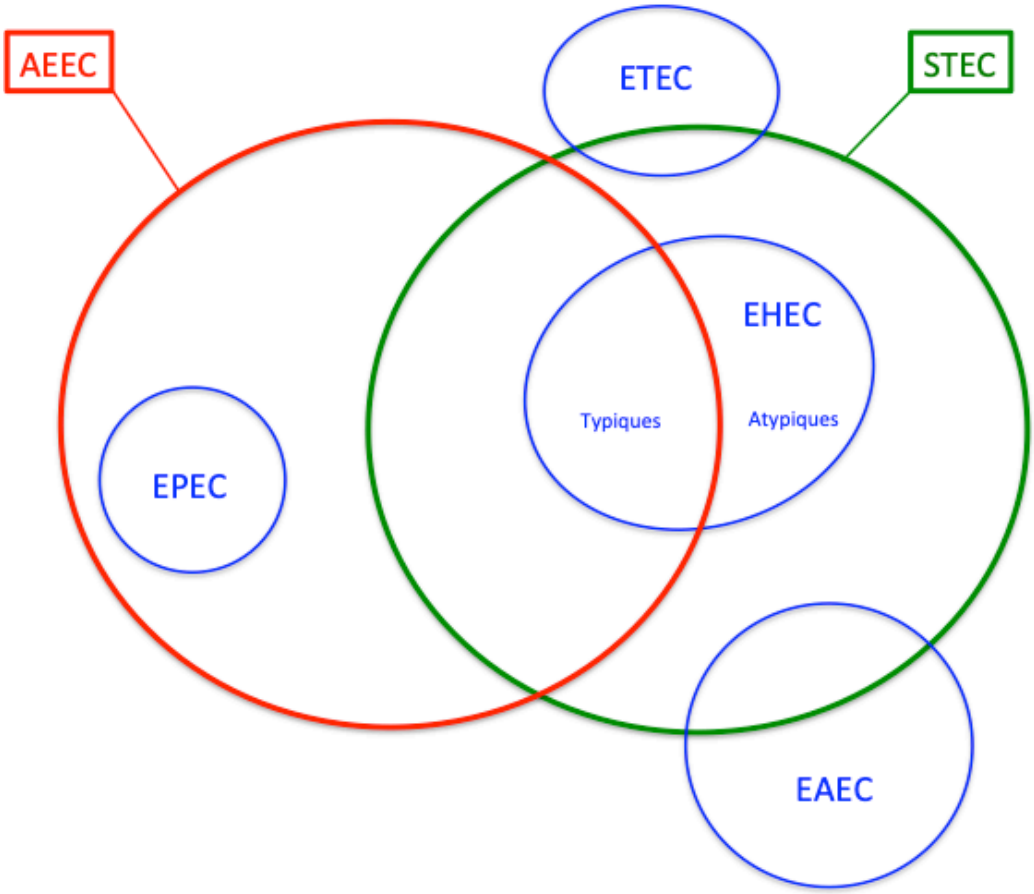

Figure 6 : Classification des EHEC à partir des facteurs de virulence associés aux souches d'*E. coli*. Les différents pathotypes d'*E. coli* (bleu) sont répartis selon leurs caractéristiques génotypiques. Les AEEC (rouges) possèdent le LEE et les STEC (vert) possèdent les gènes *stx*. (D'après (ANSES 2010)).

Il existe plus de 400 sérotypes de STEC parmi lesquels plus de 200 sont associés à des pathologies humaines. Les sérotypes des STEC peuvent également être divisés en 5 séropathotypes classés de A à E (qui ne reflètent en rien la classification en phylogroupes évoquée précédemment) selon leur incidence et la sévérité des signes cliniques provoqués par les souches (Karmali *et al.* 2003). Les séropathotypes A et B comprennent les sérotypes les plus virulents associés avec des épidémies et des cas de SHU. Les souches O157:H7 appartiennent au séropathotype A. Les souches associées avec des cas sporadiques de SHU appartiennent au séropathotype C. Le séropathotype D regroupe les souches ayant une faible incidence sur la santé humaine (diarrhées) et non associées à des cas de SHU. Enfin, le séropathotype E comprend les souches STEC non isolées de cas cliniques. L'une des explications de la différence de virulence entre ces séropathotypes est la présence ou l'absence d'OI dans le

génome des bactéries. Par exemple, l'OI-122 est présent chez toutes les souches du séropathotype A alors que qu'il est de moins en moins répandu dans les autres séropathotypes (Karmali *et al.* 2003). En 2013, l'EFSA (European Food Safety Authority) a pointé du doigt les limites de cette classification qui ne regroupe pas de manière exhaustive les sérotypes pathogènes (EFSA 2013).

I. 4. Epidémiologie des infections à EHEC

La surveillance des infections à EHEC a d'abord été basée sur la détection du sérotype O157:H7 mais a depuis été élargie aux sérotypes non-O157 qui sont également à l'origine d'infections sporadiques et sources d'épidémies. La surveillance et le contrôle des infections à EHEC représentent donc un objectif majeur de santé publique et chaque pays possède son propre système de surveillance. Cette surveillance est opérée en France par Santé publique France en collaboration avec les services hospitaliers. A l'échelle Européenne, c'est le système de surveillance des maladies et zoonoses d'origine alimentaires et liées à l'eau (FWD : Food and Waterborne Diseases and Zoonoses) coordonné par le centre européen de contrôle et de prévention des maladies (ECDC : European Center for Disease Prevention and Control) ainsi que l'EFSA. Enfin aux Etats-Unis, le réseau de surveillance active des maladies d'origine alimentaire (FoodNet : Foodborne Diseases Active Surveillance Network) a été mis en place par le centre de contrôle et de prévention des maladies (CDC : Center for Disease Control and Prevention) en collaboration avec le département de l'agriculture des Etats-Unis (USDA : United States department of Agriculture) et l'administration de l'alimentation et de la drogue (FDA : Food an Drug Administration).

Tableau 1 : Les sérogroupes principaux des EHEC et leurs séropathotypes. (D'après (Croxen *et al.* 2013) et (ANSES 2017)).

Sérogroupe	Antigène(s) H	Séropathotype	CH	SHU
O157	NM, H7	A	+	+
O26	NM, H11	B	+	+
O145	NM, H25, H28	B	+	+
O111	NM, H8	B	+	+
O103	NM, H2	B	+	+

CH : colite hémorragique; SHU : syndrome hémolytique et urémique

Si les cas sporadiques représentent la grande majorité des cas d'infections à EHEC, de nombreuses épidémies ont été reportées dans le monde, notamment dans les pays industrialisés. Pour information, les infections à STEC sont la 4e source de zoonoses humaines en Europe ces dernières années avec plus de 6000 cas d'infections détectés en 2017 (1,66 cas pour 100 000 personnes) et notamment 164 cas de SHU en France (1,40 pour 100 000 enfants de moins de 15 ans) (EFSA 2018, INVS 2018). En France, l'agence nationale de la sécurité sanitaire de l'alimentation, de l'environnement et du travail (ANSES) a identifié les cinq sérotypes d'EHEC les plus fréquemment associés à des épidémies et à des pathologies, classés selon leur importance : O157:H7, O26:H11, O145:H28, O111:H8 et O103:H2 (Tableau 1). L'ANSES attire de plus l'attention sur le sérotype O80:H2 émergent en France (ANSES 2017). Ces souches EHEC typiques majeures qualifiées d'hautement pathogènes (AFSSA 2008, AFSSA 2010) possèdent toutes les gènes *stx* et *eae*. Aux Etats-Unis, l'USDA a recensé les 7 sérogroupes d'EHEC majoritairement impliquées dans des pathologies sévères : O157, O26, O111, O103, O121, O45 et O145 (USDA 2012). Les souches appartenant à ces sérogroupes sont également positives pour les gènes *stx* et *eae*.

Le pays le plus touché par ces épidémies sont les Etats-Unis qui ont en connu plus d'une centaine depuis 1982 (Armstrong *et al.* 1996, Rangel *et al.* 2005). C'est à cette date que les EHEC ont étés décrits pour la première fois lors de deux épidémies à *E. coli* O157:H7 qui ont permis d'isoler la souche de référence EDL933. L'aliment en cause était la viande hachée peu cuite principalement retrouvée dans les hamburgers et ayant conduit à l'infection de 47 personnes (Riley *et al.* 1983). La plus grande épidémie répertoriée à ce jour a eu lieu au Japon

en 1996 suite à la consommation de pousses de radis crus contaminées par une souche de sérotype O157:H7 Sakaï : plus de 9000 malades, dont 101 cas de SHU et 12 décès, principalement des enfants ont été dénombrés (Mermin *et al.* 1999).

La dernière épidémie de grande ampleur a eu lieu en Allemagne en 2011 et impliquait des graines germées de fenugrec contaminées par le sérotype O104:H4 (Frank *et al.* 2011). Cette souche hybride possède les gènes de virulence caractéristiques des EAEC et a acquis les gènes codant pour la β-lactamase et pour la Stx2 des EHEC mais ne possède pas le LEE (Brzuszkiewicz *et al.* 2011, Rasko *et al.* 2011, Karch *et al.* 2012). Le bilan recense 3816 cas, 845 SHU et 54 morts. Ici, le SHU a frappé principalement des adultes. En juin 2011, une autre épidémie à EHEC O157:H7 survient en France dans le Nord Pas de Calais et en Picardie, suite à la consommation de steaks hachés contaminés : 31 cas, 21 SHU et 10 morts. L'épidémie la plus récente a eu lieu en France au printemps 2018, incriminant des reblochons au lait cru contaminés par des bactéries du sérogroupe O26. Parmi les 14 enfants de moins de 5 ans infectés, 11 développèrent un SHU et un en succomba (France 2018).

I. 5. Réservoirs et modes de transmission des EHEC

Le TD des bovins constitue le réservoir naturel principal des EHEC O157:H7 et non-O157 (notamment O5, O26, O103 ou O145) (Duffy *et al.* 2014). Si ces bactéries peuvent survivre dans le rumen (Chaucheyras-Durand *et al.* 2006, Chaucheyras-Durand *et al.* 2010, Free *et al.* 2012), elles transitent rapidement vers l'intestin grêle, le caecum puis le côlon et surtout la jonction recto-anale qui est leur site principal de colonisation et de prolifération (Grauke *et al.* 2002, Naylor *et al.* 2003, Lim *et al.* 2007). Les EHEC sont donc excrétés dans l'environnement par voie fécale (Omisakin *et al.* 2003). Cependant d'autres ruminants comme les ovins ou les caprins peuvent constituer un réservoir important d'EHEC O157 (La Ragione *et al.* 2009) et d'autres animaux peuvent être des hôtes accidentels comme les porcs (nombreux sérogroupes dont le O157) (Booher *et al.* 2002), les oiseaux (Pedersen 2007) ou les animaux de compagnie (Hogg *et al.* 2009). Si ces animaux sont dans la plupart des cas des porteurs asymptomatiques, probablement en raison de l'absence des récepteurs des Stx, certaines STEC peuvent également être pathogènes pour l'animal (Mainil *et al.* 2005). C'est par exemple le cas du sérogroupe O118 fréquemment associé à des cas de diarrhées chez le veau (Wieler *et al.* 1998, Beutin *et al.* 2000). De nombreux facteurs peuvent influer sur la prévalence des EHEC,

notamment l'espèce et l'âge de l'animal, le mode d'élevage, le régime alimentaire, la saison ou l'environnement (Zhao *et al.* 1995, Jenkins *et al.* 2002, Hussein 2007).

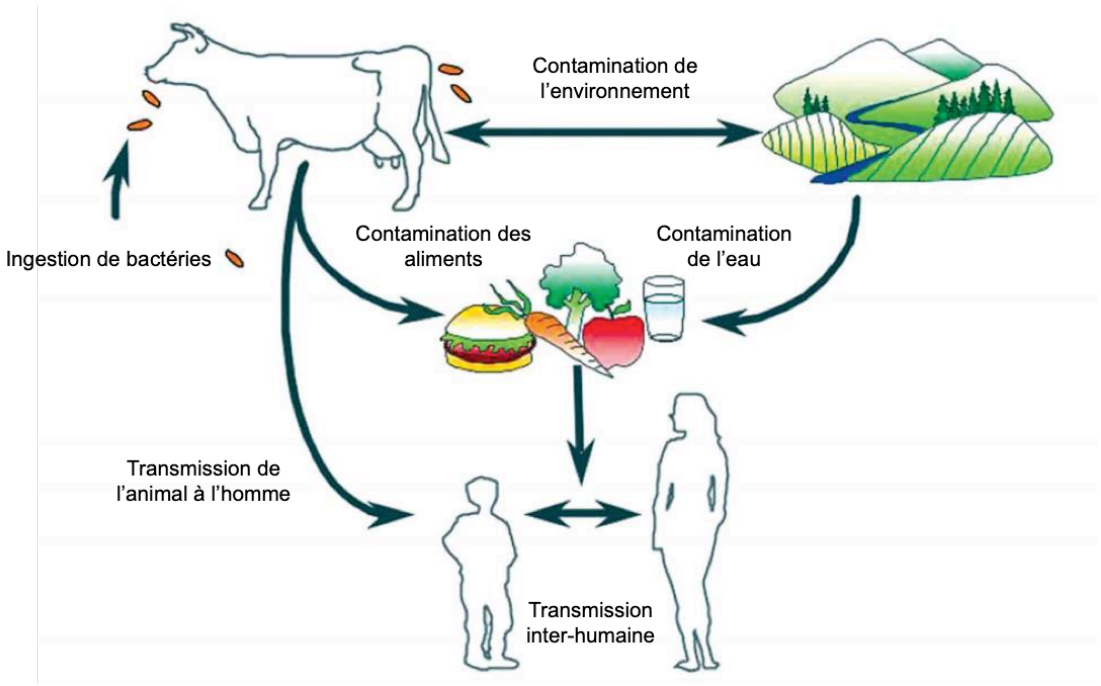

Figure 7 : Modes de transmission des EHEC à l'Homme. (D'après (Fairbrother *et al.* 2006))

La source majeure d'infection à EHEC pour l'Homme est ainsi d'origine alimentaire (Figure 7) (EFSA 2016). Les aliments responsables d'infections proviennent des animaux porteurs ou en contact direct ou indirect avec les fèces de ces porteurs tels que la viande et les produits laitiers (laits et fromages au lait cru) non pasteurisés (Kintz *et al.* 2017). Parmi ces aliments, le plus fréquemment associé à des cas d'épidémies en Europe reste la viande de bœuf hachée consommée insuffisamment cuite. L'environnement agro-alimentaire répond à des règles d'hygiène et de sécurité précises fixées par le règlement (CE) no. 853/2004 de l'Union européenne rendant très faible le risque de contamination de la viande lors de la découpe (UE 2004). Le respect de ces règles exclut le risque de contamination par contact direct avec le contenu du TD des bovins lors d'un accident d'éviscération, puisqu'après examen vétérinaire ces carcasses sont alors exclues du circuit destiné à la consommation humaine. Néanmoins, les cuirs sont reconnus comme une source principale de contamination des carcasses par les EHEC (Moxley *et al.* 2014). Après excrétion, les EHEC peuvent également être disséminés dans l'eau d'irrigation ou le traitement des sols par effluents agricoles (Fremaux *et al.* 2008). Comme dans

les fèces, les EHEC ont la capacité de survivre pendant plusieurs semaines dans l'eau, le sol ou sur les végétaux (Wang *et al.* 1998, Bolton *et al.* 1999, Fremaux *et al.* 2008). De ce fait, d'autres produits peuvent être à l'origine d'infections tels que les produits végétaux ou eaux contaminés. L'Homme peut également être contaminé par contact direct avec l'animal ou ses fèces (Caprioli *et al.* 2005). Enfin la contamination interhumaine par voie oro-fécale est un mode de transmission important, notamment en milieu familial ou dans les collectivités telles que les crèches, les maisons de retraites et les hôpitaux (AFSSA 2003, Rangel *et al.* 2005).

Chapitre II : Déterminants moléculaires de surface impliqués dans la colonisation des EHEC

La capacité des EHEC à adhérer à des surfaces aussi bien biotiques qu'abiotiques leur permet de coloniser leur environnement et favorise leur transmission dans la chaine alimentaire (Figure 7). Chez les bactéries, la colonisation de surfaces implique différents déterminants moléculaires présents à la surface des cellules (Chagnot *et al.* 2012), en particulier les exopolysaccharides (EPS), l'ADN extracellulaire (eADN) et les protéines.

Les facteurs de colonisation les plus nombreux sont les déterminants protéiques. Les protéines de surfaces peuvent donc permettre à la bactérie de coloniser des surfaces aussi bien abiotiques que biotiques, telles que la matrice extracellulaire (ECM). L'ECM est un composant ubiquiste des tissus animaux et est donc retrouvée aussi bien dans le TD bovin, l'aliment et le tractus intestinal humain. Ce complexe protéique permet le soutien et le lien entre les cellules et les tissus et peut être composé de protéines fibrillaires ou de protéoglycanes. Si la composition de l'ECM peut varier selon les tissus, son organisation reste la même et peut se résumer à des protéines immergées dans un gel constitué de protéoglycanes, le tout formant une matrice complexe et structurée (Vakonakis *et al.* 2007). Les constituants principaux sont le collagène, l'élastine, la laminine et la fibronectine (Chagnot *et al.* 2012). L'ECM pourrait ainsi servir de substrat à l'adhésion des EHEC pour coloniser les différents niveaux de la chaine alimentaire. A ce titre, les protéines de surfaces peuvent donc appartenir à la famille des protéines MSCRAMM (microbial surface components recognising adhesive matrix molecule) (Chagnot *et al.* 2012).

Pour être présentes en surface, ces protéines doivent être transportées au travers de l'enveloppe cellulaire, de l'intérieur vers l'extérieur de la cellule, via des systèmes de sécrétion. Parmi les 9 types de systèmes de sécrétion identifiés chez les bactéries didermes-LPS numérotés de I à IX (T1SS à T9SS), les EHEC en possèdent 8 (Monteiro *et al.* 2016). Ces systèmes permettent ainsi la sécrétion de certaines structures fimbrillaires ou d'adhésines afimbrillaires caractérisées dans le processus de colonisation des EHEC. Brièvement, le T1SS est impliqué dans la synthèse et l'assemblage de la dispersine, le T2SS permet la formation de pili de type 4 (T4P), le T3SS assemble les composants de l'injectisome et du flagelle, le T4SS joue un rôle dans la conjugaison bactérienne via la formation de pili conjugatifs, le T5SS sécrète une grande variété de protéines appelées autotransporteurs (AT), le T6SS est impliqué dans la formation

des tubes HCP (haemolysin-coregulated protein), le T7SS permet la synthèse et l'assemblage d'un grand nombre de pili et T8SS est responsable de la formation des curli (Monteiro *et al.* 2016) (Figure 8).

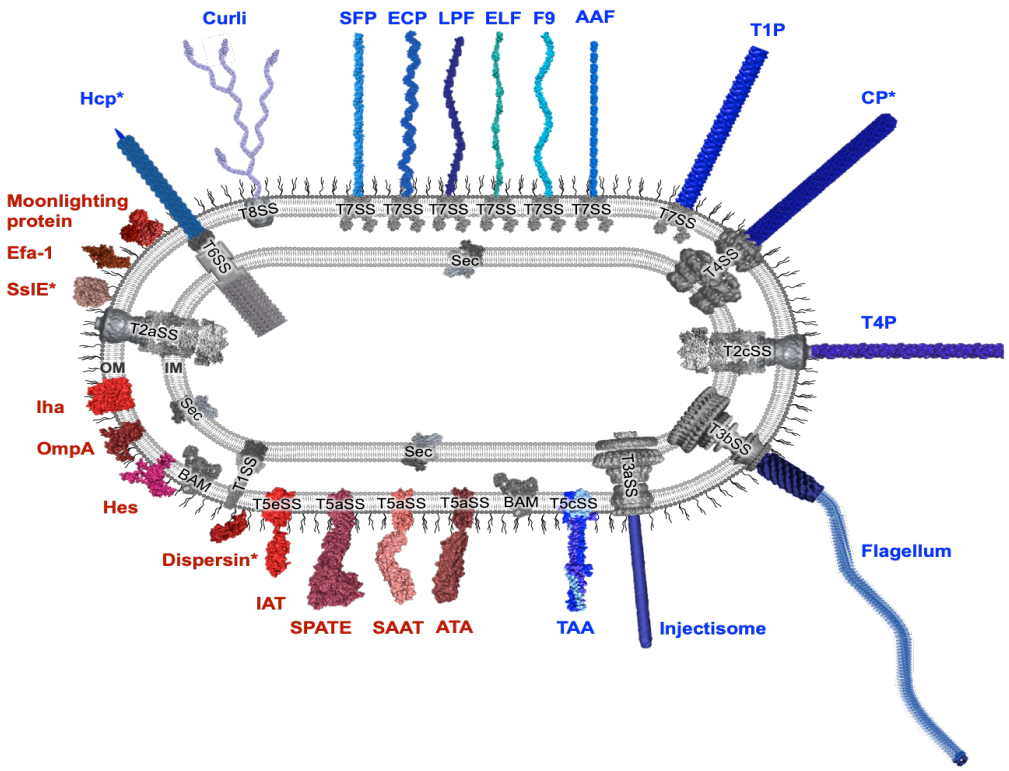

Figure 8 : Représentation schématique des déterminants moléculaires protéiques potentiellement présents à la surface de la cellule bactérienne des EHEC. Les adhésines afimbrillaires sont représentées en nuances de rouge alors que les structures fimbrillaires sont représentées en nuances de bleu. Les systèmes de sécrétion et les machineries d'exportation sont représentés en gris. Les astérisques représentent les déterminants moléculaires présents chez les EHEC mais pas encore caractérisés dans le processus de colonisation.

II. 1. Les exopolysacharides (EPS)

Les EPS sont parmi les composants principaux de la matrice d'exopolymères des biofilms formés par *E. coli* (Beloin *et al.* 2008). Les EHEC peuvent synthétiser différents types d'EPS, à savoir le LPS, le PNAG (poly-β-1,6-N-acetyl-D-glucosamine), l'acide colanique et la cellulose.

Le LPS est considéré comme une endotoxine en raison de la présence d'un composé toxique dans leur structure : le lipide A. Ce dernier est lié à la région centrale du LPS, elle-même liée à l'antigène O à la surface de la cellule (Raetz *et al.* 2002). La région centrale peut arborer 5 structures différentes, R1, R2, R3, R4 et K12. La structure R3 est la plus associée

avec les souches EHEC (Currie *et al.* 1999, Amor *et al.* 2000, Chart *et al.* 2002). Le LPS permet ainsi à EHEC O157:H7, grâce à son antigène O, de coloniser des hôtes animaux tels que la souris, le lapin et le bovin (Ho *et al.* 2007, Sheng *et al.* 2008).

Le PNAG est un polymère de polysaccharides, codé par l'opéron *pgaABCD*, qui reste attaché à la surface de la cellule bactérienne et participe à la formation de biofilms (Wang *et al.* 2004, Itoh *et al.* 2008). Chez les EHEC, cet EPS prend part dans le processus infectieux et joue également un rôle dans la formation de biofilms sur des germes et des racines de tomates (Matthysse *et al.* 2008, Al Safadi *et al.* 2012).

L'acide colanique est un polymère de glucose, galactose, fucose et acide glucuronique chargé négativement et produit par la majorité des souches d'*E. coli* (Obadia *et al.* 2007). Si sa contribution exacte à la formation de biofilms est encore mal connue, il joue un rôle dans la survie d'EHEC O157:H7 dans des conditions gastrointestinales simulées (Mao *et al.* 2006).

La cellulose est un homopolysaccharde linéaire composé d'unités D-glucopyranose, généralement associé au règne végétal mais également présent chez certaines bactéries où il participe à la formation de biofilms (Solano *et al.* 2002, Ude *et al.* 2006). Chez EHEC O157:H7, la cellulose est impliquée dans la colonisation de végétaux tels que les germes de luzerne (Matthysse *et al.* 2008).

II. 2. L'ADN extracellulaire

Dans la matrice d'exopolymères du biofilm, l'ADN extracellulaire sert principalement de composant structurel mais peut également contribuer au gradient de cations, induire une résistance à des antibiotiques et faciliter le transfert horizontal de gènes (Bockelmann *et al.* 2006, Palchevskiy *et al.* 2006, Sanchez-Torres *et al.* 2011). Si son implication dans la maturation du biofilm a déjà été démontrée chez *E. coli* K12 (Xi *et al.* 2010, Nakao *et al.* 2012), son rôle chez les DEC reste à élucider.

II. 3. Le système de sécrétion de Type 1 (T1SS)

La dispersine est une protéine d'anti-agrégation (Aap : anti-aggregation protein) impliquée dans la propagation des cellules bactériennes le long de la muqueuse intestinale de l'hôte (Sheikh *et al.* 2002). Cette protéine contribue notamment à l'adhésion et la colonisation des EAEC en empêchant l'hyper-agrégation et l'effondrement des AAF (aggregative adherence

fimbriae). La dispersine se retrouve exposée à la surface de la bactérie par liaison aux LPS par interactions électrostatiques (Velarde JJ 2007). Ce système de sécrétion et les protéines sécrétées qui y sont apparentées sont codées dans le locus *aat* (aggregative ABC transporter) du plasmide de virulence pAA (plasmide d'adhésion agrégative) de certaines EAEC (Nishi J 2003). La dispersine est également présente chez certaines STEC, comme la souche hybride O104:H4 ainsi que chez des souches commensales (Monteiro *et al.* 2009, Muniesa *et al.* 2012).

II. 4. Le système de sécrétion de Type II (T2SS)

Après leur exportation dans le périplasme via la translocase Sec ou le système Tat (twin-arginine translocation pathway), le T2SS ou SDP (secretion-dependant pathway) permet la sécrétion de protéines solubles extracellulaires *via* les sous-types a (T2aSS) et b (T2bSS) ou l'assemblage de pili via le sous-type c (T2cSS) (Ayers *et al.* 2010, Durand *et al.* 2011, Chagnot *et al.* 2013).

Précédemment connue sous le nom de YghJ (Yang *et al.* 2007, Iguchi *et al.* 2009), SslE (secreted and surface-associated lipoprotein from *E. coli*) a récemment été décrite comme une mucinase d'*E. coli* grâce à son motif zinc metallopeptidase (Luo *et al.* 2014, Nesta *et al.* 2014). Cette protéine est sécrétée par un T2aSS mais les mécanismes moléculaires de sa maturation en tant que lipoprotéine de surface restent méconnus. Cette protéine peut être divisée en deux principaux variants dont les gènes peuvent être retrouvés chez les EPEC, ETEC et EHEC (Moriel *et al.* 2010, Decanio *et al.* 2013). Des anticorps dirigés contre le variant I (de la souche ExPEC IHE3034) sont capables d'inhiber la translocation de souches d'*E. coli* à travers une matrice de mucine. De plus, l'immunisation d'animaux avec SslE I réduit significativement la colonisation intestinale par des souches de différents pathotypes exprimant Ssle II (Nesta *et al.* 2014). Ces observations font de SslE un acteur clé de la colonisation des surfaces mucosales par *E. coli* et un candidat potentiel pour le développement d'un vaccin contre les DEC (Naili *et al.* 2016, Naili *et al.* 2017).

Les T4P sont sécrétés et assemblés par le T2cSS (Chagnot *et al.* 2013). Chez les souches *E. coli* pathogènes, ces pili sont connus pour être associés à la virulence en favorisant le processus de colonisation (Craig *et al.* 2004). Plus particulièrement chez les EHEC, les T4P sont appelés HCP (haemorrhagic coli pili) et composés de sous-unités de 19 kDa codées par le gène chromosomique *hcpA*. L'inactivation de ce gène chez la souche EHEC O157:H7 réduit l'adhésion aux cellules épithéliales bovines et humaines (Xicohtencatl-Cortes *et al.* 2009). De

plus, les pili HCP peuvent se lier à la fibronectine et la laminine, s'agglutiner sur des érythrocytes de lapin, aider à la formation de biofilms et permettre la motilité saccadée (twitching motility) (Xicohtencatl-Cortes *et al.* 2009).

II. 5. Le système de sécrétion de Type III (T3SS)

Le T3SS est subdivisé en deux sous-types : le sous-type a (ou injectisome) décrit précédemment et le sous-type b (T3bSS). En plus de son rôle dans la virulence des EHEC, l'injectisome est également impliqué dans l'adhésion aux plantes avec un tropisme particulier pour les stomates (Shaw RK 2008, Berger CN 2010, Croxen *et al.* 2013).

Le T3bSS est impliqué dans la synthèse et l'assemblage des composants des flagelles, à savoir le crochet, la coiffe et la flagelline (Macnab 2004, Desvaux *et al.* 2006, Altegoer *et al.* 2018, Galan *et al.* 2018). La flagellation, ubiquiste chez *E. coli*, est péritriche avec un nombre de flagelles variant entre 6 et 10 (Macnab 1987). En mouvement, les flagelles s'unissent pour former une structure rigide en queue de cheval au pôle de la cellule (Bray 2001, DiLuzio *et al.* 2005). Le flagelle est principalement impliqué dans la motilité bactérienne par nage (swimming) et/ou essaimage (swarming) selon les conditions environnementales (Turner *et al.* 2010). La motilité active permet donc à la bactérie de surmonter les forces de répulsions électrostatiques et hydrodynamiques et de s'approcher de la surface cible et initier le processus de colonisation (Donlan 2002, Haiko *et al.* 2013). Chez les EHEC, la flagelline FliC favorise l'attachement initial et l'adhésion aux cellules épithéliales intestinales bovines (Mahajan *et al.* 2009) ainsi que la formation de biofilms sur des surfaces abiotiques et des feuilles d'épinards (Xicohtencatl-Cortes *et al.* 2009, Vikram *et al.* 2013, Nagy *et al.* 2015). De plus, des tests de vaccination ciblant FliC ont montré une diminution significative de la colonisation des EHEC chez le bovin, soulignant l'importance du flagelle dans la colonisation de l'intestin (McNeilly *et al.* 2008).

II. 6. Le système de sécrétion de type IV (T4SS)

La conjugaison bactérienne est un processus bien connu permettant notamment le transfert horizontal de facteurs de virulence ou de facteurs d'adhésion (Manwaring *et al.* 1999, Mazel *et al.* 1999, Llosa *et al.* 2002, Sorensen *et al.* 2005). Le T4SS contient la machinerie de conjugaison homologue à celle codée par l'opéron *tra* (transfert) localisé sur le plasmide conjugatif F(Manwaring *et al.* 1999, Lawley *et al.* 2003, Christie *et al.* 2014). Ainsi, les pili

conjugatifs permettent le transfert de nucléoprotéines depuis une cellule bactérienne donneuse (arborant le plasmide F) et une cellule bactérienne receveuse via le T4SS (Lawley *et al.* 2003). Ce transfert d'information génétique est d'autant plus facilité pendant le mode de vie sessile des bactéries du fait de la proximité physique entre celles-ci (Hausner *et al.* 1999, Molin *et al.* 2003, Maeda *et al.* 2006). Chez *E. coli* K12, les pili conjugatifs contribuent à la formation de biofilms en favorisant l'adhésion des cellules bactériennes (Ghigo 2001, May *et al.* 2008, May *et al.* 2011). Dans le biofilm, ce phénomène est amplifié car les bactéries possédant le plasmide conjugatif F augmentent le nombre de trans-conjugants. De ce fait l'adhésion et la formation de biofilms sont donc favorisés chez des bactéries à faible capacité de colonisation (Pereira *et al.* 2010). Un nouveau plasmide conjugatif nommé pO157-Sal codant un T4SS a été identifié chez la souche EHEC O157:H7 Xuzhou. Cependant, sa contribution au processus de colonisation n'a toujours pas été étudiée (Wang *et al.* 2011, Zhao *et al.* 2013).

II. 7. Le système de sécrétion de type V (T5SS)

Le T5SS est subdivisé en 5 sous-types (a à e) : (i) les AT (T5aSS), (ii) le système de sécrétion à deux partenaires (T5bSS), (iii) les AT trimériques (T5cSS), (iv) les AT hybrides (T5dSS), et (v) les AT inversés (T5eSS). Les protéines empruntant ces systèmes sont exportées dans le périplasme via la translocase Sec et sécrétées via un pore en tonneau β au travers de la membrane externe (Desvaux *et al.* 2003, Desvaux *et al.* 2004, Leo *et al.* 2012). Chez les EHEC, certaines protéines sécrétées via les T5aSS, T5cSS et T5eSS peuvent être impliquées dans le processus de colonisation.

II. 7. A. Les AT (autotransporters)

Les AT partagent des caractéristiques structurales communes à savoir (i) un peptide signal (SP) N-terminal qui adresse la protéine à la membrane interne (ii) un domaine passager N-proximal en hélice β, possédant le motif fonctionnel de la protéine (iii) un domaine autochaperon (AC) impliqué dans le repliement du domaine passager et (iv) un domaine translocateur C-terminal en tonneau β transmembranaire, permettant la sécrétion du domaine passager au travers de la membrane externe (Henderson *et al.* 1999, Desvaux *et al.* 2003, Desvaux *et al.* 2004, Leo *et al.* 2012, Drobnak *et al.* 2015). Les fonctions associées aux AT sont assez variées comme par exemple les AT associés à la vacuolisation (VAT), les AT phosphatase/hydrolase (PHAT) ou les AT lipase/esterase (LEAT) (Henderson *et al.* 2004,

Rojas-Lopez *et al.* 2017). En tant que facteurs de colonisation, les ATA (adhesins autotransporters), les SAAT (self-associating autotransporters) et les SPATE (serine-protease autotransporters of *Enterobacteriaceae*) doivent être pris en considération. Sur le plan de la prévalence, la distribution des AT est plutôt corrélée avec le groupe phylogénétique qu'avec le pathotype car la plupart d'entre eux peuvent être retrouvés chez les InPEC, ExPEC ou *E. coli* commensaux. A l'inverse, seul un petit nombre d'AT est spécifique d'un pathotype *sensu stricto* (Zude *et al.* 2014).

Les ATA (adhesins autotransporters)

Les ATA permettent l'adhésion directe à des supports abiotiques tels que le verre, la céramique en plastique ou inoxydable et/ou des surfaces biotiques comme des cellules de mammifères ou des protéines de l'ECM (Vo *et al.* 2017).

Parmi les protéines Eha (EHEC autotransporters) identifiées chez les EHEC (Wells *et al.* 2008), EhaB a été montrée comme favorisant l'adhésion des bactéries à la laminine et au collagène I (Wells *et al.* 2008, Wells *et al.* 2009) alors que EhaJ adhère spécifiquement à la fibronectine, au fibrinogène et aux collagènes II, III et IV (Easton *et al.* 2011). A la proximité immédiate de *ehaJ*, *egtA* code une glycosyltransferase. Ainsi la glycosylation de EhaJ provoque la formation de biofilms mais ne semble pas importante pour l'adhésion aux protéines de l'ECM (Easton *et al.* 2011). EhaD est impliquée dans la formation de biofilms mais son rôle exact dans ce phénomène ainsi que sa contribution à l'adhésion restent à déterminer (Wells *et al.* 2008). De la même manière, la contribution de EhaC dans le processus de colonisation de surface est encore mal connu (Wells *et al.* 2008).

Les SAAT (self-associating autotransporters)

Les SAAT ont la capacité de se reconnaître entre eux et d'entrainer une autoagrégation des cellules bactériennes (Klemm 2006). Les 3 principales familles de SAAT retrouvées chez *E. coli* sont l'Ag43 (Antigen 43), AIDA-I (adhesin involved in diffuse adherence phenotype) et TibA (toxigenic invasion locus b) (Trunk *et al.* 2018).

Le SAAT ayant suscité le plus d'intérêt est sans aucun doute l'Ag43 dont la plupart des informations proviennent de travaux réalisés sur la souche *E. coli* K12 (van der Woude *et al.* 2008). Cette protéine sera présentée plus en plus en détails dans le chapitre suivant. Brièvement, chez *E. coli* K12, l'Ag43 est responsable d'autoagrégation, d'adhésion et de formation de biofilms (Kjaergaard *et al.* 2000, Sherlock *et al.* 2006, de Luna *et al.* 2008). L'Ag43 est retrouvé

chez quasiment toutes les souches d'*E. coli* et notamment chez EHEC O157:H7 où il est impliqué dans l'autoagrégation et la formation de biofilms (Torres *et al.* 2002, van der Woude *et al.* 2008).

Chez les EPEC, AIDA-I permet l'autoagrégation, l'adhésion diffuse des bactéries à des cellules épithéliales et la formation de biofilms (Benz *et al.* 1989, Sherlock *et al.* 2006). Ce SAAT est présent chez les STEC d'origines porcines, mais chez EHEC O157:H7 il n'est pas impliqué dans l'adhésion à des cellules épithéliales. Plus récemment, EhaA a été identifié chez les EHEC comme étant un nouveau SAAT et possède la capacité d'autoagréger, d'adhérer à des cellules épithéliales bovines et de former des biofilms (Wells *et al.* 2008).

Les SPATE (serine-protease autotransporters of Enterobacteriaceae)

En tant que sérine protéase, la fonction première des SPATE est la dégradation de protéines diverses et variées comme la mucine ou l'hémoglobine. Toutefois, les SPATE peuvent aussi contribuer à la virulence des bactéries via leur effet cytotoxique et même être impliqués dans les processus de colonisation de surface (Dautin 2010). Ainsi, EspP (extracellular serine protease plasmid encoded) contribue à l'adhérence d'EHEC O157:H7 à des cellules épithéliales intestinales, à la colonisation de l'intestin bovin et à la formation de biofilms (Dziva *et al.* 2007, Puttamreddy *et al.* 2010, Farfan *et al.* 2012). EspP, également connu sous le nom de PssA (protease secreted by STEC), favorise aussi la colonisation de l'intestin des veaux par STEC O26 (van Diemen *et al.* 2005). A la surface de la bactérie, les domaines passagers d'EspP s'auto-assemblent pour former des structures supramoléculaires appelées cordes (rod). En plus de l'activité cytotoxique, ces cordes possèdent de fortes propriétés adhésives sur les cellules hôtes et peuvent servir de substrat pour l'adhésion bactérienne et la formation de biofilms (Xicohtencatl-Cortes *et al.* 2010).

Chez les EAEC, Pic (protein involved in colonisation) participe à l'adhésion à la mucine et à sa dégradation ainsi qu'à la colonisation intestinale et à la formation de biofilms bactéries-mucus (Gutierrez-Jimenez *et al.* 2008, Navarro-Garcia *et al.* 2011, Andrade *et al.* 2017). Pic est également exprimée par la souche hybride *E. coli* O104:H4 mais son rôle dans le processus de colonisation n'a toujours pas été caractérisé (Abreu *et al.* 2015, Abreu *et al.* 2016).

II. 7. B. Les TAA (trimeric autotransporters)

Les TAA correspondent à des AT assemblés après homotrimérisation (Cotter *et al.* 2005). Tous les TAA identifiés à ce jour sont exposés à la surface de la bactérie et sont impliqués dans l'adhésion (Linke *et al.* 2006, Szczesny *et al.* 2008). Lorsqu'il est surexprimé chez *E. coli* K12, EhaG d'*E. coli* O157:H7 est impliqué dans le phénomène d'autoagrégation conduisant à la formation de biofilms résistants. De plus, ce TAA joue un rôle dans l'adhésion aux cellules épithéliales intestinales ainsi qu'à une grande variété de protéines de l'ECM : fibrinogène, laminine, fibronectine, collagènes I, II, III et IV. Les gènes codant EhaG sont en outre retrouvés chez d'autres DEC et sont prévalents chez les STEC (Totsika *et al.* 2012, Zude *et al.* 2014).

Comme leur nom l'indique, les Eib (*E. coli* immunoglobuin-binding proteins) ont à l'origine été caractérisées pour leur capacité à se lier à la fraction immunoglobuline de l'IgA et de l'IgG (Sandt *et al.* 2000, Leo *et al.* 2009). Parmi les 7 Eibs identifiées jusqu'à présent, EibG est impliquée dans l'adhésion des STEC LEE-négatives O91 aux cellules épithéliales selon une configuration particulière appelée CLA (chain-like adhesion) (Lu *et al.* 2006). Plus précisément, la CLA correspond à la formation de longues chaînes d'agrégats bactériens sur des cellules épithéliales intestinales humaines et bovines. Les gènes codant EibG sont subdivisés en 3 sous-types : EibG-α et EibG-β impliqués dans une CLA classique et EibG-γ responsable d'une CLA atypique composée de chaines et d'agrégats plus petits (Merkel *et al.* 2010). De son côté, EibD a été montrée comme induisant l'autoagrégation des bactéries et la formation de biofilms (Leo *et al.* 2011). Au regard de leurs similarités structurelles, les autres Eibs pourraient posséder des fonctions biologiques semblables, mais des confirmations expérimentales restent à effectuer.

Certaines STEC LEE-négatives, dont des EHEC, possèdent également un TAA nommé Saa (STEC autoagglutination adhesin) jouant un rôle dans l'adhésion aux cellules épithéliales intestinales (Paton *et al.* 2001, Jenkins *et al.* 2003). Ces bactéries possèdent aussi le TAA Sab (STEC-autotransporter mediating biofilm formation) contribuant à l'adhésion diffuse aux cellules épithéliales et à la formation de biofilms sur des surfaces abiotiques comme le polystyrène (Paton *et al.* 2001, Jenkins *et al.* 2003, Herold *et al.* 2009, Farfan *et al.* 2012). La présence de ces molécules seulement chez les STEC LEE-négatives pourrait constituer un mode d'adhésion alternatif indépendant de l'établissement de lésions A/E (Herold *et al.* 2009).

II. 7. C. Les IAT (inverted autotransporters)

L'organisation modulaire des IAT est à l'exact opposé de celle des AT classiques puisque leur domaine translocateur est localisé en N-terminal tandis que le domaine passager est en C-terminal (Tsai *et al.* 2010, Oberhettinger *et al.* 2012). L'intimine, des EHEC et des EPEC, est le membre prototypique des IAT (Nataro *et al.* 1998, Leo *et al.* 2015). En plus de son rôle dans les lésions A/E, l'intimine peut contribuer à la colonisation de l'intestin d'une manière indépendante de Tir (Mallick *et al.* 2012). Enfin, cette protéine peut également se lier à des récepteurs alternatifs de l'hôte tels que les intégrines β-1 ou la nucléoline (Liu *et al.* 1999, Leo *et al.* 2015).

L'IAT FdeC (factor adherence of *E. coli*) est très répandu chez *E. coli* et peut être retrouvé chez tous les pathotypes d'InPEC (Nesta *et al.* 2012, Easton *et al.* 2014). Chez EHEC 026:H11, FdeC prend part à la formation de biofilms (Easton *et al.* 2014). L'IAT YeeJ, caractérisé chez *E. coli* K12 comme participant à la formation de biofilms, a été amplifié chez certaines souches EHEC, sans pour autant connaître son rôle exact dans la colonisation (Martinez-Gil *et al.* 2017).

II. 8. Le système de sécrétion de Type VI (T6SS)

Chez les EAEC, les tubes HCP formés par le T6SS sont suggérés comme étant important pour la formation de biofilms (Aschtgen *et al.* 2008). Des analyses bioinformatiques ont identifié plus de 10 orthologues du T6SS chez les EHEC, EPEC et EAEC (Shrivastava S 2008, Lloyd *et al.* 2009, Aschtgen *et al.* 2010, Moriel *et al.* 2010). Ceci suggère que ces composants pourraient être cruciaux dans l'agrégation bactérienne à la surface de la cellule hôte (Dudley *et al.* 2006). Toutefois, des études complémentaires sont nécessaires pour déterminer le rôle exact et les mécanismes moléculaires impliqués dans le processus de colonisation de surfaces par les HCP ou le T6SS (Monteiro *et al.* 2016).

II. 9. Le système de sécrétion de Type VII (T7SS)

Chez les bactéries didermes-LPS, le T7SS correspond à la voie chaperon-escorteur (chaperon-usher pathway) impliquée dans la sécrétion et l'assemblage d'un grand nombre de pili (Desvaux *et al.* 2009). Les LPF (long polar fimbriae) sont parmi les pili les mieux caractérisés chez les STEC. Ils sont codés par deux opérons *lpf1* et *lpf2* situés sur les OI O141

et O154 d'*E. coli* O157:H7, respectivement (Perna *et al.* 2001, Doughty *et al.* 2002) et peuvent également être retrouvés chez des souches EHEC atypiques (Doughty *et al.* 2002, Wurpel *et al.* 2013). Chez EHEC O157:H7, les LPF peuvent se lier à la fibronectine, la laminine, le collagène IV et au follicule associé à l'épithélium des plaques de Peyer de l'Homme (Fitzhenry *et al.* 2006, Farfan *et al.* 2012, McWilliams *et al.* 2014). Ils contribuent également à l'adhésion des EHEC à l'épithélium intestinal et peuvent même influer sur le tropisme bactérien envers différents tissus (Torres *et al.* 2008, Farfan *et al.* 2011). Récemment, il a été montré que des isolats STEC positifs pour *lpf2* formaient plus de biofilms que les isolats *lpf2*-négatifs (Vogeleer *et al.* 2015). Chez la souche O104:H4, LpfA contribue à l'adhésion à des cellules épithéliales intestinales, à la formation de biofilms et à la colonisation de caecum et du gros intestin (Ross *et al.* 2015).

Les pili F9, SFP (sorbitol fermenting protein) ou encore ELF (*E. coli* laminin binding fimbriae) sont spécifiques des EHEC. En effet, les pili F9 sont codés par des gènes localisés sur l'OI O161 des EHEC (Low *et al.* 2006, Wurpel *et al.* 2013). Chez la souche O157:H7, ces pili sont impliqués dans la colonisation de l'épithélium intestinal bovin ainsi que dans l'adhésion à la fibronectine (Low *et al.* 2006). En revanche, les bactéries mutées pour la sous-unité principale du pilus sont toujours capables de coloniser le rectum terminal suggérant que l'adhésine n'est pas exclusivement responsable du tropisme rectal et qu'elle pourrait contribuer à la colonisation d'autres sites, notamment chez les jeunes animaux (Low *et al.* 2006). Les gènes codant les SFP sont localisés dans le plasmide de virulence pSFO157 des EHEC (Brunder *et al.* 2006). L'expression de ces pili conduit à une meilleure adhésion aux cellules épithéliales via la piline SfpG (Musken *et al.* 2008, Bielaszewska 2009). Si l'opéron *sfp* a été identifié chez d'autres sérotypes EHEC tel que O165 (Bielaszewska 2009), sa prévalence parmi les STEC est probablement assez faible (Toma *et al.* 2004). Enfin, les ELF ont été caractérisés chez EHEC O157:H7 comme étant impliqués dans l'adhésion à la laminine et aux cellules épithéliales humaines, bovines et porcines (Samadder *et al.* 2009).

L'ECP (*E. coli* common pilus) est fortement prévalent chez *E. coli*, incluant les EHEC (Rendon *et al.* 2007). Chez ces derniers, les ECP étaient originellement appelés Mat (meningitidis-associated temperature dependent pilus) et permettent l'adhésion des bactéries aux cellules épithéliales humaines et la formation de biofilms par des liaisons inter-organelles via la piline EcpA (Rendon *et al.* 2007). De plus, ils semblent pouvoir cibler les résidus arabinosyl des parois des plantes pour permettre une adhésion aux cellules végétales (Rossez *et*

al. 2014). Cependant, leur rôle dans le processus de colonisation de l'intestin humain et bovin ne semble qu'être secondaire (Tatsuno *et al.* 2000, Dziva *et al.* 2004).

Chez *E. coli*, les pili de type 1 (T1P) sont certainement les pili les plus étudiés (Capitani *et al.* 2006). L'implication de ces pili dans le processus de colonisation a été démontrée chez de nombreux pathotypes d'*E. coli*, incluant les EPEC, les EAEC et les ETEC (Elliott *et al.* 1997, Moreira *et al.* 2003, Sheikh *et al.* 2017). En revanche, si les gènes codant ces pili sont bien présents chez de nombreuses souches de STEC et d'EHEC, dont *E. coli* O157:H7 (Li *et al.* 1997), l'expression des T1P n'a été montrée que chez un sous-ensemble des souches STEC 026 et O118 (Enami *et al.* 1999). Les T1P jouent ici un rôle dans l'adhésion à l'épithélium intestinal bovin ainsi qu'à des surfaces abiotiques (Galfi *et al.* 1998, Cookson *et al.* 2002).

Les AAF (aggregative adherence fimbriae) jouent eux un rôle important dans le phénomène d'adhérence agrégative chez les EAEC. Ces structures supramoléculaires sont associées à l'adhésion à des protéines de l'ECM telles que la fribonectine et la laminine, la formation de biofilms et l'agglutination sur les érythrocytes humains (Wakimoto *et al.* 2004, Farfan *et al.* 2008, Berry *et al.* 2014). La souche hybride *E. coli* O104:H4 possèdent également les gènes codant pour un AAF (Brzuszkiewicz *et al.* 2011, Mellmann *et al.* 2011). La présence de cet AAF, des Stx et de la dispersine chez cette souche pourrait représenter une combinaison très virulente (Boisen *et al.* 2008).

II. 10. Le système de sécrétion de type VIII (T8SS)

Le T8SS est impliqué dans la sécrétion et l'assemblage de fibres amyloides appelées curli (Desvaux *et al.* 2004, Goyal *et al.* 2014). Les curli facilitent les interactions cellule-surface et cellule-cellule et entrainent la formation de biofilms (Olsen *et al.* 1993, Cookson *et al.* 2002, Szabo *et al.* 2005, McCrate *et al.* 2013). Ils sont d'ailleurs généralement considérés comme les composants protéiques majeurs de la matrice des biofilms d'*E. coli* (Smyth *et al.* 1996, Kostakioti *et al.* 2005, Evans *et al.* 2014). Chez EHEC O157:H7, les curli sont impliqués dans l'adhésion aux feuilles d'épinards, aux cellules épithéliales et à diverses surfaces abiotiques comme le polystyrène ou l'acier inoxydable (Kim *et al.* 2004, Pawar *et al.* 2005, Macarisin *et al.* 2012). De plus, lorsque les curli sont produits conjointement avec la cellulose, ils favorisent la formation de biofilms et *de facto* la colonisation de l'hôte et la survie de la bactérie dans divers environnements (Saldana *et al.* 2009).

II. 11. Le système BAM (β-barrel assembly machinery)

Chez les bactéries didermes-LPS, les OMP (β-barrel outer membrane proteins) sont intégrées à la membrane externe par le système BAM (Knowles *et al.* 2009, Leyton *et al.* 2015, Han *et al.* 2016). Chez *E. coli*, ce système est constitué de 5 composants : l'OMP BamA qui est la sous-unité principale du système et les lipoprotéines BamB, C, D et E ancrées du côté périplasmique de la membrane externe. BamB et le complexe BamCDE se lient à BamA via les domaines POTRA (polypeptide transport associated) (Leyton *et al.* 2015, Han *et al.* 2016). En plus de certains AT décrits précédemment, d'autres OMP peuvent jouer le rôle de facteur de colonisation.

Chez EHEC O157:H7, OmpA (outer membrane protein A) est impliquée dans l'adhésion à des cellules épithéliales intestinales (Torres *et al.* 2003, Kudva *et al.* 2015). De plus, OmpA apparait être le facteur clé de l'adhésion de ces bactéries aux surfaces végétales telles que les pousses de luzerne (Torres *et al.* 2005). Cette protéine était à l'origine considérée comme formant un pore (Sugawara *et al.* 1992) avec son tonneau-β offrant un canal pour le passage continu d'eau et de solutés, mais ce rôle reste aujourd'hui controversé (Smith *et al.* 2007). De nos jours, OmpA est plutôt considérée comme une protéine multi-facettes incluant les fonctions d'adhésine et d'invasine. Cette OMP sert également de récepteur pour bactériophages et bactériocines (Smajs *et al.* 1997, Power *et al.* 2006). Concernant la formation de biofilms, l'implication d'OmpA reste plus confuse : alors qu'elle permet à *E. coli* K12 d'adhérer à des surfaces abiotiques et d'influencer significativement sur la formation de biofilms (Lower *et al.* 2005, Barrios *et al.* 2006), chez EHEC O157:H7, OmpA participe à l'adhésion à des cellules épithéliales intestinales (Torres *et al.* 2005, Kudva *et al.* 2015). Il est intéressant de constater qu'OmpA pourrait être codée par au moins deux allèles différents *ompA1* et *ompA2* (Power *et al.* 2006). La plupart des propriétés d'interaction d'OmpA émaneraient de boucles peptidiques à l'extérieur de la membrane externe et exposées à la surface de la cellule bactérienne (Smith *et al.* 2007). Des différences entre les deux allèles au niveau de ces régions pourraient donc influencer les propriétés d'adhésine ou d'invasine de la protéine.

IhA (IrgA homologue adhesin) est un facteur d'adhésion homologue de la protéine IrgA (iron-regulated protein A) de *Vibrio cholerae* (Tarr *et al.* 2000). En plus d'une structure en tonneau-β caractéristique des OMP permettant l'ancrage à la membrane, Iha possède également des domaines dirigés vers l'extérieur. Retrouvée chez de nombreuses STEC, Iha confère

notamment à EHEC O157:H7 le phénotype d'adhésion diffuse et a clairement été démontrée comme étant impliquée dans la colonisation intestinale (Yin *et al.* 2009).

Hes (hemagglutinin from STEC) est une protéine de la famille des Hra (Heat-resistant agglutinin) récemment découverte chez les STEC (Montero *et al.* 2017). Plus précisément, le gène codant cette protéine est très répandu chez les STEC LEE-négatives mais pas chez les LEE-positives. Hes est impliquée dans l'autoagrégation, la formation de biofilms, l'agglutination sur les érythrocytes et l'adhésion aux cellules épithéliales (Montero *et al.* 2017).

II. 12. Autres mécanismes de transport des protéines

Chez les EPEC et quelques EHEC non-O157, Efa-1 (EHEC factor for adherence 1) ou LifA (lymphostatin) est connue pour inhiber la prolifération des lymphocytes ainsi que la synthèse des cytokines pro-inflammatoires et des interférons gamma (Klapproth *et al.* 2000, Abu-Median *et al.* 2006). De plus, Efa-1 s'est avérée induire la colonisation de l'intestin du veau indépendamment des motifs cystéine ou glycotransférase (Deacon *et al.* 2010). Si cette protéine possède bien un domaine extra-cytoplasmique et est probablement exposée à la surface de la bactérie (Nicholls *et al.* 2002), les mécanismes moléculaires de sa sécrétion et de sa présentation à la surface restent inconnus. Chez les souches O157, l'homologue d'Efa-1 est ToxB et contribue à l'adhésion à des cellules épithéliales intestinale (Tatsuno *et al.* 2001). Cependant, cette protéine n'est pas impliquée dans la colonisation intestinale de modèles animaux et ne possède pas d'activité similaire à celle de LifA (Stevens *et al.* 2004, Abu-Median *et al.* 2006).

Certaines protéines cytoplasmiques ont pu être identifiées à la surface des cellules d'*E. coli* et peuvent jouer le rôle de facteur de colonisation. De par cette seconde fonction complètement différente de leur fonction cytoplasmique, ces protéines sont dites moonlighting (travaillant au noir) (Khan *et al.* 2014). L'une des plus communément retrouvée est la GAPDH (glyceraldehyde 3-phosphate dehydrogenase), démontrée comme pouvant lier le fibrinogène chez les EHEC et les EPEC (Egea *et al.* 2007). De plus, la GAPDH est clairement impliquée dans l'adhésion à des cellules épithéliales intestinales lors d'une infection. En fait, la plupart des enzymes glycolytiques ont pu être décrites comme ayant une seconde fonction chez diverses bactéries (Henderson *et al.* 2011). L'un des traits communs de ces protéines travaillant au noir est qu'elles ne possèdent pas de SP N-terminal pour leur exportation à travers la membrane cytoplasmique et que le système de transport permettant leur sécrétion à travers la membrane externe est inconnu. Pour les protéines dont le système de sécrétion est inconnu, on parle de

sécrétion non-classique. Dans le cas de la GAPDH, il a fortement été suggéré que la sécrétion ait lieu par un processus de piggybacking (à dos de baudet) en s'associant à un effecteur du T3aSS permettant ainsi sa sécrétion (Aguilera *et al.* 2012).

Chapitre III : L'Ag43 d'*E. coli*

L'Ag43 a été isolé pour la première fois chez la souche *E. coli* ML308-225 qui a longtemps servi de modèle pour les travaux d'architecture cellulaire et de biochimie des procédés chez *E. coli* (Winkler *et al.* 1966, Owen *et al.* 1987). Cet AT appartient à la famille des SAAT dont il partage les caractéristiques structurales évoquées précédemment (Figure 9).

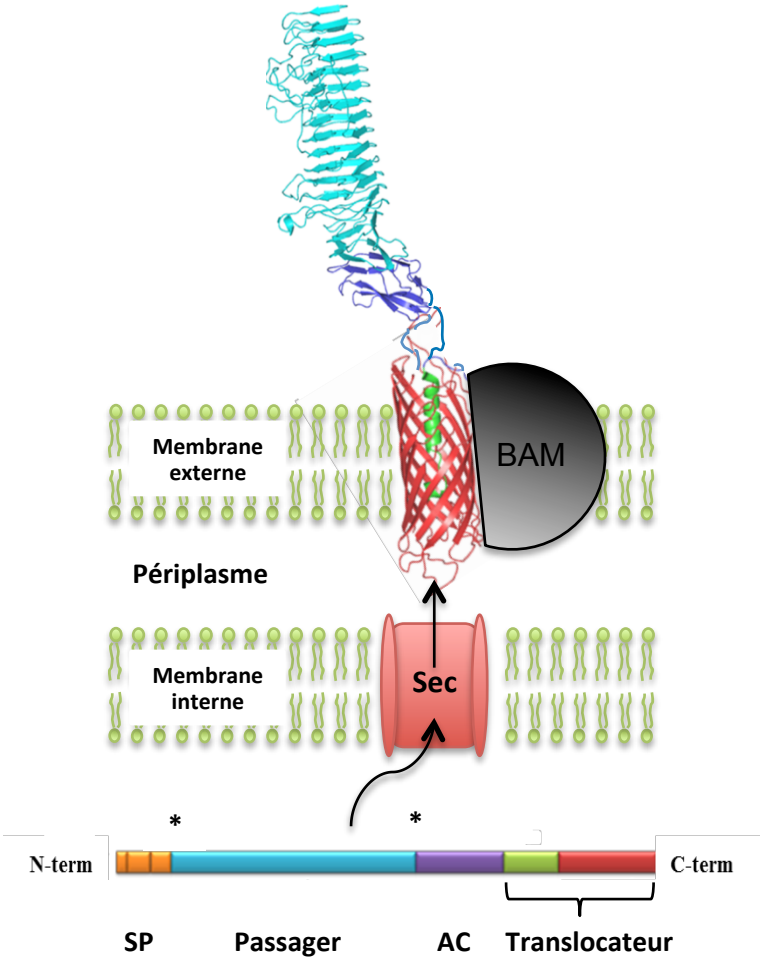

Figure 9 : Structure de l'Ag43 et sécrétion via le T5aSS. Le peptide signal (SP en orange) permet l'exportation de la protéine dans le périplasme via la translocase Sec. Le système BAM permet ensuite l'intégration dans la membrane externe du domaine translocateur constitué d'un domaine de liaison en hélice α (vert) et d'un domaine en tonneau β (rouge). Le domaine passager (bleu) contenant le motif fonctionnel de la protéine et le domaine AC (en violet) peuvent enfin passer à travers le tonneau β et être exposés à la surface de la cellule bactérienne. Les astérisques correspondent aux sites de clivage N-terminal et C-terminal.

Plusieurs allèles du gène codant cette protéine existent et sont répartis en deux sous-familles phylogénétiques distinctes (van der Woude *et al.* 2008). L'Ag43 est une des protéines régulée par variation de phase les plus abondantes de la membrane externe d'*E. coli* et est

principalement connu pour le phénotype d'autoagrégation qu'il confère à la population bactérienne (Diderichsen 1980). Cette autoagrégation des cellules et la sédimentation qu'elle entraine peut provoquer des effets collatéraux et potentiellement jouer un rôle dans la formation de biofilms, l'adhésion et la virulence des bactéries.

III. 1. Structure de la protéine Ag43

La majorité des protéines sécrétées par le T5SS possèdent dans leurs SP un domaine chargé (N), un domaine hydrophobe (H) ainsi qu'un site de clivage par la peptidase (C) et emploient la translocase Sec pour leur translocation à travers la membrane cytoplasmique via la protéine d'adressage SecB (Henderson *et al.* 1998, Randall *et al.* 2002, Henderson *et al.* 2004). Le SP de l'Ag43 possède ces 3 domaines canoniques (N2, H2 et C2) dans sa partie C-terminale et arbore de plus une extension N-terminale conservée appelée ESPR (extended signal peptide region) dans laquelle se situent deux domaines additionnels (N1 et H1) (Desvaux *et al.* 2007). Il a été suggéré que l'ESPR jouait un rôle dans l'adressage de la protéine à la translocase Sec de manière post-traductionnelle et qu'il était indirectement impliqué dans le repliement de la protéine dans le périplasme ainsi que dans sa translocation à travers la membrane externe (Chevalier *et al.* 2004, Szabady *et al.* 2005, Peterson *et al.* 2006, Desvaux *et al.* 2007). Après translocation à travers la membrane interne, le SP est clivé (Figure 9) (Henderson *et al.* 1998).

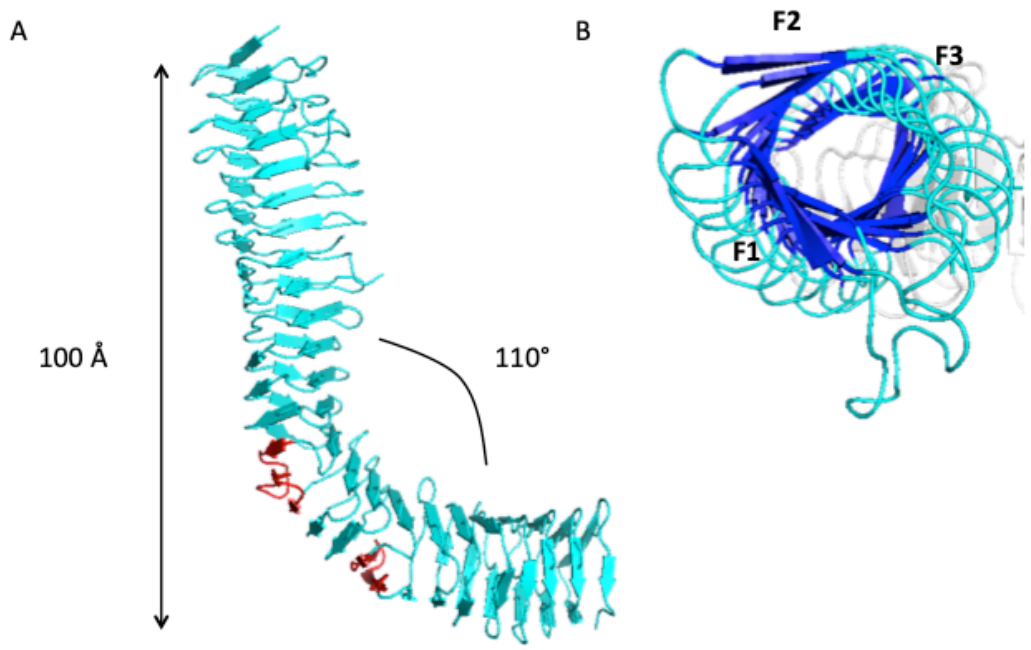

Figure 10 : Organisation et structure du domaine passager de l'Ag43a d'UPEC CFT073. (A) Représentation de la structure secondaire de l'hélice β du domaine passager de l'Ag43a avec les feuillets β (flèches) reliés par des boucles. Les motifs en épingles à cheveux β sont représentés en rouge. (B) L'hélice β vue du dessus avec les feuillets β (bleu) reliés par des boucles (cyan) formant un prisme déformé avec 3 interfaces (F1, F2 et F3). (D'après (Heras *et al.* 2014))

Le domaine passager contient le motif fonctionnel de la protéine et confère le phénotype d'autoagrégation grâce à son tiers N-terminal (Klemm *et al.* 2004, Charbonneau *et al.* 2007). Récemment, la structure du domaine passager de l'Ag43a d'UPEC CFT073 a été résolue par cristallographie et représente à ce jour la seule structure résolue d'Ag43 disponible (Heras *et al.* 2014). Ainsi, le domaine passager de l'Ag43a se replie en une hélice β tricaténaire comprenant 23 tours d'environ 20 résidus, dont chacun est composé de 3 feuillets β reliés par des régions en boucles (Figure 10). De plus, il dévoile une structure caractéristique en forme de L jamais observée jusqu'alors chez les AT. La tige du L est constituée d'un domaine N-terminal en hélice β de 13 tours parallèles, torsadé et formant un prisme déformé (Figure 10); cette conformation génère de longs sillons le long de la surface de la protéine. La courbure d'environ 110° de la structure est permise par 3 tours flanqués par 4 motifs β en épingle à cheveux. Enfin, le bas du L est formé par une hélice β C-terminale de 7 tours parallèles. L'intérieur de l'hélice β lui, est principalement constitué d'amas de résidus polaires en tours consécutifs dont les chaines latérales créent un important réseau de liaisons hydrogènes inter-brins (Heras *et al.* 2014).

Plusieurs motifs ont été originellement identifiés dans le domaine passager de l'Ag43 d'*E. coli* ML308-225, tous englobés dans l'hélice β (van der Woude *et al.* 2008): (i) le motif RGD (Arg-Gly-Asp), caractérisé chez AIDA-I comme pouvant interagir avec l'intégrine et *de facto* faciliter l'adhésion à la cellule hôte (Henderson *et al.* 1999, Charbonneau *et al.* 2007) (ii) l'aspartyl protéase et (iii) le leucine zipper impliqué dans l'interaction protéine-protéine (Junker *et al.* 2006). Cependant, aucun de ces motifs ne s'avère être conservé chez l'Ag43. En outre, des analyses structures/fonctions par mutagénèse dirigée révèlent qu'aucun de ces motifs ne semble posséder la fonction prédite *in silico* et que leur rôle serait plutôt lié à l'intégrité structurale de la protéine (van der Woude *et al.* 2008).

Le domaine passager de certains AT peut être clivé par des protéases de la membrane externe ou de manière auto-catalytique s'il contient un domaine protéolytique (Barnard *et al.* 2007, Dautin *et al.* 2007). Le domaine passager de l'Ag43 ne contient pas de tel domaine mais peut tout de même être clivé chez des souches contenant peu de protéases dans leurs membranes externes (Owen *et al.* 1996, Henderson *et al.* 1999). Le mécanisme et les conditions nécessaires pour ce clivage restent flous car le domaine passager peut expérimentalement rester accroché à la surface cellulaire (Owen *et al.* 1996). Un site de clivage de 46 résidus a été identifié entre les domaines passager et translocateur d'Ag43 similaires à celui d'*E. coli* K12, mais ce site n'est pas conservé chez tous les Ag43 (Owen *et al.* 1996, van der Woude *et al.* 2008). Une des hypothèses est donc que le clivage pourrait se faire en amont de la partie C-terminale, dans le pore du tonneau-β, entrainant une modification la structure du domaine de translocateur (Barnard *et al.* 2007, Dautin *et al.* 2007).

Les travaux étudiant la sécrétion et le repliement de l'AT BrkA (Bordetella resistance to killing) ont permis de mettre en évidence le rôle d'un domaine appelé AC localisé à la partie C-terminale du domaine passager (Oliver *et al.* 2003). La localisation exacte et la prévalence de l'AC chez les AT restent cependant controversées (Drobnak *et al.* 2015). L'implication de ce domaine AC dans la sécrétion et l'exposition à la surface cellulaire a été mise en évidence chez d'autres AT comme AIDA-I (Berthiaume *et al.* 2007), EspP (Velarde *et al.* 2004), ou Pet (pasmid encoded toxin) (Dutta *et al.* 2003). La structure de ce domaine a été résolue pour la première fois chez l'AT IcsA (intra-cellular spread protein A) et arbore un repliement en β sandwich (Kuhnel *et al.* 2011). La prévalence du domaine AC a été étudiée et a montré une distribution en 20 clusters principaux (A à Y en alphabet cyrillique), tous retrouvés chez des AT possédant un domaine passager en hélice β mais pas chez ceux arborant un domaine passager globulaire formé d'hélices α (Rojas-Lopez *et al.* 2017). Le domaine AC de l'Ag43

appartient au cluster E (Rojas-Lopez *et al.* 2017). Un motif en boucle-P impliqué dans la fourniture d'énergie via liaison à l'adénosine triphosphate a été identifié dans le domaine AC de l'Ag43 d'*E. coli* K12 mais sa mutation n'impacte pas les fonctions de la protéine (van der Woude *et al.* 2008). Chez l'Ag43, le clivage survient entre le domaine passager de la protéine et l'AC, ce dernier reste alors accroché au translocateur (Figure 9) (Konieczny *et al.* 2001, van der Woude *et al.* 2008).

L'Ag43 possède un domaine translocateur commun aux AT constitué d'un domaine de liaison en hélice α et d'un tonneau β à travers lequel le domaine passager est transporté pour être exposé à la surface de la bactérie (Oomen *et al.* 2004). L'assemblage de ce domaine requiert l'intervention des protéines BamA et BamD du complexe BAM (Figure 9) (Rossiter *et al.* 2011).

Les SAAT AIDA-I et TibA ont la particularité de pouvoir être glycosylés via l'ajout d'un heptose sur certains acides aminés du domaine passager par Aah et TIbC, respectivement. Les gènes codant ces glycosyltransferases sont adjacents aux gènes codant les AT (Charbonneau *et al.* 2007). Si les analyses en spectroscopie de masse n'ont pas montré de modification post-traductionnelles du domaine passager de l'Ag43 d'*E. col*i ML308-225 (Henderson *et al.* 1999), la protéine peut tout de même être glycosylée par Aah et TibC lorsqu'elle est en configuration *trans* (Sherlock *et al.* 2006). Par ailleurs, la purification du domaine passager de l'Ag43 d'*E. coli* K12 glycosylé par Aah révèle la présence de 16 résidus thréonine et sérine glycosylés et des analyses *in vitro* indiquent que cette glycosylation protège de la dénaturation et facilite le repliement après dénaturation (Knudsen *et al.* 2008). La glycosylation n'affecte toutefois pas la stabilité des Ag43 à d'*E. coli* K12 ou d'*E. coli* ML308-225 à la surface de la bactérie (Henderson *et al.* 1999, Charbonneau *et al.* 2007). Une glycsoylation naturelle de l'Ag43 a été observée chez deux isolats fécaux d'*E. coli,* mais pas chez *E. coli* K12 (Reidl *et al.* 2009). Ce phénomène pouvant donc varier selon les souches et les conditions, il est d'autant plus important d'évaluer l'impact naturel de la glycosylation de l'Ag43, surtout chez les souches pathogènes et en tenant compte de la diversité allélique de l'*agn43* (van der Woude *et al.* 2008).

III. 2. Phylogénie et diversité allélique du gène *agn43*

Le gène *agn43* est situé sur un phage CP4 provenant d'une excision du chromosome de *Shigella flexneri* (Luck *et al.* 2004, Sakellaris *et al.* 2004, Turner *et al.* 2004). Historiquement,

les gènes codant l'Ag43 portaient différents noms en fonction des souches où ils étaient identifiés, comme par exemple *flu* (fluffing) chez *E. coli* K12 ou *cah* (calcium binding antigen 43 homologue) chez EHEC O157:H7 (Henderson *et al.* 1999, Torres *et al.* 2002); par convention, il a été proposé d'utiliser la nomenclature *agn43* (van der Woude *et al.* 2008). L'hypothèse de l'existence de plusieurs allèles du gène *agn43* au sein d'une même souche d'*E. coli* a été formulée pour la première fois en 1996 lors d'études en western-blot (Owen *et al.* 1996). Les progrès du séquençage et l'augmentation du nombre de génomes séquencés ont par la suite permis d'avoir une meilleure idée de la distribution de ce gène chez *E. coli* et de confirmer cette hypothèse. En effet, de multiples copies d'*agn43* ont été découvertes chez *E. coli* ML308-225 ainsi que chez des *E. coli* pathogènes comme la souche EHEC O157:H7 EDL933 qui en possède deux copies identiques situées sur deux OI différents (Roche *et al.* 2001, Torres *et al.* 2002, Klemm *et al.* 2004, Restieri *et al.* 2007). Au delà d'*E. coli* (et *Shigella*), l'*agn43* a également été identifié chez une autre espèce proche : *Citrobacter rodentium* (van der Woude *et al.* 2008). Si la diversité allélique de l'*agn43* a longtemps été considérée comme résultant de la spéciation (Yen *et al.* 2002), elle serait en réalité plutôt liée à une acquisition horizontale car l'*agn43* peut être retrouvé chez différentes souches et à différentes localisations sur le chromosome (van der Woude *et al.* 2008). A l'inverse, certaines souches comme *E. coli* K12 MG1655 n'en possèdent qu'un exemplaire et d'autres comme EPEC E2348/69 peuvent ne pas en posséder du tout (Roche *et al.* 2001, van der Woude *et al.* 2008). Cette absence pourrait être expliquée par le fait que la présence de l'*agn43* serait antérieure à la différenciation entre les espèces *E. coli* et *C. rodentium* et que le gène ait été perdu au fil de l'évolution (van der Woude *et al.* 2008).

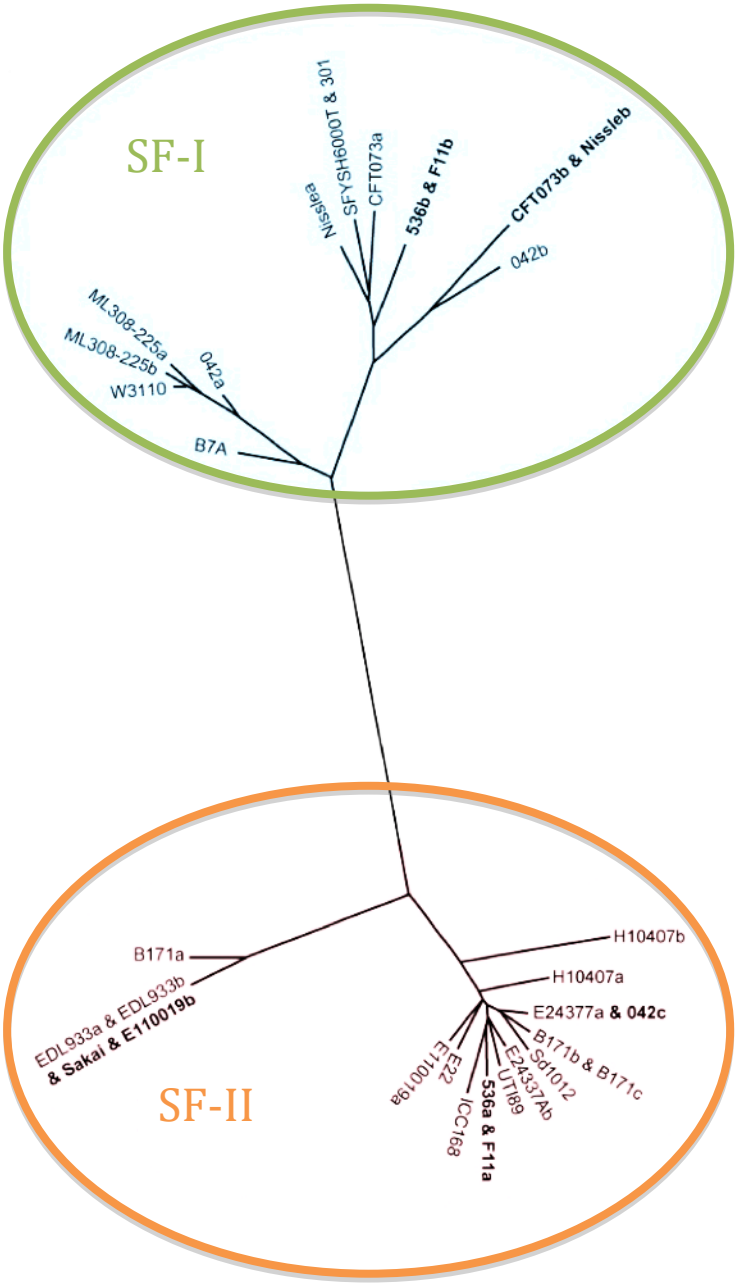

Figure 11 : Phylogénie des protéines de l'Ag43. Le profil de l'arbre montre une représentation bivalente avec deux sous-familles distinctes. La sous-famille I (SF-I) incluse des protéines codées par l'allèle d'*agn43* de la souche *E. coli* K12. La sous-famille II (SF-II) inclue des protéines codées par l'allèle d'*agn43* de la souche *E. coli* O157:H7 EDL933. (D'après (van der Woude *et al.* 2008))

Une analyse phylogénétique ciblant les relations génétiques entre les différents allèles d'*agn43* a révélé l'existence de deux sous-familles, i.e. SF-I et SF-II (Figure 11) (van der Woude *et al.* 2008). Cette distribution bivalente s'expliquerait par une diversité significative au niveau de la région qui englobe la partie C-terminale du passager et la partie N-terminale du translocateur (correspondant à la région du site de clivage entre ces 2 domaines) (Figure 9). En comparant les profils MLST (multilocus sequence typing) avec l'arbre phylogénétique

auteurs n'ont pu établir aucune corrélation entre l'évolution d'*E. coli* et la possession de l'un ou l'autre des allèles, renforçant encore plus l'hypothèse de l'acquisition horizontale de l'*agn43* (van der Woude *et al.* 2008). L'analyse des génomes d'*E. coli* séquencés suggère que les allèles appartenant à SF-II sont tous associés à des souches pathogènes et sont plus prévalents que ceux appartenant à SF-I (van der Woude *et al.* 2008). Une étude antérieure à cette analyse phylogénétique avait montré une prévalence plus importante de l'Ag43 chez les souches pathogènes (90%) que chez les souches commensales (56%) (Restieri *et al.* 2007). Ces observations pourraient néanmoins être biaisées par le fait que les génomes d'*E. coli* séquencés sont principalement ceux de souches pathogènes (van der Woude *et al.* 2008). Il n'existe cependant pas de corrélation absolue entre la présence ou l'absence d'un allèle et un pathotype particulier d'*E. coli* (van der Woude *et al.* 2008).

III. 3. Régulation de l'expression du gène *agn43*

L'une des principales caractéristiques de l'*agn43* est qu'il est régulé par variation de phase, un phénomène stochastique épigénétique, héritable et réversible (Henderson *et al.* 1997, Henderson *et al.* 1997, Henderson *et al.* 1999, van der Woude *et al.* 2004, van der Woude 2006, van der Woude *et al.* 2008). Ce phénomène est répandu chez de nombreuses protéines de surface et engendre chez les bactéries aussi bien commensales que pathogènes, une expression différente du gène au sein de la population de cellules bactériennes : les cellules peuvent exprimer le gène (phase ON) ou non (phase OFF). La variation de phase est caractérisée par la fréquence de changement d'état d'expression (nombre de cellules qui changent d'état d'expression par génération), qui est estimée à 10^{-3} pour l'*agn43* d'*E. coli* K12.

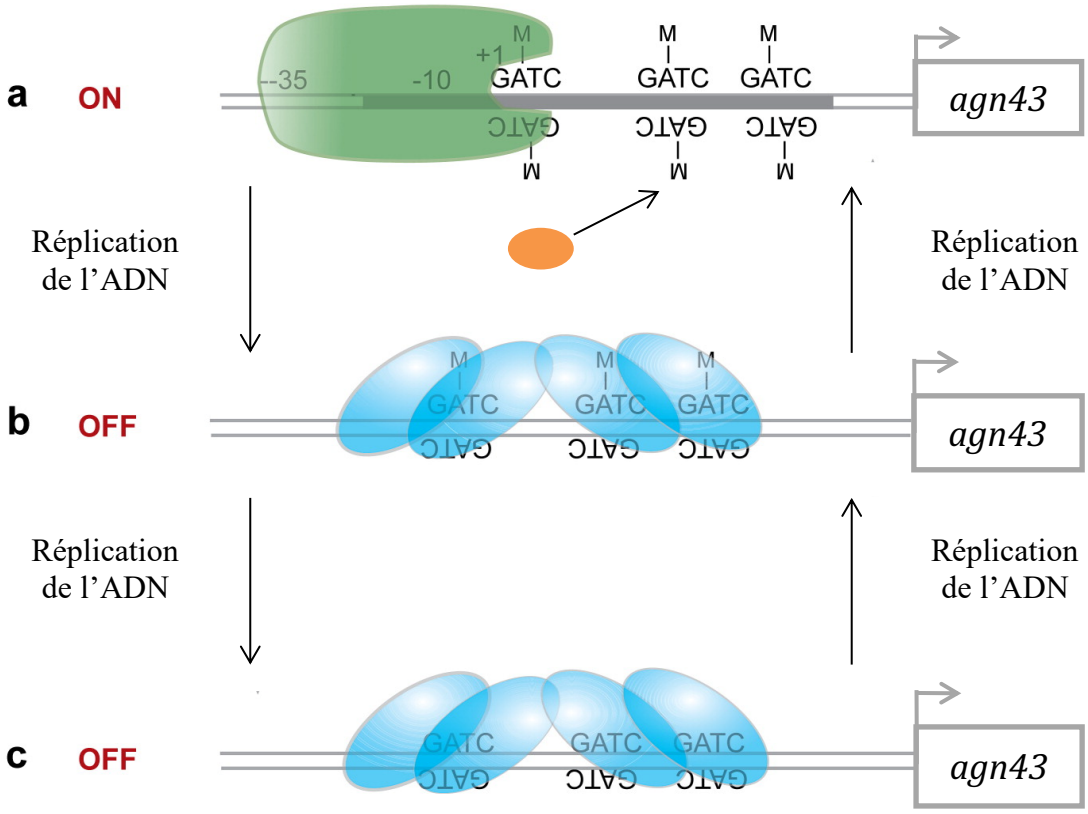

Figure 12 : Représentation schématique de la régulation de l'*agn43* par variation de phase. Les deux brins d'ADN sont représentés avec les sites GATC ciblés par Dam (orange) et la méthylation est indiquée par un M. La région promotrice s'étend de la position -10 à la position -35 et le site d'initiation de la transcription correspond à la position +1. (a) : phase ON caractérisée par les séquences GATC méthylées et la transcription du gène par l'ARN polymérase (vert). (b) : phase OFF intermédiaire avec un ADN hémi-méthylé et des dimères d'OxyR fixés (bleu). (c) : phase OFF dans laquelle l'ADN n'est pas méthylé et où OxyR possède une affinité de liaison plus grande que dans le schéma b. La réplication de l'ADN est nécessaire pour la transition d'une phase à l'autre. (D'après (van der Woude *et al.* 2008))

La régulation de l'*agn43* par variation de phase implique deux intervenants : la méthylase Dam (DNA adenine methylase) qui méthyle l'adénine au niveau des sites GATC et le régulateur de transcription OxyR (Henderson *et al.* 1997). Des études approfondies ont été menées concernant les interactions OxyR-ADN *in vitro*, les effets des mutations et de la surexpression des deux facteurs sur l'expression de l'*agn43* ainsi que l'état de méthylation de l'ADN des bactéries possédant l'*agn43* (Haagmans *et al.* 2000, Waldron *et al.* 2002, Wallecha *et al.* 2002). Ces travaux ont permis d'établir un modèle de la régulation de l'*agn43* d'*E. coli* K12 par Dam et OxyR (van der Woude *et al.* 2008) (Figure 12). La région régulatrice du gène *agn43* qui recouvre la position -10 de la séquence promotrice contient le site de fixation du régulateur OxyR. En phase ON, les 3 séquences GATC situées dans ce site sont méthylées par Dam et empêchent la fixation d'OxyR. Lorsque ces séquences sont méthylées, l'ARN polymérase peut se fixer à la région promotrice et effectuer la transcription du gène. A l'inverse en phase OFF, OxyR peut se lier à son site de fixation car les séquences GATC ne sont pas

méthylées, bloquant ainsi la fixation de l'ARN polymérase et la transcription du gène. Donc, la méthylation des séquences GATC et la fixation d'OxyR sont mutuellement exclusives et la phase d'expression du gène résulte de la compétition de ces deux facteurs pour la région régulatrice de l'*agn43*. Le retour à un état d'ADN non-méthylé chez *E. coli* requiert une synthèse d'ADN et passe forcément par un état d'ADN hémi-méthylé. Cet état joue un rôle significatif dans la variation de phase car le répresseur OxyR peut se lier à la région régulatrice hémi-méthylée de l'*agn43* et empêcher une nouvelle méthylation par Dam (Correnti *et al.* 2002). La réplication de l'ADN est également nécessaire pour faciliter le retrait des dimères d'OxyR fixés sur les séquences GATC (van der Woude 2017). L'utilisation d'une nouvelle approche utilisant l'ajout de sites de fixation auxiliaires d'OxyR en amont de la région régulatrice a de plus montré que la fréquence de passage de la phase ON à la phase OFF peut être modulée sans modifier la fréquence de passage de la phase OFF à la phase ON. Ainsi, la présence d'OxyR à la fourche de réplication lors du passage au niveau de la région régulatrice de l'*agn43* est un facteur clé pour l'amorçage d'une transition de la phase ON à OFF (Kaminska *et al.* 2010).

OxyR est originellement connu pour son implication dans la réponse au stress oxydatif qui altère son niveau d'oxydation et ses capacités d'interaction avec l'ADN (Tartaglia *et al.* 1989, Storz *et al.* 1999). De ce fait, il a été suggéré que le stress oxydatif puisse influer sur son rôle de répresseur de la transcription de l'*agn43* et *de facto* l'autoagrégation. Cependant, une étude a réfuté cette hypothèse en montrant que la variation de phase de l'Ag43 était indépendante du niveau d'oxydation d'OxyR (Wallecha *et al.* 2003). En effet, des analyses *in vitro* ont été réalisées pour comparer l'affinité de trois formes différentes d'OxyR (oxydée, naturelle et réduite) à la région régulatrice de l'ADN de l'*agn43* d'*E. coli* K12. Cette étude a montré que ces 3 formes d'OxyR possèdent la même affinité pour la région régulatrice non-méthylée de l'*agn43*. Cette affinité est par ailleurs similaire à celle observée pour d'autres gènes induits par la forme oxydée d'OxyR lors de la réponse au stress oxydatif (Wallecha *et al.* 2003). Au regard des différences observées au niveau des séquences des régions régulatrices, la question de l'effet du niveau d'oxydation d'OxyR sur les différents allèles d'*agn43* pourrait toutefois être investiguée (van der Woude *et al.* 2008). En outre, la variation de phase de l'*agn43* a été modélisée en utilisant la GFP comme rapporteur de l'expression du gène (Lim *et al.* 2007). Ces travaux ont mis en évidence l'existence d'états d'expression intermédiaires avec un faible niveau d'expression de la GFP ne dépendant pas du promoteur de l'*agn43* ni de sa région régulatrice. En s'intéressant à la distribution des cellules entre les différents états

d'expression (ON, OFF et intermédiaires), les auteurs ont démontré que les états d'expression intermédiaires étaient rarement occupés et que les cellules y accédant étaient reconduites à leur état d'origine (ON ou OFF). En conséquence, une cellule quittant la phase OFF ou la phase ON ne passe que rarement à l'état d'expression opposé (Lim *et al.* 2007).

Figure 13 : Représentation schématique de la région régulatrice de l'*agn43* d'*E. coli* K12. La région promotrice s'étend de la position -35 à -10 tandis que le site d'initiation de la transcription correspond à la position +1. La séquence leader ARN s'étend de la position +1 à la position +222, l'ATG en position +223 (ATG 2) correspondant au site prédit d'initiation de la traduction de la protéine Ag43. Un second codon d'initiation a été identifié dans la séquence leader ARN en position +67 (ATG 1). Le terminateur de transcription (T) est représenté par une tige boucle allant de la position +162 à +195. (D'après (Wallecha *et al.* 2014))

Etant donné que l'état d'oxydation d'OxyR n'affecte pas la régulation de l'*agn43* d'*E. coli* K12 et que les concentrations cellulaires respectives d'OxyR et Dam ne montrent pas de signes de variation, la seule régulation pouvant altérer la fréquence de transition doit très probablement être de nature post-transcriptionnelle. Par ailleurs, les différents allèles possèdent de longues séquences d'acides ribonucléiques (ARN) leader de tailles variables en amont de la région codante. L'analyse de la région leader de l'*agn43* d'*E. coli* K12 a ainsi révélé deux nouveaux modes de régulation de l'expression de ce gène (Wallecha *et al.* 2014). En effet, il été démontré que la région leader de ce gène arbore un terminateur de transcription *rho*-indépendant en tige-boucle fonctionnel *in vitro* et qu'un mécanisme d'anti-terminaison de transcription doit être opérationnel dans les phases précoces de croissance (Wallecha *et al.* 2014). Le second mécanisme fait intervenir deux codons d'initiation de la traduction identifiés dans la séquence régulatrice du gène : le premier en amont de ce terminateur de transcription et l'autre, correspondant au site d'initiation de la traduction de la séquence codante de la protéine, en aval (Figure 13). La mutation de l'un ou l'autre de ces codons a montré qu'ils étaient tous les deux nécessaires à une expression optimale de l'*agn43*. De plus, la traduction débutée au codon d'initiation situé en amont augmente l'expression de l'*agn43* en aval en stabilisant l'ARN messager (Wallecha *et al.* 2014). Cependant, si la régulation de l'*agn43* d'*E. coli* K12 est particulièrement bien caractérisée, ce n'est pas le cas pour les autres allèles. Bien que l'analyse phylogénétique a mis en lumière des différences au niveau des séquences des régions

régulatrices des deux sous-familles d'*agn43*, les 3 sites GATC sont retrouvés dans les deux cas avec un espacement similaire suggérant également une expression soumise à la variation de phase par Dam et OxyR (van der Woude *et al.* 2008).

La variation de phase est répandue chez *E. coli* et intervient notamment dans la régulation de l'expression de nombreux fimbriae. Le niveau d'expression de ces gènes est souvent soumis à des signaux environnementaux ainsi qu'à la régulation par d'autres régulateurs provenant d'autres opérons fimbriaux. Ainsi, il a été proposé qu'un réseau de régulation contrôle la variation de phase de différents fimbriae en permettant une expression séquentielle des gênes au sein de la population (Holden *et al.* 2004). L'implication de l'Ag43 dans un réseau de régulation a été suggérée, mais des résultats contradictoires ont été observés dans le cas de la coordination de l'expression de l'Ag43 et du T1P (Hasman *et al.* 2000, Schembri *et al.* 2002). A ce jour, la participation de la régulation de l'Ag43 à un réseau de régulation de gènes contrôlant le comportant cellulaire reste donc incertaine (van der Woude *et al.* 2008).

III. 4. Rôles de l'Ag43

Avant d'entrer dans les détails, précisons que la localisation exacte de l'Ag43 à la surface cellulaire pourrait avoir une conséquence sur les propriétés qu'il confère aux cellules bactériennes. Ainsi, selon des travaux en immunofluorescence sur des souches *E. coli* K12 et *E. coli* ML308-225, l'Ag43 serait distribué de manière homogène sur toute la surface bactérienne (Henderson *et al.* 1997). Toutefois, d'autres AT sont sécrétés aux pôles cellulaires et gardent cette localisation chez des souches possédant un LPS complet, mais pas chez les souches arborant un LPS incomplet, i.e. sans l'antigène O (Jain *et al.* 2006). Si ces observations s'appliquent à l'Ag43, la localisation de la protéine pourrait varier selon les souches d'*E. coli* et potentiellement impacter les fonctions conférées. De plus, en raison des statuts ON/OFF aléatoires des cellules au sein de la population bactérienne, il a été montré qu'ignorer la variation de phase pouvait conduire à des erreurs d'interprétation concernant les phénotypes liés associés à l'Ag43 (Chauhan *et al.* 2013). Tout cela illustre donc l'importance de tenir compte de l'allèle, du niveau d'expression et du fond génétique de la souche utilisée lorsque l'on s'intéresse au rôle de l'Ag43 (van der Woude *et al.* 2008).

III. 4. A. L'autoagrégation

Le phénotype d'autoagrégation correspond à la formation d'agrégats bactériens entrainant la sédimentation des bactéries au fond d'une culture liquide statique (Figure 14).

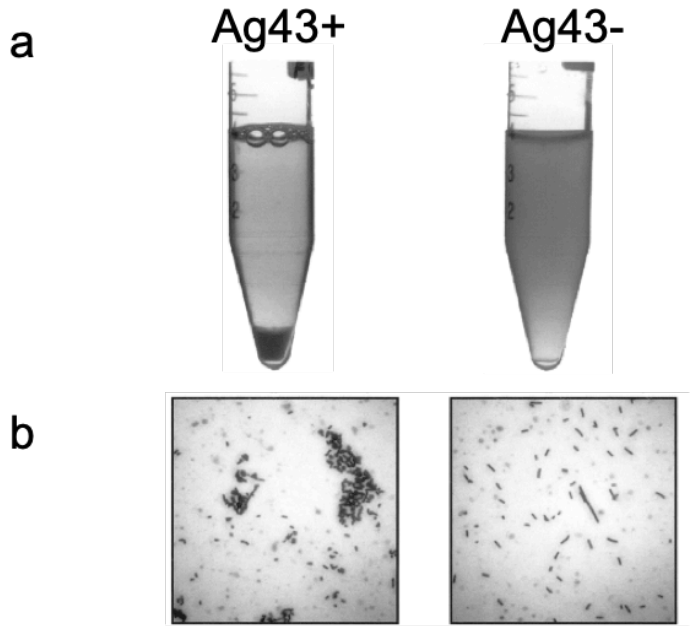

Figure 14 : Observations macroscopique et microscopique du phénotype d'autoagrégation. (a) : Observation macroscopique de la sédimentation au fond de la culture liquide des bactéries exprimant l'Ag43 (Ag43+) alors que les bactéries n'exprimant pas l'Ag43 (Ag43-) ne sédimentent pas. (b) : Observation microscopique d'agrégats bactériens formés quand l'Ag43 est exprimé et de cellules seules lorsqu'il ne l'est pas. (D'après (Torres *et al.* 2002))

Le rôle de l'Ag43 dans l'autoagrégation a été largement démontré chez de nombreuses souches aussi bien commensales que pathogènes (Henderson *et al.* 1997, Torres *et al.* 2002, Klemm *et al.* 2004). Une étude a montré que l'Ag43 pouvait également interagir avec d'autres membres de la famille des SAAT pour former des agrégats bactériens mixtes, i.e. AIDA-I et TibA (Klemm 2006). Fait intéressant, le degré d'autoagrégation pourrait dépendre de l'allèle *agn43* exprimé. En effet, des études basées sur la surexpression de différents allèles d'*agn43* via des plasmides ont été menées et des résultats contrastés ont notamment été observés pour les 2 variants de la souche UPEC CFT073. Une première étude a révélé que les allèles présentant une forte similarité avec l'Ag43 d'*E. coli* K12 provoquent plus rapidement l'agrégation alors que d'autres ne la confèrent pas, incluant les deux allèles l'Ag43 chez UPEC CFT073 (Klemm *et al.* 2004). A l'inverse, une autre étude a montré que ces deux mêmes allèles conféraient l'autoagrégation, avec une sédimentation plus rapide de l'Ag43a par rapport à l'Ag43b (Ulett *et al.* 2007). Enfin, lorsqu'ils sont surexprimés, l'Ag43 d'*E. coli* K12, les deux

variants d'UPEC 536 et les deux variants d'*E. coli* Nissle 1917 agrègent également à des degrés différents (Reidl *et al.* 2009).

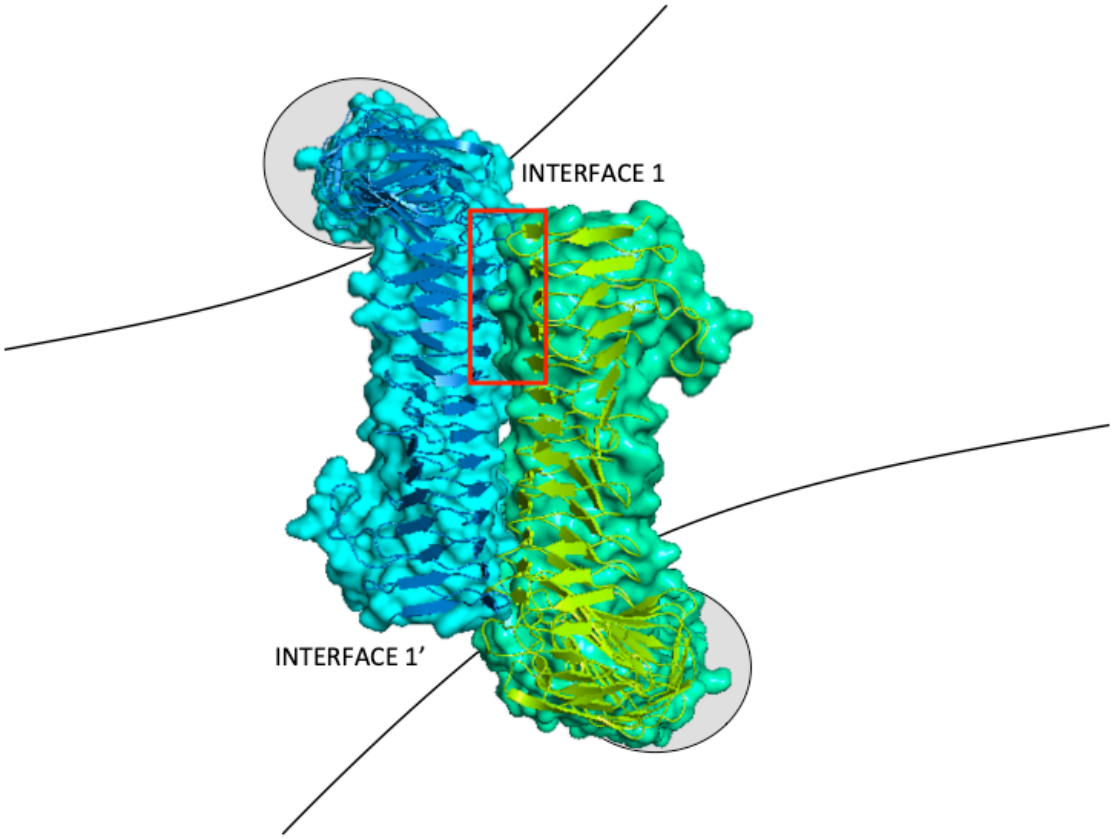

Figure 15 : Interaction entre deux domaines passager de l'Ag43a d'UPEC CFT073 via le mécanisme de velcro moléculaire. Auto-association au niveau des interfaces 1 et 1' de deux domaines passagers de l'Ag43a d'UPEC CFT073 en configuration *trans*, exposés à la surface de la bactérie grâce aux domaines translocateurs (gris). Le cadre rouge délimite l'endroit où s'établissent à chaque interface 9 liaisons hydrogènes et 1 liaison ionique (rouge). (D'après (Heras *et al.* 2014))

L'origine de ces différences pourrait être d'ordre structurale. La résolution de la structure de l'Ag43a par cristallographie a effectivement permis de mettre en lumière le mécanisme moléculaire de l'association entre 2 domaines passagers de ce variant (Heras *et al.* 2014). Les dimères formés sont ainsi compacts avec un enroulement des hélices β l'une autour de l'autre en configuration *trans*, similaire à celle observée pour l'AT Hap d'*Haemophilus influenzae* (Meng *et al.* 2011). Cette association est stabilisée grâce à 9 liaisons hydrogènes et 1 liaison électrostatique au niveau de chaque interface, ce modèle d'interaction a été qualifié de velcro moléculaire (Figure 15) (Heras *et al.* 2014). A noter que des oligomères plus complexes peuvent également être formés, comme l'interaction de deux paires de dimères. Les séquences des domaines passagers de l'Ag43a et de l'Ag43b sont identiques à 85% et la modélisation de la structure de celui du variant b suppose que l'interaction entre les deux variants est possible.

L'utilisation de souches fluorescentes a permis de valider cette hypothèse et les auteurs ont par la suite montré, grâce à un panel de mutants, que la forme en L du domaine passager de l'Ag43a était nécessaire pour l'autoagrégation (Heras *et al.* 2014).

A partir de ces observations, une étude s'est intéressée à la dynamique de l'auto-association de l'Ag43 d'*E. coli* K12 dans des conditions similaires à celles rencontrées dans l'environnement en utilisant une approche par spectroscopie de force sur molécule seule et par spectroscopie de force dynamique (Jacquot *et al.* 2014). Cette étude a révélé une extensibilité importante du domaine passager de l'Ag43 avec un dépliement des sous-domaines correspondant à une extension moléculaire de 150 nm. De plus, l'auto-association peut impliquer la formation de dimères et de tétramères d'Ag43 dirigée par des forces d'interaction *cis* faibles et rapides ainsi que des interactions *trans* lentes et fortes d'une magnitude de 100-250 pN. La dynamique de ces interactions s'est avérée être fortement influencée par le pH et les forces de cisaillement appliquées. Ces résultats suggèrent une modulation des conditions environnementales sur l'autoagrégation des bactéries induites par l'Ag43 (Jacquot *et al.* 2014).

D'autres facteurs peuvent influencer le phénomène d'autoagrégation. Il a notamment été montré que la présence de structures supramoléculaires plus grandes, comme les T1P ou les capsules, bloquerait l'autoagrégation par l'Ag43 probablement en empêchant l'interaction entre ces protéines (Hasman *et al.* 1999, Klemm 2006). De plus, la O-glycosylation du domaine passager de AIDA-I et TibA par Aah et TibC, impacte fortement leurs fonctions (Lindenthal *et al.* 1999, Benz *et al.* 2001, Cote *et al.* 2013). De ce fait, la O-glycosylation des domaines passagers de différents variants d'Ag43 par Aah a été étudiée et réduit dans la majeure partie des cas leur capacité à autoagréger, sauf pour l'Ag43 d'*E. coli* K12 (Reidl *et al.* 2009). Plus récemment, il a été mis en évidence que l'expression de l'auto-inducteur 2 (AI-2), connu pour son rôle dans le SQ, peut augmenter l'autoagrégation d'*E. coli* W3110 grâce au chimiotactisme qu'il provoque (Laganenka *et al.* 2016). En plus de la sédimentation des cellules bactériennes, l'autoagrégation aurait des conséquences importantes sur la virulence des bactéries en permettant une protection face aux stress environnementaux, la résistance aux antibiotiques, la défense contre d'éventuels prédateurs et en favorisant la formation de biofilms (Trunk *et al.* 2018).

III. 4. B. La formation de biofilms et l'adhésion

Il a été montré que l'expression de l'Ag43 augmente modérément la formation de biofilms *in vitro* (Danese *et al.* 2000, Kjaergaard *et al.* 2000, Klemm *et al.* 2004, Roux *et al.* 2005). Cet effet modéré pourrait être dû en partie à des effets opposés ou indirects de l'expression de l'*agn43* sur ce phénotype. Entre autres, la motilité est un facteur favorisant la formation d'un biofilm (Danese *et al.* 2000, Kjaergaard *et al.* 2000, Roux *et al.* 2005) alors que l'expression de l'Ag43 est accompagnée d'une diminution de la motilité bactérienne (Ulett *et al.* 2006). De plus, les T1P sont souvent requis pour initier le développement sessile des bactéries mais ils bloquent également l'autoagrégation induite par l'Ag43 (Hasman *et al.* 1999, Danese *et al.* 2000, Ulett *et al.* 2006). Il a été proposé que l'Ag43 ne participerait pas aux phases précoces d'interaction entre la bactérie et la surface chez *E. coli* K12 mais plutôt à la phase de maturation du biofilm (Beloin *et al.* 2008). L'hypothèse récurrente que l'autoagrégation induite par l'Ag43 faciliterait le développement sessile des bactéries a été démontrée grâce à des études en conditions dynamiques montrant que les cellules d'*E. coli* K12 retrouvées dans les biofilms étaient le résultat d'une sélection physique due à leurs propriétés autoagrégatives (Chauhan *et al.* 2013). Les auteurs ont également montré que l'autoagrégation induite par l'Ag43 pourrait de plus biaiser les résultats observés lors de tests d'adhésion ou de formation de biofilms en masquant les effets d'autres potentielles adhésines (Chauhan *et al.* 2013). En plus de son impact sur l'autoagrégation, le chimiotactisme par l'AI-2 induit également une augmentation de la résistance d'*E.coli* W3110 au stress et une meilleure capacité des bactéries à former des biofilms (Laganenka *et al.* 2016).

La diversité allélique de l'*agn43* impacte également la capacité des bactéries à former des biofilms. En effet, l'étude basée sur les deux allèles d'UPEC CFT073 a montré que lors de la surexpression chez *E. coli* K12, l'un des deux montre une meilleure capacité à induire le développement sessile des bactéries. En revanche dans le fond génétique de la souche UPEC CFT073, aucun des deux ne semble contribuer à ce phénotype (Ulett *et al.* 2007). Par ailleurs, les différents variants d'Ag43 testés dans une autre étude confèrent également des capacités de formation de biofilms différentes aux bactéries (Reidl *et al.* 2009). L'effet de la O-glycosylation sur la formation de biofilms a démontré des différences entre les allèles (Reidl *et al.* 2009). Très peu d'études se sont cependant intéressées au rôle de l'Ag43 dans le processus de colonisation des STEC. L'Ag43 d'EHEC O157:H7 EDL933 possède la capacité de lier le calcium et de former des biofilms sur des surfaces abiotiques en milieu minimum (Torres *et al.* 2002). D'autre part, l'expression de l'Ag43 chez des souches non O157 formant des biofilms suggère un rôle

pour cette protéine dans la colonisation de l'environnement par ces bactéries (Biscola *et al.* 2011).

En plus de son rôle dans la formation de biofilms, plusieurs faits suggèrent que l'Ag43 pourrait également fonctionner comme une adhésine, notamment à cause de sa forte homologie avec AIDA-I. Cependant, des recherches *in vitro* et *in vivo* des variants d'Ag43 d'*E. coli* ML308-225 et d'EAEC 042 ne confirment pas l'implication de l'Ag43 dans l'adhésion spécifique à différentes lignées cellulaires ou dans la colonisation de l'intestin en modèle murin (de Luna *et al.* 2008). Il a depuis été prouvé que différents variants d'Ag43 pouvaient conférer aux bactéries une adhésion spécifique à des cellules épithéliales rénales ainsi qu'au collagène et à la laminine, avec des affinités plus ou moins marquées selon les allèles (Reidl *et al.* 2009). La O-glycosylation impacte également différemment les capacités d'adhésion des différents variants (Reidl *et al.* 2009).

III. 4. C. Implication dans la virulence

La participation de l'Ag43 dans la virulence a été particulièrement étudiée chez les UPEC par des approches *in vitro* avec des cellules eucaryotes et des modèles simulant l'infection du tractus urinaire (Fexby *et al.* 2007). Ces études se sont principalement focalisées sur les souches UPEC CFT073 et UTI89 qui arborent des allèles d'*agn43* différents, rendant la comparaison des résultats et leur interprétation difficiles (van der Woude *et al.* 2008). Il a pourtant été prouvé que l'Ag43 pouvait être exprimé dans les conditions d'infection. En effet, des anticorps dirigés contre l'Ag43 d'UPEC CFT073 sont retrouvés dans le sérum dans un modèle d'infection murin (Hagan *et al.* 2007), suggérant ici un avantage de la variation de phase de l'Ag43 qui faciliterait l'évasion immune (van der Woude *et al.* 2008). En outre, bien qu'elles soit considérées comme non-invasives, des UPEC UTI89 exprimant l'Ag43 peuvent être retrouvées à l'intérieur de cellules urothéliales sous formes de groupes de bactéries étroitement serrées (Anderson *et al.* 2003). De plus, l'Ag43a induit la colonisation de la vessie à long terme lorsqu'il est exprimé constitutivement, en cohérence avec l'observation d'une colonisation diminuée à 5 jours post-infection lorsque cet allèle est muté (Ulett *et al.* 2007). Il a également été déterminé par des tests *in vitro* que l'expression de l'Ag43 depuis un plasmide multicopies entraine l'absorption et la survie dans les neutrophiles polymorphonucléaires (Fexby *et al.* 2007). Difficile toutefois d'affirmer si ces rôles proposés *in vivo* ont un rapport avec l'autoagrégation induite par l'Ag43 ou même avec la formation de biofilms.

Chapitre IV : Objectifs de recherche

Les EHEC sont des bactéries pathogènes alimentaires dont le sérotype le plus connu et le plus virulent est O157:H7. Si la très grande majorité des cas d'infections sont sporadiques, de nombreuses épidémies ont également été recensées avec de forts impacts sur la société civile et l'industrie agro-alimentaire. Le TD des bovins étant le réservoir naturel des EHEC, la source d'infection principale reste la consommation d'aliments provenant de ces animaux comme la viande insuffisamment cuite, ou les produits laitiers non pasteurisés. De plus, la capacité des EHEC à adhérer à différentes surfaces aussi bien biotiques qu'abiotiques leur permet de se disséminer dans l'environnement et être transmis à l'Homme plus facilement. Ainsi, les EHEC peuvent être retrouvés à tous les échelons de la chaine alimentaire, depuis le portage asymptomatique animal, les chaînes de production et dans l'aliment comme vecteur de contamination et *in fine* chez l'Homme lors d'une infection après ingestion d'aliments contaminés. Un des dénominateurs qui peut être retrouvé à ces différentes étapes est la présence de l'ECM avec laquelle la bactérie peut adhérer, que ce soit dans le TD de l'Homme ou de l'animal ou dans des produits alimentaires comme la viande. L'identification des déterminants moléculaires bactériens impliqués dans cette interaction apparaît donc comme un des points importants pour mieux comprendre le processus de colonisation de la chaîne alimentaire et éventuellement le développement de nouvelles stratégies permettant de limiter le portage chez l'animal et la contamination des aliments.

Des travaux antérieurs réalisés au laboratoire ont permis de mettre en évidence l'effet de différents paramètres environnementaux sur l'adhésion d'*E. coli* O157:H7 à différentes protéines de l'ECM. En effet, il s'avère que dans certaines conditions, les bactéries adhéraient préférentiellement aux collagènes I et III tandis que le milieu de culture, la température et le pH influaient sur cette adhésion (Chagnot *et al.* 2013). Dans la littérature, certaines molécules des EHEC sont connues pour prendre part dans la colonisation de l'ECM mais beaucoup d'autres acteurs restent à identifier et caractériser. Parmi les cibles potentielles, le rôle de l'Ag43 est très bien caractérisé dans le processus de colonisation d'*E. coli* K12, mais peu de données sont disponibles concernant ses fonctions et son mode de régulation chez EHEC O157:H7.

Dans ce contexte, l'objectif de ma thèse était donc de caractériser l'implication de déterminants moléculaires protéiques impliqués dans le processus de colonisation d'*E. coli* O157:H7 et plus particulièrement l'Ag43. Cette thèse s'inscrit dans le cadre du projet CoMBa (Colonisation de la Matrice par les Bactéries) financé par le Fond Régional d'Investissement

Auvergne (cluster Institut de Recherche Pharmabiotique) impliquant les sociétés BioFilm Control et Lallemand. Ces travaux ont débuté au sein de l'équipe BaDiAl (Adaptation des Bactéries aux environnements Digestif et Alimentaire) de l'UR454 Microbiologie à l'INRA (Institut National de la Recherche Agronomique) Auvergne Rhône-Alpes, sur le site de Theix. Ils se sont par la suite poursuivis et terminés dans la nouvelle UMR MEDiS (Microbiologie Environnement Digestif et Santé) sous la cotutelle de l'INRA et l'Université Clermont-Auvergne.

La première partie de ma thèse a consisté à étudier l'impact de différents déterminants protéiques de surfaces sur l'adhésion d'*E. coli* O157:H7 à des composants de l'ECM dans différentes conditions expérimentales. Ces premiers résultats m'ont conduit à me focaliser sur l'Ag43 afin de caractériser son rôle dans l'autoagrégation et son mode de régulation chez *E. coli* O157:H7. Ces résultats constituent le premier chapitre expérimental de ce manuscrit.

La deuxième partie de ma thèse a ainsi poursuivi la caractérisation de l'Ag43 d'*E. coli* en se basant sur une étude phylogénétique ayant préalablement révélé l'existence de deux sous-familles d'Ag43 (van der Woude *et al.* 2008). En s'appuyant sur de récentes données structurales (Heras *et al.* 2014), l'analyse phylogénétique a permis de mettre en évidence des différences structurales insoupçonnées et la caractérisation fonctionnelle de différentes classes d'Ag43 ayant une capacité différentielle à établir des interactions homotypiques et hétérotypiques. Enfin, la prévalence de ces différentes classes d'Ag43 chez les bactéries a été investiguée et ces résultats constituent le troisième chapitre expérimental de ce manuscrit.

RÉSULTATS

Chapitre I : Caractérisation de l'Ag43 dans l'adhésion aux protéines de l'ECM, l'autoagrégation et sa régulation chez *E. coli* O157:H7

Les EHEC peuvent être retrouvés à tous les niveaux de la chaine alimentaire : depuis le réservoir animal, ils peuvent contaminer diverses denrées alimentaires et déclencher une infection chez l'Homme après ingestion (Vogeleer *et al.* 2014). La colonisation de la chaine alimentaire par ces bactéries pourrait être étroitement liée à leur capacité d'interaction avec les différents composants de l'ECM. Certains déterminants moléculaires des bactéries didermes-LPS sont connus pour leur implication dans l'adhésion aux protéines de l'ECM et sont qualifiés de MSCRAMM (Chagnot *et al.* 2012). Pour être exposées à la surface de la cellule bactérienne, ces MSCRAMM franchissent les deux membranes biologiques des bactéries didermes-LPS par l'intermédiaire des différents systèmes de sécrétion (Desvaux *et al.* 2009). Selon le système utilisé, les protéines MSCRAMM peuvent être (i) sécrétées et assemblées en structures supramoléculaires telles que les flagelles et les pili (T3SS, T2SS, T4SS, T7SS et T8SS) ou (ii) ancrées à la membrane externe grâce à un tonneau β comme les AT secrétés par le T5SS (Chagnot *et al.* 2013). L'adhésion d'*E. coli* O157:H7 a été étudiée et s'est révélée être dépendante des conditions de culture, suggérant une expression complexe de ces différents MSCRAMM (Chagnot *et al.* 2013). En revanche, si certaines protéines bactériennes sont connues pour leur implication dans l'adhésion à l'ECM, de nombreux MSCRAMM potentiels restent à identifier et caractériser.

La première partie de mes travaux de thèse a donc consisté à identifier divers déterminants moléculaires protéiques potentiellement impliqués dans le processus d'adhésion d'*E. coli* O157:H7 aux protéines de l'ECM. Les cibles sélectionnées regroupent ainsi le flagelle H7, le pilus HCP, le curli et l'injectisome pour les structures supramoléculaires ainsi que les AT EhaA, EhaB et Ag43. Ce dernier est très bien caractérisé chez la souche de laboratoire *E. coli* K12 pour son implication dans l'autoagrégation, l'adhésion et la formation de biofilms ainsi que pour sa régulation par variation de phase impliquant les protéines Dam et OxyR (van der Woude *et al.* 2008). En revanche, aucune information n'est disponible concernant la régulation de l'Ag43 d'*E. coli* O157:H7 ni son implication dans l'autoagrégation et l'adhésion aux protéines de l'ECM. Les autres déterminants moléculaires sélectionnés n'ont soit pas été étudiés dans le fond génétique d'*E. coli* O157:H7 soit pas été étudiés dans les mêmes conditions

expérimentales. Une approche de génétique fonctionnelle a donc été mise en place pour identifier l'effet de ces différents déterminants moléculaires dans l'adhésion d'*E. coli* O157:H7 à des protéines de l'ECM dans différentes conditions de culture. Le rôle de l'Ag43 dans l'autoagrégation ainsi que la régulation de ce phénotype ont ensuite étés caractérisés (Article 1).

ARTICLE n°1

Valentin AGEORGES, Ricardo MONTEIRO, Nelly CACCIA, Sabine LEROY, Alfredo PEZZICOLI, Mariagrazia PIZZA, Frédérique CHAUCHEYRAS-DURAND, Christian PROVOT, and Mickaël DESVAUX.

Involvement of antigen 43 (Ag43) in bacterial adhesion of *Escherichia coli* O157:H7 to the extracellular matrix and characterisation of the autoaggregation phenomenon

Pour publication dans Applied and Environmental Microbiology

Involvement of antigen 43 (Ag43) in bacterial adhesion of *Escherichia coli* O157:H7 to the extracellular matrix and characterisation of the autoaggregation phenomenon

Valentin AGEORGES [1], Ricardo MONTEIRO [2], Nelly CACCIA [1], Sabine LEROY [1], Alfredo PEZZICOLI [3], Mariagrazia PIZZA [3], Frédérique CHAUCHEYRAS-DURAND [1,4], Christian PROVOT [5], and Mickaël DESVAUX [1,*].

[1] Université Clermont Auvergne, INRA, UMR454 MEDiS, 63000 Clermont-Ferrand, France

[2] Institute for Research and Innovation in Health, Host interaction and Response Unit, Molecular Microbiology group, 4200-135 Porto, Portugal.

[3] GSK, Via Fiorentina 1, 53100 Siena, Italy.

[4] Lallemand SAS, 31702 Blagnac Cedex, France.

[5] BioFilm Control, Saint-Beauzire, F-63360, France

*Corresponding author: Dr Hab Mickaël DESVAUX-LENÔTRE (Hab, PhD, Mag), INRA (Institut National de la Recherche Agronomique) Auvergne-Rhône-Alpes, Centre de Recherche de Clermont-Ferrand, UMR454 MEDiS (Microbiologie, Environnement Digestif, Santé), Site de Theix, F-63122 Saint-Genès Champanelle, France. Tel.: +33 (0)4 73 62 47 23, Fax: +33 (0)4 73 62 45 81, E-mail: mickael.desvaux@inra.fr.

ABSTRACT

Enterohaemorrhagic *E. coli* (EHEC) O157:H7 are anthropozoonotic agents leading to haemorrhagic colitis and haemolytic-uremic syndrome (HUS). From their animal reservoir, i.e. ruminants, EHEC may contaminate some foodstuffs and consequently infect humans. The presence of extracellular matrix (ECM) components along the food chain could increase their colonisation capacity. Numerous proteins can be present at the bacterial cell surface of EHEC and could participate to the bacterial adhesion and biofilm formation onto the ECM. This study aimed at identifying and characterising the involvement of cell-surface proteins in the colonisation process of EHEC to some ECM components. Following a functional genetic analysis, several colonisation factors were identified as potentially involved in bacterial adhesion to ECM components. Among them, the autotransporter Antigen 43 (Ag43) was further shown to contribute to *E. coli* O157:H7 autoaggregation. Besides, this phenotype was demonstrated to be regulated by the Dam methylase and the OxyR transcriptional regulator. These results suggest that Ag43 represents an important colonisation factor of EHEC and provide new insights about the regulation of the autoaggregation in *E. coli* O157:H7.

INTRODUCTION

Escherichia coli is a lipopolysaccharidic (LPS) diderm bacterium (archetypal Gram-negative bacterium) known as a harmless commensal of the gastrointestinal tract (GIT) of warm-blooded animals (Mason *et al.* 1981, Tenaillon *et al.* 2010). Nevertheless, some *E. coli* strains, such as enterohemorragic *E. coli* (EHEC) are diarrheagenic as they have acquired and express virulence factors (Kaper *et al.* 2004). EHEC are anthropozoonotic pathogens responsible for foodborne infections. The natural reservoir of EHEC is the gut of ruminants, in which carriage is asymptomatical (Pruimboom-Brees *et al.* 2000). Following direct or indirect faecal contamination, human infections occur through consumption of contaminated food products (Kintz *et al.* 2017). Beef meat is the most frequently incriminated but raw milk products or vegetables can also be vectors of contamination. EHEC are pathogenic Shiga-toxin (Stx) encoding *E. coli* (STEC) that are able to cause symptoms ranging from watery diarrhea up to life-threatening complications such as haemolytic haemorrhagic colitis, uremic syndrome (HUS) as well as thrombotic thrombocytopenic purpura (TTP) (Tarr *et al.* 2005). In addition to Stx, the LEE (locus of enterocyte effacement) is generally considered as the main virulence factor of EHEC responsible for attaching and effacing (A/E) lesions on epithelial cells (McDaniel TK 1995). While EHEC infections are mainly sporadic, these bacteria can also be responsible for major outbreaks worldwide (Karch *et al.* 1999). Based on epidemiologic data, *E. coli* O157:H7 is the most virulent serotype and the most commonly isolated during outbreaks (Mora *et al.* 2007, Croxen *et al.* 2013).

While Stx is the hallmark of EHEC, the LEE is not systematically present in all strains, suggesting the role of additional or alternative molecular determinants in the pathogenesis. EHEC further possess numerous colonisation factors that can contribute to virulence or act as niche factors in different environments (Hill 2012, Monteiro *et al.* 2016). Considering the ecophysiology of EHEC from animal carriage, food contamination to human infection (Vogeleer *et al.* 2014), the extracellular matrix (ECM) could serve as a support for bacterial colonisation and contribute to EHEC contamination all along the food chain (Chagnot *et al.* 2012). The ECM is a complex protein network found in animal tissues, mainly composed of proteoglycans and fibrillar proteins, namely collagens, laminin, fibronectin and elastin. The bacterial molecular determinants involved in ECM adhesion are collectively called MSCRAMM (microbial surface components recognising adhesive matrix molecules). Thus, a better understanding of the molecular and cellular mechanisms taking part in EHEC adhesion to ECM proteins is crucial to develop new prevention strategies of contamination of the food

chain by these pathogens. In diderm-LPS bacteria, numerous proteinaceous MSCRAMM can be exposed at the bacterial cell surface (Chagnot *et al.* 2013). In *E. coli* O157:H7, the HCP (haemorrhagic coli pili) pili was shown to bind fibronectin, laminin and collagens (Xicohtencatl-Cortes *et al.* 2009), curli can interact with laminin and fibronectin (Saldana *et al.* 2009) and EhaB (EHEC autotransporter B) bound to collagen I and laminin (Wells *et al.* 2009). While, *E. coli* O157:H7 was shown to adhere specifically to collagen I and collagen III (Chagnot *et al.* 2013), bacterial adhesion to ECM components is influenced by environmental conditions suggesting differential expression and contribution of various colonisation factors (Chagnot *et al.* 2013). In colonisation processes, the motility conferred by the H7 flagella was further demonstrated to be involved in bacterial adhesion to spinach leaves or to glass (Donlan 2002, Nagy *et al.* 2015) as well as the injectisome in the adhesion to plants (Shaw RK 2008). Moreover, EhaA and Ag43 (antigen 43) were shown to take part in autoaggregation and biofilm formation (Torres *et al.* 2002, Wells *et al.* 2008). From investigations in domesticated *E. coli* strains such as K12, the expression of *agn43* appeared heterogeneous, since cells in the bacterial population can randomly express the gene or not (van der Woude *et al.* 2008, Chauhan *et al.* 2013); of note, *agn43* was originally known as *flu* (fluffing) in *E. coli* K12 (Owen *et al.* 1996). Control of *agn43* expression is epigenetic and known to be prone to phase-variation (Henderson *et al.* 1999, van der Woude *et al.* 2008). It actually involves a methylase, Dam (DNA adenine methylase) and the transcription regulator OxyR; binding of OxyR to unmethylated GATC sequences in the regulator region close by the promoter represses the gene transcription (phase OFF), whereas the methylation of GATC sequences by Dam prevent binding of OxyR and consequently repression by OxyR, which results in gene transcription (phase ON) (Henderson *et al.* 1997). In *E. coli* O157:H7 EDL933, two identical copies of *agn43* were identified and originally named *cah* (calcium binding antigen 43 homologue) (Torres *et al.* 2002). When expressed in *E. coli* K12, this gene product promotes autoaggregation, biofilm formation and binding to calcium. In *E. coli* O157:H7 genetic background, expression of *cah* is induced in DMEM (Dulbecco's modified eagle medium) but not LB (lysogeny broth) and participates to biofilm formation (Torres *et al.* 2002). However, no data about the gene regulation, the involvement in autoaggregation and interaction with the ECM proteins is as yet available.

Selecting relevant colonisation factors potentially at play, this study aimed at investigating their contribution in the adhesion to ECM proteins in *E. coli* O157:H7 background under different growth conditions. We focused on Ag43 in order to better understand its role in the autoaggregation of *E. coli* O157:H7, and its regulation was further characterised.

RESULTS

Ag43, EscN and CsgA are involved in the adhesion of *E. coli* O157:H7 to collagen I and collagen III

Considering (i) the animal reservoir and initial meat contamination occurring at dehiding stage in slaughterhouse, and (ii) the intestinal site of human infection, two conditions were tested, i.e. bacterial growth at 39°C and adhesion at 25°C or growth and adhesion at 37°C, respectively. From literature survey, several surface colonisation factors were targeted as potentially involved in bacterial adhesion to ECM, namely (i) the flagellin FliC (or forming the H7 flagellar filament, (ii) CsgA, the major subunit of the curli, (iii) HcpA, the major pilin in HCP, (iv) EscN, a key ATPase for the Type III secretion system assembly, and several cell-surface autotransporters, i.e. (v) Ag43, (vi) EhaA and (vii) EhaB.

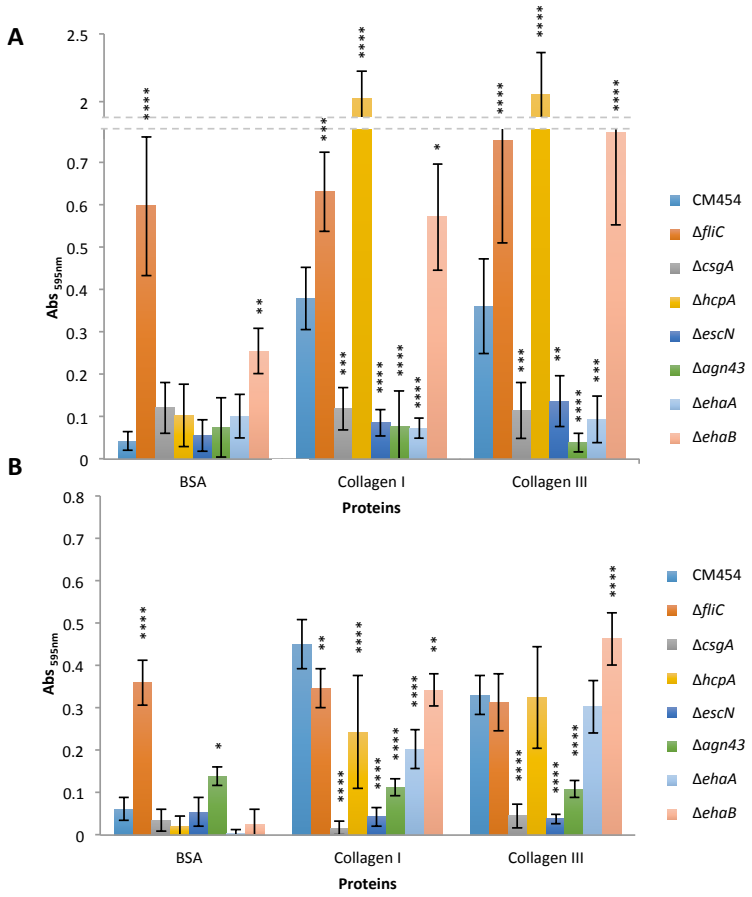

Figure 1: Role of proteinaceous MSCRAMM in adhesion of *E. coli* O157:H7 CM454 to collagen I and collagen III. The adhesion assays to coated collagen I and collagen III were performed in BHI. BSA was used as a control of adhesion specificity. Experiments were carried out with (A) growth at 39°C and adhesion at 25°C, or (B) both growth and adhesion at 37°C.

Upon deletion of targeted genes in *E. coli* O157:H7 CM454, adhesion assays were carried out on collagens I and III. In both tested conditions, *E. coli* O157:H7 CM454 adhered specifically to collagens I and III and not to BSA (used as a specificity control) (Figure 1). For bacteria grown at 39°C and left adhered at 25°C, mutants deleted in *csgA*, *escN*, *agn43, ehaA* or *ehaB* showed a significant decrease in the amount of adhered biomass to both collagens when compared to *E. coli* O157:H7 CM454 (Figure 1). This suggests the involvement of curli, injectisome, Ag43 and EhaA in bacterial adhesion to these ECM proteins. Surprisingly enough, deletion of *fliC*, *ehaB* and particularly *hcpA* resulted in a significantly increased adhesion of bacterial cells to both collagens I and III compared to *E. coli* O157:H7 CM454 (Figure 1A). This would suggest that the absence of the autotransporter adhesin EhaB and some cell-surface organelles, namely the flagella and HCP, could unmask other colonisation factors involved in ECM protein binding. In contrast to the *hcpA* mutant, the adhesion of the *fliC* and *ehaB* mutants was non-specific as it adhered to BSA, collagen I and collagen III (Figure 1A).

When bacterial cells were grown and left adhered at 37°C, mutants deleted in *csgA*, *escN* and *agn43* again shown a significant decrease in the amount of adhered biomass to both collagens when compared to *E. coli* O157:H7 CM454 whereas the adhesion of the *fliC* mutant was again non-specific but not higher than *E. coli* O157:H7 CM454 (Figure 1B). In contrast to the previous tested condition, for all the mutants, the amount of adhered biomass to collagen I was significantly lower when compared to *E. coli* O157:H7 CM454 (Figure 1B). On collagen III, though, the deletion of *ehaB* resulted in a significant increase in the amount of adhered biomass (Figure 1B).

Taken together, these results would indicate that the temperature influences the expression of some colonisation factors and/or their direct involvements in the adhesion process to collagen I and III. While further complementation and gene expression experiments would be required to confirm and decipher their respective contribution in adhesion to ECM proteins, we decided to focus on Ag43 as it was the only colonisation factor exhibiting a very highly significant decrease in bacterial adhesion to collagens in both tested conditions;

Ag43 is responsible for the autoaggregation of *E. coli* O157:H7 CM454 in DMEM medium

The involvement of Ag43 in bacterial autoaggregation of *E. coli* O157:H7 CM454 was further investigated. To determine the proportion of colonies displaying an autoaggregation phenotype, 100 isolated colonies isolated from agar plates were inoculated in LB or DMEM

broth. After 24h of incubation at 25°C, 100% of cell cultures grown in DMEM were sedimented at the bottom of the tubes whereas no sedimentation was observed in LB (Figure 2A).

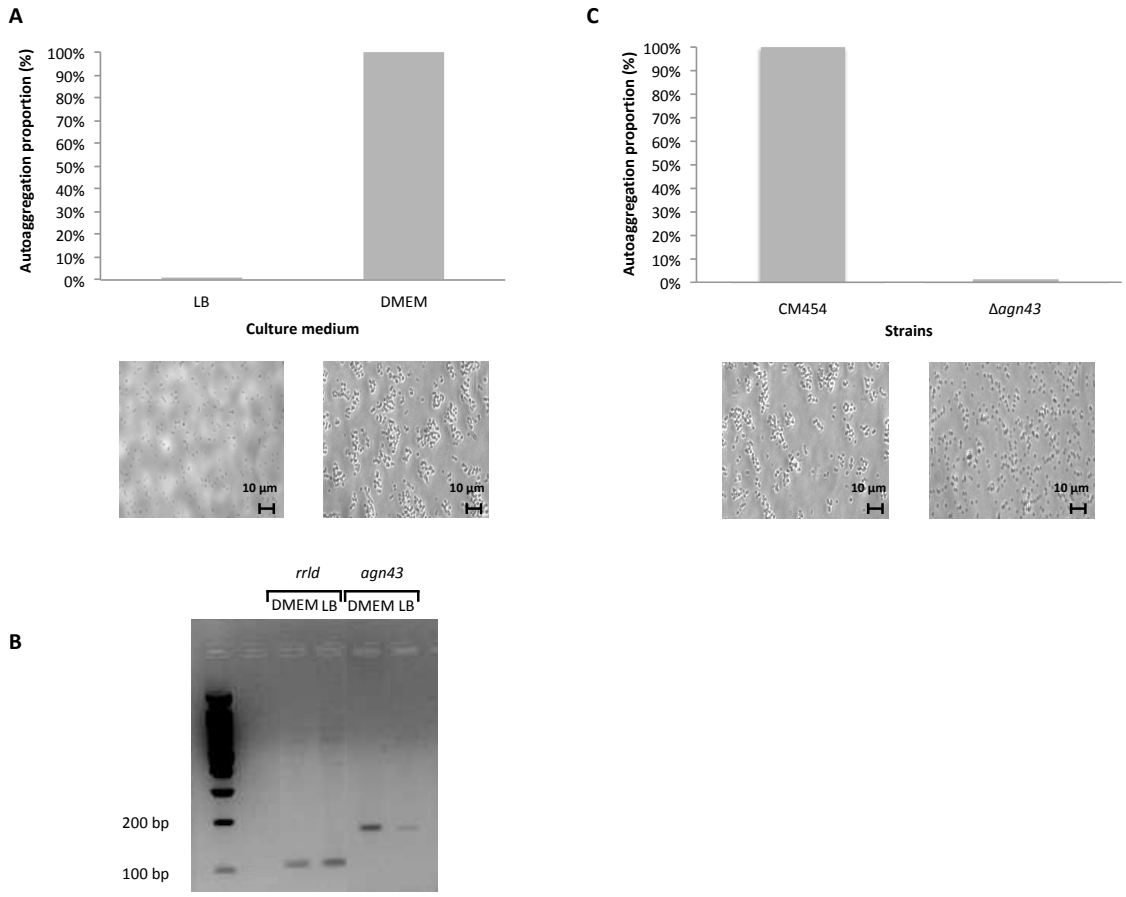

Figure 2: Role of *agn43* expression on autoaggregation of *E. coli* O157:H7 CM454 in DMEM. (A) Proportion of sedimented tubes grown in LB or DMEM from isolated colonies of *E. coli* O157:H7 CM454 colonies and corresponding microscopic observations after 24h incubation at 25°C. (B) RT-PCR amplicons of specific fragments of the housekeeping gene *rrld* and *agn43* when bacterial cells are grown in LB or DMEM. (C) Proportion of sedimented tubes grown in DMEM from isolated colonies of *E. coli* O157:H7 CM454 or *E. coli* O157:H7 CM454 Δ*agn43* (Δ*agn43*) colonies and subsequent microscopic after 24h incubation at 25°C.

Microscopic examinations confirmed that sedimentation was the result of autoaggregation of bacterial cells in DMEM and that only rare cell aggregates were observed in LB. Regarding expression of *agn43*, high expression was detected for bacterial cells grown DMEM whereas low expression was observed in LB (Figure 2B). The autoaggregation was completely abolished for *E. coli* O157:H7 CM454 Δ*agn43* grown in DMEM (Figure 2C), strongly suggesting the involvement of Ag43 in the autoaggregation phenotype. Following immunofluorescence detection in *E. coli* O157:H7 CM454, Ag43 was detected at the apex and

at the interface between two auto-associating cells (Figure 3). Taken together, these results indicate Ag43 is involved in the autoaggregation phenomenon of *E. coli* O157:H7 CM454 in DMEM and Ag43-Ag43 interaction would occur at the apex of the bacterial cells.

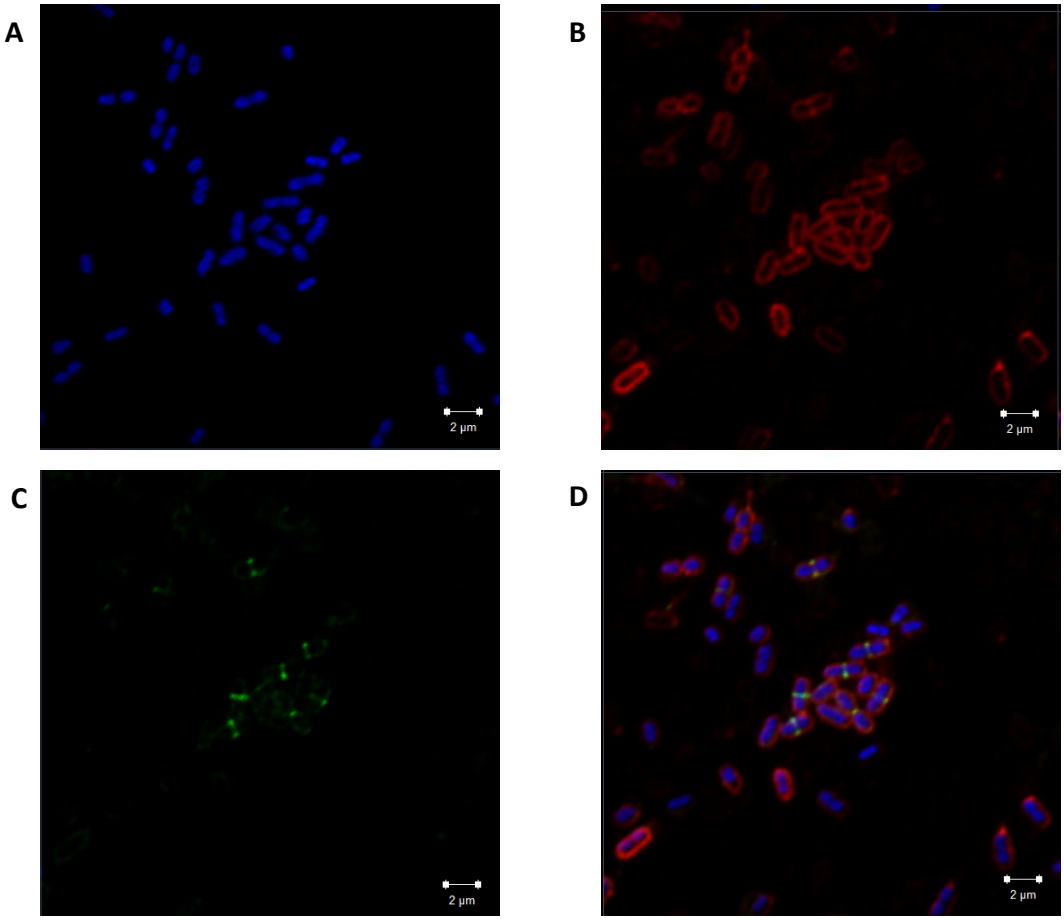

Figure 3: Subcellular localisation of Ag43 in *E. coli* O157:H7 CM454. Confocal images of *E. coli* O157:H7 CM454 grown in DMEM. (A) DAPI (blue) was used to visualize chromosomal DNA. (B) Antibodies to whole-*E. coli* bacteria and a fluorescent secondary antibody (red) were used to visualize bacterial cells. (C) Ag43 was detected using specific antibodies and visualized with a fluorescent secondary antibody (green). (D) Confocal image with all the fluorescent markers.

Dam and OxyR impact the autoaggregation phenotype of *E. coli* O157:H7 CM454

The latter results suggested that the autoaggregation phenomenon of *E. coli* O157:H7 CM454 mediated by Ag43 is regulated depending on the culture medium. The involvement of Dam and OxyR in the regulation of *agn43* by phase variation was investigated in our strain by a functional genetic analysis. While *E. coli* O157:H7 CM454 was unable to autoaggregate, a strong flocculation and sedimentation were observed with *E. coli* O157:H7 CM454 Δ*oxyR* in LB culture left standing statically during 24h (Figure 4A). Microscopic observations indicated

that the sedimentation was associated with the formation of bacterial cell aggregates contrary to *E. coli* O157:H7 CM454 or *E. coli* O157:H7 CM454 Δ*dam* (Figure 4B).

To confirm these observations, genes encoding Dam and OxyR were cloned into pBAD for protein expression in *E. coli* O157:H7 CM454. Sedimentation kinetics showed that the cultures of *E. coli* O157:H7 CM454 pBAD-Dam flocculated contrary to control *E. coli* O157:H7 CM454 pBAD or *E. coli* O157:H7 CM454 pBAD-OxyR (Figure 4C). Microscopic observations further confirmed the presence of bacterial cells aggregates for *E. coli* O157:H7 CM454 pBAD-Dam (Figure 4D).

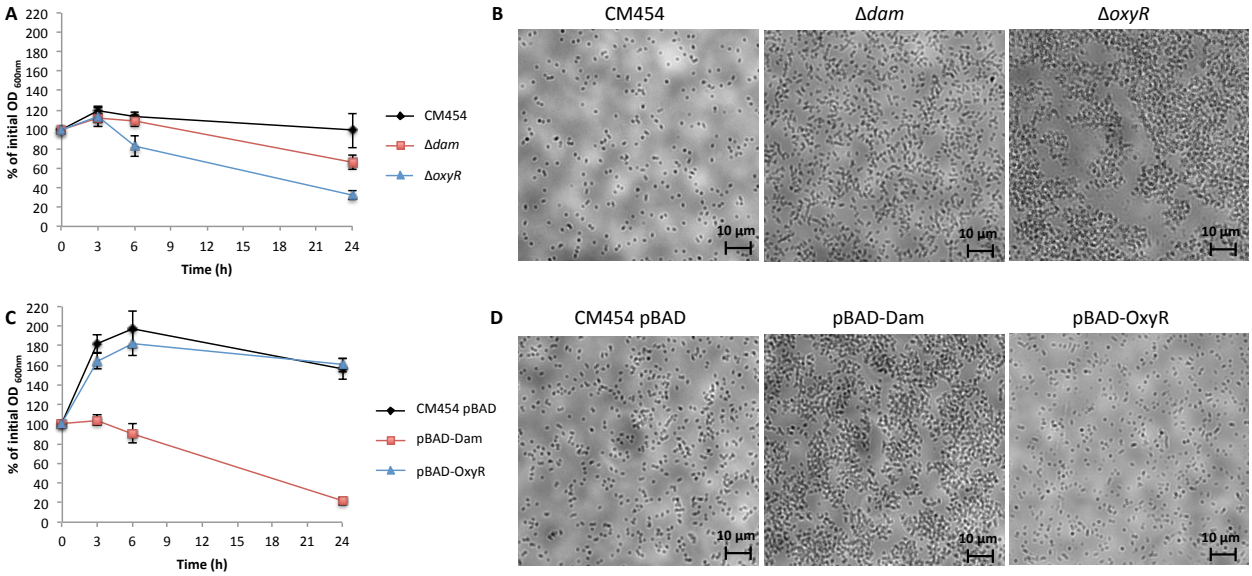

Figure 4: Impact of Dam and OxyR on the autoaggregation of *E. coli* O157:H7 CM454 in LB. (A) Sedimentation assay of *E. coli* O157:H7 CM454 (CM454), *E. coli* O157:H7 CM454 Δ*dam* (Δ*dam*) and *E. coli* O157:H7 CM454 Δ*oxyR* (Δ*oxyR*). (B) Corresponding microscopic observations (C) Sedimentation assay of *E. coli* O157:H7 CM454 pBAD, *E. coli* O157:H7 CM454 pBAD-Dam and *E. coli* O157:H7 CM454 pBAD-OxyR. (D) Corresponding microscopic observations. Sedimentation assays were performed in LB at 25°C and microscopic observations were carried out after 24h incubation.

Taken together, these results demonstrated that Dam and OxyR are involved in the regulation of the autoaggregation phenomenon of *E. coli* O157:H7 CM454, which could explain the difference in the autoaggregation phenotype between cells grown in LB and DMEM.

DISCUSSION

EHEC can adhere to different surfaces, whatever biotic or abiotic, and colonise the human food chain (Vogeleer *et al.* 2014). In a previous study, bacterial adhesion of *E. coli* O157:H7 to ECM proteins was demonstrated to be greatly influenced by the growth medium, temperature and pH as it was maximal and specific towards collagen I and III in LB at pH 7 and 25°C (Chagnot *et al.* 2013). Based on these observations, we decided to study the contribution of several selected and targeted proteinaceous MSCRAMM on adhesion of *E. coli* O157:H7 to collagen I and collagen III in two different environmental conditions (respective to temperature). Our results suggested differential involvement of some MSCRAMM in bacterial adhesion depending on environmental conditions. Although HCP and EhaB have been shown to promote bacterial adhesion to collagen (Erdem *et al.* 2007, Wells *et al.* 2009, Xicohtencatl-Cortes *et al.* 2009), our results show that *E. coli* O157:H7 cells deficient in their production were shown to better adhere to collagens. As previously demonstrated between capsule and short bacterial adhesins (Schembri *et al.* 2004) or between pili and Ag43 (Hasman *et al.* 1999), the presence/absence of HCP could mask/unmask other colonisation factors following steric shield. While flagellar motility has been demonstrated to take part in the colonisation process of bacteria (Donlan 2002, Haiko *et al.* 2013), the loss of motility may promote cell sedimentation and/or bacterial adhesion to surfaces or may be a resulting factor of the autoaggregation process (Ulett *et al.* 2006). Further investigations would be required to confirm and decipher their exact contribution, however our work suggests that EhaA, curli, injectisome and Ag43 are involved in adhesion of *E. coli* O157:H7 to collagens. As suggested here, future research directions should be worth to evaluate the differential genetic regulations and protein expressions of these colonisation factors under various environmental conditions encountered along the food chain as well as their respective roles in the colonisation process, including biofilm formation. Without *a priori* consideration of the colonisation factors involved, a proteomic analysis could also be considered to identify new proteinaceous MSCRAMM.

Ag43 is probably one of the most investigated autotransporter in *E. coli* (van der Woude *et al.* 2008), however to our knowledge only one study has been dedicated to the Ag43 variant encoded in EHEC (*E. coli* O157:H7) (Torres *et al.* 2002). Previous work had shown that autoaggregation could occur and was associated to non-specific bacterial adhesion to ECM (Chagnot *et al.* 2013). Here, we confirmed that autoaggregation could not be evidenced in complex laboratory growth media like LB or BHI (data not shown) but occurred in DMEM, a medium generally used for mammalian cell culture, and we showed that Ag43 was key in this

process. This finding corroborates with the previously observed induction of *agn43* in DMEM (Torres *et al.* 2002). Moreover, this growth medium is known to mimic optimal conditions of expression and secretion of virulence factors encoded in the LEE of A/E pathogens, particularly due to the presence of the signalling molecule sodium bicarbonate (Rosenshine *et al.* 1996, Kenny *et al.* 1997). However, the mechanism of this induction remains to be elucidated to define the exact conditions required for expression of these genes. Regarding the subcellular localisation, our results suggest that Ag43 is located at the apex and at the interface between two auto-associating cells in *E. coli* O157:H7. In *E . coli* K12, it has been shown that the Ag43 can be find all around the cell but with an higher concentration at or near the cell poles (Vejborg *et al.* 2009, Duvernoy *et al.* 2018). Further investigations will thus be required to confirm our observations. The Ag43 variants of *E. coli* O157:H7 and *E. coli* K12 have been shown to belong to two distinct phylogenetic subfamilies (van der Woude *et al.* 2008). Nevertheless, the three GATC sequences involved in the phase variation regulation were identified in both subfamilies, suggesting a potential regulation of the expression by the Dam methylase and the OxyR transcription regulator. Our results demonstrated that in the absence of OxyR or upon overexpression of Dam, *E. coli* O157:H7 CM454 autoaggregate in LB medium. This is the first evidence of the involvement of Dam and OxyR in the regulation of the autoaggregation phenomenon of *E. coli* O157:H7. Nonetheless, further work would be required to demonstrate the direct involvement of these proteins in the regulation of the transcription of *agn43*. Considering two identical copies of *agn43* are present in *E. coli* O157:H7, further functional analysis would be required to decipher their respective and potential regulation by phase variation like in *E. coli* K12 (Owen *et al.* 1996) as well as their contribution to autoaggregation. These results suggest that Ag43 represent an important colonisation factor of EHEC and provide new insights about the regulation of the autoaggregation in *E. coli* O157:H7.

Based on these data, future work could aims at identifying competitive strains for EHEC adhesion to ECM proteins in order to reduce the asymptomatic carriage in the bovine reservoir or identifying critical points of contamination of meat and carcasses in slaughterhouse. These results could lead to the development of new innovative strategies to limit the risk of contamination of the food chain by EHEC.

MATERIAL AND METHODS

Bacterial strains and culture conditions

Experiments were carried out using the non-toxigenic *E. coli* O157:H7 EDL933 *stx*$^-$, i.e. *E. coli* O157:H7 CM454 (Gobert *et al.* 2007). All bacterial strains and plasmids used are listed in Table 1. Bacteria were cultured in different relevant nutrient media, either complex undefined, i.e. BHI (brain heart infusion, Becton-Dickinson) and LB (lysogeny broth) (Bertani 2004), or chemically defined, i.e. DMEM (Dulbecco's modified eagle medium, Gifco). Briefly, strains were plated on LB agar medium and incubated overnight at 37°C and a preculture was set up from one colony in the relevant medium at 39°C (bovine temperature) or 37°C (human temperature) in an orbital shaker (150 rpm or 70 rpm) till stationary phase. When required, nutrient media were supplemented with antibiotics, i.e. chloramphenicol (25 $\mu g.ml^{-1}$) or ampicillin (100 $\mu g.ml^{-1}$), or kanamycin (40 $\mu g.ml^{-1}$). Arabinose at a final concentration of 0.2 % (w/v) was added in the medium to induce gene expression from pBAD/Myc-HisA vector (ThermoFischer Scientific).

Construction of in-frame deletion mutuants

The different isogenic mutants in genes encoding MSCRAMM proteins were obtained from *E. coli* O157:H7 CM454 using the one-step PCR-based method for inactivation of chromosomal genes described previously (Datsenko *et al.* 2000). Ten genes were targeted, namely *fliC* (z3013), *csgA* (z1676), *hcpA* (z0118), *escN* (z5119), *ehaA* (z0402), *ehaB* (z0469), *dam* (z4740), *oxyR* (z5519) and *cah1* (z1211) and *cah2* (z1651), two 100% identical genes at the nucleotide level both encoding the Ag43 autotransporter. Briefly, the chloramphenicol resistance cassette (chloramphenicol acetyltransferase) was amplified by high-fidelity PCR (LA Taq DNA polymerase, TaKaRa Bio Group) from plasmidic DNA (pKD3) using specific primers with homologous ends complementary to the flanking sequences of the target gene (Table S1). The amplicon was then purified using a QIAquick PCR Purification Kit (Qiagen) and electroporated into *E. coli* O157:H7 CM454 harbouring pKD46 the λ-red genes to promote homologous recombination. Positive colonies were selected after screening on LB agar supplemented with chloramphenicol. Primers outFw/outRv (Table S1) were further used to confirm the isogenic mutants were by colony PCR (GoTaq DNA Polymerase, Promega) and by sequencing (GATC-biotech). However, two different outRv primers were designed to confirm the deletion of *cah1* (z1211) and *cah2* (z1651) since the two genes are located on two 97% identical O-islands of approximately 87 Kb, namely OI-43 and OI-48 respectively. Therefore,

reverse primers were designed to target sequences located 11kb downstream of the genes. Colony PCR (LA Taq DNA polymerase) and sequencing were then used to confirm the double mutant.

Cloning and expression of genes encoding Dam and OxyR

Genes encoding *dam* (z4740) and *oxyR* (z5519) were amplified by PCR (LA Taq DNA polymerase) from chromosomal DNA of *E. coli* O157:H7 CM454 by introducing XhoI and HindIII (New England Biolabs) restriction sites. Primers were designed before be synthesised by Eurofins Genomics (Table S1). Digestive PCR products were ligated by the T4 DNA Ligase (Roche) into the pBAD/Myc-His A vector, hereafter called pBAD (ThermoFischer Scientific), previously digested by the same restriction enzymes. DNA sequencing (Mix2Seq Kit, Eurofins Genomics) was used to confirm the correct nucleotide sequence of the constructs. The subsequent plasmids were finally transformed into *E. coli* O157:H7 CM454 by electroporation prior to selection on LB agar supplemented ampicillin.

Coating of microtiter plates with ECM proteins

Coating of 96-wells polystyrene microplates (Falcon) with ECM proteins was realised as previously described (Hennequin *et al.* 2007). The ECM proteins consisted of and collagen I and (Millipore, 08-115) collagen III (Millipore, CC078). BSA (bovine serum albumin, Sigma, A3803) was used as an adhesion specificity control to ECM proteins. ECM proteins were solubilised in 0.1 M carbonate coating buffer (pH 9.6) and 250 µL were added in the wells at a saturating concentration of 50 µg ml^{-1} prior to overnight incubation at 4°C. The wells were then washed with a TS (tryptone salt) solution and used for the bacterial adhesion assay.

Bacterial adhesion assay

Bacterial adhesion assay was realised as previously described (Chagnot *et al.* 2013). Briefly, overnight precultures were diluted 1:100, grown and induced as described above in BHI at a pH of 7.2 (39°c or 37°C at low speed to preserve cell surface supramolecular structures). Sampling was performed during the mid-exponential growth phase at an $O.D._{600nm}$ of 0.5. Bacterial growth and de novo protein synthesis were blocked by adding 90 µg.ml^{-1} chloramphenicol. Bacterial suspension (200 µl) was then deposited in relevant protein coated-wells of the microtitre plate using wide-bore tips. Control wells were filled with sterile nutrient medium. Microtitre plates were incubated statically at 25°C for 3h. After incubation, bacterial suspension was removed by pipetting. Well were further washed with TS to remove loosely

attached cells prior to fix adherent bacterial cells with absolute ethanol. Wells were then emptied by pipetting and dried prior to staining with an aqueous-solution of crystal violet (0.1% w/v). After washing, the bound dye was solubilised from stained cells using an aqueous solution of acetic acid (33% v/v) and the absorbance at 595 nm read using a microtiter plate reader. The readings were corrected by subtracting the average absorbance from control wells.

Sedimentation assay

Sedimentation assay was performed as previously described (Chagnot *et al.* 2013). Briefly, overnight precultures were diluted 1:100 and grown in LB at 37°C under orbital shaking (150 rpm). Arabinose was added in early exponential phase (OD_{600nm} = 0.25) to induce the Dam and OxyR expression from pBAD when required. Cell suspensions were placed in conical tubes at the same OD_{600nm} of 0.5, incubated statically and vertically at 25°C and sampled from the top of the tube to measure the OD_{600nm} after 3h, 6h and 24h of incubation. Data are shown as the percentage of the initial OD_{600nm}. Observations in phase-contrast microscopy were made at the end of the experiments to confirm the bacterial cell autoaggregation.

Determination of the autoaggregation proportion

This experiment was adapted from a previous study (Chauhan *et al.* 2013). Basically, one hundred isolated colonies were selected to inoculate one hundred tubes of 3 mL of LB. These overnight precultures were diluted 1:100 and grown in 3 mL LB or DMEM at 37°C under orbital shaking (150 rpm) till stationary phase. Cultures were then incubated statically and vertically at 25°C. The autoaggregation properties of each individual culture were determined by visual observation of sedimented cells and microscopic observation. The autoaggregation proportion was then calculated as a percentage of autoaggregated cultures out of the total number of cultures.

RNA extraction

RNA extraction was performed on bacterial cells grown in LB or DMEM at 37°C. Briefly, overnight precultures were diluted 1:100 and grown in LB or DMEM at 37°C under orbital shaking (150 rpm) until the exponential growth phase. After incubation, bacterial suspensions were centrifuged at 3500g for 5 min at 4°C. The supernatants were discarded and the bacterial pellets were flash frozen in liquid nitrogen and resuspended in TE buffer (10 mM Tris-HCl, 1 mM EDTA; pH 8) in 2mL Graduated Skirted tube with tethered screw cap (BioSpec) containing 600 mg 0.1 mm diameter zirconia beads (Biospec), 500 µL of acid phenol

(pH 4), 50 µL of 10% Sodium Dodecyl Sulfate (SDS), and 3.5 µL β-mercaptoethanol (PROLABO). Bacterial cells were then lysed using a FastPrep®-24 instruments (MP Biomedicals) twice for 30 seconds with a speed of 6 m/s. Chloroform was added and tubes were gently mixed before a 20 min centrifugation at 12000g for 20 min at 4°C. The aqueous phase was recovered and total RNA was then purified using the Nucleospin® RNA (Macherey Nagel) according to the manufacturer recommendations. Remaining contaminant DNA was removed with using the TurboDNase protocol (Ambion, Thermofischer Scientific). The RNA concentration was measured using a Nanodrop ND-1000 spectrophotometer (Nanodrop technologies).

Reverse transcription and PCR

Reverse transcription of RNA samples was carried out using the SuperScript II reverse transcriptase kit (Invitrogen) and hybrid DNA was amplified by PCR (GoTaq DNA Polymerase, Promega) using primers designed to specifically amplify fragments of 90 to 200 bp (Table S1). The housekeeping gene *rrld* was used as a control of gene expression. RNA samples were also included to assess DNA contamination.

Image acquisition and analysis of phase contrast microscopic observations

Microscopic observations of bacterial cells were performed in phase contrast reflected light. Image acquisition was performed using an Axioplan 2 microscope coupled to an AxioCam MRm camera driven by the AxioVision 4.8 software (Carl Zeiss). A 20x/0.75 objective (Plan-NEOFLUAR) was used to acquire images.

Observations in confocal fluorescent microscopy

Overnight precultures were diluted 1:100 and grown in DMEM at 37°C under orbital shaking until OD_{600nm} = 0.5. Cells were then washed, resuspended in PBS and fixed in paraformaldehyde 2% (PFA) for 20 min on a poly-L-lysine-coated slide (Thermo scientific). After a blocking step in 10% normal goat serum (Invitrogen), slides were incubated with anti-Ag43 mouse serum and then with a goat anti-mouse IgG (Jackson Immuno-Research Laboratories). Rabbit polyclonal antibodies raised against whole cell *E. coli* associated with the secondary antibody Alexa Fluor488 goat anti-rabbit IgG (Life Technologies) were used to localise *E. coli* O157:H7 CM454. The samples were mounted using the Pro-Long Gold antifade reagent containing the blue-fluorescent nuclear counterstain DAPI (Invitrogen). Image acquisition was performed using a 64 x or 100 x oil objective (1.4 n.a.) mounted on a Zeiss

LSM710 confocal microscope. In the pictures the signal from Ag43 was pseudocoloured in green, while the signals from bacteria are shown in red.

Statistical analysis

Statistical analyses were performed using GraphPad Prism v6.01. At least three biological replicates were carried out for all statistical analyses. Error bars on the figures represent the 95% confidence interval for the mean of the independent experiments. The mean values obtained from the biological replicates were compared to that from the control. Data were statistically analysed following ANOVA one way or two way test, followed by the post-hoc Tukey test, with differences considered significant ($p<0.05$, *), very significant ($p<0.01$, **), highly significant ($p<0.001$, ***) or very highly significant ($p<0.0001$, ****).

ACKNOWLEDGEMENTS

This work was supported in part by INRA (Institut National de la Recherche Agronomique), by the CoMBa ("Colonisation de la Matrice extracellulaire Bacterienne") project funded by the FRI ("Fond Régional Innovation") cluster IRP ("Institut de Recherche Pharmabiotique") involving the industrial partners Lallemand SAS and Biofilm Control and by EU FP7 Marie Curie Actions ITN (Initial Training Network) EID (European Industrial Doctorates) DIS*Co* (A multidisciplinary Doctoral Industrial School on novel preventive strategies against *E. Coli* infections; http://www.discoproject.eu/; no FP7-PEOPLE-607611) project. Ricardo MONTEIRO was a Marie Curie PhD Research Fellow granted by ITN EID DIS*Co* and performed at INRA and GSK. Valentin AGEORGES is a PhD Research Fellow granted by the "Conseil Région Auvergne – FEDER (Fonds Européen de Développement Régional)".

Table 1: Plasmids and bacterial strains used in this study.

Name	Relevant genotype and/or phenotype	References
Plasmid		
pKD46	Vector containing the Red recombinase of phage λ, Amp^R	(Datsenko *et al.* 2000)
pKD3	Vector containing the *cat* gene, Cm^R	(Datsenko *et al.* 2000)
pBAD/Myc-HisA	Protein expression vector under the control of the inducible *ara*BAD promoter (P_{BAD}), Amp^R	Thermo Fischer Scientific
pBAD-Dam	Gene encoding Dam in XhoI-HindIII digested pBAD/Myc-HisA, Amp^R	This study
pBAD-OxyR	Gene encoding OxyR in XhoI-HindIII digested pBAD/Myc-HisA, Amp^R	This study
E. coli strain		
CM454	Non-toxigenic *E. coli* O157:H7 EDL933 (Δ*stx2AB*, *stx1AB*::kan), Kan^R	(Gobert *et al.* 2007)
CM454 pKD46	pKD46 in *E. coli* CM454, Amp^R	This study
CM454 pBAD	pBAD/Myc-HisA in *E. coli* CM454, Amp^R	This study
CM454 pBAD-Dam	pBAD-Dam in *E. coli* CM454, Amp^R	This study
CM454 pBAD-OxyR	pBAD-OxyR in *E. coli* CM454, Amp^R	This study
CM454 Δ*fliC*	CM454 deleted of z3013 encoding FliC, Cm^R	This study
CM454 Δ*csgA*	CM454 deleted of z1676encoding CsgA, Cm^R	This study
CM454 Δ*hcpA*	CM454 deleted of z0118 encoding HcpA, Cm^R	This study
CM454 Δ*escN*	CM454 deleted of z5119 encoding EscN, Cm^R	This study
CM454 Δ*agn43*	CM454 deleted of z1211 and z1651 encoding Ag43, Cm^R	This study
CM454 Δ*ehaA*	CM454 deleted of z0402 encoding EhaA, Cm^R	This study
CM454 Δ*ehaB*	CM454 deleted of z0469 encoding EhaB, Cm^R	This study

CM454 Δ*dam*	CM454 deleted of z4740 encoding Dam, Cm^R	This study
CM454 Δ*oxyR*	CM454 deleted of z5519 encoding OxyR, Cm^R	This study

SUPPLEMENTARY MATERIAL

Table S1: Oligonucleotides used in this study

Name	Oligonucleotide sequence (5'-3')
*dam*XhoIFw	CGCGCTCGAGATAATAAAGGAGGAATTAATCATGAAGA AAAATCGCGCTTTTTTGA
*dam*HindIIIRv	CGCGCAAGCTTTTATTTTTTCGCGGGTGAAACGACT
*oxyR*XhoIFw	CGCGCTCGAGATAATAAAGGAGGAATTAATCATGAATA TTCGTGATCTTGAGTACC
*oxyR*HindIIIRv	CGCGCAAGCTTTTAAACCGCCTGTTTTAACATTTA
pBADFw	ATGCCATAGCATTTTTATCC
pBADRv	CTGATTTAATCTGTATCAGG
*fliC*Fw	GGTGGAAACCCAATACGTAATCAACGACTTGCAATATA GGATAACGAATCGTGTAGGCTGGAGCTGCTTC
*fliC*Rv	ATCAGGCAATTTGGCGTTGCCGTCAGTCTCAGTTAATCA GGTTACAACGACATATGAATATCCTCCTTAGT
*csgA*Fw	TGCAGAACGGAATTGCACACAACAGACTGATACTTCTC TGTCTGATGAAGGTGTAGGCTGGAGCTGCTTC
*csgA*Rv	CAGAAAGTCATATTCAGCGTAGCCTGACCGTTATAGCC TTCAGCGCTGCTCATATGAATATCCTCCTTAGT
*hcpA*Fw	GCTGATGGCGTTTGCCAAAGTAGCACCAACCAAATCAA GGAGCGAAACAGGTGTAGGCTGGAGCTGCTTC
*hcpA*Rv	TGATAACGCAGGCACAGGGCCGTGAGCTGTGGAATATT CATTGCCGCTCCCATATGAATATCCTCCTTAGT
*escN*Fw	TTTTCAGGAGGTTGGGAATAATATCGAACTTAAAGTAT TAGGAACGGTAAGTGTAGGCTGGAGCTGCTTC
*escN*Rv	CAATCGGTTCGCTCTGCTTTTACGAATAGATAAAATTCT GTCCAACATACCATATGAATATCCTCCTTAGT
*cah*Fw	TGCAGAACGGAATTGCACACAACAGACTGATACTTCTC TGTCTGATGAAGGTGTAGGCTGGAGCTGCTTC
*cah*Rv	CAGAAAGTCATATTCAGCGTAGCCTGACCGTTATAGCC TTCAGCGCTGCTCATATGAATATCCTCCTTAGT
*ehaA*Fw	AATGCTATGCAAAATTTTTGAATTTTTATATGCGATGAA CAGTAACCGGAGTGTAGGCTGGAGCTGCTTC
*ehaA*Rv	AGCCGAATTATCTCCCGTTCGGCTCTCTCTCCTTTTACT CATTATTCAGACATATGAATATCCTCCTTAGT
*ehaB*Fw	CATTGATAATTGATGGATCAACTTATTACGTCCCTGAGG AGGGATGACAAGTGTAGGCTGGAGCTGCTT
*ehaB*Rv	ACGGGCATTATTTAGATTAAATTACCAGTATTTCTTCGG AGTGAAGAATACATATGAATATCCTCCTTAGT
*dam*Fw	GCTGTCGGAGCTTTCTCCACAGCCGGAGAAGGTGTAAT TAGTTAGTCAGCGTGTAGGCTGGAGCTGCTTC

*dam*Rv	CAGAATTGAGGGGGCAATCAAATACTGTTTCATCCGCT TCTCCTTGAGAACATATGAATATCCTCCTTAGT
*oxyR*Fw	TTGCTATTCTACCTATCGCCATGAACTATCGTGGCGATG GAGGATGGATAGTGTAGGCTGGAGCTGCTTC
*oxyR*Rv	TTAACTACCCGACGATGGCGGAAGCCTATCGGGTAGCT GCGTTAAACGGTCATATGAATATCCTCCTTAGT
*fliC*outFw	GTTATCGGCCTGAATTGCG
*fliC*outRv	CCACAGCGAGTGAATTGCG
*csgA*outFw	GTCTGCCAGGGGAAATGTC
*csgA*outRv	GCCGGTCGGCTTCATCACA
*hcpA*outFw	CTACTGGAAGTGTCTGGC
*hcpA*outRv	GATGCGGTAGGCATTGTC
*escN*outFw	CATGAATCCAACGGGGAACG
*escN*outRv	CAGCGATGCACAATATGCCT
*cah*outFw	GTCTGCCAGGGGAAATGTC
*cah1*outRv	TCTTCCGTAGTCGCTGATTG
*cah2*outR*v*	CTATCCACCATTGGGTATCG
*ehaA*outFw	CGGGAATCATTGACTCGCTG
*ehaA*outRv	CTTCAGGCGTGACGATATGG
*ehaB*outFw	CCATTCCTCTATAACGCCGC
*ehaB*outRv	GGCTGGTAGATGATATCCCG
*dam*outFw	CAGTTGAACAATGTGGCGGT
*dam*outRv	TACCTTTGCCAGTTCACTG
*oxyR*outFw	CGGTAAATGCGGTCGGGTGC
*oxyR*outRv	GCTGACTGCCTGCTCTATGC
RT-PCR*cah*FW	CCAATGCCAGCGGTAAGTTC
RT-PCR*cah*Rv	GTTCTGCCACTCAGACTTCC
RT-PCR*rrld*FW	CCCACCTATCCTACACATC
RT-PCR*rrld*Rv	CTGTGAAGATGCAGTGTACC

REFERENCES

Bertani, G. (2004). "Lysogeny at mid-twentieth century: P1, P2, and other experimental systems." J Bacteriol **186**(3): 595-600.

Chagnot, C., A. Agus, S. Renier, F. Peyrin, R. Talon, T. Astruc and M. Desvaux (2013). "In vitro colonization of the muscle extracellular matrix components by Escherichia coli O157:H7: the influence of growth medium, temperature and pH on initial adhesion and induction of biofilm formation by collagens I and III." PLoS One **8**(3): e59386.

Chagnot, C., A. Listrat, T. Astruc and M. Desvaux (2012). "Bacterial adhesion to animal tissues: protein determinants for recognition of extracellular matrix components." Cell Microbiol **14**(11): 1687-1696.

Chagnot, C., M. A. Zorgani, T. Astruc and M. Desvaux (2013). "Proteinaceous determinants of surface colonization in bacteria: bacterial adhesion and biofilm formation from a protein secretion perspective." Front Microbiol **4**: 303.

Chauhan, A., C. Sakamoto, J. M. Ghigo and C. Beloin (2013). "Did I pick the right colony? Pitfalls in the study of regulation of the phase variable antigen 43 adhesin." PLoS One **8**(9): e73568.

Croxen, M. A., R. J. Law, R. Scholz, K. M. Keeney, M. Wlodarska and B. B. Finlay (2013). "Recent advances in understanding enteric pathogenic Escherichia coli." Clin Microbiol Rev **26**(4): 822-880.

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

Donlan, R. M. (2002). "Biofilms: microbial life on surfaces." Emerg Infect Dis **8**(9): 881-890.

Duvernoy, M. C., T. Mora, M. Ardre, V. Croquette, D. Bensimon, C. Quilliet, J. M. Ghigo, M. Balland, C. Beloin, S. Lecuyer and N. Desprat (2018). "Asymmetric adhesion of rod-shaped bacteria controls microcolony morphogenesis." Nat Commun **9**(1): 1120.

Erdem, A. L., F. Avelino, J. Xicohtencatl-Cortes and J. A. Giron (2007). "Host protein binding and adhesive properties of H6 and H7 flagella of attaching and effacing Escherichia coli." J Bacteriol **189**(20): 7426-7435.

Gobert, A. P., M. Vareille, A. L. Glasser, T. Hindre, T. de Sablet and C. Martin (2007). "Shiga toxin produced by enterohemorrhagic Escherichia coli inhibits PI3K/NF-kappaB signaling pathway in globotriaosylceramide-3-negative human intestinal epithelial cells." J Immunol **178**(12): 8168-8174.

Haiko, J. and B. Westerlund-Wikstrom (2013). "The role of the bacterial flagellum in adhesion and virulence." Biology (Basel) **2**(4): 1242-1267.

Hasman, H., T. Chakraborty and P. Klemm (1999). "Antigen-43-mediated autoaggregation of Escherichia coli is blocked by fimbriation." J Bacteriol **181**(16): 4834-4841.

Henderson, I. R., M. Meehan and P. Owen (1997). "Antigen 43, a phase-variable bipartite outer membrane protein, determines colony morphology and autoaggregation in Escherichia coli K-12." FEMS Microbiol Lett **149**(1): 115-120.

Henderson, I. R., P. Owen and J. P. Nataro (1999). "Molecular switches--the ON and OFF of bacterial phase variation." Mol Microbiol **33**(5): 919-932.

Hennequin, C. and C. Forestier (2007). "Influence of capsule and extended-spectrum beta-lactamases encoding plasmids upon Klebsiella pneumoniae adhesion." Res Microbiol **158**(4): 339-347.

Hill, C. (2012). "Virulence or niche factors: what's in a name?" J Bacteriol **194**(21): 5725-5727.

Kaper, J. B., J. P. Nataro and H. L. Mobley (2004). "Pathogenic Escherichia coli." Nat Rev Microbiol **2**(2): 123-140.

Karch, H., M. Bielaszewska, M. Bitzan and H. Schmidt (1999). "Epidemiology and diagnosis of Shiga toxin-producing Escherichia coli infections." Diagn Microbiol Infect Dis **34**(3): 229-243.

Kenny, B., A. Abe, M. Stein and B. B. Finlay (1997). "Enteropathogenic Escherichia coli protein secretion is induced in response to conditions similar to those in the gastrointestinal tract." Infect Immun **65**(7): 2606-2612.

Kintz, E., J. Brainard, L. Hooper and P. Hunter (2017). "Transmission pathways for sporadic Shiga-toxin producing E. coli infections: A systematic review and meta-analysis." Int J Hyg Environ Health **220**(1): 57-67.

Mason, T. G. and G. Richardson (1981). "Escherichia coli and the human gut: some ecological considerations." J Appl Bacteriol **51**(1): 1-16.

McDaniel TK, J. K., Donnenberg MS, Kaper JB. (1995). "A genetic locus of enterocyte effacement conserved among diverse enterobacterial pathogens." Proceedings of the National Academy of Sciences of the United States of America 1664-1668.

Monteiro, R., V. Ageorges, M. Rojas-Lopez, H. Schmidt, A. Weiss, Y. Bertin, E. Forano, G. Jubelin, I. R. Henderson, V. Livrelli, A. P. Gobert, R. Rosini, M. Soriani and M. Desvaux (2016). "A secretome view of colonisation factors in Shiga toxin-encoding Escherichia coli (STEC): from enterohaemorrhagic E. coli (EHEC) to related enteropathotypes." FEMS Microbiol Lett **363**(16).

Mora, A., M. Blanco, J. E. Blanco, G. Dahbi, C. Lopez, P. Justel, M. P. Alonso, A. Echeita, M. I. Bernardez, E. A. Gonzalez and J. Blanco (2007). "Serotypes, virulence genes and intimin types of Shiga toxin (verocytotoxin)-producing Escherichia coli isolates from minced beef in Lugo (Spain) from 1995 through 2003." BMC Microbiol **7**: 13.

Nagy, A., J. Mowery, G. R. Bauchan, L. Wang, L. Nichols-Russell and X. Nou (2015). "Role of Extracellular Structures of Escherichia coli O157:H7 in Initial Attachment to Biotic and Abiotic Surfaces." Appl Environ Microbiol **81**(14): 4720-4727.

Owen, P., M. Meehan, H. de Loughry-Doherty and I. Henderson (1996). "Phase-variable outer membrane proteins in Escherichia coli." FEMS Immunol Med Microbiol **16**(2): 63-76.

Pruimboom-Brees, I. M., T. W. Morgan, M. R. Ackermann, E. D. Nystrom, J. E. Samuel, N. A. Cornick and H. W. Moon (2000). "Cattle lack vascular receptors for Escherichia coli O157:H7 Shiga toxins." Proc Natl Acad Sci U S A **97**(19): 10325-10329.

Rosenshine, I., S. Ruschkowski and B. B. Finlay (1996). "Expression of attaching/effacing activity by enteropathogenic Escherichia coli depends on growth phase, temperature, and protein synthesis upon contact with epithelial cells." Infect Immun **64**(3): 966-973.

Saldana, Z., J. Xicohtencatl-Cortes, F. Avelino, A. D. Phillips, J. B. Kaper, J. L. Puente and J. A. Giron (2009). "Synergistic role of curli and cellulose in cell adherence and biofilm formation of attaching and effacing Escherichia coli and identification of Fis as a negative regulator of curli." Environ Microbiol **11**(4): 992-1006.

Schembri, M. A., D. Dalsgaard and P. Klemm (2004). "Capsule shields the function of short bacterial adhesins." J Bacteriol **186**(5): 1249-1257.

Shaw RK, B. C., Feys B et al. (2008). "Enterohemorrhagic Escherichia coli exploits EspA filaments for attachment to salad leaves." Appl Environ Microbiol 2908-2914.

Tarr, P. I., C. A. Gordon and W. L. Chandler (2005). "Shiga-toxin-producing Escherichia coli and haemolytic uraemic syndrome." Lancet **365**(9464): 1073-1086.

Tenaillon, O., D. Skurnik, B. Picard and E. Denamur (2010). "The population genetics of commensal Escherichia coli." Nat Rev Microbiol **8**(3): 207-217.

Torres, A. G., N. T. Perna, V. Burland, A. Ruknudin, F. R. Blattner and J. B. Kaper (2002). "Characterization of Cah, a calcium-binding and heat-extractable autotransporter protein of enterohaemorrhagic Escherichia coli." Mol Microbiol **45**(4): 951-966.

Ulett, G. C., R. I. Webb and M. A. Schembri (2006). "Antigen-43-mediated autoaggregation impairs motility in Escherichia coli." Microbiology **152**(Pt 7): 2101-2110.

van der Woude, M. W. and I. R. Henderson (2008). "Regulation and function of Ag43 (flu)." Annu Rev Microbiol **62**: 153-169.

Vejborg, R. M. and P. Klemm (2009). "Cellular chain formation in *Escherichia coli* biofilms." Microbiology **155**(Pt 5): 1407-1417.

Vogeleer, P., Y. D. Tremblay, A. A. Mafu, M. Jacques and J. Harel (2014). "Life on the outside: role of biofilms in environmental persistence of Shiga-toxin producing Escherichia coli." Front Microbiol **5**: 317.

Wells, T. J., T. N. McNeilly, M. Totsika, A. Mahajan, D. L. Gally and M. A. Schembri (2009). "The Escherichia coli O157:H7 EhaB autotransporter protein binds to laminin and collagen I and induces a serum IgA response in O157:H7 challenged cattle." Environ Microbiol **11**(7): 1803-1814.

Wells, T. J., O. Sherlock, L. Rivas, A. Mahajan, S. A. Beatson, M. Torpdahl, R. I. Webb, L. P. Allsopp, K. S. Gobius, D. L. Gally and M. A. Schembri (2008). "EhaA is a novel autotransporter protein of enterohemorrhagic Escherichia coli O157:H7 that contributes to adhesion and biofilm formation." Environ Microbiol **10**(3): 589-604.

Xicohtencatl-Cortes, J., V. Monteiro-Neto, Z. Saldana, M. A. Ledesma, J. L. Puente and J. A. Giron (2009). "The type 4 pili of enterohemorrhagic Escherichia coli O157:H7 are multipurpose structures with pathogenic attributes." J Bacteriol **191**(1): 411-421.

Chapitre II : Etude de la diversité phylogénétique, structurale et fonctionnelle de l'Ag43

L'Ag43 est exposé à la surface de la cellule bactérienne *via* le T5SS et appartient à la famille des AT. Cette protéine partage donc les caractéristiques structurales communes aux AT : (i) un SP N-terminal, (ii) un domaine passager central et (iii) un domaine translocateur C-terminal (Desvaux *et al.* 2004). L'Ag43 possède toutefois des particularités avec la présence d'un domaine ESPR conservé dans la partie N-terminale du SP et d'un domaine AC en amont du translocateur (Desvaux *et al.* 2007, Rojas-Lopez *et al.* 2017). De plus, la cristallisation du domaine passager de l'Ag43a d'UPEC CFT073 a révélé une conformation en forme de L de l'hélice β (Heras *et al.* 2014). D'un point de vue fonctionnel, l'Ag43 est reconnu comme un membre des SAAT pour sa capacité à induire l'autoagrégation de la population bactérienne grâce à des interactions homotypiques entre Ag43 (Klemm 2006). Par ailleurs, des interactions hétérotypiques avec les SAAT TibA et AIDA-I ont aussi été observées (Klemm *et al.* 2004, Klemm 2006). La très grande majorité des études fonctionnelles consacrées à l'Ag43 a été réalisée chez *E. coli* K12 chez qui il est impliqué dans l'autoagrégation, l'adhésion et la formation de biofilms (van der Woude *et al.* 2008, Reidl *et al.* 2009). De plus rares études se sont par la suite intéressées à la caractérisation de cet AT chez des souches pathogènes telles que les EHEC et les UPEC (Torres *et al.* 2002, Ulett *et al.* 2007). Si EHEC O157:H7 EDL933 possède deux copies identiques du gène, UPEC CFT073 en arbore deux variants différents associés à des caractéristiques phénotypiques distinctes (Ulett *et al.* 2007). De plus, une analyse phylogénétique a révélé une distribution de l'Ag43 en deux sous-familles (SF-I et SF-II) (van der Woude *et al.* 2008). Ces observations suggèrent donc des différences structurales et fonctionnelles entre les variants de l'Ag43.

La seconde partie de mes travaux de thèse a donc premièrement porté sur l'étude approfondie de la phylogénie de l'Ag43 avec l'identification de nouvelles classes phylogénétiques caractérisées par des différences structurales. Un représentant prototypique de chaque classe a été sélectionné afin de comparer les capacités d'autoagrégation et d'interactions homotypiques et hétérotypiques de chaque variant (Article 2). Dans un second temps, l'identification de ces nouvelles classes phylogénétiques et la présence possible de plusieurs exemplaires d'Ag43 au sein d'un même génome ont orienté la dernière partie de mon travail vers l'étude de la diversité et de la prévalence des différentes classes d'Ag43 chez les bactéries (Article 3).

ARTICLE n°2

Valentin AGEORGES, Marion SCHIAVONE, Grégory JUBELIN, Nelly CACCIA, Philippe RUIZ, Xavier BAILLY, Etienne DAGUE, Sabine LEROY, Jason PAXMAN, Begoña HERAS, Frédérique CHAUCHEYRAS-DURAND, Amanda E. ROSSITER, Ian R. HENDERSON and Mickaël DESVAUX.

Differential homotypic and heterotypic interactions of antigen 43 (Ag43) variants in autotransporter-mediated bacterial autoaggregation

Publié dans Scientific Reports (2019), 9:11100

SCIENTIFIC REPORTS
natureresearch

OPEN

Differential homotypic and heterotypic interactions of antigen 43 (Ag43) variants in autotransporter-mediated bacterial autoaggregation

Received: 5 March 2019
Accepted: 18 June 2019
Published online: 31 July 2019

Valentin Ageorges[1], Marion Schiavone[2,3], Grégory Jubelin[1], Nelly Caccia[1], Philippe Ruiz[1], Ingrid Chafsey[1], Xavier Bailly[5], Etienne Dague[4], Sabine Leroy[1], Jason Paxman[6], Begoña Heras[6], Frédérique Chaucheyras-Durand[1,2], Amanda E. Rossiter[7], Ian R. Henderson[7,8] & Mickaël Desvaux[1]

Antigen 43 (Ag43) is a cell-surface exposed protein of *Escherichia coli* secreted by the Type V, subtype a, secretion system (T5aSS) and belonging to the family of self-associating autotransporters (SAATs). These modular proteins, comprising a cleavable N-terminal signal peptide, a surface-exposed central passenger and an outer membrane C-terminal translocator, self-recognise in a Velcro-like handshake mechanism. A phylogenetic network analysis focusing on the passenger revealed for the first time that they actually distribute into four distinct classes, namely C1, C2, C3 and C4. Structural alignment and modelling analyses demonstrated these classes arose from shuffling of two different subdomains within the Ag43 passengers. Functional analyses revealed that homotypic interactions occur for all Ag43 classes but significant differences in the sedimentation kinetics and aggregation state were present when Ag43^{C3} was expressed. In contrast, heterotypic interaction occurred in a very limited number of cases. Single cell-force spectroscopy demonstrated the importance of specific as well as nonspecific interactions in mediating Ag43-Ag43 recognition. We propose that structural differences in the subdomains of the Ag43 classes account for different autoaggregation dynamics and propensities to co-interact.

In lipopolysaccharidic (LPS) diderm bacteria (archetypical Gram-negative bacteria), nine protein secretion systems (T1SS-T9SS) are currently recognised as enabling protein transport across the outer membrane (OM), whereas the Sec and Tat export pathways enable protein translocation across the cytoplasmic membrane, also called the inner membrane (IM)[1–3]. Among those, the T5SS represents the simplest and most widely used mechanism for protein secretion in LPS-diderm bacteria[4–7], where proteins are translocated across the OM via a β-barrel with the assistance of some periplasmic chaperones (e.g. Skp, SurA or DegP), the BAM (β-barrel assembly machinery) and TAM (translocation and assembly module) complexes along the process[8–11]. Based on phylogenetic analyses and structural features, the T5SS is subdivided into the (i) classical autotransporters (T5aSS), (ii) two-partner passenger-translocators (T5bSS), (iii) trimeric autotransporters (T5cSS), (iv) hybrid autotransporters (T5dSS), and (v) inverted autotransporters (T5eSS)[11–14].

[1]Université Clermont Auvergne, INRA, UMR454 MEDiS, 63000, Clermont-Ferrand, France. [2]Lallemand Animal Nutrition, 31702, Blagnac, Cedex, France. [3]LISBP, Université de Toulouse, CNRS UMR5504, INRA UMR792, INSA, 31077, Toulouse, Cedex, France. [4]LAAS-CNRS, Université de Toulouse, 31077, Toulouse, Cedex, France. [5]INRA UR346 Epidémiologie Animale, 63122, Saint Genès Champanelle, France. [6]Department of Biochemistry and Genetics, La Trobe Institute for Molecular Science, La Trobe University, Bundoora, Victoria, 3086, Australia. [7]Institute of Microbiology and Infection, University of Birmingham, B152TT, Birmingham, United Kingdom. [8]Institute for Molecular Biosciences, University of Queensland, St. Lucia, Queensland, 4067, Australia. Correspondence and requests for materials should be addressed to M.D. (email: mickael.desvaux@inra.fr)

A classical autotransporter (AT) is a modular protein constituted of three main regions, i.e. (i) a N-terminal signal peptide (SP), (ii) a central passenger, and (iii) a translocator at the C-terminus[15,16]. As such, AT export via the Sec translocase is mediated by the SP before the passenger is secreted through the β-barrel pore formed by the translocator in the OM. Once secreted, the passenger can have three different fates as it may be (i) displayed at the bacterial cell surface, (ii) cleaved from the translocator and released as an exoprotein into the extracellular milieu, or (iii) processed but remain associated to the cognate translocator[10]. The functions of ATs are quite diverse and can be broadly categorised as (i) self-associating ATs (SAATs), (ii) protease ATs, (iii) phosphatase/hydrolase ATs, (iv) lipase/esterase ATs, (v) vacuolating ATs, (vi) adhesin ATs, and (vii) ATs of unknown function[3,10]. SAATs promote bacterial aggregation by cell self-association, a phenomenon called autoaggregation[17,18]. Besides flocculation and sedimentation of bacterial cells, autoaggregation is an important mechanism with consequences for bacterial virulence and infection due to biofilm formation, protection from environmental stresses, antibiotic resistance and defence against predators[18]. The antigen 43 (Ag43) is a prototypical member of the SAATs; this category of ATs also include AIDA (adhesin involved in diffuse adherence) and TibA (toxigenic invasion locus b)[17]. Besides homotypic interactions, heterotypic interactions were demonstrated between AIDA-Ag43, AIDA-TibA and Ag43-TibA[17,19]. By promoting interspecies cell-to-cell contact, these interactions would have important ecophysiological roles in inducing co-operative multispecies biofilms that allows for more successful microbial colonisation of various niches[20].

Besides the basic and common modular architecture of an AT, Ag43 exhibits some additional features. At the N-terminus, Ag43 displays a highly conserved extended SP region (ESPR) that slows the rate of export across the IM through the Sec-YidC pathway to prevent protein misfolding in the periplasm and consequently sustain the secretion-competent state of the passenger across the OM through the translocator[21,22]. Of note, the ESPR is not exclusive to Ag43 as it occurs in many different, but not all, ATs. N-terminal to the translocator, Ag43 exhibits a type 1 autochaperone domain (AC1), which is considered as essential for folding of β-helical passengers[3,23]. Like most ATs, the Ag43 passenger folds into a single-stranded right-handed parallel β-helix with triangular-shaped coil cross-sections[3,24]. Most recently, the crystal structure of the Ag43 from uropathogenic *E. coli* (UPEC) CFT073 (Ag43a) was solved and revealed the passenger displayed a twisted L-shape β-helical architecture[25]. The L-shaped conformation of the passenger further appeared as a requirement for driving the formation of cell aggregates and self-association in a Velcro-like handshake mechanism. While the function of Ag43 was primarily investigated in *Escherichia coli* K12[24], its implication in autoaggregation, adhesion and biofilm formation was further confirmed in some pathogenic species, namely enterohaemorrhagic *E. coli* (EHEC) and UPEC[26,27]. The Ag43 is encoded by two identical alleles in EHEC O157:H7 EDL933 but by two different alleles in UPEC CFT073; the UPEC Ag43a variant mediates a strong autoaggregation and biofilm formation phenotypes, whereas the Ag43b variant is a less efficient aggregating and biofilm forming factor[27]. Phylogenetic analyses had revealed the existence of two subfamilies (SF) of Ag43, namely SF-I and SF-II[24].

While homotypic Ag43 interactions have previously been noted, heterotypic interactions of Ag43 variants had not been extensively characterised as yet, it prompted us to investigate these aspects in a systematic study. In a cooperative evolution hypothesis that "birds of a feather flock together", different combinations of Ag43 variants might select bacterial interactions with siblings to the exclusion of others. From phylogenetic analyses of Ag43 passengers, it first appeared that the grouping was not dichotomous but clustered into four different classes associated to structural differences. Following the expression of representative Ag43 members, different propensities of the variants to induce autoaggregation were evidenced. By monitoring the formation of cell aggregates by epifluorescence microscopy and flow cytometry, homotypic and heterotypic interactions between different Ag43 variants were for the first time identified and characterised. Moreover, biophysical studies by atomic force microscopy (AFM) revealed different dynamics of Ag43 interactions suggesting that structural differences in the N-terminal and C-terminal subdomains of the Ag43 passengers can modulate the association of the different variants.

Results

Ag43 passengers cluster into four distinct phylogenetic classes with combination of two different subdomains. The functional properties of the modular protein Ag43 are determined by the passenger, which protrudes from the bacterial OM upon translocation. To apprehend the diversity of Ag43, a phylogenetic network analysis of the passenger was applied to a dataset of 106 distinct passengers generated from previously reported Ag43 sequences in the literature[24,28] (Supplementary Table S1). The network obtained defined for four main classes of passengers (C1 to C4) supported by robust splits (Fig. 1A).

C1 harbours the Ag43 passenger from *E. coli* K12 MG1655 (UPI00003B25EE), which is among the first identified and the most investigated Ag43[24]. C2 includes the Ag43 passenger from EHEC O157:H7 EDL933 (UPI00000D07E5), which is the first Ag43 investigated in a pathogenic *E. coli*[26]. C3 harbours the Ag43a (UPI00000E4A66) and Ag43b (UPI00000E49D2) from UPEC CFT073, the former being the first and only structurally characterised Ag43 passenger (PDB: 4KH3)[25]. To date, no Ag43 homologues from C4 have been functionally or structurally characterised. The original SF-I grouping corresponds to the gathering of C1 and C3, whereas SF-II grouping reflects C2 and C4[24]. The initial clustering correlated with the most obvious difference among the Ag43 passengers, their sizes; while Ag43 passengers from C1 and C3 contain around 500 amino acids, C2 and C4 homologues are smaller with about 430 amino acids (Supplementary Table S1).

Our updated phylogenetic network further revealed a strong ambiguity in the relatedness patterns among Ag43 passengers from C1, C2, C3, and C4, illustrated by the diamond shape of the network with almost no specific divergence. Therefore, a more detailed analysis of subdomains of the passengers was performed to identify the source of this ambiguity. Homologies to pectin lyase fold (IPR011050) overlapping with P22 tailspike protein (IPR012332) and AT adhesin (IPR030930) domains were systematically predicted for all Ag43 passengers, highlighting the β-helix topology encountered in SAATs (Supplementary Table S1). Modelling of the 3D structure

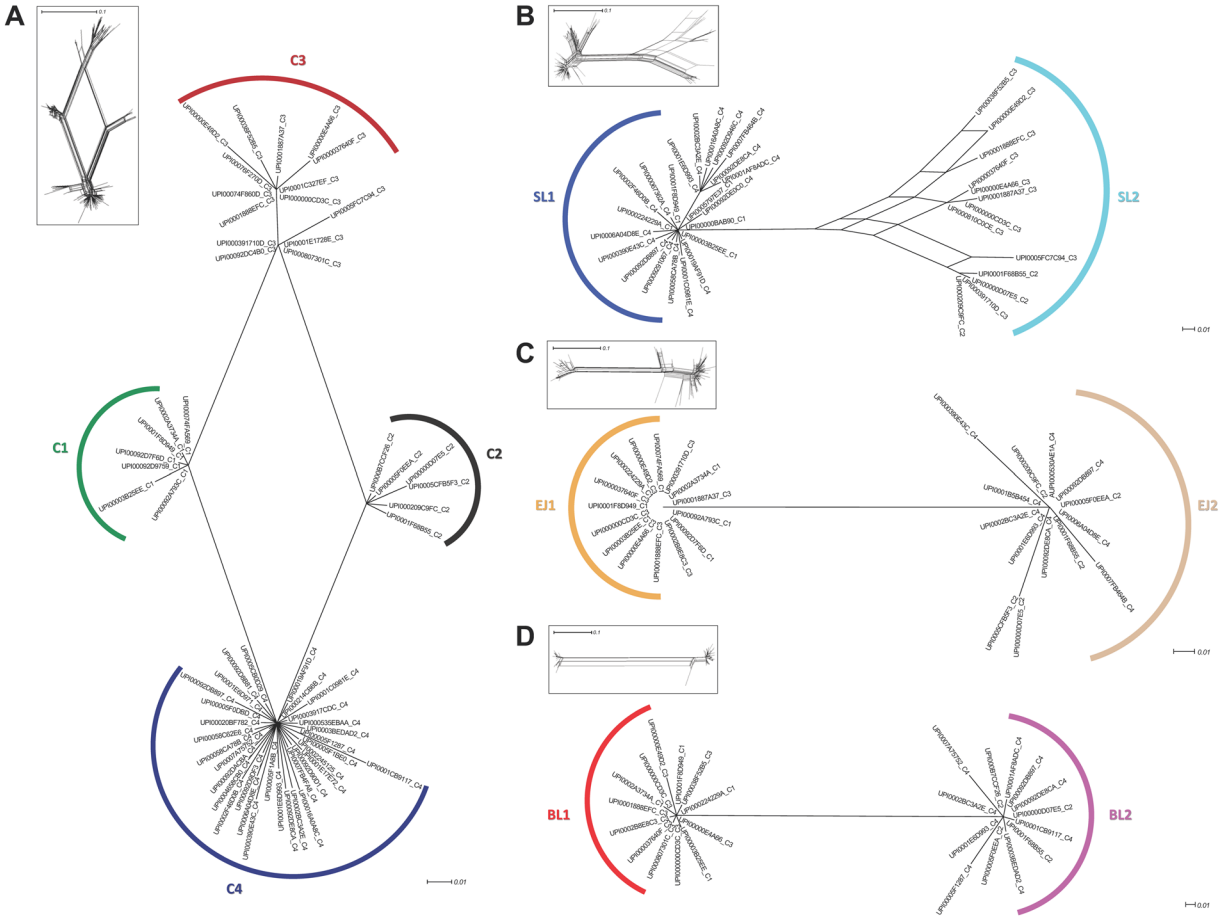

Figure 1. Phylogenetic analysis of Ag43 passengers and their subdomains. The amino acid sequences of the passengers were aligned using T-Coffee and structural information of Ag43$^{C3\text{-}UPI00000E4A66}$ (PDB: 4KH3) as template. Based on the L-shaped Ag43 structure (Fig. 2), the amino acid sequences of the SL, EJ and BL subdomains were recovered for alignment. Trees were generated using SplitsTree applying bootstrapping and 95% confidence for the observed nodes. Inserts represent the trees before applying bootstrapping and confidence threshold as provided in Supplementary Fig. S1. (**A**) Tree of the Ag43 passengers highlighting the four classes, C1, C2, C3 and C4. (**B**) Tree of the SL subdomain highlighting the two subtypes, SL1 and SL2. (**C**) Tree of the EJ subdomain highlighting the two subtypes, EJ1 and EJ2. (**D**) Tree of the BL subdomain highlighting the two subtypes, BL1 and BL2. For clarity, some sequences were removed from the trees but are readily available in Supplementary Table S1 and Fig. S1. The scale bars represent the number of substitutions per site.

of the Ag43 passengers confirmed that they are all predicted to exhibit a right-handed parallel β-helix fold with an L-shaped structure as originally observed in Ag43$^{C3\text{-}UPI00000E4A6625}$ (Fig. 2A). Following the analysis of the coil cross-sections along the Ag43 passengers[29], it appeared the N-terminal part, namely the SL (for stem of the letter L) region, systematically exhibited triangular (T) β-solenoids of subtype 4 (T4) (Fig. 2 and Supplementary Table S1). A phylogenetic network analysis of the SL subdomain identified two robust and unambiguously differentiated clusters of sequences, SL of type 1 (SL1) and SL2 (Fig. 1B). The two subtypes were related by a single branch, and remaining splits were located within SL2. Sequences from SL1 and SL2 subdomains had mean percentages of similarity of 93 and 88%, respectively, but these percentages fell at 75% when comparing SL1 versus SL2 (Supplementary Table S1). While SL1 clustered subdomains corresponding to Ag43^{C1} and Ag43^{C4} passengers, SL2 clustered subdomains from Ag43^{C2} and Ag43^{C3} passengers. Nevertheless, the SL region appeared slightly longer by 15–17 amino acids in Ag43^{C1} and Ag43^{C3} passengers.

The elbow region joining (EJ) the upper and lower parts of the L-shape was always bordered by β-hairpin motifs with three-rung parallel T5 β-solenoids in-between, which exhibited a highly conserved sequence in the first rung, namely VVLE[N/S]GGRLDVL. The EJ region reveals a remarkably conserved size of 85 amino acids. The phylogenetic network analysis of the EJ subdomain showed a different relatedness pattern compared to the SL subdomain. Two robust sequence subtypes were clustered, namely (i) EJ of type 1 (EJ1), which included Ag43 passenger subdomains from C1 and C3, and (ii) EJ2, which clustered subdomains from the C2 and C4 (Fig. 1C). In average, sequences from EJ1 and EJ2 subdomains had percentages of similarity of 97 and 94%, respectively, but these percentages fell at 53% when comparing EJ1 versus EJ2 (Supplementary Table S1).

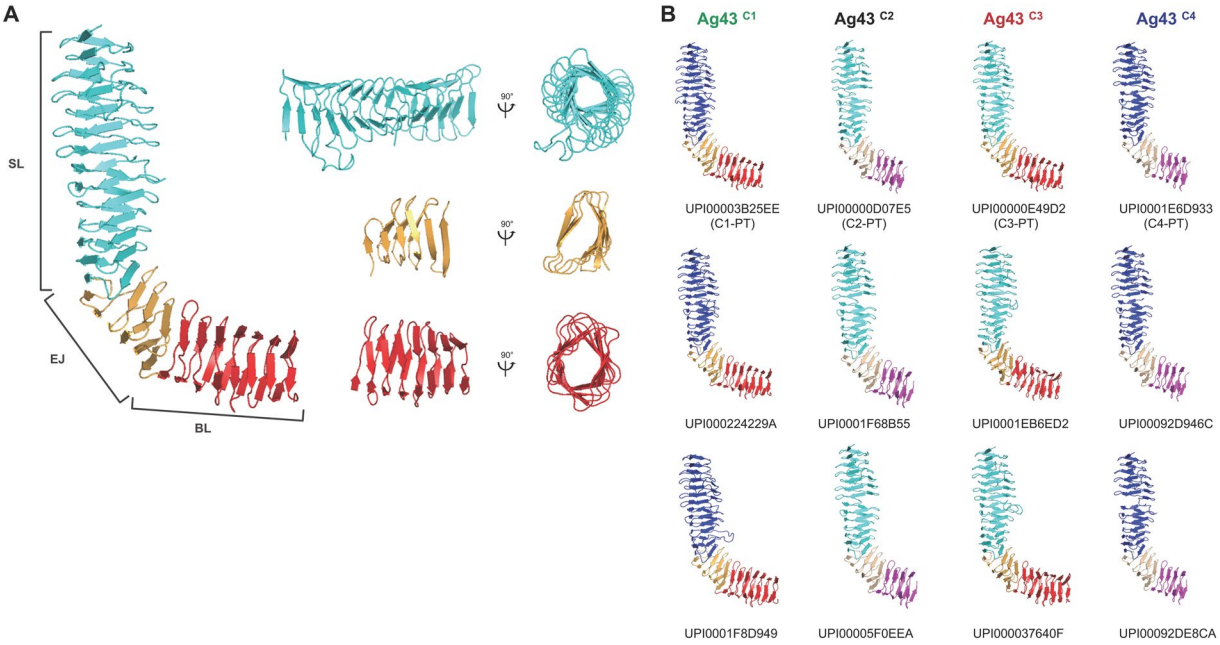

Figure 2. Modular structure of the Ag43 passengers belonging to classes 1, 2, 3 and 4. (**A**) Based on the experimentally determined L-shaped structure of Ag43$^{C3\text{-}UPI00000E4A66}$ (PDB: 4KH3), three subdomains could be defined SL (for stem of the letter L), EJ (for elbow joint), and BL (for bottom of the letter L)[25]. Both SL and BL subdomains exhibit a right-handed parallel β-helix, the cross-sections of each rung along SL form triangular (T) β-solenoids of subtype 4 (T4) and of T5-type β-solenoids for BL (Supplementary Table S1). The EJ subdomain is flanked by β-hairpin motifs with three-rung parallel T5 β-solenoids. (**B**) Molecular modelling of the structures of Ag43 passengers from C1 to C4 using Phyre[54,55]. The SL domain is shown in shades of blue, with electric blue corresponding to SL1 and cyan to SL2. The EJ domain is depicted in shades of orange, with bright orange corresponding to EJ1 and light orange to EJ2. The BL domain is represented in shades of red, with bright red corresponding to BL1 and magenta to BL2. PT: prototypical variant.

The BL (for bottom of the letter L) subdomain in the C-terminal part of the passengers is composed of T5-type β-solenoids (Fig. 2 and Supplementary Table S1). The phylogenetic network analysis of the BL was coherent with the analysis of the EJ subdomain. Two robust sequence subtypes were clustered: (i) BL of type 1 (BL1), which included Ag43 passengers subdomains from C1 and C3, and (ii) BL2, which clustered subdomains from C2 and C4 (Fig. 1D). Sequences from BL1 and BL2 subdomains had percentages of similarity from and above 95 and 93% (mean values of 99 and 98%), respectively, but these percentages fell below 30% (28% in average) when comparing SL1 versus SL2 (Supplementary Table S1). Furthermore, a difference of 53 amino-acids among the length of BL subdomains was observed. It accounted for most of the observed variation in the length of Ag43 passengers between C1/C3 over C2/C4 (Supplementary Table S1). This difference indeed resulted from a truncation at the C-terminus of BL2 compared to BL1.

These data suggest that the two SL phylogenetic clusters and the two coherent phylogenetic clusters of EJ and BL were shuffled during the evolution of Ag43 proteins explaining the observed ambiguity in relatedness pattern of the passengers (Fig. 1A). It resulted in four combinations of Ag43 passengers subdomains: (i) SL1-EJ1-BL1 corresponding to Ag43^{C1}, (ii) SL2-EJ2-BL2 corresponding to Ag43^{C2}, (iii) SL2-EJ1-BL1 corresponding to Ag43^{C3}, and (iv) SL1-EJ2-BL2 corresponding to Ag43^{C4}.

Autoaggregation occurs through homotypic interactions for all Ag43 variants but the size of the aggregates is significantly much smaller with Ag43^{C3}. Our new classification of Ag43 delineated the proteins into four classes based on sequence similarity and consequently structural differences. We hypothetised that these variations would result in functional differences. Therefore, we investigated the autoaggregation phenotype of three selected Ag43 homologues from each of the four classes (Fig. 2 and Table 1) when expressed in the non-aggregative *E. coli* SRUD[30]. When bacterial cell cultures were left standing statically, all bacterial cells expressing the Ag43 flocculated and sedimented at the bottom of tubes contrary to the control SRUD pBAD (with empty vector pBAB/Myc-HisA, i.e. Ag43^{-}) (Fig. 3A,B) (Table 1). No significant differences could be observed for the sedimentation kinetics of bacterial cells expressing Ag43 from C1, C2 and C4 with all three having compact sedimentation pellets. Microscopic observations indicated the flocculation and sedimentation of the bacterial cultures resulted from bacterial cell autoaggregation as no cell aggregates could be observed with SRUD pBAD, control (Fig. 3C).

In contrast, strains expressing the three C3 homologues (Fig. 2), namely Ag43$^{C3\text{-}UPI00000E4A66}$, Ag43$^{C3\text{-}UPI00000E49D2}$ and Ag43$^{C3\text{-}UPI0001EB6ED2}$, showed a significant delay in the autoaggregation kinetics compared to other Ag3 expressing strains (Fig. 3A). In addition, the sedimentation pellets were looser and cell aggregates appeared smaller, with the presence of individual single cells (Fig. 3B,C). Within C3, some significant differences were observed as bacterial cultures of SRUD expressing Ag43$^{C3\text{-}UPI00000E4A66}$ sedimented significantly quicker than

Name of *E. coli* strain[a]	Relevant genotype and/or phenotype
SRUD	*E. coli* K12 MG1655 Δ*fim* Δ*agn43*[30]
SRUD-Green pBAD	Empty pBAD/Myc-HisA in *E. coli* SRUD-Green
SRUD-Red pBAD	Empty pBAD/Myc-HisA in *E. coli* SRUD-Red
SRUD-Blue pBAD	Empty pBAD/Myc-HisA in *E. coli* SRUD-Blue
SRUD Ag43$^{C1\text{-}PT}$	pBAD-Ag43$^{C1\text{-}UPI00003B25EE}$ in *E. coli* SRUD
SRUD-Green Ag43$^{C1\text{-}PT}$	pBAD-Ag43$^{C1\text{-}UPI00003B25EE}$ in *E. coli* SRUD-Green
SRUD-Red Ag43$^{C1\text{-}PT}$	pBAD-Ag43$^{C1\text{-}UPI00003B25EE}$ in *E. coli* SRUD-Red
SRUD-Blue Ag43$^{C1\text{-}PT}$	pBAD-Ag43$^{C1\text{-}UPI00003B25EE}$ in *E. coli* SRUD-Blue
SRUD Ag43$^{C2\text{-}PT}$	pBAD-Ag43$^{C2\text{-}UPI00000D07E5}$ in *E. coli* SRUD
SRUD-Green Ag43$^{C2\text{-}PT}$	pBAD-Ag43$^{C2\text{-}UPI00000D07E5}$ in *E. coli* SRUD-Green
SRUD-Red Ag43$^{C2\text{-}PT}$	pBAD-Ag43$^{C2\text{-}UPI00000D07E5}$ in *E. coli* SRUD-Red
SRUD-Blue Ag43$^{C2\text{-}PT}$	pBAD-Ag43$^{C2\text{-}UPI00000D07E5}$ in *E. coli* SRUD-Blue
SRUD Ag43$^{C3\text{-}PT}$	pBAD-Ag43$^{C3\text{-}UPI00000E49D2}$ in *E. coli* SRUD
SRUD-Green Ag43$^{C3\text{-}PT}$	pBAD-Ag43$^{C3\text{-}UPI00000E49D2}$ in *E. coli* SRUD-Green
SRUD-Red Ag43$^{C3\text{-}PT}$	pBAD-Ag43$^{C3\text{-}UPI00000E49D2}$ in *E. coli* SRUD-Red
SRUD-Blue Ag43$^{C3\text{-}PT}$	pBAD-Ag43$^{C3\text{-}UPI00000E49D2}$ in *E. coli* SRUD-Blue
SRUD Ag43$^{C4\text{-}PT}$	pBAD-Ag43$^{C4\text{-}UPI0001E6D993}$ in *E. coli* SRUD
SRUD-Green Ag43$^{C4\text{-}PT}$	pBAD-Ag43$^{C4\text{-}UPI0001E6D993}$ in *E. coli* SRUD-Green
SRUD-Red Ag43$^{C4\text{-}PT}$	pBAD-Ag43$^{C4\text{-}UPI0001E6D993}$ in *E. coli* SRUD-Red
SRUD-Blue Ag43$^{C4\text{-}PT}$	pBAD-Ag43$^{C4\text{-}UPI0001E6D993}$ in *E. coli* SRUD-Blue

Table 1. Genotype and phenotype of most relevant *E. coli* strains used in the study. [a]The complete list of *E. coli* strains and plasmids used in this study is available in Supplementary Table S2. PT: prototypical variant.

those expressing Ag43$^{C3\text{-}UPI00000E49D2}$ or Ag43$^{C3\text{-}UPI0001EB6ED2}$. To confirm this slower autoaggregation phenotype for bacterial cells expressing Ag43^{C3}, three additional homologues from this class were tested; in each case, autoaggregation was slower (Fig. 3A).

To further characterise the homotypic interactions that are responsible for the autoaggreagtion phenotype, one prototypical (PT) representative was selected from each class (Table 1), namely Ag43$^{C1\text{-}PT}$ (UPI00003B25EE), Ag43$^{C2\text{-}PT}$ (UPI00000D07E5), Ag43$^{C3\text{-}PT}$ (UPI00000E49D2) and Ag43$^{C4\text{-}PT}$ (UPI0001E6D993). Fluorescence microscopy revealed that in the absence of Ag43 (Ag43^{-}, i.e. SRUD-Green pBAD and SRUD-Red pBAD) no cell aggregates could be observed (Fig. 4A). In contrast aggregates were detected for cells expressing Ag43 variants from each group (Fig. 4B–E). The combination of Ag43^{-} red-fluorescent-tagged cells mixed with Ag43^{+} green-fluorescent-tagged cells (i.e. SRUD-Green pBAD-Ag43) systematically exhibited only one type of bacterial cell aggregate, consisting uniquely of green-fluorescent-tagged cells where only a few Ag43^{-} cells could be observed as trapped within the aggregates (Fig. 4B–E). When green- and red-fluorescent cells expressing the same Ag43 were mixed equally, both fluorescent cells were found in the aggregates (Fig. 4F–I). These results confirmed that all tested Ag43 proteins, Ag43$^{C1\text{-}PT}$, Ag43$^{C2\text{-}PT}$, Ag43$^{C3\text{-}PT}$ or Ag43$^{C4\text{-}PT}$, enable homotypic interactions inducing bacterial autoaggregation. However, examination of the obtained micrographs seemed to indicate that more individual cells were present when Ag43$^{C3\text{-}PT}$ was expressed (Fig. 4H).

To gain quantitative insight, fluorescent nonaggregative and aggregative *E. coli* cells were analysed using flow cytometry. SSC-A (side scatter area) and FSC-A (forward scatter area) parameters are commonly used to discriminate different types of cells within a heterogeneous population with respect to granularity and cell size, respectively. Here, SSC-A over FSC-A dot plots allowed differentiating single cells from aggregated cells. As shown in Supplementary Fig. S3, flow cytometry analysis of the non-aggregative *E. coli* SRUD pBAD revealed a FCS-A/SSC-A pattern that is clearly distinct from the one obtained using the aggregative *E. coli* SRUD expressing Ag43$^{C1\text{-}PT}$. Comparison of these 2 samples allowed discriminating single cells from cell aggregates, the latter forming a comet tail in the upper part of the SSC-A over FSC-A representation (Supplementary Fig. S3). In agreement with microscopic examination, no cell aggregates were observed in the absence of Ag43, whereas cells expressing Ag43$^{C1\text{-}PT}$, Ag43$^{C2\text{-}PT}$, Ag43$^{C3\text{-}PT}$ and Ag43$^{C4\text{-}PT}$ formed aggregates (Supplementary Fig. S4). Furthermore, when green- with blue-fluorescent-tagged SRUD cells expressing the same Ag43 allele were combined, mixed aggregates formed (Fig. 5A); these represented around 90% of the total aggregates (Fig. 5B). As determined by flow cytometry analyses, quantification of the concentration of single individual cells showed a significant difference for cells expressing Ag43$^{C3\text{-}PT}$ compared to other Ag43 variants (Fig. 5C). Further, the shorter comet tail for *E. coli* SRUD Ag43$^{C3\text{-}PT}$ was indicative of smaller aggregates (Fig. 5A) and the mean size of the aggregates was confirmed to be significantly smaller for cells expressing Ag43$^{C3\text{-}PT}$ compared to any of the other Ag43 variants (Fig. 5D). To ensure the differences reported for Ag43$^{C3\text{-}PT}$ were not influenced by the protein expression levels, Western blot analyses were performed and indicated the Ag43 variants were expressed at similar levels in *E. coli* SRUD and displayed at the OM (Fig. S5).

Autoaggregation occurs only for some combinations of heterotypic interactions. To study heterotypic interactions, all six possible combinations of interactions between the prototypical Ag43 proteins from C1 to C4 were tested (Fig. 6). Fluorescent microscopic examinations revealed three major trends (i) large clumps

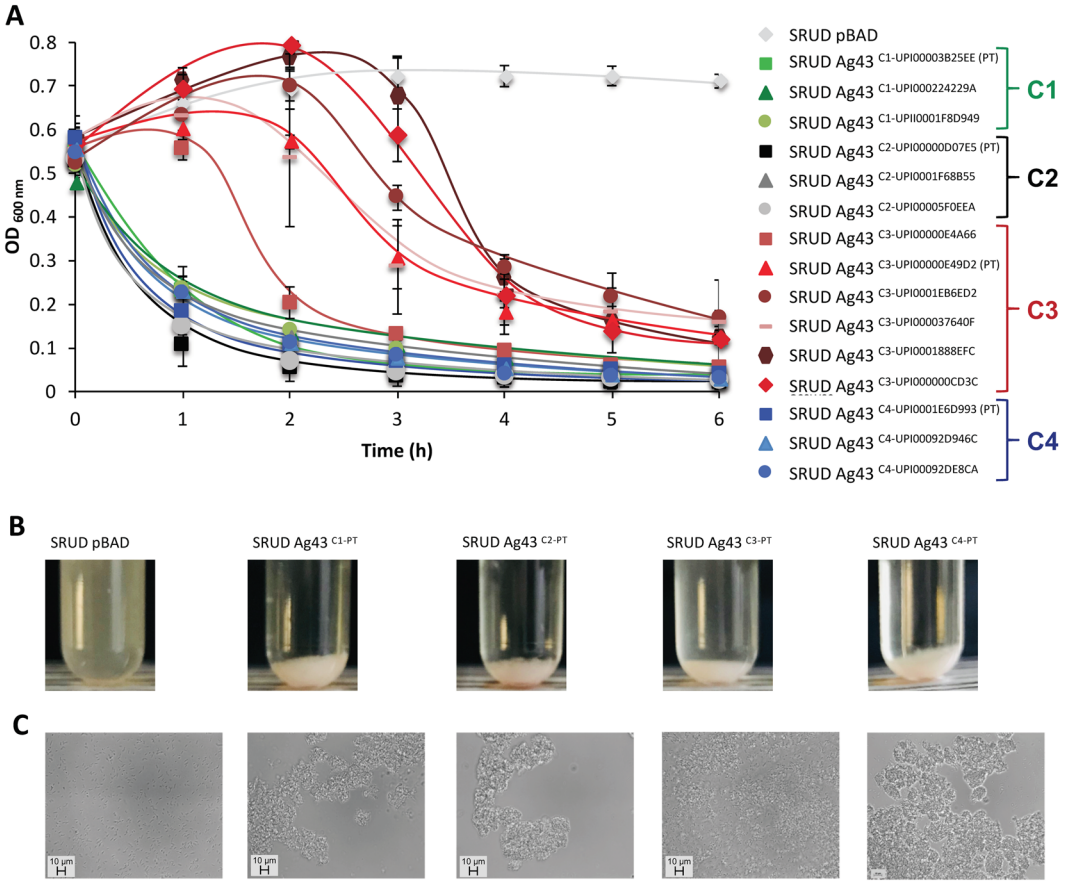

Figure 3. Autoaggregation of *E. coli* SRUD expressing different Ag43 variants. (**A**) The sedimentation assay was performed in LB at 37 °C over 6 h with SRUD expressing variants of Ag43 from C1 (shades of green), Ag43^{C2} (shades of grey), (iii) Ag43^{C3} (shades of red), Ag43^{C4} (shades of blue). The nonaggregative SRUD pBAD was used as negative control. (**B**) Sedimentation was visualised after 6 h of static incubation. (**C**) Phase-contrast microscopy observation from sampling at 6 h of static incubation. PT: prototypical variant (Table 1).

of bacteria containing only red- or green-fluorescent-tagged cells, suggesting only homotypic interactions occur (Fig. 6A,B), (ii) large aggregates containing only one type of fluorescent-tagged cell with smaller aggregates of the other tagged cells associated with the periphery, suggesting a weak propensity for heterotypic interactions (Fig. 6C,D), and (iii) clumps containing both green and red fluorescent cells suggesting the occurrence of heterotypic interactions (Fig. 6E,F).

Flow cytometry analyses confirmed the absence of heterotypic interaction between cells expressing Ag43$^{C1\text{-}PT}$ and Ag43$^{C2\text{-}PT}$ or cells expressing Ag43$^{C2\text{-}PT}$ and Ag43$^{C4\text{-}PT}$ (Fig. 7A) as the percentage of mixed aggregates were systematically estimated at less than 3% (Fig. 7B). Similarly, flow cytometry analyses confirmed the absence of significant heterotypic interactions as observed by epifluorescent microscopy for combinations of cells expressing Ag43$^{C1\text{-}PT}$/Ag43$^{C3\text{-}PT}$ or Ag43$^{C2\text{-}PT}$/Ag43$^{C3\text{-}PT}$, with a percentage of mixed aggregates estimated at less than 5% (Fig. 7A,B). However, combinations of cells expressing Ag43$^{C1\text{-}PT}$/Ag43$^{C4\text{-}PT}$ or Ag43$^{C3\text{-}PT}$/Ag43$^{C4\text{-}PT}$ showed significant heterotypic interactions with mixed cell aggregates estimated at just above 50% and 12%, respectively (Fig. 7A,B). In both cases, the percentage of mixed cell aggregates were significantly lower than the corresponding homotypic combinations (Fig. 5B). For all heterotypic combinations, no significant differences could be observed in the concentration of single cells (Fig. 7C). Heterotypic combinations resulted in aggregates with different sizes, some showing a mean size significantly larger than the corresponding homotypic combinations (i.e. Ag43$^{C1\text{-}PT}$/Ag43$^{C2\text{-}PT}$, Ag43$^{C1\text{-}PT}$/Ag43$^{C3\text{-}PT}$, and Ag43$^{C1\text{-}PT}$/Ag43$^{C4\text{-}PT}$) (Figs 5D and 7D). Interestingly, the association of cells expressing Ag43$^{C3\text{-}PT}$ with cells expressing any other class of Ag43 resulted in significantly lower concentration of single individual cells (Figs 5C and 7C) as well as significantly larger cell aggregates (Figs 5D and 7D) when compared to homotypic interactions.

Overall, heterotypic interactions between Ag43 from different classes is rather compromised, since no heterotypic interactions were observed in four out of the six possible combinations of prototypical Ag43 from C1 to C4. When heterotypic interactions occurred, the percentages of mixed aggregates were significantly lower than for homotypic interactions. Nonetheless, the highest percentage of mixed aggregates resulted from cells expressing Ag43$^{C1\text{-}PT}$/Ag43$^{C4\text{-}PT}$, both passengers consisting of the SL1 subdomain (Fig. 2B). Conversely, and despite exhibiting the same SL2 subdomain, no heterotypic interactions could be observed for cells expressing Ag43$^{C2\text{-}PT}$/Ag43$^{C3\text{-}PT}$. Surprisingly, some heterotypic interactions were observed between cells expressing Ag43$^{C3\text{-}PT}$/Ag43$^{C4\text{-}PT}$, which would involve interactions between SL1 and SL2 subdomains.

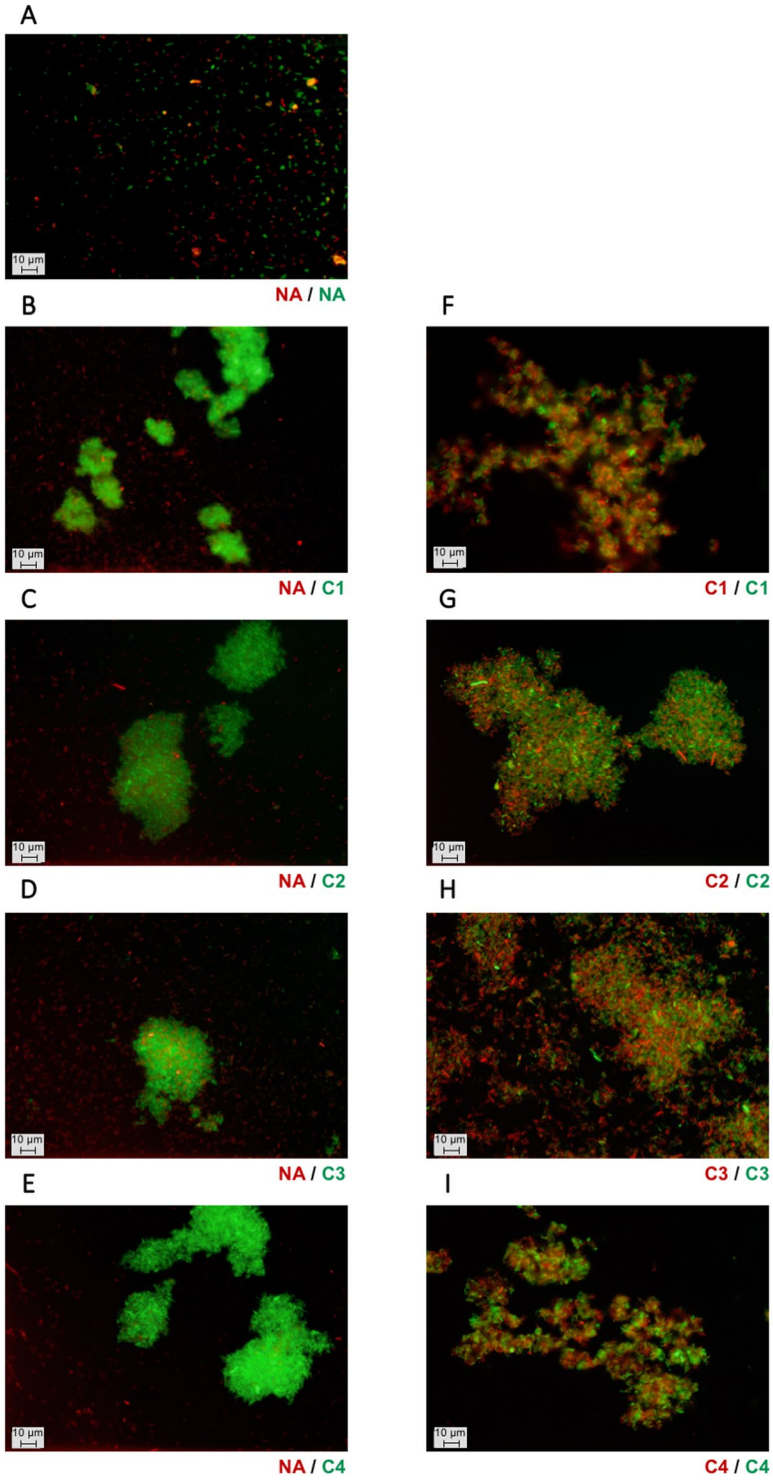

Figure 4. Fluorescent microscopy analysis of homotypic interactions of bacterial cells expressing different variants of Ag43. Epifluorescence microscopic observations were performed from bacterial cells grown in LB and after 6 h of static incubation at 37 °C. Co-culture of nonaggregative (NA) SRUD-Red pBAD with (**A**) SRUD-Green pBAD, (**B**) SRUD-Green Ag43$^{C1\text{-}PT}$, (**C**) SRUD-Green Ag43$^{C2\text{-}PT}$, (**D**) SRUD-Green Ag43$^{C3\text{-}PT}$, or (**E**) SRUD-Green Ag43$^{C4\text{-}PT}$. Co-culture of SRUD-Red and SRUD-Green both expressing (**F**) Ag43$^{C1\text{-}PT}$, (**G**) Ag43$^{C2\text{-}PT}$, (**H**) Ag43$^{C3\text{-}PT}$, or (**I**) Ag43$^{C4\text{-}PT}$. The NA, C1, C2, C3 and C4 labels match the colour of the fluorescent bacterial strains.

Ag43-mediated bacterial autoaggregation involves both specific and nonspecific interactions. Our data suggest the Ag43-Ag43 interactions could be stronger or weaker for some pairings. To quantify this and reinforce the hypothesis, single cell-force spectroscopy experiments were carried out to gain further insight

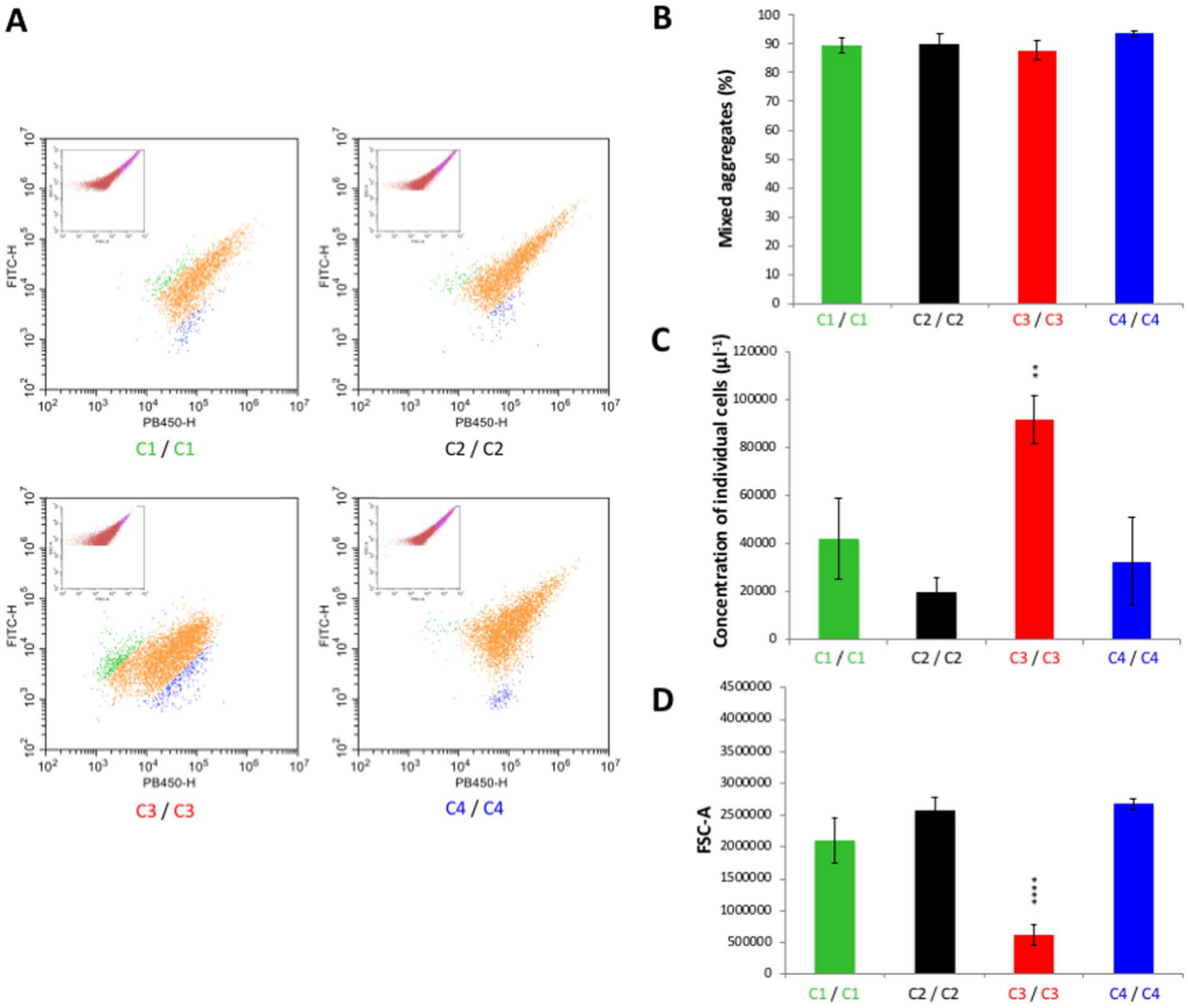

Figure 5. Characterisation of homotypic interactions of Ag43 variants by flow cytometry. Flow cytometry analyses were performed by mixing SRUD-Green and SRUD-Blue both expressing the same Ag43, namely Ag43 $^{C1\text{-}PT}$ (C1), Ag43$^{C2\text{-}PT}$ (C2), Ag43$^{C3\text{-}PT}$ (C3) or Ag43$^{C4\text{-}PT}$ (C4). (**A**) Based on the SSC-A over FSC-A representation (pictured in the insert), the fluorescent FITC-H over PB450-H representation focused on the gate corresponding to the cell aggregates as previously defined (Supplementary Fig. S3). The green cloud corresponds to cell aggregates of SRUD-Green, the blue cloud to cell aggregates of SRUD-Blue, and the orange cloud to mixed cell aggregates of SRUD-Green and SRUD-Blue. Graphs correspond to one representative experiment. (**B**) The percentage of mixed cell aggregates was calculated from the events detected in orange cloud over the total number of events corresponding to aggregates. (**C**) The concentration of cells was calculated from the gates corresponding to single individual cells as previously defined (Supplementary Fig. S3). (**D**) The size of aggregates was recovered from the gate corresponding to cell aggregates as previously defined (Supplementary Fig. S3).

in the adhesive forces at play in the autoaggregation process. First, the self-interaction between cells expressing different Ag43 variants was investigated by recording multiple force curves to measure the frequencies of adhesive events and adhesion forces between two *E. coli* cells (Table 2). For homotypic interactions, Ag43$^{C2\text{-}PT}$ expressing cells exhibited the highest percentage of adhesive events, whereas bacterial cells expressing Ag43$^{C3\text{-}PT}$ exhibited the lowest (approximately 72 and 28%, respectively). For homotypic associations of Ag43$^{C1\text{-}PT}$ cells, the distribution of the adhesion forces show the highest most probable rupture force with 113 pN (Table 2 and Supplementary Fig. S6). Interestingly, adhesion forces for Ag43$^{C2\text{-}PT}$ homotypic interactions showed a bimodal distribution with two maxima at 91 and 223 pN, respectively, which suggest that two different interaction conformations can exist for this protein (Table 2 and Supplementary Fig. S6). Force curves show that multiple adhesion events took place during the self-interaction of Ag43$^{C1\text{-}PT}$ or Ag43$^{C2\text{-}PT}$ expressing cells, whereas single interaction events were mostly recorded for Ag43$^{C3\text{-}PT}$ or Ag43$^{C4\text{-}PT}$ self-interaction (Fig. 8A). For cells expressing Ag43$^{C1\text{-}PT}$ or Ag43$^{C2\text{-}PT}$ but not Ag43$^{C3\text{-}PT}$ or Ag43$^{C4\text{-}PT}$, statistical analysis of the number of rupture events detected per force curves evidenced a maximum of 10 rupture events, indicating sequential unfolding of these homotypic interactions (Fig. 8B). Using the WLC (worm-like chain) model (allowing to measure the contour length of the

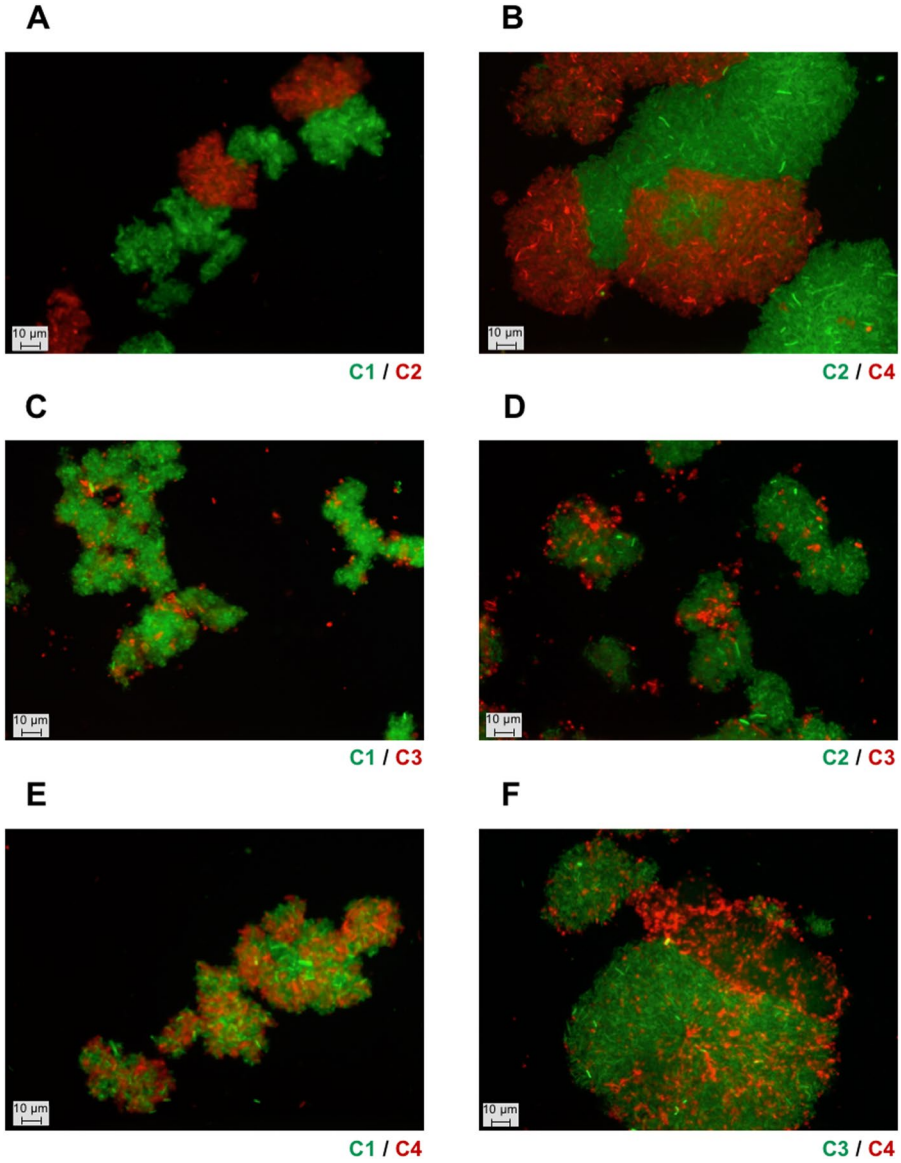

Figure 6. Fluorescent microscopy analysis of heterotypic interactions of *E. coli* cells expressing different variants of Ag43. (**A**) Co-culture of SRUD-Green $Ag43^{C1\text{-}PT}$ (C1) and SRUD-Red $Ag43^{C2\text{-}PT}$ (C2). (**B**) Co-culture of SRUD-Green $Ag43^{C2\text{-}PT}$ and SRUD-Red $Ag43^{C4\text{-}PT}$ (C4). (**C**) Co-culture of SRUD-Green $Ag43^{C1\text{-}PT}$ and SRUD-Red $Ag43^{C3\text{-}PT}$ (C3). (**D**) Co-culture of SRUD-Green $Ag43^{C2\text{-}PT}$ and SRUD-Red $Ag43^{C3\text{-}PT}$. (**E**) Co-culture of SRUD-Green $Ag43^{C1\text{-}PT}$ and SRUD-Red $Ag43^{C4\text{-}PT}$. (**F**) Co-culture of SRUD-Green $Ag43^{C3\text{-}PT}$ and SRUD-Red $Ag43^{C4\text{-}PT}$. The NA, C1, C2, C3 and C4 labels match the colour of the fluorescent bacterial strains.

unfolded molecule, which is defined as the length of the fully extended protein), the distribution of adhesion forces as a function of the contour length of unfolded Ag43 displayed different type of interactions (Fig. 8C). For $Ag43^{C2\text{-}PT}$, $Ag43^{C3\text{-}PT}$ and $Ag43^{C4\text{-}PT}$, the presence of adhesive events associated with unfolding at contour length of >400 nm and rupture forces (50–200 pN) suggests the presence of specific interactions between two cells via multiple electrostatic and/or hydrogen bonds. At the same time, adhesive events with contour length of <100 nm were observed for all Ag43 variants and suggests non-specific hydrophobic interactions. Taking together, these observations could explain the highest frequency of adhesive events measured for $Ag43^{C2\text{-}PT}$.

As mixed cell aggregates were observed between cells expressing $Ag43^{C4\text{-}PT}$ and (i) $Ag43^{C1\text{-}PT}$, as well as at lesser extent with (ii) $Ag43^{C3\text{-}PT}$, but not with (iii) $Ag43^{C2\text{-}PT}$, these heterotypic interactions were further investigated. A high frequency of adhesive events reaching approximately 74% was recorded between cells expressing $Ag43^{C4\text{-}PT}$ and $Ag43^{C1\text{-}PT}$ (Table 2). Conversely, this frequency remained moderate between cells expressing $Ag43^{C4\text{-}PT}$ and $Ag43^{C3\text{-}PT}$ (approximately 32%) and as expected extremely poor between $Ag43^{C4\text{-}PT}$ and $Ag43^{C2\text{-}PT}$ (approximately 4%), the latter indicative of an absence of significant interactions. Interestingly, the distribution of adhesion forces in heterotypic interactions showed two maxima corresponding to a bimodal distribution indicating the presence of two Ag43 binding conformations (Supplementary Fig. S7). For heterotypic interactions

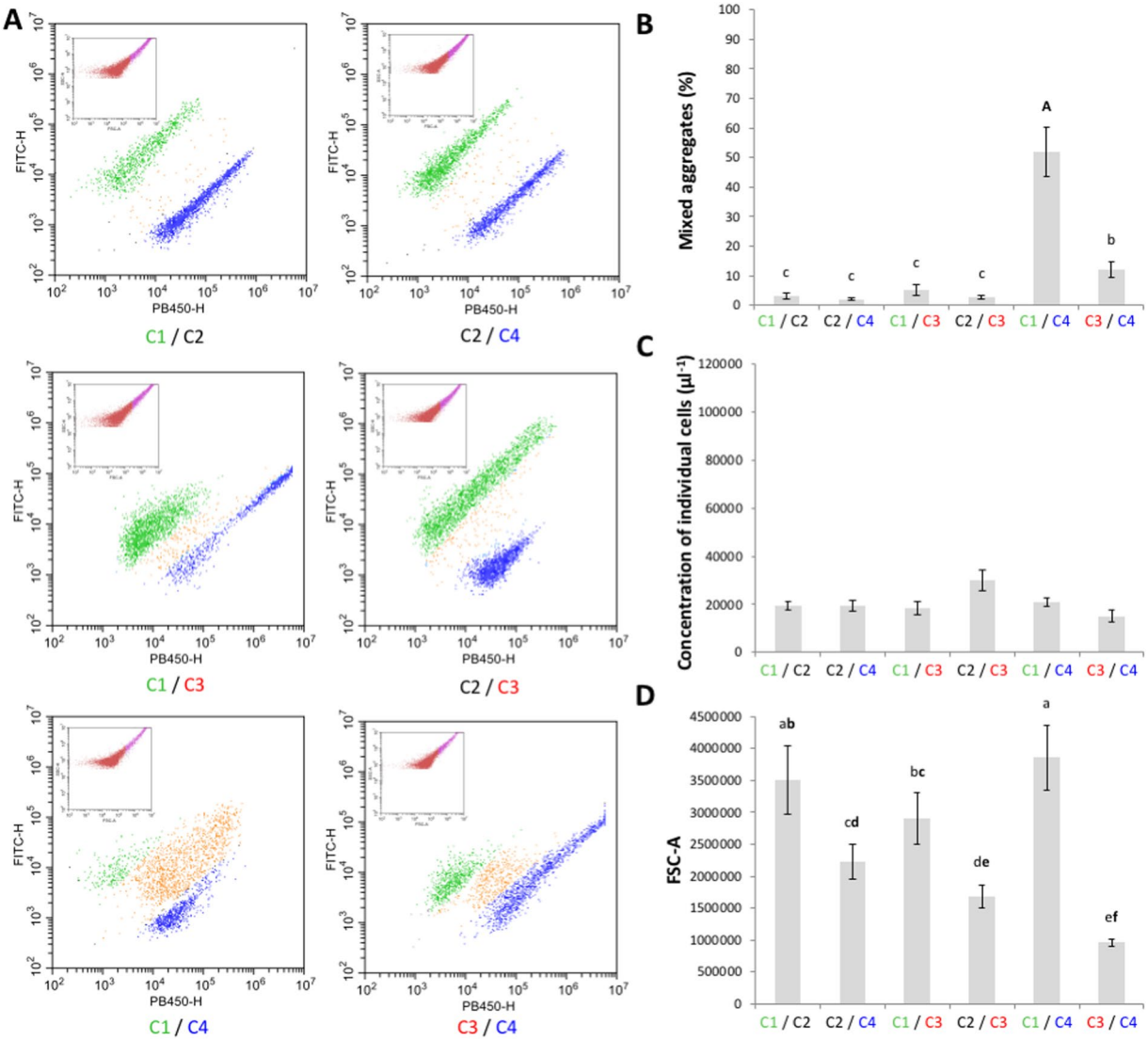

Figure 7. Characterisation of heterotypic interactions of Ag43 variants by flow cytometry. Flow cytometry analyses were performed by mixing SRUD-Green and SRUD-Blue expressing respectively Ag43$^{C1\text{-}PT}$ (C1) and Ag43$^{C2\text{-}PT}$ (C2), Ag43$^{C2\text{-}PT}$ and Ag43$^{C4\text{-}PT}$ (C4), Ag43$^{C1\text{-}PT}$ and Ag43$^{C3\text{-}PT}$ (C3), Ag43$^{C2\text{-}PT}$ and Ag43$^{C3\text{-}PT}$, Ag43$^{C1\text{-}PT}$ and Ag43$^{C4\text{-}PT}$, or Ag43$^{C3\text{-}PT}$ and Ag43$^{C4\text{-}PT}$. (**A**) Based on the SSC-A over FSC-A representation (pictured in the insert), the fluorescent FITC-H over PB450-H representation focused on the gate corresponding to the cell aggregates as previously defined (Supplementary Fig. S3). Graphs correspond to one representative experiment. (**B**) The percentage of mixed cell aggregates was calculated from the events detected in orange cloud over the total number of events corresponding to bacterial cell aggregates. (**C**) The concentration of single individual cells was calculated from the corresponding gate as previously defined (Supplementary Fig. S3). (**D**) The size of the aggregates was recovered from the gate for cell aggregates as previously defined (Supplementary Fig. S3).

between cells expressing Ag43$^{C4\text{-}PT}$ and Ag43$^{C1\text{-}PT}$, multiple rupture events were observed with a maximum of 21 rupture events (Fig. 9A,B). Additionally, Ag43$^{C1\text{-}PT}$ and Ag43$^{C4\text{-}PT}$ heterotypic association seemed to also exhibit both non-specific hydrophobic and several specific interactions via multiple electrostatic and/or hydrogen bonds (Fig. 9C). In summary, whenever the occurrence of homotypic or heterotypic interactions, AFM analyses emphasised the important role of both specific and nonspecific interactions in the Ag43-mediated bacterial autoaggregation.

Discussion

Previous phylogenetic analysis had simply divided Ag43 homologues into two distinct clusters, namely SF-I and SF-II[24]. Our further phylogenetic network analysis on the functional Ag43 passenger domains responsible for bacterial autoaggregation, has now revealed for the first time the existence of four distinct Ag43 phylogenetic classes, C1 to C4. Despite these divisions, structural modelling predicted that all Ag43 passengers form a similar L-shaped turn, the passenger structure themselves being subdivided into SL, EJ and BL subdomains. Ag43 passengers consistently displayed overall right-handed parallel β-helix fold with the exception of the two hairpins

Binding Couple[A]	Adhesive Events (%)[B]	Rupture Force (PN)[C]
Homotypic Interactions		
C1/C1	44.0 ± 9.0	113 ± 13
C2/C2	71.9 ± 14.2	91 ± 30; 223 ± 175
cC3/C3	27.8 ± 1.5	78 ± 16
C4/C4	37.3 ± 10.3	84 ± 108
Heterotypic Interactions		
C4/C2	4.2 ± 2.6	149 ± 90
C4/c1	73.8 ± 6.1	67 ± 30; 350 ± 200
C4/c3	31.9 ± 6.2	68 ± 40; 315 ± 100

Table 2. Percentage of adhesive events and values of the rupture force measured by single-cell force spectroscopy from bacteria expressing different Ag43 variants. [A]*E. Coli* SRUD was used to express the prototypical Ag43 variants from the different classes, namely Ag43$^{C1\text{-}PT}$ (C1), Ag43$^{C2\text{-}PT}$ (C2), Ag43$^{C3\text{-}PT}$ (C3) and Ag43$^{C4\text{-}PT}$ (C4). [B]Data correspond to mean value ± standard deviation and were extracted from AFM force-distance curves analysed with WLC model as described in the Methods section (Figs S4 and S5). [C]Data correspond to the most probable rupture force (most frequent and maximum value) ± the range of the distribution of the adhesion forces and were extracted from AFM force-distance curves analysed with WLC model as described in the Methods section (Figs S4 and S5). From bimodal distributions, two maxima could be calculated.

(each formed of two antiparallel β-sheets) delimiting the EJ subdomain. While there was evidence of structural variation reported in the literature[30,31], the present investigation evidenced that each SL, EJ and BL subdomain belong to two distinct phylogenetic types, whereby combinations of SL with EJ-BL subdomains divide the four different classes of Ag43 passengers. This new finding represents a major breakthrough in the field as it provides a rational framework for directed functional investigations of Ag43.

With this new understanding of the diversity of Ag43, the autoaggregation phenotype was reinvestigated based on this updated Ag43^{C1}, Ag43^{C2}, Ag43^{C3} and Ag43^{C4} classification. Following bacterial cell expression of several Ag43 from each of these classes, the most striking difference was based in the significant delay of the sedimentation kinetics for Ag43 proteins belonging to the class 3, compared to the Ag43^{C1}, Ag43^{C2} and Ag43^{C4}. While phenotypical analyses using different Ag43 have been reported in the scientific literature[30,31], no rationale had been previously described. Further, some contradictory results have also been conveyed[24] that could be attributed to differences in critical parameters in the experimental set-up, such as the culture media (e.g. complex over chemically defined) or the bacterial genetic background (e.g. the use of *E. coli* laboratory strain for Ag43 expression still containing *agn43* and/or *fim*). Here, using the nonaggregative *E. coli* SRUD strain and the same growth medium, the Ag43a from UPEC CFT073 (UPI00000E4A66) appeared to exhibit a slow sedimentation as originally reported[31]. Together with Ag43b from UPEC CFT073 (UPI00000E49D2) and other Ag43^{C3}, their sedimentation kinetics were significantly slower than Ag43^{C1}, Ag43^{C2} or Ag43^{C4}. Further, these aggregation experiments also revealed differences within C3 with Ag43a showing the strongest sedimentations properties in this class compared to other proteins like Ag43b, which showed weaker cell settling properties. Therefore, the latter Ag43 homologue better represents the phenotypical properties of this class and could be considered a prototypical Ag43^{C3}. Most of the current understanding on Ag43 stems from the work done on the prototypical Ag43$^{C1\text{-}UPI00003B25EE}$ from *E. coli* K12 MG1655, which has a SL1-EJ1-BL1 passenger. To date, only one article has been reported characterising Ag43$^{C2\text{-}UPI00000D07E5}$ from EHEC O157:H7 EDL933, which presents a SL2-EJ2-BL2 passenger[26]. Our work is the first to report the characterisation of Ag43 homologues from C4, such as Ag43$^{C4\text{-}UPI0001E6D993}$ from enterotoxigenic *E. coli* O78:H11 H10407, which exhibits a SL1-EJ2-BL2 passenger[32]. We show that bacterial cells expressing different Ag43 variants were able to induce autoaggregation through homotypic interactions as supported by microscopic examination and flow cytometry analyses (Fig. 10). However, differences were observed among the four Ag43 classes, with Ag43$^{C3\text{-}PT}$ showing slower sedimentation kinetics, looser autoaggregation pellets and significantly smaller aggregates associated with lower rupture force revealed by AFM analyses. With regard to single-molecule force spectroscopy analysis, it is worth noting that for Ag43$^{C1\text{-}PT}$ similar force curves and rupture forces of 100–200 pN were previously recorded[33].

Considering that heterologous associations have been reported between very different members of the SAATs family, namely Ag43, AIDA-I and TibA[17,19], the paucity of heterotypic interactions between Ag43 variants evidenced in this work was unexpected. While the recognition and involvement of binding interfaces in both homotypic and heterotypic interactions were suggested in the case of co-interacting SAATs, the restricted number of binding pairs respective to the different Ag43 variants suggest the involvement of different binding interfaces. Structural studies of homotypic interactions in Ag43$^{C3\text{-}UPI00000E4A66}$ revealed a Velcro-like handshake mechanism via trans-association of the SL subdomains stabilised by multiple hydrogen bonds and electrostatic interactions[25]. Ag43$^{C1\text{-}PT}$ and Ag43$^{C4\text{-}PT}$ share the SL1 subdomain and we could observe some heterotypic interactions between cells expressing those Ag43 homologues. Nonetheless, the level and force of this heterotypic association are significantly weaker than the corresponding homotypic interactions. Differences at amino acid level between Ag43$^{C1\text{-}PT}$ and Ag43$^{C4\text{-}PT}$ proteins probably result in a reduced number of interactions stabilising the hetero-dimers compared to the homodimers. Additionally, the EJ-BL subdomains could also play an indirect role in the Ag43-Ag43 interaction. Indeed, previous work showed that modification of the EJ subdomain altered the angle between the SL and BL subdomains and the overall L-shaped structure, completely abolishing the ability of cells expressing

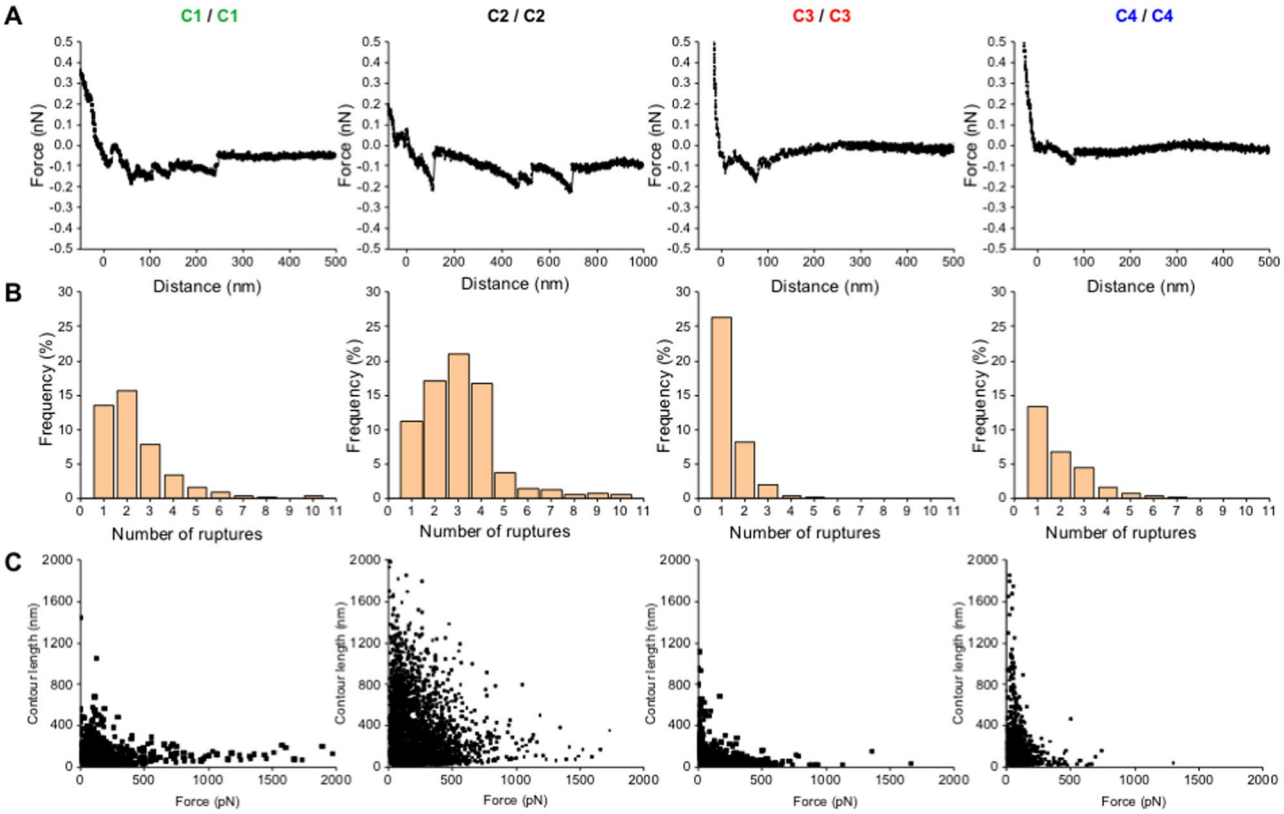

Figure 8. Retraction force curve and analysis of the rupture events during homotypic interactions of *E. coli* cells expressing different Ag43 variants. (**A**) Retraction force curves. (**B**) Rupture events measured per force curve. (**C**) Distribution of adhesion force measured between two *E. coli* cells expressing different Ag43 as function of the contour length for each force-distance curves. Analyses were performed using SRUD expressing the Ag43 from the different classes, namely $Ag43^{C1\text{-}PT}$ (C1), $Ag43^{C2\text{-}PT}$ (C2), $Ag43^{C3\text{-}PT}$ (C3) or $Ag43^{C4\text{-}PT}$ (C4).

Ag43 to form aggregates[25]. The absence of aggregates between cells expressing $Ag43^{C2\text{-}PT}$ and $Ag43^{C3\text{-}PT}$, despite both having a SL2, suggests this subdomain is less tolerant towards heterotypic associations and may be due to the binding interface regions being more selective in forming interactions. Similarly, no heterotypic interactions could be detected for combinations of Ag43 exhibiting different SL subdomain, namely SL1/SL2, except in the case of cells expressing $Ag43^{C3\text{-}PT}$ and $Ag43^{C4\text{-}PT}$. Considering the only two observed heterotypic interactions involved $Ag43^{C4\text{-}PT}$, it is tempting to hypothesise that the amino acid composition of SL1 and the EJ2-BL2 architecture in the passenger are more permissive to heterotypic interactions. Diverse EJ-BL subdomains may induce different bending angles to the L-shape Ag43 structure and modulate the ability to form heterologous associations. However, the molecular modelling approach used here to model the structure of different Ag43 homologues was not able to detect any differences as the angle between SL and BL biased the bending towards the $Ag43^{C3\text{-}UPI00000E49D2}$ template. Solving the experimental structures of Ag43 belonging to classes 1, 2 and 4, as well as other SAATs is critical to further inform the role of the different passenger subdomains in inducing autoaggregation and potentially predict heterologous associations among different SAATs.

Autoaggregation is a widespread phenomenon in bacteria but its ecophysiological importance is not fully elucidated. Several direct and side effects of autoaggregation have been reported in the literature, *e.g.* mechanism of adaptation against different stresses such as host defences as well as promoting cell invasion and colonisation[18]. The differential autoaggregation as well as the paucity of heterotypic interactions suggests that the Ag43 variants could play a role in structuring the *E. coli* population in an ecosystem by rallying or excluding some bacterial strains and thus stratifying the bacterial community. Considering that up to three *agn43* alleles can be present in one bacterial genome[24], further genomic investigations would be required to characterise the prevalence, distribution and diversity of Ag43, such as the combinations of classes and domain shuffling, among *E. coli* and within the Bacteria kingdom, as well the correlations with some features of the bacterial strains, such as the origin, the level of virulence or serotype. The reason why *E. coli* can carry different *agn43* alleles in one genome and express so many different variants of Ag43 is still unclear[24] but their implication in social behaviour especially through selective altruism following a greenbeard effect is an attractive concept[34–36]. Considering *agn43* can identify the presence of copies of themselves in other cells through Ag43-Ag43 interactions, this might induce some kind of nepotism behaviour toward those individual cells. The diversity of Ag43 variants and multiplicity of alleles could then favour differential and beneficial interactions between individual cells in a bacterial population.

These different aspects would undoubtedly require further investigations to understand how it could contribute to colonisation process and specific adaptation of *E. coli* within an ecosystem, both considering commensal

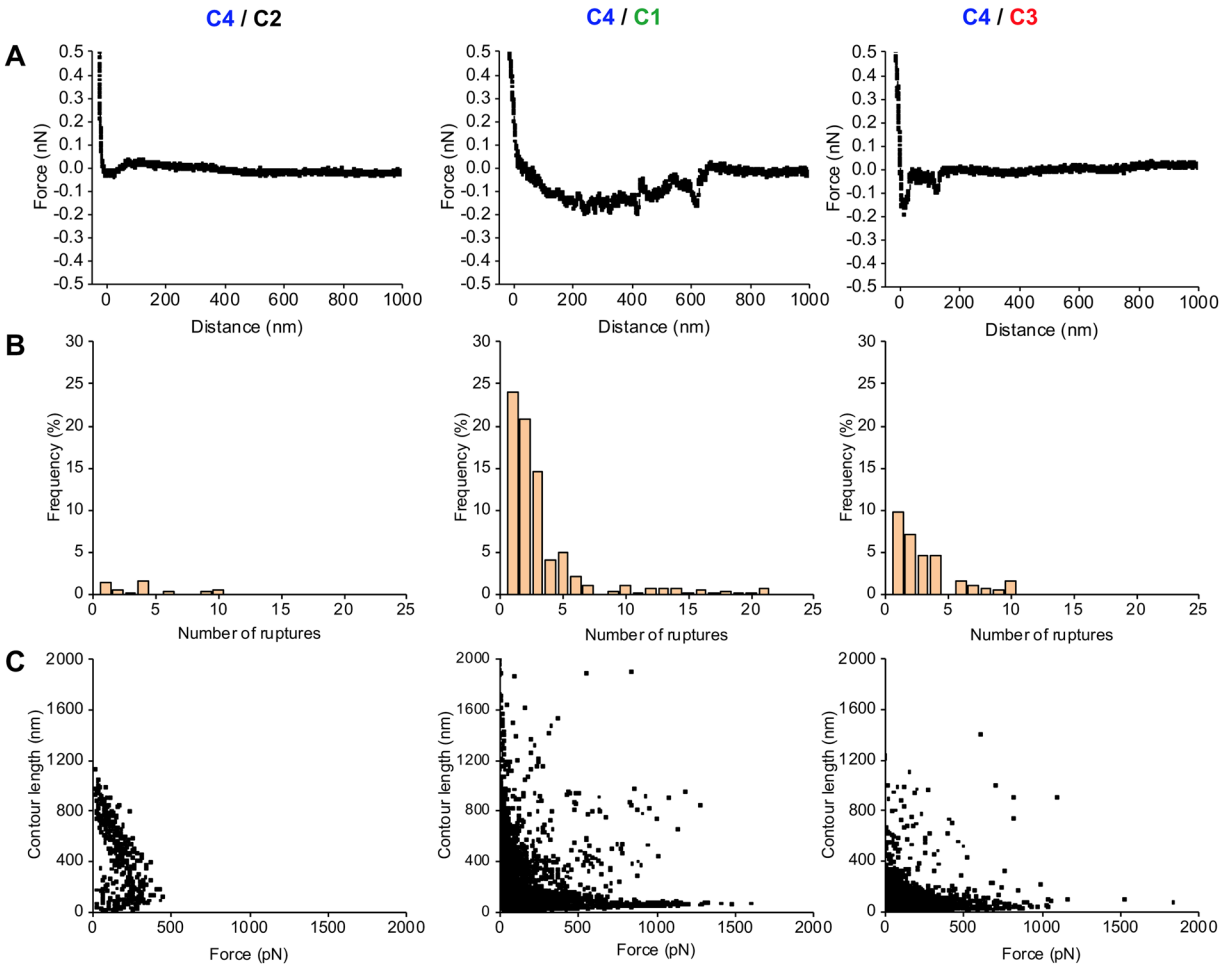

Figure 9. Retraction force curve and analysis of the rupture events during heterotypic interactions of *E. coli* cells expressing different Ag43 variants. (**A**) Retraction force curves. (**B**) Rupture events measured per force curve. (**C**) Distribution of adhesion force measured between two *E. coli* cells expressing different Ag43 as function of the contour length for each force-distance curves. Analyses were performed by mixing SRUD expressing Ag43 from different classes, respectively $Ag43^{C1\text{-}PT}$ (C1), $Ag43^{C2\text{-}PT}$ (C2), $Ag43^{C3\text{-}PT}$ (C3) or $Ag43^{C4\text{-}PT}$ (C4).

and pathogenic *E. coli* strains but also the complexity of bacterial communities, including the microbiota and the rest of the biotope as well as the biocenosis.

Methods

Bacterial strains and culture conditions. Experiments were carried out in an *E. coli* K12 MG1655 Δ*fim* Δ*agn43* deficient in type 1 fimbriae and Ag43 expression making the cells unable to autoaggregate, hereafter referred as *E. coli* SRUD (in honour to the inventors Sebastian Reidl and Ulrich Dobrindt)[30]. The most relevant bacterial strains used in this study are listed in Table 1; the exhaustive list of strains and plasmids is presented in details in Supplementary Table S1. Bacteria were cultured in LB (lysogeny broth) medium[37,38]. Briefly, strains were plated on LB agar medium and incubated overnight at 37 °C and a preculture was set up from one colony in LB broth at 37 °C in an orbital shaker (150 rpm). When required, LB medium was supplemented with antibiotics, i.e. ampicillin (100 $\mu g.ml^{-1}$), chloramphenicol (25 $\mu g.ml^{-1}$) or kanamycin (40 $\mu g.ml^{-1}$). Arabinose at a final concentration of 0.2% (w/v) was added in the medium to induce gene expression from pBAD/Myc-HisA vector (ThermoFischer Scientific).

Cloning and expression of genes encoding Ag43 variants. CDS encoding the different Ag43 variants (namely UniParc: UPI003B25EE, UPI000224229A, UPI0001F8D949, UPI00000D07E5, UPI0001F68B55, UPI00005F0EEA, UPI00000E4A66, UPI00000E49D2, UPI0001EB6ED2, UPI000037640F, UPI0001888EFC, UPI000000CD3C, UPI0001E6D993, UPI00092D946C, UPI00092DE8CA) were synthesised *de novo* by GeneArt (ThermoFischer Scientific) and cloned into the pBAD/Myc-HisA expression vector at the XhoI and HindIII restriction sites. Plasmid constructs were transformed into *E. coli* SRUD by electroporation using a Gene Pulser II and Pulse Controler Plus (Bio-Rad) (12.5 kV/cm, 200Ω, 25 μF) prior to selection on LB agar supplemented with ampicillin.

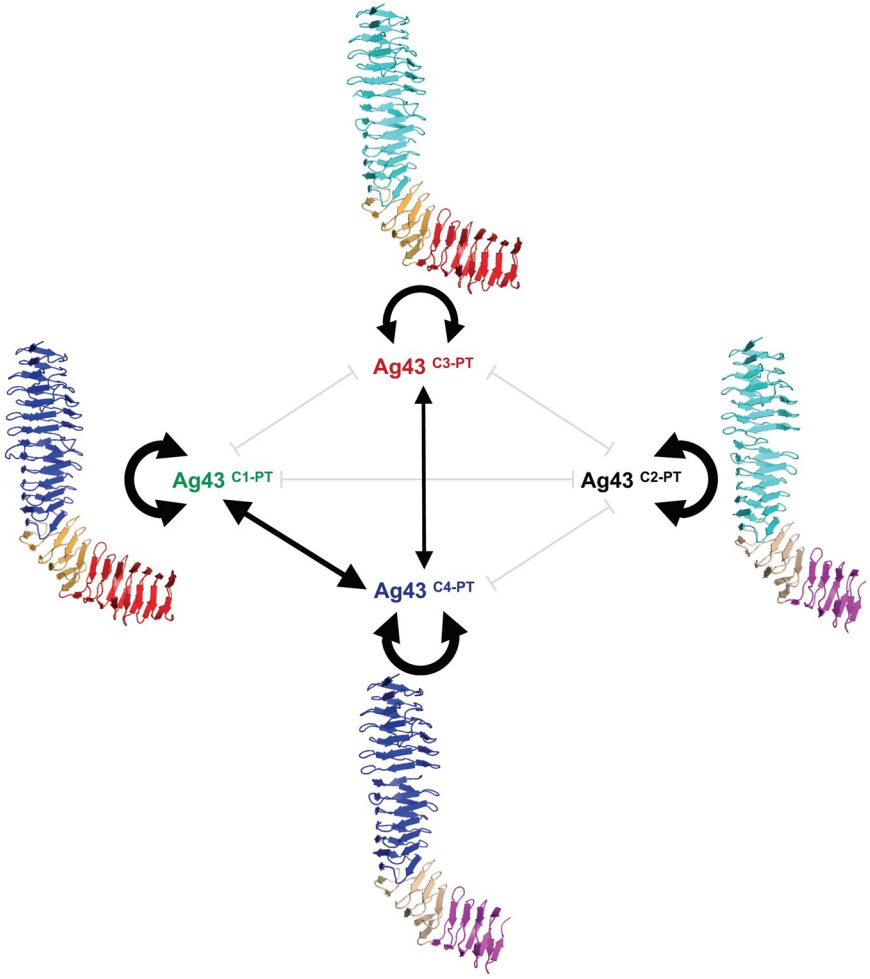

Figure 10. Summary of the different combinations of homotypic and heterotypic interactions of *E. coli* cells expressing Ag43 from the different classes in relation with the model structure of the passengers. Size of the black arrows schematically represent the force of the interactions. Grey lines indicate the tested combinations where no interaction could be experimentally evidenced. All structures of the Ag43 passengers were modelled using Phyre[54,55].

To obtain fluorescent bacterial strains, genes encoding for mRuby2, mTagBFP2 and mEmerald fluorescent proteins (GeneBank: AFR60232.1, AIQ82697.1 and AIA24532.1, respectively) were synthesised by GeneArt including a constitutive promoter (iGEM: BBa-J23119) and ribosome binding site (RBS) from pBAD/Myc-HisA, prior to cloning into the pACYC184 *Ava*I and *Bsu*36I/*Tth*111I restriction sites to generate pVaAgRed, pVaAgBlue and pVaAgGreen respectively, whose p15a replication origin belongs to a different incompatibility group from the pBR322 replication origin in pBAD-vector series. Plasmid constructs were transformed by electroporation in *E. coli* SRUD with or without a pVaAg-vector series and selection was carried on LB agar supplemented with chloramphenicol and when required with the additional supplementation of ampicillin.

Immunodetection and subcellular localisation of Ag43 variants. Bacterial cells were grown with arabinose till late exponential phase as described above. Cells were harvested by centrifugation (7 000 *g*, 5 min, 4 °C), washed and resuspended in PBS (0.01 M, 1 mM PMSF, pH8). Cells were disrupted with glass microbeads using FastPrep (20 s twice at 6 m.s^{-1}). Part of lysed whole cells was reserved and the other part subjected to cell fractionation using a modified sarkosyl (sodium lauroyl sarcosinate) method[39,40]. Briefly, unbroken cells and cellular debris were removed by centrifugation (12 000 *g*, 10 min, 4 °C) before the supernatant was ultracentrifuged (100 000 *g*, 40 min, 4 °C). The supernatant was retained as the soluble fraction (cytoplasm and periplasm), and the pellet was resuspended in 2% sarkosyl (20 min, room temperature) followed by another ultracentrifugation. The pellet corresponded to the sarcosyl-insoluble fraction gathering the OM. Cold acetone was added to the supernatant (overnight at −20 °C), centrifuged and the pellet corresponding to the sarcosyl-soluble fraction gathering the IM. Protein concentrations were determined using the Bradford method[41].

Proteins from the different fractions were separated by SDS-PAGE. Cell fractions were diluted in Laemmli buffer and heated up at 95 °C for 5 min. From prior testing with increasing amounts of whole cell fractions, 1 µg of total proteins was determined as a non-saturating concentration and equal amounts of each fraction based on the starting material were loaded per lane using Mini-PROTEAN Precast TGX Gels 4–15% (Bio-Rad). Precision Plus

Protein Dual Color Standards (Bio-Rad) was used to enable molecular mass estimation as well as transfer control. For Western blotting, proteins were dry transferred from SDS-PAGE on iBlot Transfer Stacks using iBlot 2 Dry Blotting System (Invitrogen). After saturating unoccupied protein-binding sites with blocking agents, membranes were probed with specific anti-Ag43 primary antibodies (dilution 1/5000 in PBS with 0.1% tween and 0.1% milk, 1 h, room temperature). Rabbit Ag43 antibodies were obtained as previously described[42,43] and are crossed reactive between Ag43 variants[44]. Goat anti-Rabbit IgG antibodies conjugated with horseradish peroxidase was used as secondary antibodies (dilution of 1/10 000, 1 h, room temperature) and TMB (3,3′,5,5′-tetramethylbenzidine) was used to detect the enzymatic activity. Enzymatic reaction was stopped by rinsing the membranes with milliQ water. Western blots were imaged under white light conversion screen using E-BOX (VILBER).

Phylogenetic analysis of protein sequences. The Ag43 protein sequences were retrieved from the literature[24,28] and the corresponding passengers recovered by removing the SP, AC1 and translocator regions. For identification of the SPs in Ag43 proteins, the sequences submitted to SignalP v4.1[45] were deleted from the ESPR (InterPro: IPR024973), which was first identified using InterProScan v5.22[46]. The AC1 domains were identified based on a structural alignment as previously described[3]. The translocators were identified using the twin HMM (Hidden Markov Model) strategy for AT detection[47]. Finally, truncated and redundant Ag43 passenger sequences were removed, resulting in a dataset of 106 distinct Ag43 passengers (Supplementary Table S1). Then, the amino acid sequences were aligned with T-Coffee in the Expresso mode for automatic incorporation of structural information[48,49] using the solved structure of Ag43 (PDB: 4KH3) as template. The relatedness among protein sequences was studied using a Neighbor-Net phylogenetic network approach with SplitsTree v4.14.8 using the Hamming uncorrected-P distance[50]. The most robust splits of the network were identified by bootstrap using a 95% threshold over 1,000 pseudo-replicates. A similar phylogenetic analysis was repeated for the different passenger subdomains. Pairwise sequence similarities and identities were obtained using Needleman-Wunsch alignment of two sequences with BLAST[51,52] (Supplementary Table S1).

Identification of conserved motifs and structural modelling. Conserved motifs were searched using InterProScan v5.22[46] and interrogating InterPro (IPR) v61.0 database[53]. Three-dimensional structures were modelled with Phyre v2.0[54,55] as well as I-TASSER v5.1[56,57]. The types of β-solenoid motif present in the β-helices were predicted using PfScan[29]. In β-helices, rungs were defined as adjacent parallel β-strands. For selected variants as presented in Fig. 2, Ag43 homolog models were checked and modified in the program COOT[58] using its Ramachandran, geometry and rotamer analysis validation software. The quality of the models were then assessed using ProCheck[59] with its Ramachandran errors reported as percentages of generously allowed and disallowed amino acids (Supplementary Table S3). Protein structures were visualised using PyMOL v1.7.4[60,61].

Autoaggregation assay. Autoaggregation assays were performed as previously described[62]. Briefly, overnight precultures were diluted 1:100 and grown in LB at 37 °C under orbital shaking. Arabinose was added in early exponential phase ($OD_{600\,nm} = 0.25$) to induce the Ag43 expression from pBAD-vector series (Supplementary Table S2). Cell suspensions were placed in conical tubes at the same $OD_{600\,nm}$, incubated statically and sampled every hour from the top of the tube to measure the $OD_{600\,nm}$. Observations in phase-contrast microscopy were made at the end of the experiments from sediment at the bottom of the tube to confirm the bacterial cell autoaggregation.

Image acquisition and analysis of phase contrast and epifluorescence microscopic observations. Microscopic observations of bacterial cells were performed in phase contrast and phase contrast epifluorescence reflected light. Image acquisition was performed using an Axioplan 2 microscope coupled to an AxioCam MRm camera driven by the AxioVision 4.8 software (Carl Zeiss). For phase contrast microscopy, a 40 × /0.75 objective (Plan-NEOFLUAR) was used and the fluorescence light source was a mercury lamp (HBO100W/2, AttoArc, Carl Zeiss). For fluorescence image acquisition a SYTO9-AO-FITC cube was fitted with a 450–490 nm exciter, 515–565 nm emitter, and 510 nm dichroic mirror as well as a Propidium iodide-Rhodamine (Con A) cube containing 546/12 nm exciter, 590 nm emitter, and 580 nm dichroic mirror (Carl Zeiss).

Quantification of bacterial cell interactions by flow cytometry. To investigate interactions of bacterial cells expressing Ag43 variants, *E. coli* SRUD-Green (carrying pVaAgGreen) or *E. coli* SRUD-Blue (carrying pVagAgBlue) were transformed with pBAD/Myc-HisA vectors encoding the different Ag43 variants or empty vector (negative control with non-aggregative *E. coli* SRUD bacterial cells) (Supplementary Table S2). Overnight *E. coli* SRUD-Green and *E. coli* SRUD-Blue precultures were diluted 1:100 and grown as described above in LB until early exponential phase. At this point, samples from these two different bacterial cultures were adjusted at an $OD_{600\,nm}$ of 0.25, mixed together at a 50/50 (v/v) ratio with arabinose and grown at 37 °C under shaking until an $OD_{600\,nm}$ of 0.5. Conical tubes were then incubated statically for 6 h and bacterial cells fixed with paraformaldehyde (2% final concentration) for 15 min. Cells were collected in the bottom of each tube and diluted in PBS (phosphate-buffered saline).

Samples were filtered on a 40-μm-nylon cell strainer prior to flow cytometry analysis (CytoFLEX, Beckman Coulter). FSC (forward scatter), SSC (side scatter), FITC (fluorescein isothiocyanate; 488 nm exciter, 525/540 nm emitter) and PB450 (Pacific Blue 450 nm; 405 nm exciter, 450/445 nm emitter) axes were set to log display. Areas for FSC and SSC (FSC-A and SSC-A, respectively) were used to discriminate aggregates from single individual cells, whereas heights for FITC and PB450 (FITC-H and PB450-H, respectively) were used to quantify co-interactions within aggregates (Supplementary Fig. S3). At least 3000 events/sec were recorded for each sample and data were analysed with CytExpert software v2.2 (Beckman Coulter). Results were expressed as percentages of mixed aggregates over the total aggregates.

Atomic force microscopy analyses. For cell probe preparation, bacterial cells were handled as for the autoaggregation assay prior to dilution in PBS buffer to an $OD_{600\ mn}$ of 0.2. One mL of bacterial suspension was directly deposited onto glass slide pre-coated with polyethylenimin (PEI) as previously described[63]. Briefly, cleaned glass slides are immersed in 0.2% PEI solution for 1 h, then extensively rinsed with Milli-Q water and finally dried under nitrogen) for 30 min and then the surface was rinsed twice with PBS buffer to get rid of un-adhered free bacteria and remaining residues from growth medium.

A Nanowizard III AFM (JPK-Bruker) was used to attach to atipless silicon nitride cantilevers (NP-010, Bruker) 6 µm borosilicate microspheres using UV curing resin (Norland optical adhesive 63 Glue, Norland Edmund Optics). Sphere attached cantilevers were cured under UV light ($\lambda = 400$ nm) for 3 min. The cantilever was then immersed for 1 h in a 0.2% PEI solution and washed. Functionalised cantilevers were calibrated before analysis by the thermal noise method[64] and spring constants were found to be 0.048 ± 0.011 N/m. Colloidal probe[65] was brought into contact with an isolated cell for 3 min, and the obtained cell probe was then moved to another cell to perform single cell-force spectroscopy measurements. Suitable parameters for cell attachment are a contact force of 1 nN, a contact time of 2 s and an approach and retract rate of 1 μms^{-1}.

Single cell-force spectroscopy experiments were carried out with a Nanowizard III atomic force microscope (JPK-Bruker). Measurements were performed at room temperature in PBS buffer at pH 7.2. Force-distance curves were recorded on different cells, using a maximum applied force of 1 nN and a constant approach and retraction speed of 1 $\mu m.s^{-1}$. The time of contact was set at 2 s to enhance specific interactions as described previously[66,67]. The type of interactions, *i.e.* specific or non-specific, could be deduced from the patterns of the force-distance curves[68–70].

The obtained force versus distance curves were analysed with the JPK Data Processing software on the basis of the worm-like chain (WLC) model, which is suitable to describe the extension of protein. The extension x of the molecule versus the pulling force (or adhesion force) is given by the following equation:

$$F(x) = \frac{k_B T}{l_p}\left[\frac{1}{4}\left(1 - \frac{x}{L_c}\right)^{-2} + \frac{x}{L_c} - \frac{1}{4}\right]$$

where k_B is the Boltzmann constant, T the absolute temperature, l_p is the persistent length and L_c is the contour length. The persistent length gives information about the degree of structural rigidity of the polymer chain and is defined as the longest segment below which the chain can be considered as a rigid rod, the extension of the polymer until the point at which the force necessary to extend further and in our case to break the interaction with the functionalised AFM tip corresponds to the contour length.

Adhesion peaks were generated using Origin 8 software. Statistical differences of AFM data were assessed by analysis of the variance (one-way ANOVA test) using Origin 8. To check the specificity of the measured adhesion forces and avoid artifacts that can be associated with cell probe preparation, control experiments were performed. Use of PEI-coated probes without bacterial cell interacting on PEI-coated glass slide led to a dramatic decrease of the adhesion frequency with only 3% of the force curves presenting an adhesion peak.

Statistical analysis. Statistical analyses were performed using GraphPad Prism v6.01. At least five biological replicates were carried out for all statistical analyses. Error bars on the figures represent the 95% confidence interval for the mean of the independent experiments. Data were statistically analysed following ANOVA one way or two way test, followed by the post-hoc Tukey test, with differences considered significant ($p < 0.05$, * or lowercase letter), very significant ($p < 0.01$, ** or bold lowercase letter), highly significant ($p < 0.001$, *** or capital letter) or very highly significant ($p < 0.0001$, **** or bold capital letter).

Data Availabiliy

All data generated or analysed during this study are included in this published article and its Supplementary Material files. Other data are available from the corresponding author upon reasonable request.

References

1. Desvaux, M., Parham, N. J., Scott-Tucker, A. & Henderson, I. R. The general secretory pathway: a general misnomer? *Trends in microbiology* **12**, 306–309, https://doi.org/10.1016/j.tim.2004.05.002 (2004).
2. Desvaux, M., Hébraud, M., Talon, R. & Henderson, I. R. Secretion and subcellular localizations of bacterial proteins: a semantic awareness issue. *Trends in microbiology* **17**, 139–145, https://doi.org/10.1016/j.tim.2009.01.004 (2009).
3. Rojas-Lopez, M. *et al*. Identification of the Autochaperone Domain in the Type Va Secretion System (T5aSS): Prevalent Feature of Autotransporters with a β-Helical Passenger. *Frontiers in microbiology* **8**, 2607, https://doi.org/10.3389/fmicb.2017.02607 (2017).
4. Yen, M. R. *et al*. Protein-translocating outer membrane porins of Gram-negative bacteria. *Biochimica et biophysica acta* **1562**, 6–31, https://doi.org/10.1016/S0005-2736(02)00359-0 (2002).
5. Abby, S. S. *et al*. Identification of protein secretion systems in bacterial genomes. *Scientific reports* **6**, 23080, https://doi.org/10.1038/srep23080 (2016).
6. Henderson, I. R. & Lam, A. C. Polymorphic proteins of *Chlamydia* spp.–autotransporters beyond the *Proteobacteria*. *Trends in microbiology* **9**, 573–578, S0966-842X(01)02234-X (2001).
7. Desvaux, M., Khan, A., Beatson, S. A., Scott-Tucker, A. & Henderson, I. R. Protein secretion systems in *Fusobacterium nucleatum*: genomic identification of Type 4 piliation and complete Type V pathways brings new insight into mechanisms of pathogenesis. *Biochim Biophys Acta* **1713**, 92–112, https://doi.org/10.1016/j.bbamem.2005.05.002 (2005).
8. Leyton, D. L., Rossiter, A. E. & Henderson, I. R. From self sufficiency to dependence: mechanisms and factors important for autotransporter biogenesis. *Nat Rev Microbiol* **10**, 213–225, https://doi.org/10.1038/nrmicro2733 (2012).
9. Selkrig, J. *et al*. Discovery of an archetypal protein transport system in bacterial outer membranes. *Nat Struct Mol Biol* **19**(506–510), S501, https://doi.org/10.1038/nsmb.2261 (2012).

10. Henderson, I. R., Navarro-Garcia, F., Desvaux, M., Fernandez, R. C. & Ala'Aldeen, D. Type V protein secretion pathway: the autotransporter story. *Microbiology and molecular biology reviews: MMBR* **68**, 692–744, https://doi.org/10.1128/MMBR.68.4.692-744.2004 (2004).
11. Leo, J. C., Grin, I. & Linke, D. Type V secretion: mechanism(s) of autotransport through the bacterial outer membrane. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **367**, 1088–1101, https://doi.org/10.1098/rstb.2011.0208 (2012).
12. Desvaux, M., Parham, N. J. & Henderson, I. R. Le système de sécrétion de type V chez les bactéries Gram-négatives (Type V secretion system in Gram-negative bacteria). *Biofutur* **237**, 34–37 (2003).
13. Desvaux, M., Parham, N. J. & Henderson, I. R. Type V protein secretion: simplicity gone awry? *Curr Issues Mol Biol* **6**, 111–124 (2004).
14. Tsai, J. C. *et al*. The bacterial intimins and invasins: a large and novel family of secreted proteins. *PloS one* **5**, e14403, https://doi.org/10.1371/journal.pone.0014403 (2010).
15. Desvaux, M., Parham, N. J. & Henderson, I. R. The autotransporter secretion system. *Research in microbiology* **155**, 53–60, https://doi.org/10.1016/j.resmic.2003.10.002 (2004).
16. Drobnak, I. *et al*. Of linkers and autochaperones: an unambiguous nomenclature to identify common and uncommon themes for autotransporter secretion. *Molecular microbiology* **95**, 1–16, https://doi.org/10.1111/mmi.12838 (2015).
17. Klemm, P., Vejborg, R. M. & Sherlock, O. Self-associating autotransporters, SAATs: functional and structural similarities. *Int J Med Microbiol* **296**, 187–195, https://doi.org/10.1016/j.ijmm.2005.10.002 (2006).
18. Trunk, T., Khalil, H. S. & Leo, J. C. Bacterial autoaggregation. *Aims Microbiol* **4**, 140–164, https://doi.org/10.3934/microbiol.2018.1.140 (2018).
19. Sherlock, O., Schembri, M. A., Reisner, A. & Klemm, P. Novel roles for the AIDA adhesin from diarrheagenic *Escherichia coli*: cell aggregation and biofilm formation. *Journal of bacteriology* **186**, 8058–8065, https://doi.org/10.1128/JB.186.23.8058-8065.2004 (2004).
20. Kjaergaard, K., Schembri, M. A., Ramos, C., Molin, S. & Klemm, P. Antigen 43 facilitates formation of multispecies biofilms. *Environ Microbiol* **2**, 695–702, https://doi.org/10.1046/j.1462-2920.2000.00152.x (2000).
21. Desvaux, M. *et al*. A conserved extended signal peptide region directs posttranslational protein translocation via a novel mechanism. *Microbiology* **153**, 59–70, https://doi.org/10.1099/mic.0.29091-0 (2007).
22. Jong, W. S. *et al*. YidC is involved in the biogenesis of the secreted autotransporter hemoglobin protease. *The Journal of biological chemistry* **285**, 39682–39690, https://doi.org/10.1074/jbc.M110.167650 (2010).
23. Oliver, D. C., Huang, G., Nodel, E., Pleasance, S. & Fernandez, R. C. A conserved region within the *Bordetella pertussis* autotransporter BrkA is necessary for folding of its passenger domain. *Molecular microbiology* **47**, 1367–1383, https://doi.org/10.1046/j.1365-2958.2003.03377.x (2003).
24. van der Woude, M. W. & Henderson, I. R. Regulation and function of Ag43 (*flu*). *Annual review of microbiology* **62**, 153–169, https://doi.org/10.1146/annurev.micro.62.081307.162938 (2008).
25. Heras, B. *et al*. The antigen 43 structure reveals a molecular Velcro-like mechanism of autotransporter-mediated bacterial clumping. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 457–462, https://doi.org/10.1073/pnas.1311592111 (2014).
26. Torres, A. G. *et al*. Characterization of Cah, a calcium-binding and heat-extractable autotransporter protein of enterohaemorrhagic *Escherichia coli*. *Molecular microbiology* **45**, 951–966, https://doi.org/10.1046/j.1365-2958.2002.03094.x (2002).
27. Ulett, G. C. *et al*. Functional analysis of antigen 43 in uropathogenic *Escherichia coli* reveals a role in long-term persistence in the urinary tract. *Infection and immunity* **75**, 3233–3244, https://doi.org/10.1128/IAI.01952-06 (2007).
28. Vo, J. L. *et al*. Autotransporter Adhesins in *Escherichia coli* Pathogenesis. *Proteomics* **17**, https://doi.org/10.1002/pmic.201600431 (2017).
29. Kajava, A. V. & Steven, A. C. The turn of the screw: variations of the abundant beta-solenoid motif in passenger domains of Type V secretory proteins. *J Struct Biol* **155**, 306–315, https://doi.org/10.1016/j.jsb.2006.01.015 (2006).
30. Reidl, S., Lehmann, A., Schiller, R., Salam Khan, A. & Dobrindt, U. Impact of O-glycosylation on the molecular and cellular adhesion properties of the *Escherichia coli* autotransporter protein Ag43. *Int J Med Microbiol* **299**, 389–401, https://doi.org/10.1016/j.ijmm.2009.01.001 (2009).
31. Klemm, P., Hjerrild, L., Gjermansen, M. & Schembri, M. A. Structure-function analysis of the self-recognizing Antigen 43 autotransporter protein from *Escherichia coli*. *Molecular microbiology* **51**, 283–296, https://doi.org/10.1046/j.1365-2958.2003.03833.x (2004).
32. Crossman, L. C. *et al*. A commensal gone bad: complete genome sequence of the prototypical enterotoxigenic Escherichia coli strain H10407. *Journal of bacteriology* **192**, 5822–5831, https://doi.org/10.1128/JB.00710-10 (2010).
33. Jacquot, A. *et al*. The dynamics and pH-dependence of Ag43 adhesins' self-association probed by atomic force spectroscopy. *Nanoscale* **6**, 12665–12681, https://doi.org/10.1039/c4nr03312d (2014).
34. Gardner, A. & West, S. A. Greenbeards. *Evolution* **64**, 25–38, https://doi.org/10.1111/j.1558-5646.2009.00842.x (2010).
35. West, S. A. & Gardner, A. Altruism, spite, and greenbeards. *Science* **327**, 1341–1344, https://doi.org/10.1126/science.1178332 (2010).
36. West, S. A., Griffin, A. S. & Gardner, A. Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection. *J Evol Biol* **20**, 415–432, https://doi.org/10.1111/j.1420-9101.2006.01258.x (2007).
37. Bertani, G. Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*. *Journal of bacteriology* **62**, 293–300 (1951).
38. Bertani, G. Lysogeny at mid-twentieth century: P1, P2, and other experimental systems. *Journal of bacteriology* **186**, 595–600, https://doi.org/10.1128/jb.186.3.595-600.2004 (2004).
39. Filip, C., Fletcher, G., Wulff, J. L. & Earhart, C. F. Solubilization of the cytoplasmic membrane of *Escherichia coli* by the ionic detergent sodium-lauryl sarcosinate. *Journal of bacteriology* **115**, 717–722 (1973).
40. Huang, C. Z. *et al*. Systematic identification of the subproteome of *Escherichia coli* cell envelope reveals the interaction network of membrane proteins and membrane-associated peripheral proteins. *J Proteome Res* **5**, 3268–3276, https://doi.org/10.1021/pr060257h (2006).
41. Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**, 248–254, https://doi.org/10.1016/0003-2697(76)90527-3 (1976).
42. Owen, P., Caffrey, P. & Josefsson, L. G. Identification and partial characterization of a novel bipartite protein antigen associated with the outer membrane of *Escherichia coli*. *Journal of bacteriology* **169**, 3770–3777, https://doi.org/10.1128/jb.169.8.3770-3777.1987 (1987).
43. Henderson, I. R., Meehan, M. & Owen, P. Antigen 43, a phase-variable bipartite outer membrane protein, determines colony morphology and autoaggregation in *Escherichia coli* K-12. *FEMS microbiology letters* **149**, 115–120, S0378-1097(97)00065-7 (1997).
44. Roche, A., McFadden, J. & Owen, P. Antigen 43, the major phase-variable protein of the *Escherichia coli* outer membrane, can exist as a family of proteins encoded by multiple alleles. *Microbiology* **147**, 161–169, https://doi.org/10.1099/00221287-147-1-161 (2001).
45. Nielsen, H. Predicting Secretory Proteins with SignalP. *Methods Mol Biol* **1611**, 59–73, https://doi.org/10.1007/978-1-4939-7015-5_6 (2017).
46. Jones, P. *et al*. InterProScan 5: genome-scale protein function classification. *Bioinformatics* **30**, 1236–1240, https://doi.org/10.1093/bioinformatics/btu031 (2014).

47. Celik, N. *et al*. A bioinformatic strategy for the detection, classification and analysis of bacterial autotransporters. *PloS one* **7**, e43245, https://doi.org/10.1371/journal.pone.0043245 (2012).
48. Armougom, F. *et al*. Expresso: automatic incorporation of structural information in multiple sequence alignments using 3D-Coffee. *Nucleic acids research* **34**, W604–608, https://doi.org/10.1093/nar/gkl092 (2006).
49. Di Tommaso, P. *et al*. T-Coffee: a web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension. *Nucleic acids research* **39**, W13–17, https://doi.org/10.1093/nar/gkr245 (2011).
50. Huson, D. H. & Bryant, D. Application of phylogenetic networks in evolutionary studies. *Molecular biology and evolution* **23**, 254–267, https://doi.org/10.1093/molbev/msj030 (2006).
51. Needleman, S. B. & Wunsch, C. D. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of molecular biology* **48**, 443–453, https://doi.org/10.1016/0022-2836(70)90057-4 (1970).
52. Tatusova, T. A. & Madden, T. L. Blast 2 sequences - a new tool for comparing protein and nucleotide sequences. *FEMS microbiology letters* **174**, 247–250, https://doi.org/10.1111/j.1574-6968.1999.tb13575.x (1999).
53. Finn, R. D. *et al*. InterPro in 2017-beyond protein family and domain annotations. *Nucleic acids research* **45**, D190–D199, https://doi.org/10.1093/nar/gkw1107 (2017).
54. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* **10**, 845–858, https://doi.org/10.1038/nprot.2015.053 (2015).
55. Kelley, L. A. & Sternberg, M. J. Protein structure prediction on the Web: a case study using the Phyre server. *Nat Protoc* **4**, 363–371, https://doi.org/10.1038/nprot.2009.2 (2009).
56. Zhang, Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* **9**, 40, https://doi.org/10.1186/1471-2105-9-40 (2008).
57. Yang, J. *et al*. The I-TASSER Suite: protein structure and function prediction. *Nat Methods* **12**, 7–8, https://doi.org/10.1038/nmeth.3213 (2015).
58. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126–2132, https://doi.org/10.1107/S0907444904019158 (2004).
59. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. PROCHECK - a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283–291, https://doi.org/10.1107/S0021889892009944 (1993).
60. Yuan, S., Chan, H. C. S., Filipek, S. & Vogel, H. PyMOL and Inkscape Bridge the Data and the Data Visualization. *Structure* **24**, 2041–2042, https://doi.org/10.1016/j.str.2016.11.012 (2016).
61. Rigsby, R. E. & Parker, A. B. Using the PyMOL application to reinforce visual understanding of protein structure. *Biochem Mol Biol Educ* **44**, 433–437, https://doi.org/10.1002/bmb.20966 (2016).
62. Chagnot, C. *et al*. *In vitro* colonization of the muscle extracellular matrix components by *Escherichia coli* O157:H7: the influence of growth medium, temperature and pH on initial adhesion and induction of biofilm formation by collagens I and III. *PloS one* **8**, e59386, https://doi.org/10.1371/journal.pone.0059386 (2013).
63. Vadillo-Rodriguez, V. *et al*. Comparison of atomic force microscopy interaction forces between bacteria and silicon nitride substrata for three commonly used immobilization methods. *Applied and environmental microbiology* **70**, 5441–5446, https://doi.org/10.1128/AEM.70.9.5441-5446.2004 (2004).
64. Hutter, J. L. & Bechhoefer, J. Calibration of Atomic-Force Microscope Tips. *Rev Sci Instrum* **64**, 3342–3342, https://doi.org/10.1063/1.1144449 (1993).
65. Ducker, W. A., Senden, T. J. & Pashley, R. M. Direct Measurement of Colloidal Forces Using an Atomic Force Microscope. *Nature* **353**, 239–241, https://doi.org/10.1038/353239a0 (1991).
66. Beaussart, A. *et al*. Quantifying the forces guiding microbial cell adhesion using single-cell force spectroscopy. *Nature Protocols* **9**, 1049–1055, https://doi.org/10.1038/nprot.2014.066 (2014).
67. Zeng, G. H., Muller, T. & Meyer, R. L. Single-Cell Force Spectroscopy of Bacteria Enabled by Naturally Derived Proteins. *Langmuir: the ACS journal of surfaces and colloids* **30**, 4019–4025, https://doi.org/10.1021/la404673q (2014).
68. Gaboriaud, F. & Dufrene, Y. F. Atomic force microscopy of microbial cells: application to nanomechanical properties, surface forces and molecular recognition forces. *Colloids Surf B Biointerfaces* **54**, 10–19, https://doi.org/10.1016/j.colsurfb.2006.09.014 (2007).
69. Ahimou, F., Denis, F. A., Touhami, A. & Dufrêne, Y. F. Probing microbial cell surface charges by atomic force microscopy. *Langmuir*, https://doi.org/10.1021/la026273k (2002).
70. Alsteens, D., Dague, E., Rouxhet, P. G., Baulard, A. R. & Dufrene, Y. F. Direct measurement of hydrophobic forces on cell surfaces using AFM. *Langmuir* **23**, 11977–11979, https://doi.org/10.1021/la702765c (2007).

Acknowledgements

This work was supported in part by INRA (Institut National de la Recherche Agronomique) and by the CoMBa ("Colonisation de la Matrice extracellulaire Bacterienne") project funded by the "Région Auvergne" FRI ("Fond Régional Innovation") cluster IRP ("Institut de Recherche Pharmabiotique"). Support from the Australian Research Council (ARC) (DP150102287, DP180102987, FT130100580) is also acknowledged. Valentin AGEORGES is a PhD Research Fellow granted by the "Conseil Régional d'Auvergne – FEDER (Fonds Européen de Développement Régional)". The authors thank Ulrich Dobrindt for providing with *E. coli* K12 MG1655 Δ*fim* Δ*agn43*. The authors are grateful to the INRA MIGALE bioinformatics platform (http://migale.jouy.inra.fr) for providing computational resources as well as to the INRA MEDiS microscopy platform for access to the epifluorescence microscope.

Author Contributions

V.A. and M.D. conceptualise the overarching aims of the research study. V.A., M.S., E.D., G.J., B.H., I.R.H. and M.D. conceived and designed the experiments. V.A., M.S., N.C., P.R., I.C., X.B. and A.E.R. performed the experiments and data acquisition. V.A., M.S., G.J., P.R., I.C., X.B., S.L., J.P., B.H., F.C.D., I.R.H. and M.D. analysed and interpreted the data. M.D. had management as well as coordination responsibility for the execution of the research work. M.D. and B.H. contributed to the acquisition of the financial supports and resources leading to this publication. V.A., M.S., G.J., X.B., E.D., S.L., J.P., B.H., F.C.D., I.R.H. and M.D. wrote the article, including drafting and revising critically the manuscript for important intellectual content.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-019-47608-4.

Competing Interests: F.C.D. and M.S. are employees of the Lallemand SAS Company, which provided support in the form of salaries and partly financed their work. All other authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2019

ARTICLE n°3

Valentin AGEORGES, Mohamed A. ZORGANI, Jason PAXMAN, Begoña HERAS, Ian R. HENDERSON, Frédérique CHAUCHEYRAS-DURAND, Sabine LEROY, Xavier BAILLY, Eric PEYRETAILLADE and Mickaël DESVAUX.

Diversity, distribution and prevalence of antigen 43 (Ag43) in Bacteria

Pour publication dans DNA Research

Diversity, distribution and prevalence of antigen 43 (Ag43) in Bacteria

Valentin AGEORGES [1], Eric PEYRETAILLADE [2], Mohamed A. ZORGANI [1], Jason PAXMAN [3], Begoña HERAS [3], Ian R. HENDERSON [4,5], Sabine LEROY [1], Xavier BAILLY [6] and Mickaël DESVAUX [1,*].

[1] Université Clermont Auvergne, INRA, UMR454 MEDiS, 63000 Clermont-Ferrand.

[2] Université Clermont Auvergne, CNRS, LMGE, 63000 Clermont-Ferrand, France.

[3] Department of Biochemistry and Genetics, La Trobe Institute for Molecular Science, La Trobe University, Bundoora, Victoria 3086, Australia.

[4] Institute of Microbiology and Infection, University of Birmingham, B152TT Birmingham, United Kingdom.

[5] Institute for Molecular Biosciences, University of Queensland, St. Lucia, Queensland 4067, Australia.

[6] INRA, UR346 Epidémiologie animale, Saint Genès Champanelle, France.

*Corresponding author: Dr Hab Mickaël DESVAUX-LENÔTRE (Hab, PhD, Mag), INRA (Institut National de la Recherche Agronomique) Auvergne-Rhône-Alpes, Centre de Recherche de Clermont-Ferrand, UMR454 MEDiS (Microbiologie, Environnement Digestif, Santé), Site de Theix, F-63122 Saint-Genès Champanelle, France. Tel.: +33 (0)4 73 62 47 23, Fax: +33 (0)4 73 62 45 81, E-mail: mickael.desvaux@inra.fr.

ABSTRACT

The antigen 43 (Ag43) is a prototypical member of the family of self-associating autotransporters (SAATs) considered as the major phase-variable outer membrane protein in *Escherichia coli*. As an autotransporter (AT) secreted through the Type V, subtype a, secretion system (T5aSS), Ag43 has a modular architecture comprising from N- to C-terminal, (i) a signal peptide harbouring a highly conserved extended region (ESPR), (ii) a passenger subdivided into SL (stem of the letter L), EJ (elbow joining) and BL (bottom of the letter L) subdomains, (iii) an autochaperone domain of type 1 (AC1), and (iv) a translocator. Ag43 was originally thought to be distributed in into two subfamilies, namely SF-I and SF-II, but most recent phylogenetic analyses disclosed the existence of four distinct Ag43 phylogenetic classes, namely C1, C2, C3 and C4, exhibiting different propensities for autoaggregation and co-interactions. Ag43 is considered to have a quite ubiquitous distribution in *E. coli* and many strains can harbour *agn43* in multiple copies. Considering the recent progress in our understanding of the diversity of Ag43, the prevalence and distribution of this SAAT in the Bacteria Kingdom was reinvestigated. Following data mining, it first appeared Ag43 passengers cluster into two additional phylogenetic classes associated with different subtypes of SL subdomain, which is directly involved in the Velcro-handshake mechanism resulting in bacterial autoaggregation. The ESPR and translocator appeared remarkably conserved as belonging to monophylectic group whatever the class of the Ag43. The full diversity in Ag43 actually arose from the association of the SL subtypes with two different EJ-BL-AC1 modules. In the Bacteria Kingdom, *agn43* appeared to be essentially present at 96 % in *E. coli* (including *Shigella*) but not in all strains (64 %). Single copy of gene encoding Ag43 ^{C6} appeared the most prevalent situation but up to 5 copies of *agn43* with different combinations of classes could be observed. This new understanding of the diversity of Ag43 provides a rational framework for investigating its role in *E. coli* physiology, ecophysiology and physiopathology.

INTRODUCTION

Autotransporters (ATs) belong to the Type V, subtype a, secretion system (T5aSS) found in some lipopolysaccharidic (LPS) diderm bacteria (archetypal Gram-negative bacteria) (Desvaux *et al.* 2003, Desvaux *et al.* 2004, Leo *et al.* 2012). These modular proteins are constituted of three main regions, i.e. (i) a N-terminal signal peptide (SP), (ii) a central passenger, and (iii) a translocator at the C-terminus (Henderson *et al.* 2004, Drobnak *et al.* 2015). While targeted to the inner membrane (IM) by the SP and exported across the Sec translocase, the passenger is secreted across the outer membrane (OM) through a β-barrel pore formed by the translocator. The passenger thus corresponds to the effector, which can be classified into distinct functional categories, among which the self-associating ATs (SAATs) (Henderson *et al.* 2004, Klemm *et al.* 2006, Armbruster *et al.* 2017, Rojas-Lopez *et al.* 2018). The hallmark and primary function of SAATs is to confer bacterial aggregation and sedimentation via self-association (Klemm *et al.* 2006). This phenomenon called auotaggregation can have important consequences on bacterial physiology, ecophysiology and/or physiopathology since it can, for instance, influence biofilm formation, gather siblings in bacterial communities or affect bacterial virulence and infection (Klemm *et al.* 2006, Trunk *et al.* 2018).

Prototypical members of SAATs include AIDA (adhesin involved in diffuse adherence), TibA (toxigenic invasion locus b) and Ag43 (antigen 43) as well as the most recently uncovered EhaA (enterohemorrhagic *Escherichia coli* autotransporter A) (Klemm *et al.* 2006, Wells *et al.* 2008). While SAATs belong to the family of AIDA-type AT, which passengers are all predicted to be based upon right-handed β-helix structures (Henderson *et al.* 2004), members of this AT family are not systematically SAATs as they can be AT adhesins (ATAs), like EhaB or UpaB (uropathogenic *E. coli* autotransporter adhesin B), which do not trigger autoaggregation but specifically bind some extracellular matrix molecules (Armbruster *et al.* 2017). Of note, some additional ATs promote autoaggregation but not as their primary function (Henderson *et al.* 2004), *e.g.* Hap (*Haemophilus influenzae* adhesion and penetration protein) is primarily a serine protease AT with adhesive and aggregative properties harboured in its C-terminal region (Meng *et al.* 2011) or AutA (autotransporter A) is most probably an ATA mediating autoaggregation through direct or indirect interaction with extracellular DNA (Arenas *et al.* 2015).

Ag43 is a renowned antigen from *E. coli* considered as the major phase-variable outer membrane protein and undoubtedly the most investigated SAAT (Owen *et al.* 1996, van der Woude *et al.* 2008). Besides the shared and common structural features of ATs secreted by the T5aSS, Ag43 exhibits some additional conserved domains (Figure 1).

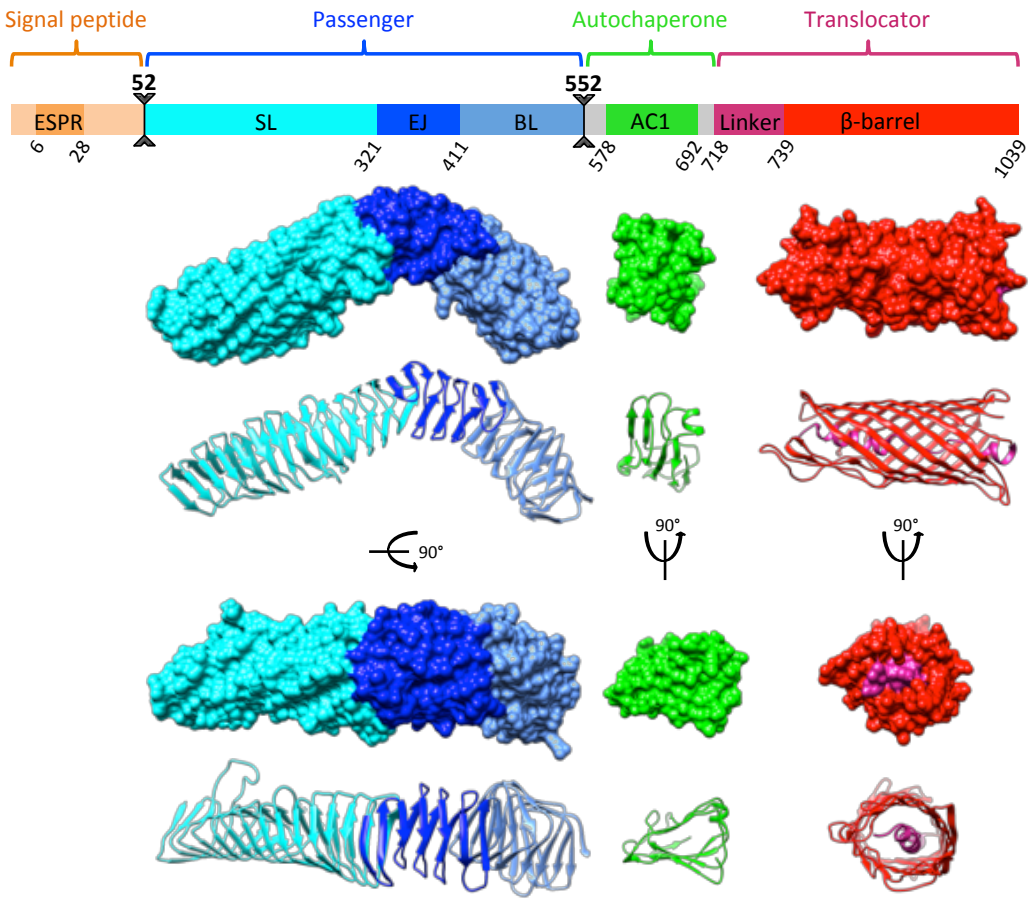

Figure 1: Modular organisation of the SAAT Ag43. The Ag43 ^{C3} from UPEC CFT073 is given as an example, which comprises a N-terminal SP exhibiting a conserved ESPR (IPR024973), a passenger domain corresponding to the effector region responsible for the autoaggregation function, a structurally conserved AC1 domain involved in folding, and a C-terminal translocator comprising a α-helical linker region and a β-barrel. The SP is processed between amino acid position 51 and 52 and the passenger is cleaved off from the AC and Translocator regions between amino acid 551 and 552. 3D structures are given as surface (upper part) and ribbon diagrams (lower part), which were generated from UCSF Chimera (Pettersen *et al.* 2004). The passenger corresponds the resolved structure of Ag43 ^{C3} from UPEC CFT073 (PDB: 4KH3), whereas the AC1 and translocator were modelled from Phyre2 (Kelley *et al.* 2015) with the AC1 from IcsA (PDB: 3ML3) (Kuhnel *et al.* 2011) and the translocator from AIDA-I (PDB: 4MEE) (Gawarzewski *et al.* 2014) used as templates.

The SP harbours a highly conserved extended region (ESPR) that would regulate the rate of export through the Sec translocase at the IM to withstand the secretion-competent state of the passenger through the translocator at the OM (Desvaux *et al.* 2007, Jong *et al.* 2010). A structurally conserved autochaperone domain of type 1 (AC1) present between the passenger

and the translocator is considered as essential for folding of β-helical passengers (Oliver *et al.* 2003, Rojas-Lopez *et al.* 2018). Actually, the Ag43 is synthesised as a preproprotein where the SP is cleaved off after export across the IM by a signal peptidase of type 1 and the passenger is cleaved off from the translocator and AC1 after secretion across the OM by an as yet unidentified protease (Figure 1). Despite being processed, the Ag43 passenger is displayed at the bacterial cell surface as it would remain associated to the cognate translocator (Henderson *et al.* 2004). While glycosylated forms of Ag43 could be detected in some faecal *E. coli* isolates (Reidl *et al.* 2009), the endogenous glycosyltransferase at play, the extent and biological relevance of glycosylation occurring naturally in Ag43 remain unknown. While two subfamilies of Ag43 were originally considered, namely SF-I and SF-II, latest phylogenetic network analysis focusing on the passenger revealed for the first time that they actually distributed into four distinct functional classes, i.e. C1, C2, C3 and C4 (Ageorges *et al.* 2019). The crystal structure of the Ag43 ^{C3} passenger from UPEC CFT073 was solved and revealed a twisted L-shape β-helical conformation, which appeared as a requirement for handshake self-association (Heras *et al.* 2014). Further structural and phylogenetic analyses demonstrated the L-shaped Ag43 passengers could be subdivided into SL (stem of the letter L), EJ (elbow joining) and BL (bottom of the letter L) subdomains each belonging to two distinct subtypes, which actually explained the occurrence of the four Ag43 classes through domain shuffling (Ageorges *et al.* 2019).

Ag43 is considered to have a quite ubiquitous distribution in *E. coli* and many strains can harbour *agn43* in multiple copies (Owen *et al.* 1996, Roche *et al.* 2001). While the presence of *agn43* appeared to be more frequent among UPEC and diarrheagenic *E. coli* than commensal isolates (Roche *et al.* 2001, Restieri *et al.* 2007), this must be interpreted with caution as it could simply reflect a bias toward sequencing pathogenic *E. coli* rather than nonpathogens (van der Woude *et al.* 2008). The new understanding of the modular architecture of Ag43 prompted us to reinvestigate the phylogenetic diversity of this SAAT with respect to the four different classes recently uncovered. Besides the distribution of Ag43 in the Bacteria Kingdom, the presence of multiple alleles and their prevalence were examined.

RESULTS

Ag43 passengers cluster into 6 different phylogenetic classes associated to different subtypes of SL subdomains.

Following our data mining approach to recover Ag43 homologues in Bacteria, 7,219 full-length Ag43 sequences were recovered and a dataset of 243 sequences representative of the diversity of Ag43 was generated for further phylogenetic studies, which increased very significantly the previous dataset of 106 Ag43 sequences (van der Woude *et al.* 2008, Vo *et al.* 2017, Ageorges *et al.* 2019). Interestingly, all uncovered *agn43* were identified on bacterial chromosomes as none appeared to be harboured by plasmids. In order to confirm they belong to the Ag43 family and not to other SAATs, a phylogenetic network analysis was applied and clearly demonstrated that all the Ag43 newly uncovered by our approach clustered separately from TibA, AIDA-I and EhaA (Figure 2A). Focusing on the Ag43 only, the phylogenetic network analysis first suggested an ambiguity in the relatedness patterns among the Ag43 illustrated by the diamond shape of the network (Figure 2B). However, bootstrapping analysis with a stringent confidence threshold (90 %) dispelled the doubts and revealed that the splits were not robust. As previously reported (van der Woude *et al.* 2008), a dichotomous grouping of Ag43 in the two subfamilies was confirmed, i.e. SF-I and SF-II.

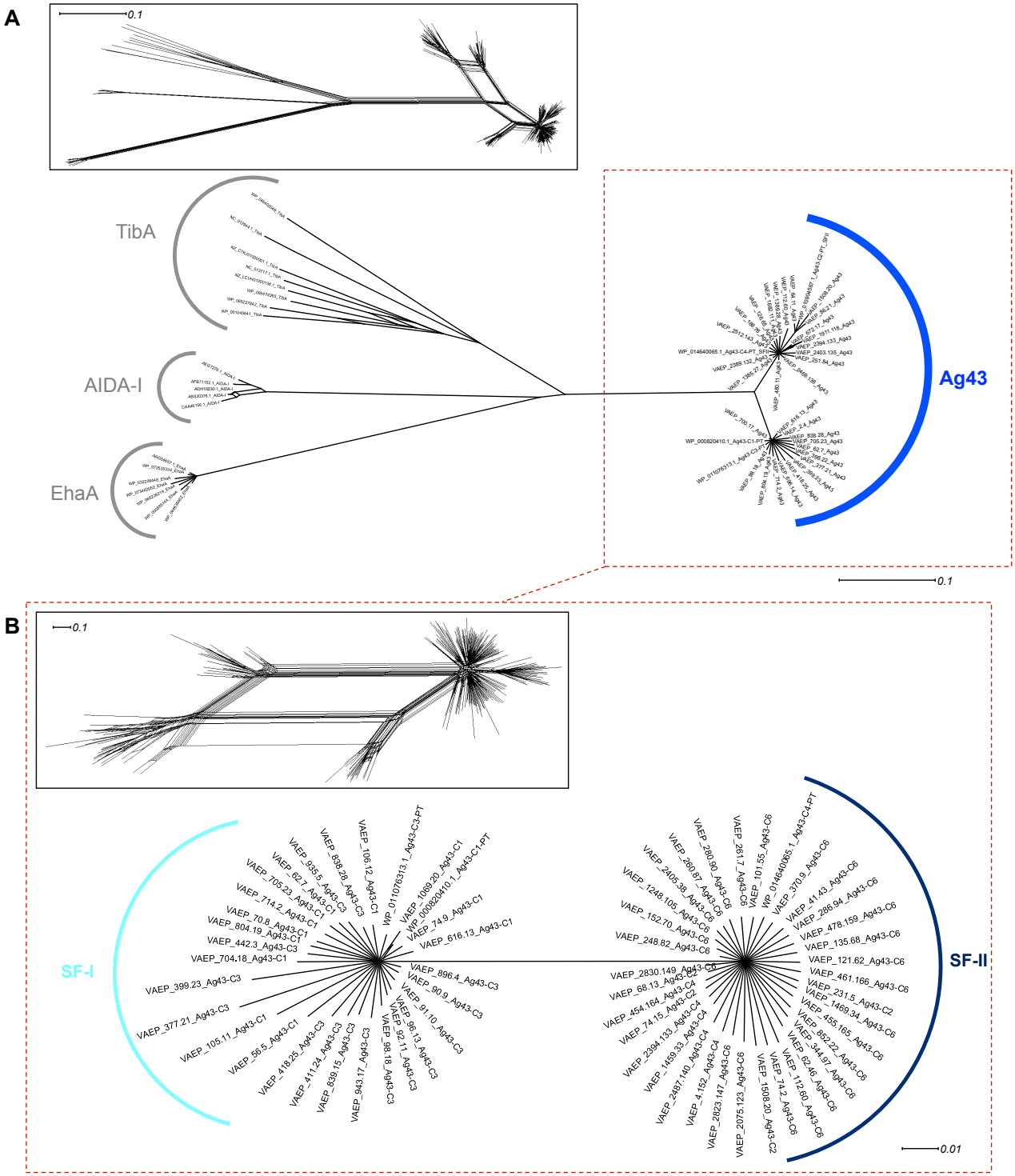

Figure 2: Phylogenetic analysis of Ag43 and other SAATs. (A) Clustering of the full-length amino acid sequences of Ag43 respective to other SAATs, namely TibA, AIDA-I and EhaA (Henderson *et al.* 2004, Klemm *et al.* 2006, Armbruster *et al.* 2017). (B) Focus on Ag43 only highlighting the subfamilies previously reported, namely SF-I and SF-II (van der Woude *et al.* 2008). Trees were generated using SplitsTree and inserts represent the trees before applying bootstrapping and confidence threshold. For legibility, some sequences were removed from the trees. The scale bars represent the evolutionary distance, i.e. the number of substitutions per site.

Considering in ATs the effector corresponds to the passenger, further split decomposition analyses were applied focusing the Ag43 functional region (Figure 3).

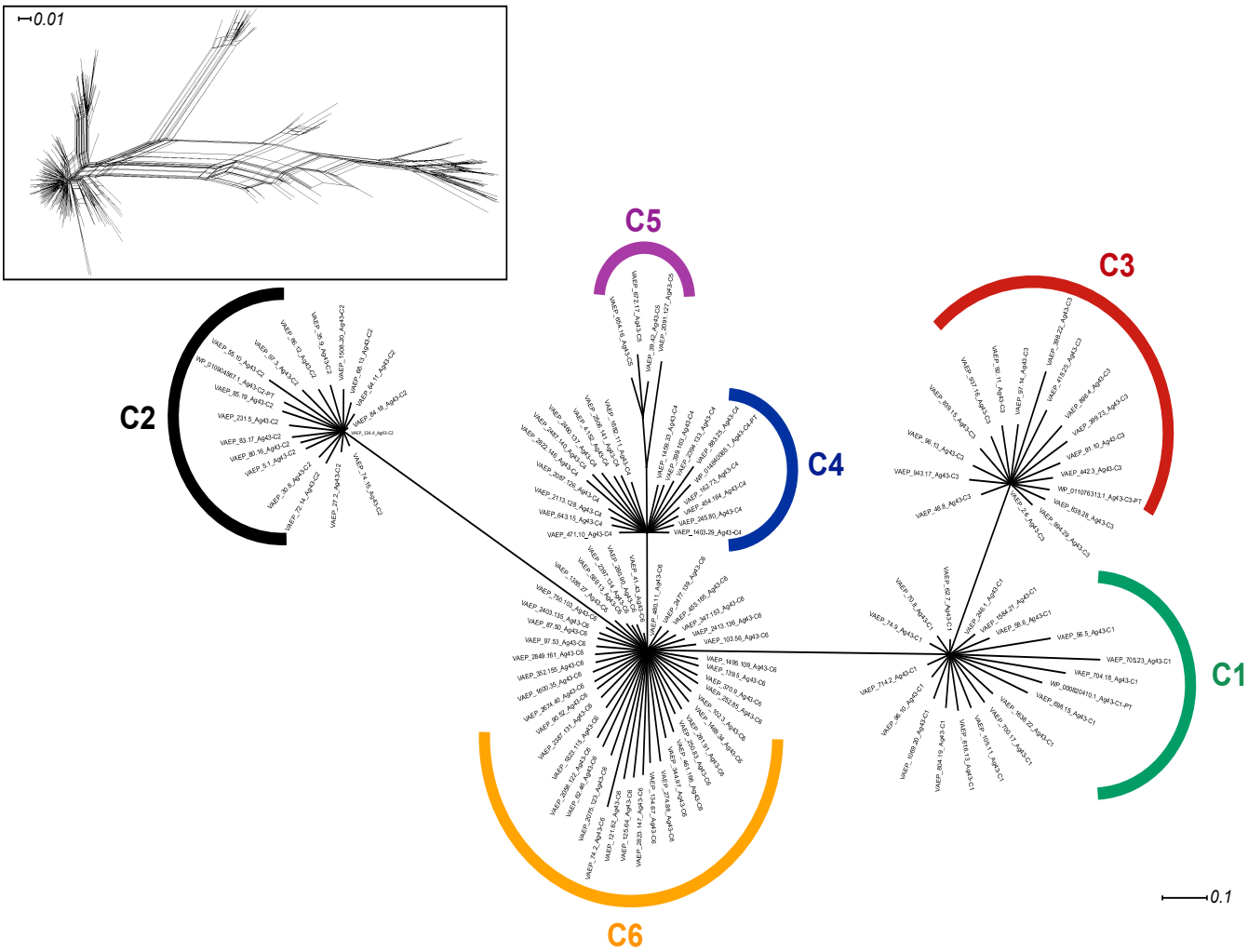

Figure 3: Phylogenetic analysis of the effector region of Ag43. Clustering of the Ag43 passengers respective to the different functional classes previously reported, namely from C1 to C4 (Ageorges *et al.* 2019). Trees were generated using SplitsTree and inserts represent the trees before applying bootstrapping and confidence threshold. For legibility, some sequences were removed from the trees. The scale bars represent the evolutionary distance, i.e. the number of substitutions per site.

Again, a diamond shape network was obtained in the first instance but bootstrapping analysis resolved the splits and revealed instead the existence of 6 classes of Ag43 passengers. While previous network analysis on a much smaller dataset of Ag43 evidenced the existence of 4 main passenger classes supported by robust splits with a strong ambiguity in their relatedness patterns with no specific divergence between C1 to C4 (Ageorges *et al.* 2019), the present analysis on a larger dataset alleviates the uncertainties between C3 and C2, C1 and C4 as well as C2 and C4 since it clearly indicated direct links only between C4 and the newly identified C5, C2 and the newly identified C6, C6 and C4, C1 and C3, as well as C6 and C1.

Respective to the two subfamilies of Ag43 based on the full-length amino acid sequences, SF-I gathers C1 and C3, whereas SF-II regroups C2, C4, C5 and C6 (Figure 2B and 3). As previously reported and with a very few exceptions, this grouping is associated with the size of Ag43 passengers, which range around 500 amino acids for C1 and C3 versus around 430 amino acids for the four remaining classes. While C1 to C4 harbour prototypical (PT) Ag43, namely from *E. coli* K12 MG1655 (UPI00003B25EE), enterohaemorrhagic *E. coli* (EHEC) O157:H7 EDL933 (UPI00000D07E5), uropathogenic *E. coli* (UPEC) CFT073 (UPI00000E49D2) and enterotoxigenic *E. coli* (ETEC) O78:H11 H10407, respectively (Ageorges *et al.* 2019), no Ag43 homologues from C5 and C6 have been functionally or structurally characterised to date.

Regarding the passenger subdomains as defined by 3D structure analysis (Heras *et al.* 2014, Ageorges *et al.* 2019), further phylogenetic network analyses were performed (Figure 4). While two robust clusters of SL sequences of type 1 and 2 were still unambiguously differentiated, it further appeared SL1 could be discriminated into clusters, named the subtypes a, b and c (Figure 4A).

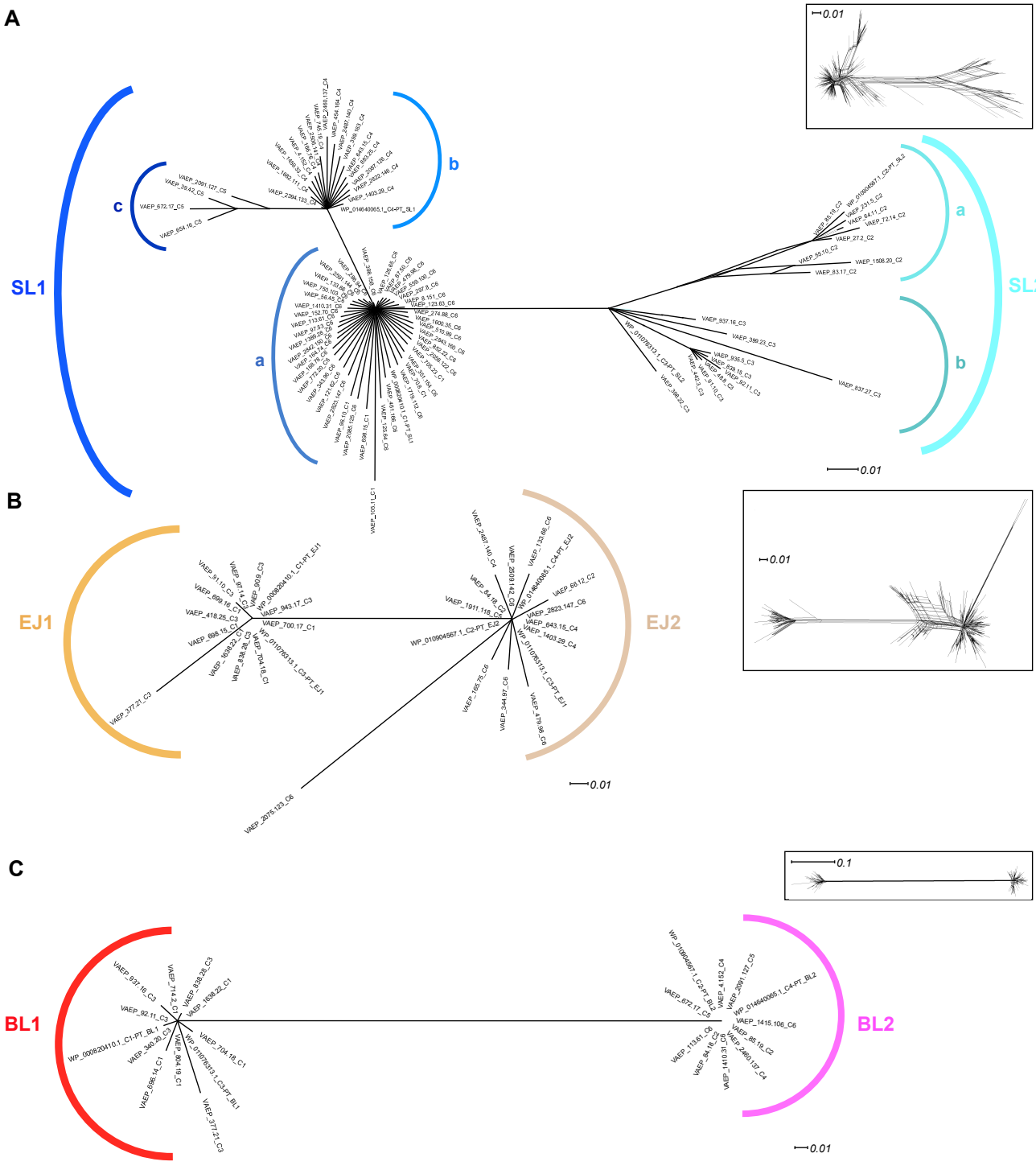

Figure 4: Phylogenetic analysis of the subdomains in the Ag43 passenger. (A) Clustering of the SL subdomain from the Ag43 passenger respective to the subtypes previously reported, namely SL1 and SL2 (Ageorges *et al.* 2019). (B) Clustering of the EJ subdomain from the Ag43 passenger respective to the subtypes previously reported, namely EJ1 and EJ2 (Ageorges *et al.* 2019). (C) Clustering of the BL subdomain from the Ag43 passenger respective to the subtypes previously reported, namely BL1 and BL2 (Ageorges *et al.* 2019). Trees were generated using SplitsTree and inserts represent the trees before applying bootstrapping and confidence threshold. For legibility, some sequences were removed from the trees. The scale bars represent the evolutionary distance, i.e. the number of substitutions per site.

SL1a corresponded to the SL1 originally identified in Ag43 $^{C1\text{-}PT}$ and clustered all Ag43 belonging to C1 as well as the newly identified C6. SL1b exclusively clustered subdomains corresponding to Ag43 ^{C4}, whereas SL1c clustered subdomains from the newly identified Ag43 ^{C5} passengers. Similarly, two subtypes could be further differentiated into SL2, namely SL2a clustering subdomains from Ag43 ^{C2} and SL2b for Ag43 ^{C3} passengers (Figure 4A). In other words, there is a direct correlation between the Ag43 of the classes C2, C3, C4 and C5 with the SL2a, SL2b, SL1b, SL1c subtypes, respectively but not for Ag43 ^{C1} and Ag43 ^{C6}, which both exhibit a SL1a.

As previously observed, two robust sequence subtypes were clustered and defined two subtypes of EJ and BL subdomains, which were systemically associated one with another (Figures 4B and 4C). Indeed, the phylogenetic network analysis of the BL subdomain was coherent with the analysis of the EJ subdomain as EJ1 was always found together with BL1 whereas EJ2 was found together with BL2. EJ1-BL1 were systematically found in Ag43 belonging to SF-I, i.e. Ag43 ^{C1} and Ag43 ^{C3}, whereas EJ2-BL2 were present in Ag43 belonging to SF-II, i.e. Ag43 ^{C2}, Ag43 ^{C4}, Ag43 ^{C5} and Ag43 ^{C6}. It consequently appears the Ag43 ^{C1} and Ag43 ^{C6} differ by the association of their SL1a with EJ1-BL1 or EJ2-BL2 modules, respectively. The EJ region reveals a remarkably conserved size of 85 amino acids whenever belonging to the type 1 or 2. On contrast, BL1 was 143 amino-acids long versus a length of 90 amino acids for BL2.

Diversity in Ag43 essentially arises from different SL subtypes in association with either EJ1-BL1-AC1 [E] or EJ2-BL2-AC1 [Ж] modules.

Fully taking into account the modular architecture of an AT such as Ag43, the diversity of the ESPR, AC1 and translocator was further examined. Respective to the distribution of the ESPR in the 10 clusters previously observed (Desvaux *et al.* 2006, Desvaux *et al.* 2007), it appears the ESPR from Ag43 all belong to cluster 5 whatever the subfamily or class of the Ag43 (Figure 5A); subsequent phylogenetic network analysis could not revealed any further relevant ESPR-5 subtypes in Ag43 (Figure 5B).

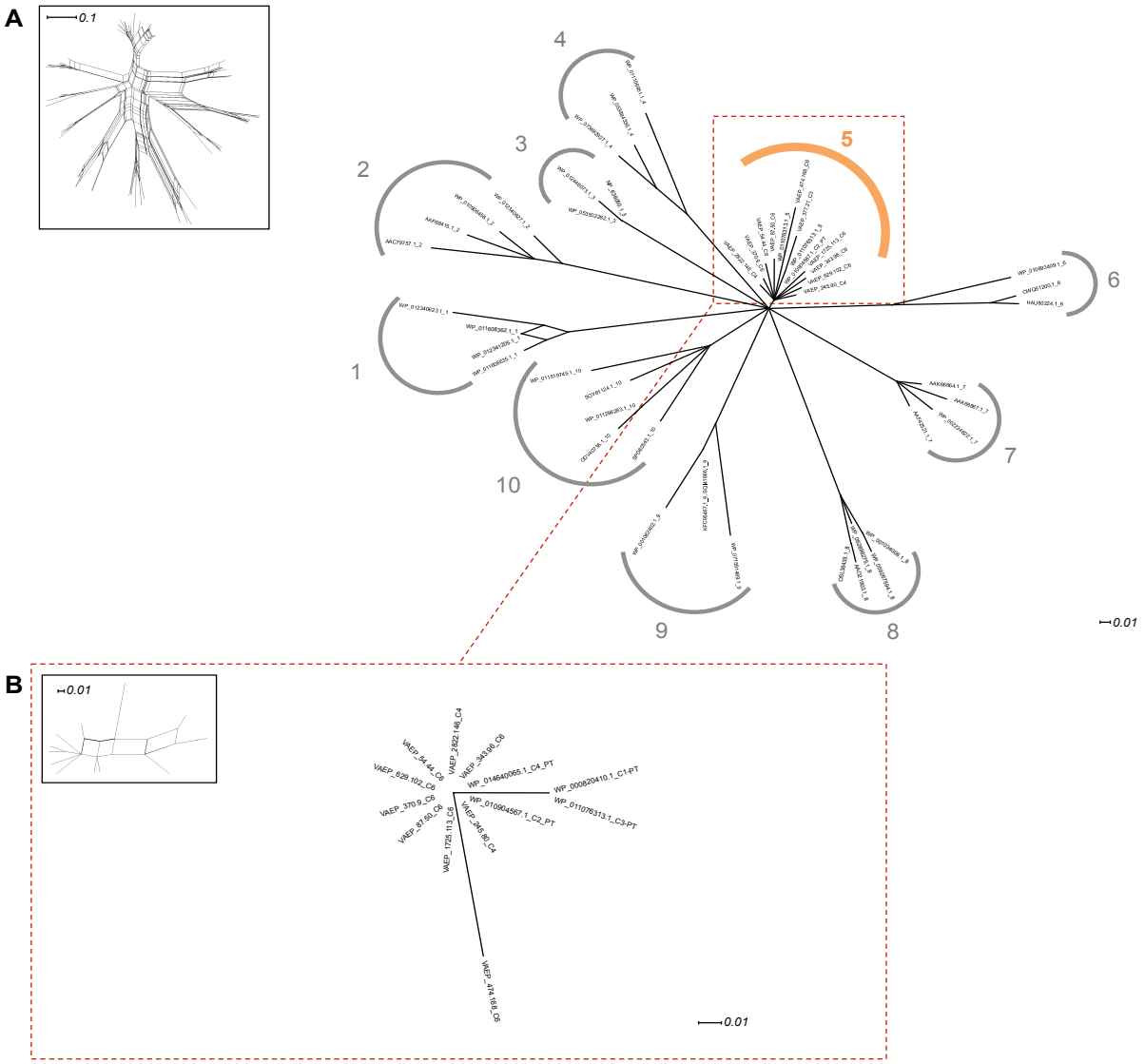

Figure 5: Phylogenetic analysis of the ESPR in Ag43. (A) Clustering of the ESPR from Ag43 respective to the ten different ESPR clusters previously reported (Desvaux *et al.* 2006). (B) Focus on the clustering of the EPSR from Ag43 only. Trees were generated using SplitsTree and inserts represent the trees before applying bootstrapping and confidence threshold. For legibility, some sequences were removed from the trees. The scale bars represent the evolutionary distance, i.e. the number of substitutions per site.

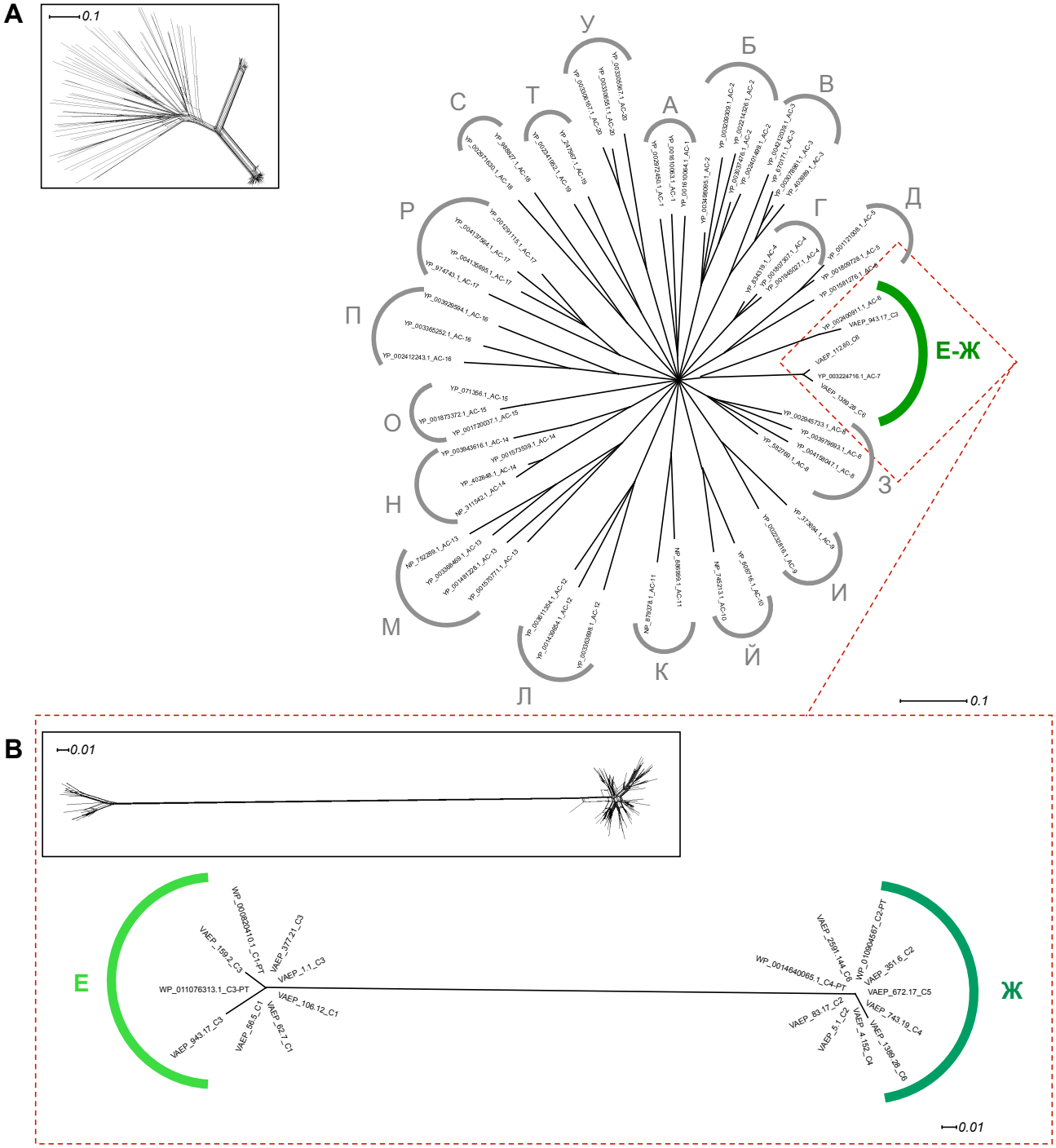

Figure 6: Phylogenetic analysis of the AC1 in Ag43. (A) Clustering of the AC1 from Ag43 respective to the different clusters previously reported in ATs, namely from A to У (cyrillic alphabet) (Rojas-Lopez *et al.* 2018). (B) Focus on the clustering of the AC1 from Ag43 only. Trees were generated using SplitsTree and inserts represent the trees before applying bootstrapping and confidence threshold. For legibility, some sequences were removed from the trees. The scale bars represent the evolutionary distance, i.e. the number of substitutions per site.

In Ag43, ESPR displayed a constant size of 23 amino acids. The AC1 was shown to distribute into 20 main clusters, named from A to У according to the cyrillic alphabet (Rojas-Lopez *et al.* 2018). Seeking for the distribution of AC1 from Ag43, it appears they all belong to cluster E-Ж previously described as harbouring AC1 from Ag43 and AIDA-I (Figure 6A). Additional phylogenetic network analysis showed a different relatedness pattern compared to the ESPR. Two robust sequence subtypes were clustered as (i) AC1 ^{E} was systematically found among Ag43 belonging to SF-I, i.e. Ag43 ^{C1} and Ag43 ^{C3}, and (ii) AC1 ^{Ж} was present in Ag43 belonging to SF-II, i.e. Ag43 ^{C2}, Ag43 ^{C4}, Ag43 ^{C5} and Ag43 ^{C6} (Figure 6B). Consequently, it means AC1 ^{E} is systematically associated with EJ1-BL1 subdomains, whereas AC1 ^{Ж} is present together with EJ2-BL2 subdomains. While AC1 ^{E} displayed a conserved size of 114 amino acids, AC1 ^{Ж} was slightly longer with a size of 117 amino acids. Latest cluster analysis has revealed the presence of fourteen distinct main groups of translocators, or translocation units (TU), in ATs (Celik *et al.* 2012). In agreement with this previous study, TU from Ag43 all clustered within G13, described as harbouring long-type β-barrel domains including AIDA-I (Figure 7A); subsequent phylogenetic network analysis could not revealed any further relevant TU ^{G13} subtypes in Ag43 (Figure 7B). In Ag43, the TU displayed a constant size of 330 amino acids.

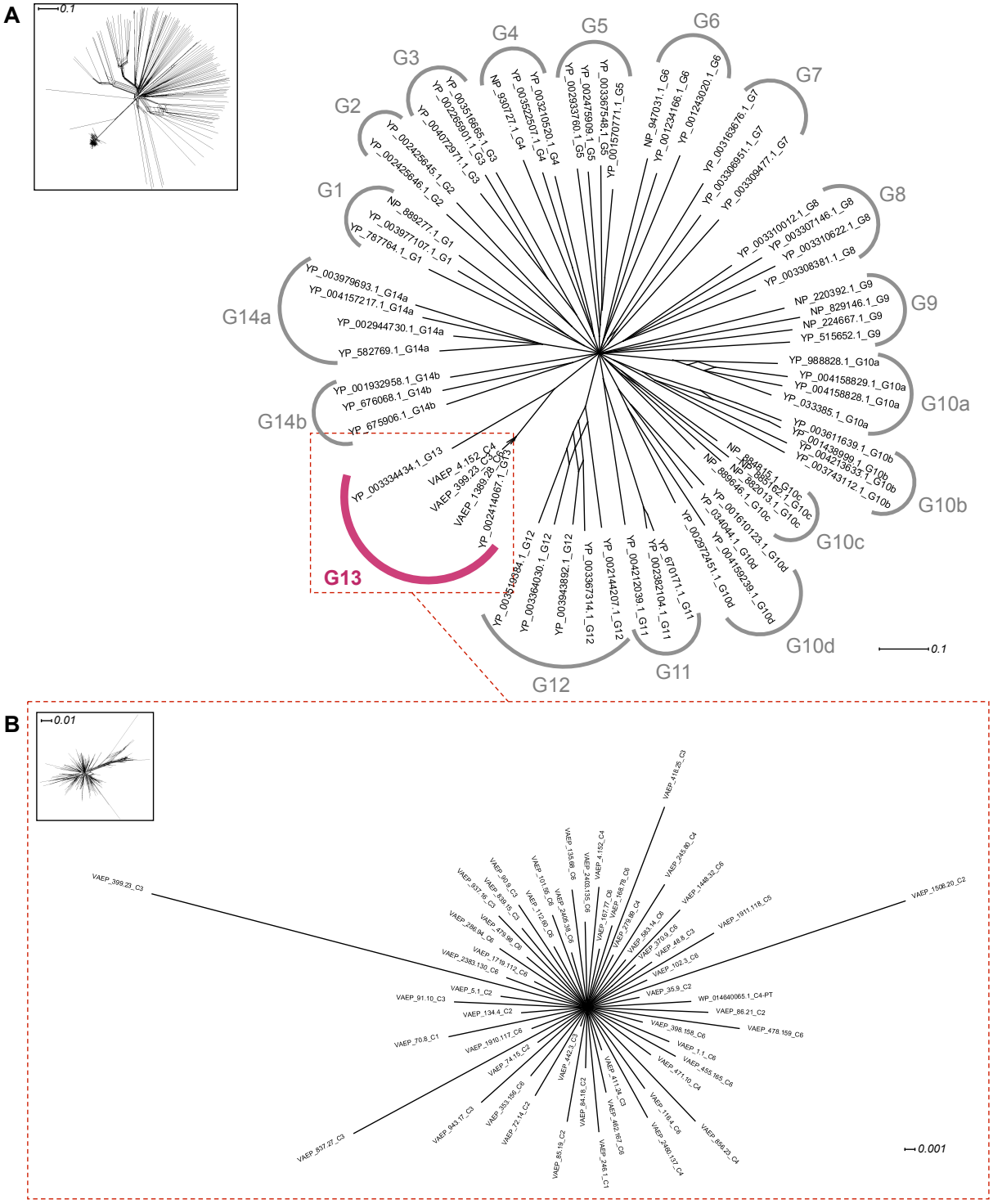

Figure 7: Phylogenetic analysis of the translocator in Ag43. (A) Clustering of the translocator from Ag43 respective to the other translocators previously reported in ATs (Celik *et al.* 2012). (B) Focus on the clustering of the translocator from Ag43 only. Trees were generated using SplitsTree and inserts represent the trees before applying bootstrapping and confidence threshold. For legibility, some sequences were removed from the trees. The scale bars represent the evolutionary distance, i.e. the number of substitutions per site.

Taken together and summarising the diversity of Ag43 (Figure 8), it appears variations among Ag43 arise from the passenger subdomains and the AC1 since the ESPR and TU belong to the same phylogenetic groups (Figure 8). Remarkably, the association of the same phylogenetic groups are found for EJ, BL and AC1 and further defined Ag43 subfamilies originally described (van der Woude *et al.* 2008). Indeed, the association of EJ1-BL1-AC1 ^{E} domains corresponds to the configuration of Ag43 belonging to SF-I and the association of EJ2-BL2-AC1 ^{Ж} domains to Ag43 belonging to SF-II. Considering the cumulative size of these modules in addition to the monophyletic ESPR and TU domains versus the SL subdomains, it further explains the dichotomous grouping obtained with the full-length Ag43. Nonetheless, it appears the Ag43 sequences essentially arise from the 4 different SL subdomains, which actually define 4 out of the classes of Ag43. The SL actually corresponds to the most functional part of Ag43 as it is directly involved the Velcro-handshake mechanism responsible for self-association (Heras *et al.* 2014). While SL1a could be found in combination with EJ1-BL1-AC1 ^{E} or EJ2-BL2-AC1 ^{Ж} modules, giving rise to Ag43 ^{C1} and Ag43 ^{C6} respectively, no alternative combination of the other SL1 or SL2 subtypes with EJ-BL-AC1 subdomains could be observed.

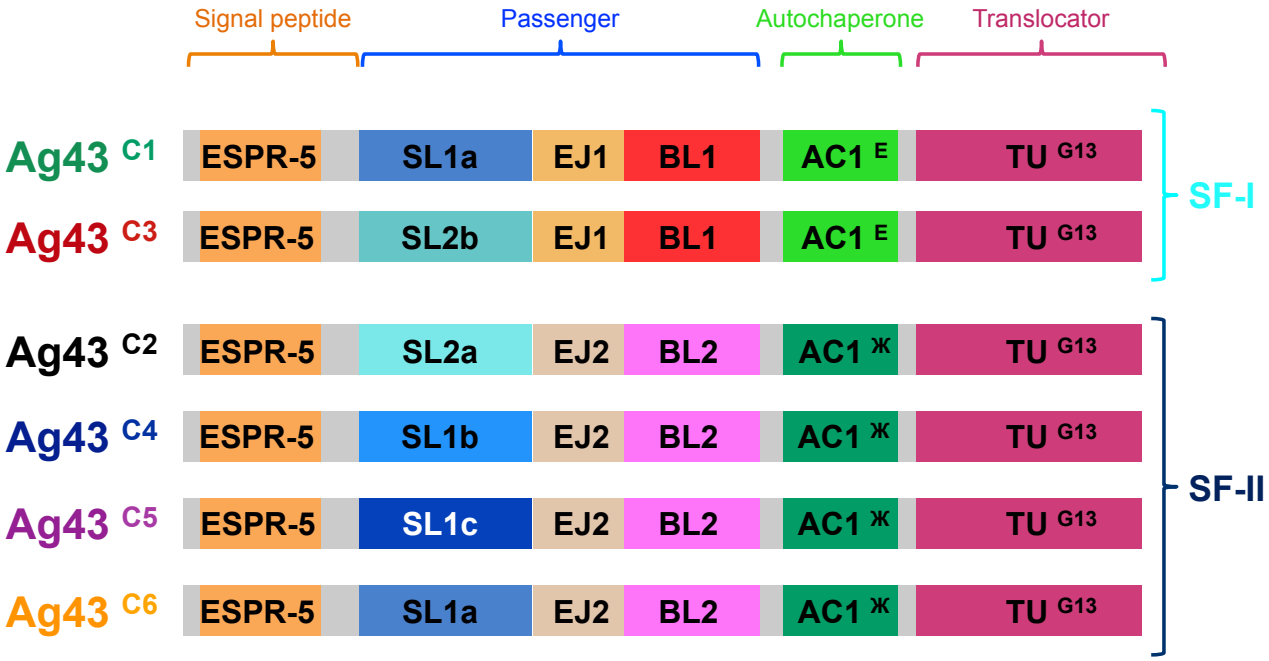

Figure 8: Diversity of Ag43 as revealed by in-depth phylogenetic analyses. Schematic representation of the variations in the modular architecture of Ag43 from C1 to C6 respective to the ESPR, passenger, autochaperone and translocation unit (TU) regions where the different types and subtypes of subdomains are highlighted as defined from the phylogenetic analyses (Figures 2-7).

The Ag43 is quite restricted to the *E. coli* genus where single copy of gene encoding Ag43 ^{C6} appeared as the most prevalent.

As no definitive conclusion can be drawn from unfinished genomes (i.e. at the stage of scaffolds or contigs) about the presence or absence of genes as well as pseudogenes, analysis of the prevalence and distribution of Ag43 was only performed on complete genomes. In fact, final assembly of unfinished genomic sequences may reveal the presence of homologues at a later date. In complete genomes, though, a few *agn43* pseudogenes were observed (52 pseudogenes out of 47 genomes) for all classes of Ag43, except for C3. These pseudogenes would result either from insertion of DNA sequence, including some mobile elements, disrupting the CDS, or the presence of internal stop codon resulting in truncated CDS; alternatively, it could result from sequencing errors and would require further confirmations.

Almost all detected *agn43* sequences appeared to belong to bacteria from the *Enterobacteriaceae* family (99.4%) as only 3 sequences were found in *Achromobacter sp.* from the *Alcaligenaceae* family. Respective to the *Enterobacteriaceae* family, almost all identified *agn43* belonged to bacteria from the *Escherichia* genus (98.6%) with only 3 sequences belonging to *Klebsiella pneumoniae* and 1 to *Citrobacter Rodentium*. Within the *Escherichia* genus, *agn43* sequences were recovered at 97.7% from the *E. coli* species (including *Shigella* (Chaudhuri *et al.* 2012)) and only a few sequences were identified in *E. albertii* and *E. fergusonnii*. Overall, it appears *agn43* is highly prevalent in *E. coli*, representing 96.0 % of all recovered *agn43* sequences in complete bacterial genomes.

Nonetheless, the presence of *agn43* was not ubiquitous in *E. coli* as it was present in only 64.0 % of genomes (i.e. out of 613 *E. coli* and 128 *Shigella* complete genomes at the time). While the large majority of these genomes contained only one *agn43* (64 %), up to 5 *agn43* copies could be identified. While the latter was observed for two genomes, 24 % of the genomes displayed two copies, 9 % possessed 3 copies and 3 % 4 copies of *agn43*. Respective to the allelic diversity, the genes encoding Ag43 were further categorised according to the 6 classes uncovered by phylogenetic analyses. Taking into consideration all the *agn43* sequences identified in *E. coli* (i.e. 719 sequences), C6 and C1 appeared as the most prevalent Ag43 classes, reaching 37 % and 30 % respectively, whereas in a hierarchical order C3, C2, C4 and C5 represented 15 %, 10 %, 4 % and 0.5 % of the uncovered sequences, respectively. However, when only one *agn43* was identified in *E. coli* genome it mostly belonged to C1 (43%) and C6

(27%), hierarchically followed by C2, C3 and C4, ranging from 13% to 4%, respectively (Figure 9A); in contrast, gene encoding Ag43 ^{C5} was never detected alone.

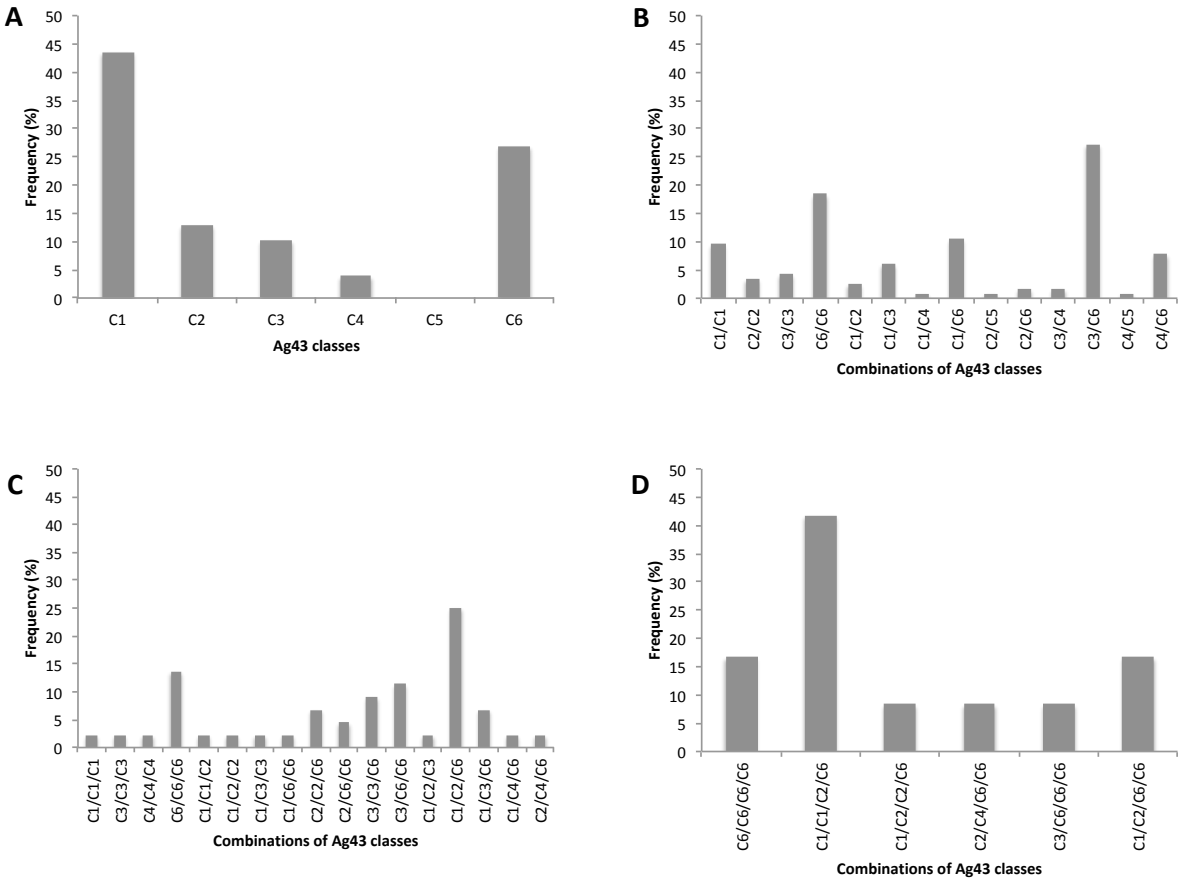

Figure 9: Prevalence and distribution for the different combinations of the 6 Ag43 classes in *E. coli* genomes. Representation of the frequency of detection of the different Ag43 classes or combinations of classes in (A) genomes encoding only one copy of Ag43, (B) genomes encoding two copies of Ag43, (C) genomes encoding three copies of Ag43 and (D) genomes encoding four copies of Ag43.

Regarding the 6 classes of Ag43, 21 possible combinations are theoretically and mathematically possible when two gene copies are present, 56 combinations for three gene copies, 126 for four copies and 252 copies for genes copies. Actually, 19 different combinations could be identified from genomic data for two *agn43* copies, the association of C3/C6 classes was the most frequent (27%), followed by C6/C6, C1/C6, C1/C1 and C4/C6 combinations of classes in this order, whereas the remaining combinations represented each less than 6% of the cases (Figure 9B). Interestingly, all genes encoding Ag43 ^{C5} representatives were identified in genomes harbouring two *agn43*. For genomes with three *agn43* copies, only 17 out of the 56 theoretical combinations could be identified, where C1/C2/C6, C3/C6/C6, C6/C6/C6 and C3/C3/C6 classes were the most prevalent (ranging from 25% to 9%), whereas the remaining

ones represented each less than 7 % (Figure 9C). When four *agn43* were present in bacterial chromosomes only 6 combinations of classes could be observed from genomic data, with the association of C1/C2/C2/C6 classes appearing as the most frequent (i.e. 41 %) as the following C1/C2/C6/C6 and C6/C6/C6/C6 combination dropped at 17% (Figure 9D). Finally, just one combination of 5 copies of *agn43* could be recovered that is C1/C1/C2/C6/C6. Taken together, it appears that whatever the combination of 2 to 5 *agn43*, gene encoding Ag43 ^{C6} is the most prevalent, whereas genes encoding Ag43 ^{C4} is among the rarest.

Because of the lack of readily available information for more than the two-thirds of the bacterial genomes where *agn43* was identified, no correlation between its presence, the number of copies or the different classes with *E. coli* pathotypes, serotypes, virulence levels or the isolation sources could be questioned.

DISCUSSION

Most recent phylogenetic analyses disclosed the existence of four distinct Ag43 phylogenetic classes exhibiting different functions (Ageorges *et al.* 2019). These different classes appeared to have aroused upon subdomain shuffling within the passenger, resulting in the following four combinations (i) SL1-EJ1-BL1 corresponding to Ag43 ^{C1}, (ii) SL2-EJ2-BL2 corresponding to Ag43 ^{C2}, (iii) SL2-EJ1-BL1 corresponding to Ag43 ^{C3}, and (iv) SL1-EJ2-BL2 corresponding to Ag43 ^{C4}. Following data mining to exhaustively identify Ag43 homologues in genomic database, a much larger dataset was here obtained enabling to resolve the strong ambiguity previously observed in the relatedness patterns among Ag43 passengers and further revealing two additional classes of Ag43, namely C5 and C6. The SL1 and SL2 actually appeared to split into a total of 5 different subtypes, which together with just the two different EJ1-BL1-AC1 ^{E} and EJ2-BL2-AC1 ^{Ж} modules explained the six different Ag43 classes here uncovered. While combinations of SL1a with the two EJ-BL-AC1 modules were spotted and gave rise to Ag43 ^{C1} and Ag43 ^{C6}, neither the association of SL1b, SL1c or SL2a with EJ1-BL1-AC1 ^{E} nor the combination of SL2b with EJ2-BL2-AC1 ^{Ж} could be observed. It can be either hypothesised such combinations exist naturally but have not been sequenced as yet or they simply do not exist, which would suggest some genotypic or phenotypic constraints. The constructions of some chimera and their phenotypic characterisation could provide a beginning of answer. While the ESPR and TU belong to the same monophyletic group, the AC1 appeared as the only one additional region displaying distinct phylogenetic clusters beyond the passenger. The strict association of AC1 ^{E} with EJ1-BL1 and AC1 ^{Ж} with EJ2-BL2 would suggest these AC1 variants are necessary for correct folding of these EJ-BL modules present upstream when they emerge from the translocator in the course of the secretion. Again, genetic constructs generating chimera could indicate whether these two AC1 variants can function interchangeably or not for the secretion of Ag43 belonging SF-I and SF-II.

To date, the most consistent phenotypic characterisations are related to Ag43 ^{C1} (van der Woude *et al.* 2008) as only a few investigations have been dedicated to Ag43 ^{C2} (Torres *et al.* 2002), Ag43 ^{C3} and most recently Ag43 ^{C4} (Ageorges *et al.* 2019). The identification of the novel and prominent Ag43 ^{C6} variant, as well as Ag43 ^{C5}, would require further phenotypic characterisation to assay their ability to autoaggregate, adhere and form biofilm as well as their propensity to co-interact with other Ag43 classes. Recent investigation of homotypic interactions of prototypical Ag43 variants has revealed the autoaggregation kinetics was slower and the aggregates smaller for Ag43 $^{C3\text{-}PT}$ compared to other classes (Ageorges *et al.* 2019).

Respective to our new findings this corresponds to Ag43 exhibiting a SL2b-EJ1-BL1-AC1 ^{E} module suggesting this association and especially SL2b would be less prone to induce a strong autoaggregation phenotype. Besides, heterotypic interactions occur only for a limited number of combinations, namely between Ag43 $^{C4\text{-}PT}$ and Ag43 $^{C1\text{-}PT}$ as well as, at lesser extent, between Ag43 $^{C4\text{-}PT}$ and Ag43 $^{C3\text{-}PT}$. This indicates Ag43 with a SL1b-EJ2-BL2-AC1 ^{Ж} module can co-interact with Ag43 exhibiting SL1a-EJ1-BL1-AC1 ^{E} or, at a lower degree, SL2b-EJ1-BL1-AC1 ^{E} modules. As a common feature, Ag43 ^{C1} and Ag43 ^{C3} are the only Ag43 belonging to SF-I and exhibiting a EJ1-BL1-AC1 ^{E} module. In a hierarchical order, this would suggest SL1b is more permissive to heterotypic interactions (two heteroptypic interactions reported) than SL1a (one heteroptypic interaction), SL2b (one heteroptypic interaction) or SL2a (no heteroptypic interaction). Alternatively or additionally, the EJ1-BL1 module could play an indirect role in the heterotypic interaction by altering the angle between the SL and BL subdomains and the overall L-shaped structure, thus modulating the hydrogen bonds and electrostatic interactions (Heras *et al.* 2014). Considering Ag43 ^{C1} and Ag43 ^{C6} are the only ones to harbour SL1a with different EJ-BL modules, homotypic interactions experiments with Ag43 ^{C4} could provide a beginning of answer to the role of SL domain over EJ-BL module. Besides, only one Ag43 belonging to the class 3 has been structurally characterised (Heras *et al.* 2014). Solving the experimental structures of Ag43 belonging to 5 other classes is necessary to decipher the role of the different passenger subdomains in inducing autoaggregation and fully understand the possibility or impossibility for heterotypic associations between Ag43 variants.

While *agn43* is largely prevalent in *E. coli* (including *Shigella*), it did not appear to have an ubiquitous distribution in this species as sometimes previously assumed (Henderson *et al.* 2004). Besides *E. coli*, the existence of *agn43* in closely related *Citrobacter rodentium* was already described in the scientific literature (van der Woude *et al.* 2008) but the present investigation is the first to report the presence of this gene in other *Escherichia* species, such as *E. albertii* or *E. fergusoni.* Interestingly, all these *Escherichia* species can be present in various environments including the soil, water and/or gastro-intestinal digestive tract where they could potentially interact (Maheux *et al.* 2014). Beyond the genus *Escherichia*, *agn43* was further identified in *Klebsiella pneumoniae* as well as *Achromobacter sp*., which is the only species outside the *Enterobacteriaceae* family where it could be found. *K. pneumoniae* is part of the normal flora of the mouth, skin and intestines, whereas bacteria from the genus *Achromobacter* are found in the soil and water but they can be both opportunistic pathogens involved in various infections, including the urinary and respiratory tracts (Swenson *et al.* 2015, Chung 2016). In

E. coli, Ag43 was suggested to be more prevalent in pathogenic *E. coli* than in commensal strains, especially members of the SF-II (Restieri *et al.* 2007). However, this interpretation is still a matter of debate as phylogenetic analyses had not revealed conserved features for Ag43 as a virulence factor and such a correlation may simply reflect the bias toward sequencing pathogenic rather than non-pathogenic *E. coli* (van der Woude *et al.* 2008). While the present study could have provided additional clue about the possible correlation between pathogenic *E. coli* and the presence or absence of particular variants of *agn43*, such an analyse was hindered by absence or lack of any information about the pathotype, serotype/serogroup, and/or origin of the isolates for a large number of the sequenced *E. coli* genomes in GenBank. Additional work combining literature survey for each sequenced *E. coli* strain with information from supplementary databases such as PATRIC (Pathosystems Resource Integration Center) (Wattam *et al.* 2017) or ClermonTyping for *in silico* phylotyping (Beghain *et al.* 2018) would be necessary to draw possible correlations.

The role of Ag43 and the autoaggregation in general have been questioned on numerous occasions in the scientific literature but their actual contribution for *E. coli* biology are far from clear (Trunk *et al.* 2018). Besides its cryptic role as a virulence factor, an additional function in bacterial physiopathology was proposed where Ag43 would promote uptake and survival in polymorphonuclear neutrophils (Fexby *et al.* 2007) but its significance remain to be evaluated (van der Woude *et al.* 2008). Respective to colonisation process, Ag43 is generally considered as participating to bacterial adhesion and biofilm formation (van der Woude *et al.* 2008) but its significant contribution to intestinal colonisation has never been evidenced (de Luna *et al.* 2008). However, the effect of the different classes of Ag43 remains to be evaluated with this respect. In fact, our new understanding of the diversity of Ag43 provides a rational framework for investigating its role in *E. coli* physiology, ecophysiology and physiopathology. For instance, its involvement in stratifying *E. coli* community in a complex ecosystem such as within the intestinal microbiota of human or animals is a hypothesis of particular interest that would request further investigations.

MATERIAL & METHODS

Data mining for identification of Ag43

Genomic data were retrieved from the NCBI assembly database that provides information on the structure of assembled genomes (Kitts *et al.* 2016). Ag43 homologues were recovered using a personal database of 106 protein sequences corresponding to the passenger domain and respective to the four classes of Ag43 recently uncovered (Ageorges *et al.* 2019). BLASTX analysis (E-value threshold 1^{e-100} and low-complexity filter disabled) was carried out with BLAST+ command-line tools version 2.7.1 (Camacho *et al.* 2009). All HSP (high scoring pairs) alignments with more than 90 % of similarity and encompassing more than 400 amino acids were considered for the following studies. In the first instance, a total of 10,587 genome sequences have been downloaded using *Escherichia coli* as keywords. According to the degree to which they have been assembled such sequences corresponded to 579 genomes, 58 chromosomes, 6,375 contigs and 3,575 scaffolds, respectively. These DNA sequences were used to extract the complete Ag43 protein sequences without considering annotation information from genomic data. Such approach allowed us to highlight mispredicted translation initiation sites (TISs) and incomplete sequences due to sequencing errors introducing frameshift, which were collectively considered as pseudogenes. In order to recover Ag43 homologues from *Enterobacteriaceae* other than *E. coli*, sequences from *Escherichia* genus but not referenced as *E. coli* species (147 sequences) and *Enterobacteriaceae* sequences excluding *Escherichia* genus (1211 sequences) have been also extracted from the NCBI assembly database, respectively. Finally, BLASTX analysis was performed against the NCBI assembly database excluding *Enterobacteriaceae* in order to recover Ag43 homologues in the Bacteria Kingdom other the ones identified here above.

Excluding pseudogenes, the 7,219 CDS (coding DNA sequence) for Ag43 recovered by this approach were clustered with a threshold of 98 % using Cd-hit software (Li *et al.* 2006, Li *et al.* 2006) and a reference sequence from each cluster was kept to generate a representative Ag43 dataset used for the phylogenetic studies as described below. Considering the modular architecture of Ag43, the Ag43 protein sequences were split as follows. For the identification of the SP, the sequences were submitted to SignalP v5.0 (Almagro Armenteros *et al.* 2019). From InterPro (IPR) database (Finn *et al.* 2017), the ESPR (IPR024973) was identified using InterProScan v5.22 (Jones *et al.* 2014). The passenger was identified following structural alignment as explained below and the SL, EJ and BL subdomains were identified from three-

dimensional (3D) structure using UCSF Chimera visualisation (Pettersen *et al.* 2004) as previously described (Heras *et al.* 2014, Ageorges *et al.* 2019). The AC1 domains were identified based on a structural alignment using Phyre v2.0 for automated modelling (Kelley *et al.* 2015) and UCSF Chimera for 3D structure visualisation as previously described (Rojas-Lopez *et al.* 2018). The translocators were identified using the twin HMM (Hidden Markov Model) strategy for AT detection (Celik *et al.* 2012). From full length Ag43, the sequences were then be split into modules corresponding to the ESPR, passenger, SL, EJ, BL, AC1 and translocator, also called translocation unit (TU).

Phylogenetic network analysis of protein sequences

For multiple sequence alignment of the full length Ag43 including the (i) passenger and its subdomains, (ii) AC1, and (iii) translocator, a special mode of T-Coffee (Notredame *et al.* 2000, Di Tommaso *et al.* 2011, Magis *et al.* 2014) was used to incorporate structural information, i.e. Expresso (Armougom *et al.* 2006), which is an extension of 3D-Coffee where structure based alignment is used as a template (O'Sullivan *et al.* 2004). Thus, the amino acid sequences were aligned using solved structures from the protein data bank (PDB) as templates (Berman *et al.* 2000), namely (i) the passenger from Ag43 (PDB: 4KH3), (ii) AC1 from IcsA (PDB: 3ML3), and (iii) the translocators comprising both the α-linker and β-barrel from AIDA-I (PDB: 4MEE) (Gawarzewski *et al.* 2014), EspP (PDB: 3SLJ) (Barnard *et al.* 2012), EstA (PDB: 3KVN) (van den Berg 2010), Hbp (PDB: 3AEH) (Tajima *et al.* 2010), and NalP (PDB: 1UYN) (Oomen *et al.* 2004), respectively. The relatedness among protein sequences was studied using a Neighbor-Net phylogenetic network approach with SplitsTree v4.14.8 using the Hamming uncorrected-P distance (Kloepper *et al.* 2008). The most relevant clusters, i.e. monophyletic groups or clades, were identified and selected based on splits or branches showing bootstrap values above 90 % over 1000 pseudo-replicates.

ACKNOWLEDGEMENTS

This work was supported in part by INRA (Institut National de la Recherche Agronomique) and by the CoMBa ("Colonisation de la Matrice extracellulaire Bacterienne") project funded by the FRI ("Fond Régional Innovation") cluster IRP ("Institut de recherche Pharmabiotique") involving the industrial partners Lallemand SAS and Biofilm Control. Valentin AGEORGES is a PhD Research Fellow granted by the "Conseil Région Auvergne – FEDER (Fonds Européen de Développement Régional)". The authors are grateful to the INRA MIGALE bioinformatics platform (http://migale.jouy.inra.fr) for providing computational resources.

REFERENCES

Ageorges, V., M. Schiavone, G. Jubelin, N. Caccia, P. Ruiz, X. Bailly, E. Dague, S. Leroy, J. Paxman, B. Heras, F. Chaucheyras-Durand, A. E. Rossiter, I. R. Henderson and M. Desvaux (2019). "Differential homotypic and heterotypic interactions of antigen 43 (Ag43) variants in autotransporter-mediated bacterial autoaggregation." *in submission*.

Almagro Armenteros, J. J., K. D. Tsirigos, C. K. Sonderby, T. N. Petersen, O. Winther, S. Brunak, G. von Heijne and H. Nielsen (2019). "SignalP 5.0 improves signal peptide predictions using deep neural networks." Nat Biotechnol **37**(4): 420-423.

Arenas, J., S. Cano, R. Nijland, V. van Dongen, L. Rutten, A. van der Ende and J. Tommassen (2015). "The meningococcal autotransporter AutA is implicated in autoaggregation and biofilm formation." Environ Microbiol **17**(4): 1321-1337.

Armbruster, K. M. and T. C. Meredith (2017). "Identification of the Lyso-Form N-acyl Intramolecular Transferase in Low-GC Firmicutes." J Bacteriol.

Armougom, F., S. Moretti, O. Poirot, S. Audic, P. Dumas, B. Schaeli, V. Keduas and C. Notredame (2006). "Expresso: automatic incorporation of structural information in multiple sequence alignments using 3D-Coffee." Nucleic Acids Res **34**(Web Server issue): W604-608.

Barnard, T. J., J. Gumbart, J. H. Peterson, N. Noinaj, N. C. Easley, N. Dautin, A. J. Kuszak, E. Tajkhorshid, H. D. Bernstein and S. K. Buchanan (2012). "Molecular basis for the activation of a catalytic asparagine residue in a self-cleaving bacterial autotransporter." J Mol Biol **415**(1): 128-142.

Beghain, J., A. Bridier-Nahmias, H. Le Nagard, E. Denamur and O. Clermont (2018). "ClermonTyping: an easy-to-use and accurate in silico method for *Escherichia* genus strain phylotyping." Microb Genom **4**(7).

Berman, H. M., J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov and P. E. Bourne (2000). "The Protein Data Bank." Nucleic Acids Res **28**(1): 235-242.

Camacho, C., G. Coulouris, V. Avagyan, N. Ma, J. Papadopoulos, K. Bealer and T. L. Madden (2009). "BLAST+: architecture and applications." BMC Bioinformatics **10**: 421.

Celik, N., C. T. Webb, D. L. Leyton, K. E. Holt, E. Heinz, R. Gorrell, T. Kwok, T. Naderer, R. A. Strugnell, T. P. Speed, R. D. Teasdale, V. A. Likic and T. Lithgow (2012). "A bioinformatic strategy for the detection, classification and analysis of bacterial autotransporters." PLoS One **7**(8): e43245.

Chaudhuri, R. R. and I. R. Henderson (2012). "The evolution of the *Escherichia coli* phylogeny." Infect Genet Evol **12**(2): 214-226.

Chung, P. Y. (2016). "The emerging problems of Klebsiella pneumoniae infections: carbapenem resistance and biofilm formation." FEMS Microbiol Lett **363**(20).

de Luna, M. G., A. Scott-Tucker, M. Desvaux, P. Ferguson, N. P. Morin, E. G. Dudley, S. Turner, J. P. Nataro, P. Owen and I. R. Henderson (2008). "The *Escherichia coli* biofilm-promoting protein Antigen 43 does not contribute to intestinal colonization." FEMS Microbiol Lett **284**(2): 237-246.

Desvaux, M., L. M. Cooper, N. A. Filenko, A. Scott-Tucker, S. M. Turner, J. A. Cole and I. R. Henderson (2006). "The unusual extended signal peptide region of the Type V secretion system is phylogenetically restricted." FEMS Microbiol Lett **264**(1): 22-30.

Desvaux, M., N. J. Parham and I. R. Henderson (2003). "Le système de sécrétion de type V chez les bactéries Gram-négatives (Type V secretion system in Gram-negative bacteria)." Biofutur **237**: 34-37.

Desvaux, M., N. J. Parham and I. R. Henderson (2004). "Type V protein secretion: simplicity gone awry?" Curr Issues Mol Biol **6**(2): 111-124.

Desvaux, M., A. Scott-Tucker, S. M. Turner, L. M. Cooper, D. Huber, J. P. Nataro and I. R. Henderson (2007). "A conserved extended signal peptide region directs posttranslational protein translocation via a novel mechanism." Microbiology **153**: 59-70.

Di Tommaso, P., S. Moretti, I. Xenarios, M. Orobitg, A. Montanyola, J. M. Chang, J. F. Taly and C. Notredame (2011). "T-Coffee: a web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension." Nucleic Acids Res **39**(Web Server issue): W13-17.

Drobnak, I., E. Braselmann, J. L. Chaney, D. L. Leyton, H. D. Bernstein, T. Lithgow, J. Luirink, J. P. Nataro and P. L. Clark (2015). "Of linkers and autochaperones: an unambiguous nomenclature to identify common and uncommon themes for autotransporter secretion." Mol Microbiol **95**(1): 1-16.

Fexby, S., T. Bjarnsholt, P. O. Jensen, V. Roos, N. Hoiby, M. Givskov and P. Klemm (2007). "Biological Trojan horse: Antigen 43 provides specific bacterial uptake and survival in human neutrophils." Infect Immun **75**(1): 30-34.

Finn, R. D., T. K. Attwood, P. C. Babbitt, A. Bateman, P. Bork, A. J. Bridge, H. Y. Chang, Z. Dosztanyi, S. El-Gebali, M. Fraser, J. Gough, D. Haft, G. L. Holliday, H. Huang, X. Huang, I. Letunic, R. Lopez, S. Lu, A. Marchler-Bauer, H. Mi, J. Mistry, D. A. Natale, M. Necci, G. Nuka, C. A. Orengo, Y. Park, S. Pesseat, D. Piovesan, S. C. Potter, N. D. Rawlings, N. Redaschi, L. Richardson, C. Rivoire, A. Sangrador-Vegas, C. Sigrist, I. Sillitoe, B. Smithers, S. Squizzato, G. Sutton, N. Thanki, P. D. Thomas, S. C. Tosatto, C. H. Wu, I. Xenarios, L. S. Yeh, S. Y. Young and A. L. Mitchell (2017). "InterPro in 2017-beyond protein family and domain annotations." Nucleic Acids Res **45**(D1): D190-D199.

Gawarzewski, I., F. DiMaio, E. Winterer, B. Tschapek, S. H. J. Smits, J. Jose and L. Schmitt (2014). "Crystal structure of the transport unit of the autotransporter adhesin involved in diffuse adherence from *Escherichia coli*." J Struct Biol **187**(1): 20-29.

Henderson, I. R., F. Navarro-Garcia, M. Desvaux, R. C. Fernandez and D. Ala'Aldeen (2004). "Type V protein secretion pathway: the autotransporter story." Microbiol Mol Biol Rev **68**(4): 692-744.

Heras, B., M. Totsika, K. M. Peters, J. J. Paxman, C. L. Gee, R. J. Jarrott, M. A. Perugini, A. E. Whitten and M. A. Schembri (2014). "The antigen 43 structure reveals a molecular Velcro-like mechanism of autotransporter-mediated bacterial clumping." Proc Natl Acad Sci U S A **111**(1): 457-462.

Jones, P., D. Binns, H. Y. Chang, M. Fraser, W. Li, C. McAnulla, H. McWilliam, J. Maslen, A. Mitchell, G. Nuka, S. Pesseat, A. F. Quinn, A. Sangrador-Vegas, M. Scheremetjew, S. Y. Yong, R. Lopez and S. Hunter (2014). "InterProScan 5: genome-scale protein function classification." Bioinformatics **30**(9): 1236-1240.

Jong, W. S., C. M. ten Hagen-Jongman, E. Ruijter, R. V. Orru, P. Genevaux and J. Luirink (2010). "YidC is involved in the biogenesis of the secreted autotransporter hemoglobin protease." J Biol Chem **285**(51): 39682-39690.

Kelley, L. A., S. Mezulis, C. M. Yates, M. N. Wass and M. J. Sternberg (2015). "The Phyre2 web portal for protein modeling, prediction and analysis." Nat Protoc **10**(6): 845-858.

Kitts, P. A., D. M. Church, F. Thibaud-Nissen, J. Choi, V. Hem, V. Sapojnikov, R. G. Smith, T. Tatusova, C. Xiang, A. Zherikov, M. DiCuccio, T. D. Murphy, K. D. Pruitt and A. Kimchi (2016). "Assembly: a resource for assembled genomes at NCBI." Nucleic Acids Res **44**(D1): D73-80.

Klemm, P., R. M. Vejborg and O. Sherlock (2006). "Self-associating autotransporters, SAATs: functional and structural similarities." Int J Med Microbiol **296**(4-5): 187-195.

Kloepper, T. H. and D. H. Huson (2008). "Drawing explicit phylogenetic networks and their integration into SplitsTree." BMC Evol Biol **8**: 22.

Kuhnel, K. and D. Diezmann (2011). "Crystal structure of the autochaperone region from the *Shigella flexneri* autotransporter IcsA." J Bacteriol **193**(8): 2042-2045.

Leo, J. C., I. Grin and D. Linke (2012). "Type V secretion: mechanism(s) of autotransport through the bacterial outer membrane." Philosophical Transactions of the Royal Society B-Biological Sciences **367**(1592): 1088-1101.

Li, W. and A. Godzik (2006). "Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences." Bioinformatics **22**(13): 1658-1659.

Li, W. and A. Godzik (2006). "VISSA: a program to visualize structural features from structure sequence alignment." Bioinformatics **22**(7): 887-888.

Magis, C., J. F. Taly, G. Bussotti, J. M. Chang, P. Di Tommaso, I. Erb, J. Espinosa-Carrasco and C. Notredame (2014). "T-Coffee: Tree-based consistency objective function for alignment evaluation." Methods Mol Biol **1079**: 117-129.

Maheux, A. F., D. K. Boudreau, M. G. Bergeron and M. J. Rodriguez (2014). "Characterization of *Escherichia fergusonii* and *Escherichia albertii* isolated from water." J Appl Microbiol **117**(2): 597-609.

Meng, G., N. Spahich, R. Kenjale, G. Waksman and J. W. St Geme, 3rd (2011). "Crystal structure of the *Haemophilus influenzae* Hap adhesin reveals an intercellular oligomerization mechanism for bacterial aggregation." EMBO J **30**(18): 3864-3874.

Notredame, C., D. G. Higgins and J. Heringa (2000). "T-Coffee: A novel method for fast and accurate multiple sequence alignment." J Mol Biol **302**(1): 205-217.

O'Sullivan, O., K. Suhre, C. Abergel, D. G. Higgins and C. Notredame (2004). "3DCoffee: combining protein sequences and structures within multiple sequence alignments." J Mol Biol **340**(2): 385-395.

Oliver, D. C., G. Huang, E. Nodel, S. Pleasance and R. C. Fernandez (2003). "A conserved region within the *Bordetella pertussis* autotransporter BrkA is necessary for folding of its passenger domain." Mol Microbiol **47**(5): 1367-1383.

Oomen, C. J., P. van Ulsen, P. van Gelder, M. Feijen, J. Tommassen and P. Gros (2004). "Structure of the translocator domain of a bacterial autotransporter." EMBO J **23**(6): 1257-1266.

Owen, P., M. Meehan, H. de Loughry-Doherty and I. Henderson (1996). "Phase-variable outer membrane proteins in *Escherichia coli*." FEMS Immunol Med Microbiol **16**(2): 63-76.

Pettersen, E. F., T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin (2004). "UCSF Chimera--a visualization system for exploratory research and analysis." J Comput Chem **25**(13): 1605-1612.

Reidl, S., A. Lehmann, R. Schiller, A. Salam Khan and U. Dobrindt (2009). "Impact of O-glycosylation on the molecular and cellular adhesion properties of the *Escherichia coli* autotransporter protein Ag43." Int J Med Microbiol **299**(6): 389-401.

Restieri, C., G. Garriss, M. C. Locas and C. M. Dozois (2007). "Autotransporter-encoding sequences are phylogenetically distributed among *Escherichia coli* clinical isolates and reference strains." Appl Environ Microbiol **73**(5): 1553-1562.

Roche, A., J. McFadden and P. Owen (2001). "Antigen 43, the major phase-variable protein of the *Escherichia coli* outer membrane, can exist as a family of proteins encoded by multiple alleles." Microbiology **147**(Pt 1): 161-169.

Rojas-Lopez, M., M. A. Zorgani, L. A. Kelley, X. Bailly, A. V. Kajava, I. R. Henderson, F. Polticelli, M. Pizza, R. Rosini and M. Desvaux (2018). "Identification of the autochaperone domain in the Type Va secretion system (T5aSS): Prevalent feature of autotransporters with a β-helical passenger." Front Microbiol **8**: 2607.

Swenson, C. E. and R. T. Sadikot (2015). "Achromobacter respiratory infections." Ann Am Thorac Soc **12**(2): 252-258.

Tajima, N., F. Kawai, S. Y. Park and J. R. Tame (2010). "A novel intein-like autoproteolytic mechanism in autotransporter proteins." J Mol Biol **402**(4): 645-656.

Torres, A. G., N. T. Perna, V. Burland, A. Ruknudin, F. R. Blattner and J. B. Kaper (2002). "Characterization of Cah, a calcium-binding and heat-extractable autotransporter protein of enterohaemorrhagic *Escherichia coli*." Mol Microbiol **45**(4): 951-966.

Trunk, T., H. Salah Khalil and J. C. Leo (2018). "Bacterial autoaggregation." AIMS Microbiol **4**(1): 140-164.

van den Berg, B. (2010). "Crystal structure of a full-length autotransporter." J Mol Biol **396**(3): 627-633.

van der Woude, M. W. and I. R. Henderson (2008). "Regulation and function of Ag43 (*flu*)." Annu Rev Microbiol **62**: 153-169.

Vo, J. L., G. C. Martinez Ortiz, P. Subedi, S. Keerthikumar, S. Mathivanan, J. J. Paxman and B. Heras (2017). "Autotransporter Adhesins in *Escherichia coli* Pathogenesis." Proteomics.

Wattam, A. R., J. J. Davis, R. Assaf, S. Boisvert, T. Brettin, C. Bun, N. Conrad, E. M. Dietrich, T. Disz, J. L. Gabbard, S. Gerdes, C. S. Henry, R. W. Kenyon, D. Machi, C. Mao, E. K. Nordberg, G. J. Olsen, D. E. Murphy-Olson, R. Olson, R. Overbeek, B. Parrello, G. D. Pusch, M. Shukla, V. Vonstein, A. Warren, F. Xia, H. Yoo and R. L. Stevens (2017). "Improvements to PATRIC, the all-bacterial Bioinformatics Database and Analysis Resource Center." Nucleic Acids Res **45**(D1): D535-D542.

Wells, T. J., O. Sherlock, L. Rivas, A. Mahajan, S. A. Beatson, M. Torpdahl, R. I. Webb, L. P. Allsopp, K. S. Gobius, D. L. Gally and M. A. Schembri (2008). "EhaA is a novel autotransporter protein of enterohemorrhagic *Escherichia coli* O157:H7 that contributes to adhesion and biofilm formation." Environ Microbiol **10**(3): 589-604.

DISCUSSION ET PERSPECTIVES

Identifications de déterminants moléculaires impliqués dans la colonisation de l'ECM par *E. coli* O157:H7

Les EHEC sont des bactéries pathogènes capables de coloniser tous les maillons de la chaine alimentaire : depuis leur réservoir bovin, elles peuvent contaminer les denrées alimentaires et infecter l'Homme via ingestion de ces aliments. Cette colonisation est notamment permise par l'interaction des bactéries avec l'ECM sous influence de facteurs environnementaux tels que la température et les conditions de croissance, suggérant l'implication de différents déterminants moléculaires appelés MSCRAMM (Chagnot *et al.* 2012, Chagnot *et al.* 2013). Le premier objectif de ma thèse était donc d'identifier dans la littérature de potentiels MSCRAMM et de caractériser leur implication dans l'adhésion d'*E. coli* O157:H7 à des protéines de l'ECM dans différentes conditions environnementales par une approche de génétique fonctionnelle. Ces travaux ont montré l'implication des curli, de l'injectisome, d'EhaA et de l'Ag43 dans l'adhésion des bactéries aux collagènes I et III quelles que soient les températures de croissance et d'adhésion. Si l'interaction des curli avec la laminine et la fibronectine a déjà été montrée (Saldana *et al.* 2009), ces résultats constituent la première mise en évidence de l'implication de l'injectisome, de l'Ag43 et d'EhaA dans le processus de colonisation de l'ECM par *E. coli* O157:H7. D'autres protéines de l'ECM devraient être testées afin de déterminer la spécificité de l'adhésion observée pour les différents MSCRAMM. A l'inverse, l'implication d'autres MSCRAMM comme le pilus HCP, le flagelle ou EhaB s'est révélée différente selon les conditions environnementales testées. Ceci suggère que les conditions environnementales peuvent affecter l'expression des différents MSCRAMM et/ou leur capacité d'interactions avec les protéines de l'ECM suite par exemple à un encombrement stérique. Des études complémentaires mimant d'autres conditions environnementales rencontrées aux différents niveaux de la chaine alimentaire sont nécessaires afin de déterminer la contribution exacte de ces MSCRAMM au processus de colonisation d'*E. coli* O157:H7 et d'identifier les conditions optimales d'expression et de régulation des gènes impliqués. Ces résultats pourraient permettre d'identifier les points critiques de risques de contamination à l'abattoir et/ou en milieu agroalimentaire et donc de mettre en place des bonnes pratiques ou des stratégies de prévention de contamination des aliments. De plus, le développement de stratégies antagonistes des interactions entre protéines MSCRAMM et ECM, incluant l'utilisation de souches microbiennes compétitrices, permettrait de limiter la

colonisation d'*E. coli* O157:H7 de certains environnements de la chaîne alimentaire, comme le TD du ruminant, en vue de limiter le portage asymptomatique.

Diversité et prévalence de l'Ag43 chez les bactéries

Par rapport aux deux sous-familles précédemment identifiées (SF-I et SF-II) (van der Woude *et al.* 2008), l'étude phylogénétique réalisée sur le domaine passager a permis d'apporter un regard nouveau sur la diversité de l'Ag43 avec la découverte de 4 classes, i.e. C1 à C4. Si tous ces Ag43 arborent un domaine passager β-hélicoïdal en forme de L similaire à celui résolu par cristallographie (Heras *et al.* 2014), les différences s'expliquent au niveau des sous domaines SL (tige de la lettre L), EJ (région de raccordement coudé) et BL (assise de la lettre L) appartenant chacun à deux types phylogénétiques distincts. Les 4 différentes classes correspondent donc à des combinaisons particulières des sous-domaines SL et des modules EJ-BL : (i) SL1-EJ1-BL1 pour Ag43 ^{C1}, (ii) SL2-EJ2-BL2 pour Ag43 ^{C2}, (iii) SL2-EJ1-BL1 pour Ag43 ^{C3}, et (iv) SL1-EJ2-BL2 pour Ag43 ^{C4}. Ces résultats nous ont amené à étudier la diversité phylogénétique de l'Ag43 sur l'ensemble des génomes bactériens disponibles dans les banques de données génomiques. Cette analyse a révélé l'existence de deux nouvelles classes d'Ag43, C5 et C6, associées à la division en nouveaux sous-types des sous-domaines SL1 et SL2 (SLla, SL1b, SL1c, SL2a et SL2b), leur association aux modules EJ1-BL1 ou EJ2-BL2 et la répartition du domaine AC dans le cluster E-Ж décrit précédemment (Rojas-Lopez *et al.* 2017). Au final, la diversité de l'Ag43 dépend de l'association des différents sous domaines SL, responsables du mécanisme de Velcro moléculaire (Heras *et al.* 2014), avec les sous domaines EJ1-BL1-AC ^{E} ou EJ2-BL2-AC ^{Ж}. Alors que des variations structurales entre différents Ag43 avaient déjà été observées (Klemm *et al.* 2004, Reidl *et al.* 2009), ces travaux constituent une réelle avancée dans la compréhension de la diversité phylogénétique et structurale de l'Ag43.

L'étude de la distribution et de la prévalence de l'*agn43* chez les bactéries a montré que si le gène est bel et bien présent essentiellement chez *E. coli*, sa présence n'est pas ubiquiste comme supposé jusqu'à présent (Henderson *et al.* 2004). Alors que l'*agn43* a été majoritairement retrouvé en un seul exemplaire, jusqu'à 5 copies du gène ont pu être identifiées dans un même génome avec différentes combinaisons de classes. Parmi ces différentes classes, les Ag43 de la classe C6 se sont avérés être les plus prévalents. Si la précédente étude phylogénétique suggérait une distribution plus importante de l'Ag43 chez des souches d'*E. coli* pathogènes (van der Woude *et al.* 2008), de futurs travaux basés sur nos données sont

nécessaires afin d'identifier la prévalence réelle des différentes classes d'Ag43 entre les souches commensales et pathogènes ainsi que l'association à des pathotypes spécifiques. Ces travaux offrent de nouvelles directions de recherche afin de comprendre le rôle biologique d'une telle diversité d'Ag43.

Caractérisation des capacités d'interaction homotypiques et hétérotypiques des différentes classes d'Ag43

La plupart des études consacrées au rôle de l'Ag43 ont pris pour modèle l'Ag43 d'*E. coli* K12 appartenant à la classe C1. L'objectif était donc de déterminer si les différences structurales observées entre les classes C1 à C4 étaient associées à des différences au niveau de la fonction d'autoagrégation. Alors que tous les représentants testés des Ag43 ^{C1} à Ag43 ^{C4} sont capables d'autoagréger grâce à des interactions homotypiques, les Ag43 ^{C3} ont montré une cinétique de sédimentation plus lente que les autres comme observé précédemment (Klemm *et al.* 2004) mais aussi contredit (Ulett *et al.* 2007), accompagnée d'agrégats plus petits et plus fragiles. L'Ag43 $^{C3\text{-}PT}$ arborant un module SL2b-EJ1-BL1-AC1 ^{E}, ces résultats suggèrent que ce module, et plus particulièrement le sous-domaine SL2b serait moins enclin à induire l'autoagrégation.

Si l'interaction hétérotypique de l'Ag43 avec les SAAT AIDA-I et TibA a déjà été démontrée, ce travail s'intéressait pour la première fois à l'étude des interactions hétérotypiques entre différentes classes d'Ag43. Ainsi, des interactions ont pu être observées entre l'Ag43 $^{C4\text{-}PT}$ et l'Ag43 $^{C1\text{-}PT}$ et à un degré moindre entre l'Ag43 $^{C4\text{-}PT}$ et l'Ag43 $^{C3\text{-}PT}$ suggérant des interactions des Ag43 possédant un module SL1b-EJ2-BL2-AC1 ^{Ж} avec ceux possédant des modules SL1a-EJ1-BL1-AC1 ^{E} et SL2b-EJ1-BL1-AC1 ^{E}, respectivement. Ceci suggère une meilleure capacité du sous-domaine SL1b à permettre l'autoagrégation. Les mêmes analyses devront être faites avec des représentants prototypiques des nouvelles classes C5 et C6 afin de déterminer leur capacité d'autoagrégation et d'interaction avec les autres classes d'Ag43. De plus, le seul Ag43 dont la structure a été caractérisée appartient à la classe 3 (Heras *et al.* 2014) ; la résolution des structures d'Ag43 appartenant aux 5 autres classes permettrait donc d'avoir une meilleure compréhension du rôle des différents modules dans l'autoagrégation. Par ailleurs, ces données permettraient d'étudier l'impact de ces modules sur l'établissement d'interactions hétérotypiques en s'intéressant à leur effet sur la structure en L du domaine passager et donc sa

capacité à s'auto-associer avec d'autres classes selon le modèle de Velcro moléculaire (Heras *et al.* 2014). De futures études devraient également caractériser le rôle de chacune des classes d'Ag43 dans l'adhésion bactérienne et à la formation de biofilms en conditions statiques et dynamiques, notamment à des protéines de l'ECM et des cellules épithéliales intestinales bovines ou humaines. Bien que l'Ag43 soit presque exclusivement retrouvé chez les *Enterobacteriacea*e, suggérant un rôle dans l'environnement digestif, il a été montré que l'Ag43 $^{C1\text{-}PT}$ n'était pas impliqué dans la colonisation de l'intestin (de Luna *et al.* 2008). Néanmoins, le rôle des Ag43 appartenant aux autres classes devrait être investigué ainsi que l'impact des interactions hétérotypiques sur ces différents phénotypes. Ces recherches permettraient d'éclaircir le rôle de l'Ag43 et de l'autoagrégation dans la physiologie, l'écophysiologie et la physiopathologie d'*E. coli*. Enfin, ces travaux ayant été réalisés *via* l'expression des différents Ag43 chez *E. coli* K12, le rôle et la régulation de ces protéines devront être explorés dans le fond génétique de leurs souches d'origine.

Rôle et régulation de l'Ag43 d'*E. coli* O157:H7

Le rôle de l'Ag43 $^{C2\text{-}PT}$ dans le phénomène d'autoagrégation dans la souche *E. coli* O157:H7 a été caractérisé. Nos résultats ont d'abord confirmé que l'autoagrégation de cette souche s'avère dépendante du milieu de culture en survenant dans un milieu riche chimiquement défini (DMEM) alors qu'elle n'était pas observable dans des milieux complexes, tels que le LB ou le BHI (Chagnot *et al.* 2013). En démontrant que c'est bien l'Ag43 qui est responsable de ce phénotype d'autoagrégation en milieu DMEM, nos résultats complètent les précédentes études indiquant que l'expression de l'*agn43* était induite dans ce milieu de culture alors que nous avons montré qu'il n'était que très faiblement exprimé en LB (Torres *et al.* 2002). Tandis qu'il serait nécessaire d'estimer quantitativement le niveau d'expression d'*agn43* dans différentes conditions environnementales en RT-qPCR, il est aussi nécessaire de considérer l'autoagrégation à deux échelles différentes : la population et la cellule. En effet, la forte expression de l'Ag43 en milieu DMEM permet la formation d'innombrables agrégats permettant à la population bactérienne de sédimenter alors que sa faible expression en LB pourrait n'être traduite que par l'autoagrégation de quelques bactéries, insuffisante pour être détectable à l'échelle de la population (Chauhan *et al.* 2013). De plus, cette souche possédant deux copies du gène, la mise en place d'analyses fonctionnelles permettrait de déterminer leurs contributions respectives aux phénomènes d'autoagrégation, d'adhésion à des protéines de l'ECM et des

cellules épithéliales intestinales ainsi qu'à la formation de biofilms afin de mieux caractériser les capacités de colonisation d'*E. coli* O157:H7 (Torres *et al.* 2002).

Les différences observées entre les milieux de culture suggèrent une régulation de l'expression de l'Ag43 d'*E. coli* O157:H7 différente. Nous avons pu démontrer que le phénomène d'autoagrégation de cette souche était régulé par la méthylase Dam et le régulateur de transcription OxyR comme observé chez *E. coli* K12 (van der Woude *et al.* 2008). Plus précisément, la surexpression de Dam ou l'absence d'OxyR ont été associées à l'autoagrégation d'*E. coli* O157:H7 en milieu LB. Il est donc probable que les conditions rencontrées dans le milieu DMEM pourraient conduire à un niveau d'expression de Dam plus important et favoriser *de facto* l'autoagrégation de la population bactérienne. En effet, ce milieu est déjà connu comme mimant les conditions optimales d'expression des facteurs de virulence codés par le LEE, en particulier le sodium bicarbonate mais dont le mécanisme d'action reste inconnu (Rosenshine *et al.* 1996, Kenny *et al.* 1997). Ainsi, le niveau d'expression de Dam et OxyR et l'état de méthylation de l'ADN d'*E. coli* O157:H7 pourraient être comparés entre le milieu DMEM et le milieu LB afin d'étudier l'impact naturel de cette régulation. De plus, une étude protéomique permettrait d'évaluer l'impact de la méthylation de l'ADN sur le protéome de surface et donc l'expression des protéines par des bactéries exprimant, ou non, Dam. Si OxyR et Dam régulent le phénomène d'autoagrégation, des travaux complémentaires sont nécessaires pour s'assurer que ces protéines influent bien directement sur la transcription des deux copies du gène d'*E. coli* O157:H7 et/ou s'ils sont régulés différemment ainsi que leurs niveaux d'expression dans différentes conditions environnementales. Compte-tenu du statut épigénétique de la variation de phase, la régulation de l'expression de l'Ag43 selon l'état de méthylation des brins d'ADN dans des conditions environnementales nécessiterait d'être plus précisément étudié chez les *E. coli* diarrhéiques.

En conclusion, ces travaux offrent un nouveau regard sur le rôle de l'Ag43 et sa régulation chez *E. coli* O157:H7 et suggèrent qu'il pourrait constituer un facteur de colonisation important des EHEC dans la chaîne alimentaire. A défaut d'une implication dans la physiopathologie des EHEC, la meilleure compréhension du rôle de l'Ag43 dans la physiologie et l'écophysiologie de ces bactéries et des *E. coli* diarrhéiques en général pourraient permettre à plus long terme de développer de nouvelles stratégies de prévention en vue de limiter le portage asymptomatique dans le réservoir animal ou la contamination d'aliments à risque et de diminuer *de facto* le risque d'infection humaine.

RÉFÉRENCES BIBLIOGRAPHIQUES

Abreu, A. G., C. M. Abe, K. O. Nunes, C. T. Moraes, L. Chavez-Duenas, F. Navarro-Garcia, A. S. Barbosa, R. M. Piazza and W. P. Elias (2016). "The serine protease Pic as a virulence factor of atypical enteropathogenic Escherichia coli." Gut Microbes **7**(2): 115-125.
Abreu, A. G., T. R. Fraga, A. P. Granados Martinez, M. Y. Kondo, M. A. Juliano, L. Juliano, F. Navarro-Garcia, L. Isaac, A. S. Barbosa and W. P. Elias (2015). "The Serine Protease Pic From Enteroaggregative Escherichia coli Mediates Immune Evasion by the Direct Cleavage of Complement Proteins." J Infect Dis **212**(1): 106-115.
Abu-Median, A. B., P. M. van Diemen, F. Dziva, I. Vlisidou, T. S. Wallis and M. P. Stevens (2006). "Functional analysis of lymphostatin homologues in enterohaemorrhagic Escherichia coli." FEMS Microbiol Lett **258**(1): 43-49.
AFSSA (2003). "Agence française de sécurité sanitaire des aliments). Bilan des connaissances relatives aux Escherichia coli producteurs de shiga-toxines (STEC).".
AFSSA (2008). "Agence française de sécurité sanitaire des aliments). Avis relatif aux souches d'Escherichia coli productrices de shiga-toxines considérées comme pathogènes pour l'homme.".
AFSSA (2010). "(Agence française de sécurité sanitaire des aliments). Avis relatif à la pertinence d'une révision de la définition des STEC pathogènes, précisée par l'avis Afssa du 15 juillet 2008.".
Aguilera, L., E. Ferreira, R. Gimenez, F. J. Fernandez, M. Taules, J. Aguilar, M. C. Vega, J. Badia and L. Baldoma (2012). "Secretion of the housekeeping protein glyceraldehyde-3-phosphate dehydrogenase by the LEE-encoded type III secretion system in enteropathogenic Escherichia coli." Int J Biochem Cell Biol **44**(6): 955-962.
Al Safadi, R., G. S. Abu-Ali, R. E. Sloup, J. T. Rudrik, C. M. Waters, K. A. Eaton and S. D. Manning (2012). "Correlation between in vivo biofilm formation and virulence gene expression in Escherichia coli O104:H4." PLoS One **7**(7): e41628.
Alnajar, S. and R. S. Gupta (2017). "Phylogenomics and comparative genomic studies delineate six main clades within the family Enterobacteriaceae and support the reclassification of several polyphyletic members of the family." Infect Genet Evol **54**: 108-127.
Altegoer, F., S. Mukherjee, W. Steinchen, P. Bedrunka, U. Linne, D. B. Kearns and G. Bange (2018). "FliS/flagellin/FliW heterotrimer couples type III secretion and flagellin homeostasis." Sci Rep **8**(1): 11552.
Amor, K., D. E. Heinrichs, E. Frirdich, K. Ziebell, R. P. Johnson and C. Whitfield (2000). "Distribution of core oligosaccharide types in lipopolysaccharides from Escherichia coli." Infect Immun **68**(3): 1116-1124.
Anderson, G. G., J. J. Palermo, J. D. Schilling, R. Roth, J. Heuser and S. J. Hultgren (2003). "Intracellular bacterial biofilm-like pods in urinary tract infections." Science **301**(5629): 105-107.
Andrade, F. B., A. G. Abreu, K. O. Nunes, T. A. T. Gomes, R. M. F. Piazza and W. P. Elias (2017). "Distribution of serine protease autotransporters of Enterobacteriaceae in typical and atypical enteroaggregative Escherichia coli." Infect Genet Evol **50**: 83-86.
Andreoli, S. P., H. Trachtman, D. W. Acheson, R. L. Siegler and T. G. Obrig (2002). "Hemolytic uremic syndrome: epidemiology, pathophysiology, and therapy." Pediatr Nephrol **17**(4): 293-298.
ANSES (2010). "Avis de l'Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail relatif à la révision de la définition des E. coli entéro-hémorragiques (EHEC) majeurs typiques, à l'appréciation quantitative des risques liés à ces bactéries à différentes étapes de la chaîne alimentaire, selon les différents modes de consommation des steaks hachés, et à la prise en compte du danger lié aux E. coli entéro-pathogènes (EPEC) dans les aliments. ."

ANSES (2017). "Avis de l'Agence nationale sanitaire de l'alimentation, de l'environnement et du travail relatif à la détection des E. coli producteurs de shiga-toxines (STEC) considérés comme hautement pathogènes en filière viande hachée bovine.".
ANSES (2017). "Les Escherichia coli entérohémorragiques (EHEC). Connaître la bactérie pour éviter les intoxications.".
Armstrong, G. L., J. Hollingsworth and J. G. Morris, Jr. (1996). "Emerging foodborne pathogens: Escherichia coli O157:H7 as a model of entry of a new pathogen into the food supply of the developed world." Epidemiol Rev **18**(1): 29-51.
Aschtgen, M. S., C. S. Bernard, S. De Bentzmann, R. Lloubes and E. Cascales (2008). "SciN is an outer membrane lipoprotein required for type VI secretion in enteroaggregative Escherichia coli." J Bacteriol **190**(22): 7523-7531.
Aschtgen, M. S., M. Gavioli, A. Dessen, R. Lloubes and E. Cascales (2010). "The SciZ protein anchors the enteroaggregative Escherichia coli Type VI secretion system to the cell wall." Mol Microbiol **75**(4): 886-899.
Ayers, M., P. L. Howell and L. L. Burrows (2010). "Architecture of the type II secretion and type IV pilus machineries." Future Microbiol **5**(8): 1203-1218.
Barnard, T. J., N. Dautin, P. Lukacik, H. D. Bernstein and S. K. Buchanan (2007). "Autotransporter structure reveals intra-barrel cleavage followed by conformational changes." Nat Struct Mol Biol **14**(12): 1214-1220.
Barrios, A. F., R. Zuo, D. Ren and T. K. Wood (2006). "Hha, YbaJ, and OmpA regulate Escherichia coli K12 biofilm formation and conjugation plasmids abolish motility." Biotechnol Bioeng **93**(1): 188-200.
Bauwens, A., J. Betz, I. Meisen, B. Kemper, H. Karch and J. Muthing (2013). "Facing glycosphingolipid-Shiga toxin interaction: dire straits for endothelial cells of the human vasculature." Cell Mol Life Sci **70**(3): 425-457.
Beghain, J., A. Bridier-Nahmias, H. Le Nagard, E. Denamur and O. Clermont (2018). "ClermonTyping: an easy-to-use and accurate in silico method for Escherichia genus strain phylotyping." Microb Genom.
Beloin, C., A. Roux and J. M. Ghigo (2008). "Escherichia coli biofilms." Curr Top Microbiol Immunol **322**: 249-289.
Benz, I. and M. A. Schmidt (1989). "Cloning and expression of an adhesin (AIDA-I) involved in diffuse adherence of enteropathogenic Escherichia coli." Infect Immun **57**(5): 1506-1511.
Benz, I. and M. A. Schmidt (2001). "Glycosylation with heptose residues mediated by the aah gene product is essential for adherence of the AIDA-I adhesin." Mol Microbiol **40**(6): 1403-1413.
Berger, C. N., V. F. Crepin, K. Baruch, A. Mousnier, I. Rosenshine and G. Frankel (2012). "EspZ of enteropathogenic and enterohemorrhagic Escherichia coli regulates type III secretion system protein translocation." MBio **3**(5).
Berger CN, S. S., Shaw RK et al. (2010). "Fresh fruit and vegetables as vehicles for the transmission of human pathogens." Environmental microbiology: 2385-2397.
Berry, A. A., Y. Yang, N. Pakharukova, J. A. Garnett, W. C. Lee, E. Cota, J. Marchant, S. Roy, M. Tuittila, B. Liu, K. G. Inman, F. Ruiz-Perez, I. Mandomando, J. P. Nataro, A. V. Zavialov and S. Matthews (2014). "Structural insight into host recognition by aggregative adherence fimbriae of enteroaggregative Escherichia coli." PLoS Pathog **10**(9): e1004404.
Berthiaume, F., N. Rutherford and M. Mourez (2007). "Mutations affecting the biogenesis of the AIDA-I autotransporter." Res Microbiol **158**(4): 348-354.
Beutin, L., M. Bulte, A. Weber, S. Zimmermann and K. Gleier (2000). "Investigation of human infections with verocytotoxin-producing strains of Escherichia coli (VTEC) belonging to serogroup O118 with evidence for zoonotic transmission." Epidemiol Infect **125**(1): 47-54.

Bielaszewska, M., R. Prager, L. Vandivinit, A. Müsken, A. Mellmann, N. J. Holt, P. I. Tarr, H. Karch, and W. Zhang. (2009). "Detection and Characterization of the Fimbrial sfp Cluster in Enterohemorrhagic Escherichia coli O165:H25/NM Isolates from Humans and Cattle." Appl. Environ. Microbiol.

Biscola, F. T., C. M. Abe and B. E. Guth (2011). "Determination of adhesin gene sequences in, and biofilm formation by, O157 and non-O157 Shiga toxin-producing Escherichia coli strains isolated from different sources." Appl Environ Microbiol **77**(7): 2201-2208.

Bockelmann, U., A. Janke, R. Kuhn, T. R. Neu, J. Wecke, J. R. Lawrence and U. Szewzyk (2006). "Bacterial extracellular DNA forming a defined network-like structure." FEMS Microbiol Lett **262**(1): 31-38.

Boisen, N., C. Struve, F. Scheutz, K. A. Krogfelt and J. P. Nataro (2008). "New adhesin of enteroaggregative Escherichia coli related to the Afa/Dr/AAF family." Infect Immun **76**(7): 3281-3292.

Bolton, D. J., C. M. Byrne, J. J. Sheridan, D. A. McDowell and I. S. Blair (1999). "The survival characteristics of a non-toxigenic strain of Escherichia coli O157:H7." J Appl Microbiol **86**(3): 407-411.

Booher, S. L., N. A. Cornick and H. W. Moon (2002). "Persistence of Escherichia coli O157:H7 in experimentally infected swine." Vet Microbiol **89**(1): 69-81.

Boudeau, J., A. L. Glasser, E. Masseret, B. Joly and A. Darfeuille-Michaud (1999). "Invasive ability of an Escherichia coli strain isolated from the ileal mucosa of a patient with Crohn's disease." Infect Immun **67**(9): 4499-4509.

Bray, D. (2001). "Cell movements: from molecules to motility." Second Edition. Taylor & Francis Group, Garland Publishing, New York, USA.

Brenner, D. J., B. R. Davis, A. G. Steigerwalt, C. F. Riddle, A. C. McWhorter, S. D. Allen, J. J. Farmer, 3rd, Y. Saitoh and G. R. Fanning (1982). "Atypical biogroups of Escherichia coli found in clinical specimens and description of Escherichia hermannii sp. nov." J Clin Microbiol **15**(4): 703-713.

Brenner, D. J., G. R. Fanning, F. J. Skerman and S. Falkow (1972). "Polynucleotide sequence divergence among strains of Escherichia coli and closely related organisms." J Bacteriol **109**(3): 953-965.

Brenner, D. J., A. C. McWhorter, J. K. Knutson and A. G. Steigerwalt (1982). "Escherichia vulneris: a new species of Enterobacteriaceae associated with human wounds." J Clin Microbiol **15**(6): 1133-1140.

Brunder, W., H. Karch and H. Schmidt (2006). "Complete sequence of the large virulence plasmid pSFO157 of the sorbitol-fermenting enterohemorrhagic Escherichia coli O157:H- strain 3072/96." Int J Med Microbiol **296**(7): 467-474.

Bryan, A., I. Youngster and A. J. McAdam (2015). "Shiga Toxin Producing Escherichia coli." Clin Lab Med **35**(2): 247-272.

Brzuszkiewicz, E., A. Thurmer, J. Schuldes, A. Leimbach, H. Liesegang, F. D. Meyer, J. Boelter, H. Petersen, G. Gottschalk and R. Daniel (2011). "Genome sequence analyses of two isolates from the recent Escherichia coli outbreak in Germany reveal the emergence of a new pathotype: Entero-Aggregative-Haemorrhagic Escherichia coli (EAHEC)." Arch Microbiol **193**(12): 883-891.

Burgess, N. R., S. N. McDermott and J. Whiting (1973). "Laboratory transmission of enterobacteriaceae by the oriental cockroach, Blatta orientalis." J Hyg (Lond) **71**(1): 9-14.

Campellone, K. G., D. Robbins and J. M. Leong (2004). "EspFU is a translocated EHEC effector that interacts with Tir and N-WASP and promotes Nck-independent actin assembly." Dev Cell **7**(2): 217-228.

Capitani, G., O. Eidam, R. Glockshuber and M. G. Grutter (2006). "Structural and functional insights into the assembly of type 1 pili from Escherichia coli." Microbes Infect **8**(8): 2284-2290.
Caprioli, A., S. Morabito, H. Brugere and E. Oswald (2005). "Enterohaemorrhagic Escherichia coli: emerging issues on virulence and modes of transmission." Vet Res **36**(3): 289-311.
Chagnot, C., A. Agus, S. Renier, F. Peyrin, R. Talon, T. Astruc and M. Desvaux (2013). "In vitro colonization of the muscle extracellular matrix components by Escherichia coli O157:H7: the influence of growth medium, temperature and pH on initial adhesion and induction of biofilm formation by collagens I and III." PLoS One **8**(3): e59386.
Chagnot, C., A. Listrat, T. Astruc and M. Desvaux (2012). "Bacterial adhesion to animal tissues: protein determinants for recognition of extracellular matrix components." Cell Microbiol **14**(11): 1687-1696.
Chagnot, C., M. A. Zorgani, T. Astruc and M. Desvaux (2013). "Proteinaceous determinants of surface colonization in bacteria: bacterial adhesion and biofilm formation from a protein secretion perspective." Front Microbiol **4**: 303.
Charbonneau, M. E. and M. Mourez (2007). "Functional organization of the autotransporter adhesin involved in diffuse adherence." J Bacteriol **189**(24): 9020-9029.
Chart, H., T. Cheasty and G. A. Willshaw (2002). "Production of serum antibodies that recognise epitopes located on the R3 region of Escherichia coli core lipopolysaccharides by patients infected with strains of enterohaemorrhagic E. coli." J Med Microbiol **51**(12): 1050-1054.
Chaucheyras-Durand, F., F. Faqir, A. Ameilbonne, C. Rozand and C. Martin (2010). "Fates of acid-resistant and non-acid-resistant Shiga toxin-producing Escherichia coli strains in ruminant digestive contents in the absence and presence of probiotics." Appl Environ Microbiol **76**(3): 640-647.
Chaucheyras-Durand, F., J. Madic, F. Doudin and C. Martin (2006). "Biotic and abiotic factors influencing in vitro growth of Escherichia coli O157:H7 in ruminant digestive contents." Appl Environ Microbiol **72**(6): 4136-4142.
Chaudhuri, R. R. and I. R. Henderson (2012). "The evolution of the Escherichia coli phylogeny." Infect Genet Evol **12**(2): 214-226.
Chauhan, A., C. Sakamoto, J. M. Ghigo and C. Beloin (2013). "Did I pick the right colony? Pitfalls in the study of regulation of the phase variable antigen 43 adhesin." PLoS One **8**(9): e73568.
Chevalier, N., M. Moser, H. G. Koch, K. L. Schimz, E. Willery, C. Locht, F. Jacob-Dubuisson and M. Muller (2004). "Membrane targeting of a bacterial virulence factor harbouring an extended signal peptide." J Mol Microbiol Biotechnol **8**(1): 7-18.
Christie, P. J., N. Whitaker and C. Gonzalez-Rivera (2014). "Mechanism and structure of the bacterial type IV secretion systems." Biochim Biophys Acta **1843**(8): 1578-1591.
Clermont, O., S. Bonacorsi and E. Bingen (2000). "Rapid and simple determination of the Escherichia coli phylogenetic group." Appl Environ Microbiol **66**(10): 4555-4558.
Clermont, O., J. K. Christenson, E. Denamur and D. M. Gordon (2013). "The Clermont Escherichia coli phylo-typing method revisited: improvement of specificity and detection of new phylo-groups." Environ Microbiol Rep **5**(1): 58-65.
Clermont, O., D. M. Gordon, S. Brisse, S. T. Walk and E. Denamur (2011). "Characterization of the cryptic Escherichia lineages: rapid identification and prevalence." Environ Microbiol **13**(9): 2468-2477.
Cookson, A. L., W. A. Cooley and M. J. Woodward (2002). "The role of type 1 and curli fimbriae of Shiga toxin-producing Escherichia coli in adherence to abiotic surfaces." Int J Med Microbiol **292**(3-4): 195-205.

Coombes, B. K., M. E. Wickham, M. Mascarenhas, S. Gruenheid, B. B. Finlay and M. A. Karmali (2008). "Molecular analysis as an aid to assess the public health risk of non-O157 Shiga toxin-producing Escherichia coli strains." Appl Environ Microbiol **74**(7): 2153-2160.
Correnti, J., V. Munster, T. Chan and M. Woude (2002). "Dam-dependent phase variation of Ag43 in Escherichia coli is altered in a seqA mutant." Mol Microbiol **44**(2): 521-532.
Cote, J. P., M. E. Charbonneau and M. Mourez (2013). "Glycosylation of the Escherichia coli TibA self-associating autotransporter influences the conformation and the functionality of the protein." PLoS One **8**(11): e80739.
Cotter, S. E., N. K. Surana and J. W. St Geme, 3rd (2005). "Trimeric autotransporters: a distinct subfamily of autotransporter proteins." Trends Microbiol **13**(5): 199-205.
Craig, L., M. E. Pique and J. A. Tainer (2004). "Type IV pilus structure and bacterial pathogenicity." Nat Rev Microbiol **2**(5): 363-378.
Croxen, M. A. and B. B. Finlay (2010). "Molecular mechanisms of Escherichia coli pathogenicity." Nat Rev Microbiol **8**(1): 26-38.
Croxen, M. A., R. J. Law, R. Scholz, K. M. Keeney, M. Wlodarska and B. B. Finlay (2013). "Recent advances in understanding enteric pathogenic Escherichia coli." Clin Microbiol Rev **26**(4): 822-880.
Currie, C. G. and I. R. Poxton (1999). "The lipopolysaccharide core type of Escherichia coli O157:H7 and other non-O157 verotoxin-producing E. coli." FEMS Immunol Med Microbiol **24**(1): 57-62.
Danese, P. N., L. A. Pratt, S. L. Dove and R. Kolter (2000). "The outer membrane protein, antigen 43, mediates cell-to-cell interactions within Escherichia coli biofilms." Mol Microbiol **37**(2): 424-432.
Dautin, N. (2010). "Serine protease autotransporters of enterobacteriaceae (SPATEs): biogenesis and function." Toxins (Basel) **2**(6): 1179-1206.
Dautin, N., T. J. Barnard, D. E. Anderson and H. D. Bernstein (2007). "Cleavage of a bacterial autotransporter by an evolutionarily convergent autocatalytic mechanism." EMBO J **26**(7): 1942-1952.
de Luna, M., A. Scott-Tucker, M. Desvaux, P. Ferguson, N. P. Morin, E. G. Dudley, S. Turner, J. P. Nataro, P. Owen and I. R. Henderson (2008). "The Escherichia coli biofilm-promoting protein Antigen 43 does not contribute to intestinal colonization." FEMS Microbiol Lett **284**(2): 237-246.
Deacon, V., F. Dziva, P. M. van Diemen, G. Frankel and M. P. Stevens (2010). "Efa-1/LifA mediates intestinal colonization of calves by enterohaemorrhagic Escherichia coli O26 : H- in a manner independent of glycosyltransferase and cysteine protease motifs or effects on type III secretion." Microbiology **156**(Pt 8): 2527-2536.
Decanio, M. S., R. Landick and R. J. Haft (2013). "The non-pathogenic Escherichia coli strain W secretes SslE via the virulence-associated type II secretion system beta." BMC Microbiol **13**: 130.
Delmas, Y., B. Vendrely, B. Clouzeau, H. Bachir, H. N. Bui, A. Lacraz, S. Helou, C. Bordes, A. Reffet, B. Llanas, S. Skopinski, P. Rolland, D. Gruson and C. Combe (2014). "Outbreak of Escherichia coli O104:H4 haemolytic uraemic syndrome in France: outcome with eculizumab." Nephrol Dial Transplant **29**(3): 565-572.
Deng, W., Y. Li, P. R. Hardwidge, E. A. Frey, R. A. Pfuetzner, S. Lee, S. Gruenheid, N. C. Strynakda, J. L. Puente and B. B. Finlay (2005). "Regulation of type III secretion hierarchy of translocators and effectors in attaching and effacing bacterial pathogens." Infect Immun **73**(4): 2135-2146.
Desvaux, M., M. Hebraud, I. R. Henderson and M. J. Pallen (2006). "Type III secretion: what's in a name?" Trends Microbiol **14**(4): 157-160.

Desvaux, M., M. Hebraud, R. Talon and I. R. Henderson (2009). "Secretion and subcellular localizations of bacterial proteins: a semantic awareness issue." Trends Microbiol **17**(4): 139-145.
Desvaux, M., N. J. Parham and I. R. Henderson (2003). "Type-V secretion system in gram-negative bacteria." Biofutur(237): 34-+.
Desvaux, M., N. J. Parham and I. R. Henderson (2004). "The autotransporter secretion system." Res Microbiol **155**(2): 53-60.
Desvaux, M., N. J. Parham, A. Scott-Tucker and I. R. Henderson (2004). "The general secretory pathway: a general misnomer?" Trends Microbiol **12**(7): 306-309.
Desvaux, M., A. Scott-Tucker, S. M. Turner, L. M. Cooper, D. Huber, J. P. Nataro and I. R. Henderson (2007). "A conserved extended signal peptide region directs posttranslational protein translocation via a novel mechanism." Microbiology **153**(Pt 1): 59-70.
Diderichsen, B. (1980). "flu, a metastable gene controlling surface properties of Escherichia coli." J Bacteriol **141**(2): 858-867.
DiLuzio, W. R., L. Turner, M. Mayer, P. Garstecki, D. B. Weibel, H. C. Berg and G. M. Whitesides (2005). "Escherichia coli swim on the right-hand side." Nature **435**(7046): 1271-1274.
Donlan, R. M. (2002). "Biofilms: microbial life on surfaces." Emerg Infect Dis **8**(9): 881-890.
Doughty, S., J. Sloan, V. Bennett-Wood, M. Robertson, R. M. Robins-Browne and E. L. Hartland (2002). "Identification of a novel fimbrial gene cluster related to long polar fimbriae in locus of enterocyte effacement-negative strains of enterohemorrhagic Escherichia coli." Infect Immun **70**(12): 6761-6769.
Drobnak, I., E. Braselmann, J. L. Chaney, D. L. Leyton, H. D. Bernstein, T. Lithgow, J. Luirink, J. P. Nataro and P. L. Clark (2015). "Of linkers and autochaperones: an unambiguous nomenclature to identify common and uncommon themes for autotransporter secretion." Mol Microbiol **95**(1): 1-16.
Dudley, E. G., N. R. Thomson, J. Parkhill, N. P. Morin and J. P. Nataro (2006). "Proteomic and microarray characterization of the AggR regulon identifies a pheU pathogenicity island in enteroaggregative Escherichia coli." Mol Microbiol **61**(5): 1267-1282.
Duffy, G., C. M. Burgess and D. J. Bolton (2014). "A review of factors that affect transmission and survival of verocytotoxigenic Escherichia coli in the European farm to fork beef chain." Meat Sci **97**(3): 375-383.
Durand, E., S. Alphonse, C. Brochier-Armanet, G. Ball, B. Douzi, A. Filloux, C. Bernard and R. Voulhoux (2011). "The assembly mode of the pseudopilus: a hallmark to distinguish a novel secretion system subtype." J Biol Chem **286**(27): 24407-24416.
Dutta, P. R., B. Q. Sui and J. P. Nataro (2003). "Structure-function analysis of the enteroaggregative Escherichia coli plasmid-encoded toxin autotransporter using scanning linker mutagenesis." J Biol Chem **278**(41): 39912-39920.
Dziva, F., A. Mahajan, P. Cameron, C. Currie, I. J. McKendrick, T. S. Wallis, D. G. Smith and M. P. Stevens (2007). "EspP, a Type V-secreted serine protease of enterohaemorrhagic Escherichia coli O157:H7, influences intestinal colonization of calves and adherence to bovine primary intestinal epithelial cells." FEMS Microbiol Lett **271**(2): 258-264.
Dziva, F., P. M. van Diemen, M. P. Stevens, A. J. Smith and T. S. Wallis (2004). "Identification of Escherichia coli O157 : H7 genes influencing colonization of the bovine gastrointestinal tract using signature-tagged mutagenesis." Microbiology **150**(Pt 11): 3631-3645.
Easton, D. M., L. P. Allsopp, M. D. Phan, D. G. Moriel, G. K. Goh, S. A. Beatson, T. J. Mahony, R. N. Cobbold and M. A. Schembri (2014). "The Intimin-Like Protein FdeC Is Regulated by H-NS and Temperature in Enterohemorrhagic Escherichia coli." Appl Environ Microbiol **80**(23): 7337-7347.

Easton, D. M., M. Totsika, L. P. Allsopp, M. D. Phan, A. Idris, D. J. Wurpel, O. Sherlock, B. Zhang, C. Venturini, S. A. Beatson, T. J. Mahony, R. N. Cobbold and M. A. Schembri (2011). "Characterization of EhaJ, a New Autotransporter Protein from Enterohemorrhagic and Enteropathogenic Escherichia coli." Front Microbiol **2**: 120.
Eckburg, P. B., E. M. Bik, C. N. Bernstein, E. Purdom, L. Dethlefsen, M. Sargent, S. R. Gill, K. E. Nelson and D. A. Relman (2005). "Diversity of the human intestinal microbial flora." Science **308**(5728): 1635-1638.
EFSA (2013). "Scientific Opinion on the public health hazards to be covered by inspection of meat (bovine animals)." EFSA Journal 2013;11(6):3266.
EFSA (2015). "Public health risks associated with Enteroaggregative Escherichia coli (EAEC) as a food-borne pathogen." EFSA Journal 2015;13(12):4330.
EFSA (2018). "The European Union summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in 2017." EFSA Journal 2018;16(12):5500.
EFSA, E. (2016). "(European Food Safety Authority and European Center for Disease Prevention and Control). The European Union summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in 2015." Journal 2016, 14(12), 4634,4231.
Egea, L., L. Aguilera, R. Gimenez, M. A. Sorolla, J. Aguilar, J. Badia and L. Baldoma (2007). "Role of secreted glyceraldehyde-3-phosphate dehydrogenase in the infection mechanism of enterohemorrhagic and enteropathogenic Escherichia coli: interaction of the extracellular enzyme with human plasminogen and fibrinogen." Int J Biochem Cell Biol **39**(6): 1190-1203.
Elliott, S. J. and J. B. Kaper (1997). "Role of type 1 fimbriae in EPEC infections." Microb Pathog **23**(2): 113-118.
Elliott, S. J., V. Sperandio, J. A. Giron, S. Shin, J. L. Mellies, L. Wainwright, S. W. Hutcheson, T. K. McDaniel and J. B. Kaper (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**(11): 6115-6126.
Enami, M., N. Nakasone, Y. Honma, S. Kakinohana, J. Kudaka and M. Iwanaga (1999). "Expression of type I pili is abolished in verotoxin-producing Escherichia coli O157." FEMS Microbiol Lett **179**(2): 467-472.
Escherich, T. (1988). "The intestinal bacteria of the neonate and breast-fed infant. 1884." Rev Infect Dis **10**(6): 1220-1225.
Etienne-Mesmin, L., B. Chassaing, P. Sauvanet, J. Denizot, S. Blanquet-Diot, A. Darfeuille-Michaud, N. Pradel and V. Livrelli (2011). "Interactions with M Cells and Macrophages as Key Steps in the Pathogenesis of Enterohemorragic Escherichia coli Infections." Plos One **6**(8).
Evans, M. L. and M. R. Chapman (2014). "Curli biogenesis: order out of disorder." Biochim Biophys Acta **1843**(8): 1551-1558.
Fairbrother, J. M. and E. Nadeau (2006). "Escherichia coli: on-farm contamination of animals." Rev Sci Tech **25**(2): 555-569.
Farfan, M. J., L. Cantero, R. Vidal, D. J. Botkin and A. G. Torres (2011). "Long polar fimbriae of enterohemorrhagic Escherichia coli O157:H7 bind to extracellular matrix proteins." Infect Immun **79**(9): 3744-3750.
Farfan, M. J., K. G. Inman and J. P. Nataro (2008). "The major pilin subunit of the AAF/II fimbriae from enteroaggregative Escherichia coli mediates binding to extracellular matrix proteins." Infect Immun **76**(10): 4378-4384.
Farfan, M. J. and A. G. Torres (2012). "Molecular mechanisms that mediate colonization of Shiga toxin-producing Escherichia coli strains." Infect Immun **80**(3): 903-913.
Farmer, J. J., 3rd, G. R. Fanning, B. R. Davis, C. M. O'Hara, C. Riddle, F. W. Hickman-Brenner, M. A. Asbury, V. A. Lowery, 3rd and D. J. Brenner (1985). "Escherichia fergusonii and

Enterobacter taylorae, two new species of Enterobacteriaceae isolated from clinical specimens." J Clin Microbiol **21**(1): 77-81.
Fexby, S., T. Bjarnsholt, P. O. Jensen, V. Roos, N. Hoiby, M. Givskov and P. Klemm (2007). "Biological Trojan horse: Antigen 43 provides specific bacterial uptake and survival in human neutrophils." Infect Immun **75**(1): 30-34.
Fitzhenry, R., S. Dahan, A. G. Torres, Y. Chong, R. Heuschkel, S. H. Murch, M. Thomson, J. B. Kaper, G. Frankel and A. D. Phillips (2006). "Long polar fimbriae and tissue tropism in Escherichia coli O157:H7." Microbes Infect **8**(7): 1741-1749.
France, S. P. (2018). "Epidémie de syndrome hémolytique et urémique pédiatrique à Escherichia coli O26 en France métropolitaine en lien avec la consommation de reblochon : point au 15 juin 2018."
Frank, C., D. Werber, J. P. Cramer, M. Askar, M. Faber, M. an der Heiden, H. Bernard, A. Fruth, R. Prager, A. Spode, M. Wadl, A. Zoufaly, S. Jordan, M. J. Kemper, P. Follin, L. Muller, L. A. King, B. Rosner, U. Buchholz, K. Stark, G. Krause and H. U. S. I. Team (2011). "Epidemic profile of Shiga-toxin-producing Escherichia coli O104:H4 outbreak in Germany." N Engl J Med **365**(19): 1771-1780.
Free, A. L., H. A. Duoss, L. V. Bergeron, S. A. Shields-Menard, E. Ward, T. R. Callaway, J. A. Carroll, T. B. Schmidt and J. R. Donaldson (2012). "Survival of O157:H7 and non-O157 serogroups of Escherichia coli in bovine rumen fluid and bile salts." Foodborne Pathog Dis **9**(11): 1010-1014.
Fremaux, B., C. Prigent-Combaret and C. Vernozy-Rozand (2008). "Long-term survival of Shiga toxin-producing Escherichia coli in cattle effluents and environment: an updated review." Vet Microbiol **132**(1-2): 1-18.
Galan, J. E. and G. Waksman (2018). "Protein-Injection Machines in Bacteria." Cell **172**(6): 1306-1318.
Galfi, P., S. Neogrady, G. Semjen, S. Bardocz and A. Pusztai (1998). "Attachment of different Escherichia coli strains to cultured rumen epithelial cells." Vet Microbiol **61**(3): 191-197.
Gaytan, M. O., V. I. Martinez-Santos, E. Soto and B. Gonzalez-Pedrajo (2016). "Type Three Secretion System in Attaching and Effacing Pathogens." Front Cell Infect Microbiol **6**: 129.
Ghigo, J. M. (2001). "Natural conjugative plasmids induce bacterial biofilm development." Nature **412**(6845): 442-445.
Goyal, P., P. V. Krasteva, N. Van Gerven, F. Gubellini, I. Van den Broeck, A. Troupiotis-Tsailaki, W. Jonckheere, G. Pehau-Arnaudet, J. S. Pinkner, M. R. Chapman, S. J. Hultgren, S. Howorka, R. Fronzes and H. Remaut (2014). "Structural and mechanistic insights into the bacterial amyloid secretion channel CsgG." Nature **516**(7530): 250-253.
Grauke, L. J., I. T. Kudva, J. W. Yoon, C. W. Hunt, C. J. Williams and C. J. Hovde (2002). "Gastrointestinal tract location of Escherichia coli O157:H7 in ruminants." Appl Environ Microbiol **68**(5): 2269-2277.
Gutierrez-Jimenez, J., I. Arciniega and F. Navarro-Garcia (2008). "The serine protease motif of Pic mediates a dose-dependent mucolytic activity after binding to sugar constituents of the mucin substrate." Microb Pathog **45**(2): 115-123.
Haagmans, W. and M. van der Woude (2000). "Phase variation of Ag43 in Escherichia coli: Dam-dependent methylation abrogates OxyR binding and OxyR-mediated repression of transcription." Mol Microbiol **35**(4): 877-887.
Hagan, E. C. and H. L. Mobley (2007). "Uropathogenic Escherichia coli outer membrane antigens expressed during urinary tract infection." Infect Immun **75**(8): 3941-3949.
Haiko, J. and B. Westerlund-Wikstrom (2013). "The role of the bacterial flagellum in adhesion and virulence." Biology (Basel) **2**(4): 1242-1267.

Han, L., J. Zheng, Y. Wang, X. Yang, Y. Liu, C. Sun, B. Cao, H. Zhou, D. Ni, J. Lou, Y. Zhao and Y. Huang (2016). "Structure of the BAM complex and its implications for biogenesis of outer-membrane proteins." Nat Struct Mol Biol **23**(3): 192-196.
Hasman, H., T. Chakraborty and P. Klemm (1999). "Antigen-43-mediated autoaggregation of Escherichia coli is blocked by fimbriation." J Bacteriol **181**(16): 4834-4841.
Hasman, H., M. A. Schembri and P. Klemm (2000). "Antigen 43 and type 1 fimbriae determine colony morphology of Escherichia coli K-12." J Bacteriol **182**(4): 1089-1095.
Hata, H., T. Natori, T. Mizuno, I. Kanazawa, I. Eldesouky, M. Hayashi, M. Miyata, H. Fukunaga, S. Ohji, A. Hosoyama, E. Aono, A. Yamazoe, K. Tsuchikane, N. Fujita and T. Ezaki (2016). "Phylogenetics of family Enterobacteriaceae and proposal to reclassify Escherichia hermannii and Salmonella subterranea as Atlantibacter hermannii and Atlantibacter subterranea gen. nov., comb. nov." Microbiol Immunol **60**(5): 303-311.
Hausner, M. and S. Wuertz (1999). "High rates of conjugation in bacterial biofilms as determined by quantitative in situ analysis." Appl Environ Microbiol **65**(8): 3710-3713.
Henderson, B. and A. Martin (2011). "Bacterial virulence in the moonlight: multitasking bacterial moonlighting proteins are virulence determinants in infectious disease." Infect Immun **79**(9): 3476-3491.
Henderson, I. R., M. Meehan and P. Owen (1997). "Antigen 43, a phase-variable bipartite outer membrane protein, determines colony morphology and autoaggregation in Escherichia coli K-12." FEMS Microbiol Lett **149**(1): 115-120.
Henderson, I. R., M. Meehan and P. Owen (1997). "A novel regulatory mechanism for a novel phase-variable outer membrane protein of Escherichia coli." Adv Exp Med Biol **412**: 349-355.
Henderson, I. R., F. Navarro-Garcia, M. Desvaux, R. C. Fernandez and D. Ala'Aldeen (2004). "Type V protein secretion pathway: the autotransporter story." Microbiol Mol Biol Rev **68**(4): 692-744.
Henderson, I. R., F. Navarro-Garcia and J. P. Nataro (1998). "The great escape: structure and function of the autotransporter proteins." Trends Microbiol **6**(9): 370-378.
Henderson, I. R. and P. Owen (1999). "The major phase-variable outer membrane protein of Escherichia coli structurally resembles the immunoglobulin A1 protease class of exported protein and is regulated by a novel mechanism involving Dam and oxyR." J Bacteriol **181**(7): 2132-2141.
Henderson, I. R., P. Owen and J. P. Nataro (1999). "Molecular switches--the ON and OFF of bacterial phase variation." Mol Microbiol **33**(5): 919-932.
Heras, B., M. Totsika, K. M. Peters, J. J. Paxman, C. L. Gee, R. J. Jarrott, M. A. Perugini, A. E. Whitten and M. A. Schembri (2014). "The antigen 43 structure reveals a molecular Velcro-like mechanism of autotransporter-mediated bacterial clumping." Proc Natl Acad Sci U S A **111**(1): 457-462.
Herold, S., J. C. Paton and A. W. Paton (2009). "Sab, a novel autotransporter of locus of enterocyte effacement-negative shiga-toxigenic Escherichia coli O113:H21, contributes to adherence and biofilm formation." Infect Immun **77**(8): 3234-3243.
Herold, S., J. Siebert, A. Huber and H. Schmidt (2005). "Global expression of prophage genes in Escherichia coli O157:H7 strain EDL933 in response to norfloxacin." Antimicrob Agents Chemother **49**(3): 931-944.
Ho, T. D. and M. K. Waldor (2007). "Enterohemorrhagic Escherichia coli O157:H7 gal mutants are sensitive to bacteriophage P1 and defective in intestinal colonization." Infect Immun **75**(4): 1661-1666.
Hogg, R. A., J. P. Holmes, S. Ghebrehewet, K. Elders, J. Hart, C. Whiteside, G. A. Willshaw, T. Cheasty, A. Kay, K. Lynch and G. C. Pritchard (2009). "Probable zoonotic transmission of verocytotoxigenic Escherichia coli O 157 by dogs." Vet Rec **164**(10): 304-305.

Holden, N. J. and D. L. Gally (2004). "Switches, cross-talk and memory in Escherichia coli adherence." J Med Microbiol **53**(Pt 7): 585-593.
Hussein, H. S. (2007). "Prevalence and pathogenicity of Shiga toxin-producing Escherichia coli in beef cattle and their products." J Anim Sci **85**(13 Suppl): E63-72.
Huys, G., M. Cnockaert, J. M. Janda and J. Swings (2003). "Escherichia albertii sp. nov., a diarrhoeagenic species isolated from stool specimens of Bangladeshi children." Int J Syst Evol Microbiol **53**(Pt 3): 807-810.
Iguchi, A., N. R. Thomson, Y. Ogura, D. Saunders, T. Ooka, I. R. Henderson, D. Harris, M. Asadulghani, K. Kurokawa, P. Dean, B. Kenny, M. A. Quail, S. Thurston, G. Dougan, T. Hayashi, J. Parkhill and G. Frankel (2009). "Complete genome sequence and comparative genome analysis of enteropathogenic Escherichia coli O127:H6 strain E2348/69." J Bacteriol **191**(1): 347-354.
Imamovic, L., J. Jofre, H. Schmidt, R. Serra-Moreno and M. Muniesa (2009). "Phage-mediated Shiga toxin 2 gene transfer in food and water." Appl Environ Microbiol **75**(6): 1764-1768.
INVS (2018). "Surveillance du syndrome hémolytique et urémique post-diarrhéique

chez l'enfant de moins de 15 ans en France en 2017 ".
Itoh, Y., J. D. Rice, C. Goller, A. Pannuri, J. Taylor, J. Meisner, T. J. Beveridge, J. F. Preston, 3rd and T. Romeo (2008). "Roles of pgaABCD genes in synthesis, modification, and export of the Escherichia coli biofilm adhesin poly-beta-1,6-N-acetyl-D-glucosamine." J Bacteriol **190**(10): 3670-3680.
Jacquot, A., C. Sakamoto, A. Razafitianamarahavo, C. Caillet, J. Merlin, A. Fahs, J. M. Ghigo, J. F. L. Duval, C. Beloin and G. Francius (2014). "The dynamics and pH-dependence of Ag43 adhesins' self-association probed by atomic force spectroscopy." Nanoscale **6**(21): 12665-12681.
Jain, S., P. van Ulsen, I. Benz, M. A. Schmidt, R. Fernandez, J. Tommassen and M. B. Goldberg (2006). "Polar localization of the autotransporter family of large bacterial virulence proteins." J Bacteriol **188**(13): 4841-4850.
Jaureguy, F., L. Landraud, V. Passet, L. Diancourt, E. Frapy, G. Guigon, E. Carbonnelle, O. Lortholary, O. Clermont, E. Denamur, B. Picard, X. Nassif and S. Brisse (2008). "Phylogenetic and genomic diversity of human bacteremic Escherichia coli strains." BMC Genomics **9**: 560.
Jenkins, C., H. Chart, T. Cheasty, G. A. Willshaw, M. C. Pearce, G. Foster, G. J. Gunn, H. R. Smith, G. Dougan, B. A. Synge and G. Frankel (2002). "Verocytotoxin-producing Escherichia coli (VTEC) other than serogroup O157 from Scottish cattle." Vet Rec **151**(2): 58-60.
Jenkins, C., N. T. Perry, T. Cheasty, D. J. Shaw, G. Frankel, G. Dougan, G. J. Gunn, H. R. Smith, A. W. Paton and J. C. Paton (2003). "Distribution of the saa gene in strains of Shiga toxin-producing Escherichia coli of human and bovine origins." J Clin Microbiol **41**(4): 1775-1778.
Junker, M., C. C. Schuster, A. V. McDonnell, K. A. Sorg, M. C. Finn, B. Berger and P. L. Clark (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**(13): 4918-4923.
Kaminska, R. and M. W. van der Woude (2010). "Establishing and maintaining sequestration of Dam target sites for phase variation of agn43 in Escherichia coli." J Bacteriol **192**(7): 1937-1945.
Kaper, J. B., J. P. Nataro and H. L. Mobley (2004). "Pathogenic Escherichia coli." Nat Rev Microbiol **2**(2): 123-140.
Karch, H., E. Denamur, U. Dobrindt, B. B. Finlay, R. Hengge, L. Johannes, E. Z. Ron, T. Tonjum, P. J. Sansonetti and M. Vicente (2012). "The enemy within us: lessons from the 2011 European Escherichia coli O104:H4 outbreak." EMBO Mol Med **4**(9): 841-848.
Karmali, M. A., M. Mascarenhas, S. Shen, K. Ziebell, S. Johnson, R. Reid-Smith, J. Isaac-Renton, C. Clark, K. Rahn and J. B. Kaper (2003). "Association of genomic O island 122 of

Escherichia coli EDL 933 with verocytotoxin-producing Escherichia coli seropathotypes that are linked to epidemic and/or serious disease." J Clin Microbiol **41**(11): 4930-4940.
Kawano, K., M. Okada, T. Haga, K. Maeda and Y. Goto (2008). "Relationship between pathogenicity for humans and stx genotype in Shiga toxin-producing Escherichia coli serotype O157." Eur J Clin Microbiol Infect Dis **27**(3): 227-232.
Kendall, M. M., C. C. Gruber, C. T. Parker and V. Sperandio (2012). "Ethanolamine controls expression of genes encoding components involved in interkingdom signaling and virulence in enterohemorrhagic Escherichia coli O157:H7." MBio **3**(3).
Kenny, B., A. Abe, M. Stein and B. B. Finlay (1997). "Enteropathogenic Escherichia coli protein secretion is induced in response to conditions similar to those in the gastrointestinal tract." Infect Immun **65**(7): 2606-2612.
Khan, I., Y. Chen, T. Dong, X. Hong, R. Takeuchi, H. Mori and D. Kihara (2014). "Genome-scale identification and characterization of moonlighting proteins." Biol Direct **9**: 30.
Kim, S. H. and Y. H. Kim (2004). "Escherichia coli O157:H7 adherence to HEp-2 cells is implicated with curli expression and outer membrane integrity." J Vet Sci **5**(2): 119-124.
Kintz, E., J. Brainard, L. Hooper and P. Hunter (2017). "Transmission pathways for sporadic Shiga-toxin producing E. coli infections: A systematic review and meta-analysis." Int J Hyg Environ Health **220**(1): 57-67.
Kjaergaard, K., M. A. Schembri, C. Ramos, S. Molin and P. Klemm (2000). "Antigen 43 facilitates formation of multispecies biofilms." Environ Microbiol **2**(6): 695-702.
Klapproth, J. M., I. C. Scaletsky, B. P. McNamara, L. C. Lai, C. Malstrom, S. P. James and M. S. Donnenberg (2000). "A large toxin from pathogenic Escherichia coli strains that inhibits lymphocyte activation." Infect Immun **68**(4): 2148-2155.
Klemm, P., L. Hjerrild, M. Gjermansen and M. A. Schembri (2004). "Structure-function analysis of the self-recognizing Antigen 43 autotransporter protein from Escherichia coli." Mol Microbiol **51**(1): 283-296.
Klemm, P., Vejborg, R., Sherlock, O. (2006). "Self-associating autotransporters, SAATs: Functional and structural similarities." International Journal of Medical Microbiology: 187–195.
Knowles, T. J., A. Scott-Tucker, M. Overduin and I. R. Henderson (2009). "Membrane protein architects: the role of the BAM complex in outer membrane protein assembly." Nat Rev Microbiol **7**(3): 206-214.
Knudsen, S. K., A. Stensballe, M. Franzmann, U. B. Westergaard and D. E. Otzen (2008). "Effect of glycosylation on the extracellular domain of the Ag43 bacterial autotransporter: enhanced stability and reduced cellular aggregation." Biochem J **412**(3): 563-577.
Konieczny, M. P. J., I. Benz, B. Hollinderbaumer, C. Beinke, M. Niederweis and M. A. Schmidt (2001). "Modular organization of the AIDA autotransporter translocator: the N-terminal beta1-domain is surface-exposed and stabilizes the transmembrane beta2-domain." Antonie Van Leeuwenhoek **80**(1): 19-34.
Kostakioti, M., C. L. Newman, D. G. Thanassi and C. Stathopoulos (2005). "Mechanisms of protein export across the bacterial outer membrane." J Bacteriol **187**(13): 4306-4314.
Kruger, A. and P. M. Lucchesi (2015). "Shiga toxins and stx phages: highly diverse entities." Microbiology **161**(Pt 3): 451-462.
Kudva, I. T., B. Krastins, A. G. Torres, R. W. Griffin, H. Sheng, D. A. Sarracino, C. J. Hovde, S. B. Calderwood and M. John (2015). "The Escherichia coli O157:H7 cattle immunoproteome includes outer membrane protein A (OmpA), a modulator of adherence to bovine rectoanal junction squamous epithelial (RSE) cells." Proteomics **15**(11): 1829-1842.
Kuhnel, K. and D. Diezmann (2011). "Crystal structure of the autochaperone region from the Shigella flexneri autotransporter IcsA." J Bacteriol **193**(8): 2042-2045.

La Ragione, R. M., A. Best, M. J. Woodward and A. D. Wales (2009). "Escherichia coli O157:H7 colonization in small domestic ruminants." FEMS Microbiol Rev **33**(2): 394-410.
Laganenka, L., R. Colin and V. Sourjik (2016). "Chemotaxis towards autoinducer 2 mediates autoaggregation in Escherichia coli." Nat Commun **7**: 12984.
Lapeyraque, A. L., M. Malina, V. Fremeaux-Bacchi, T. Boppel, M. Kirschfink, M. Oualha, F. Proulx, M. J. Clermont, F. Le Deist, P. Niaudet and F. Schaefer (2011). "Eculizumab in severe Shiga-toxin-associated HUS." N Engl J Med **364**(26): 2561-2563.
Lawley, T. D., W. A. Klimke, M. J. Gubbins and L. S. Frost (2003). "F factor conjugation is a true type IV secretion system." FEMS Microbiol Lett **224**(1): 1-15.
Lawrence, J. G. and D. L. Hartl (1991). "Unusual codon bias occurring within insertion sequences in Escherichia coli." Genetica **84**(1): 23-29.
Leclerc, H. (1962). "Biochemical Studies on Pigmented Enterobacteriaeeae." Ann. Inst. Pasteur 1962 **102**: 726-741.
Légifrance (2008). "Article R4421-3." Code du Travail.
Leo, J. C. and A. Goldman (2009). "The immunoglobulin-binding Eib proteins from Escherichia coli are receptors for IgG Fc." Mol Immunol **46**(8-9): 1860-1866.
Leo, J. C., I. Grin and D. Linke (2012). "Type V secretion: mechanism(s) of autotransport through the bacterial outer membrane." Philos Trans R Soc Lond B Biol Sci **367**(1592): 1088-1101.
Leo, J. C., A. Lyskowski, K. Hattula, M. D. Hartmann, H. Schwarz, S. J. Butcher, D. Linke, A. N. Lupas and A. Goldman (2011). "The structure of E. coli IgG-binding protein D suggests a general model for bending and binding in trimeric autotransporter adhesins." Structure **19**(7): 1021-1030.
Leo, J. C., P. Oberhettinger, M. Schutz and D. Linke (2015). "The inverse autotransporter family: intimin, invasin and related proteins." Int J Med Microbiol **305**(2): 276-282.
Levine, M. M. (1987). "Escherichia coli that cause diarrhea: enterotoxigenic, enteropathogenic, enteroinvasive, enterohemorrhagic, and enteroadherent." J Infect Dis **155**(3): 377-389.
Leyton, D. L., M. J. Belousoff and T. Lithgow (2015). "The beta-Barrel Assembly Machinery Complex." Methods Mol Biol **1329**: 1-16.
Li, B., W. H. Koch and T. A. Cebula (1997). "Detection and characterization of the fimA gene of Escherichia coli O157:H7." Mol Cell Probes **11**(6): 397-406.
Lim, H. N. and A. van Oudenaarden (2007). "A multistep epigenetic switch enables the stable inheritance of DNA methylation states." Nat Genet **39**(2): 269-275.
Lim, J. Y., J. Yoon and C. J. Hovde (2010). "A brief overview of Escherichia coli O157:H7 and its plasmid O157." J Microbiol Biotechnol **20**(1): 5-14.
Lindenthal, C. and E. A. Elsinghorst (1999). "Identification of a glycoprotein produced by enterotoxigenic Escherichia coli." Infect Immun **67**(8): 4084-4091.
Linke, D., T. Riess, I. B. Autenrieth, A. Lupas and V. A. Kempf (2006). "Trimeric autotransporter adhesins: variable structure, common function." Trends Microbiol **14**(6): 264-270.
Liu, H., L. Magoun and J. M. Leong (1999). "beta1-chain integrins are not essential for intimin-mediated host cell attachment and enteropathogenic Escherichia coli-induced actin condensation." Infect Immun **67**(4): 2045-2049.
Llosa, M., F. X. Gomis-Ruth, M. Coll and F. de la Cruz Fd (2002). "Bacterial conjugation: a two-step mechanism for DNA transport." Mol Microbiol **45**(1): 1-8.
Lloyd, A. L., T. A. Henderson, P. D. Vigil and H. L. Mobley (2009). "Genomic islands of uropathogenic Escherichia coli contribute to virulence." J Bacteriol **191**(11): 3469-3481.
Los, J. M., M. Los, A. Wegrzyn and G. Wegrzyn (2010). "Hydrogen peroxide-mediated induction of the Shiga toxin-converting lambdoid prophage ST2-8624 in Escherichia coli O157:H7." FEMS Immunol Med Microbiol **58**(3): 322-329.

Low, A. S., F. Dziva, A. G. Torres, J. L. Martinez, T. Rosser, S. Naylor, K. Spears, N. Holden, A. Mahajan, J. Findlay, J. Sales, D. G. Smith, J. C. Low, M. P. Stevens and D. L. Gally (2006). "Cloning, expression, and characterization of fimbrial operon F9 from enterohemorrhagic Escherichia coli O157:H7." Infect Immun **74**(4): 2233-2244.
Lower, B. H., R. Yongsunthon, F. P. Vellano, 3rd and S. K. Lower (2005). "Simultaneous force and fluorescence measurements of a protein that forms a bond between a living bacterium and a solid surface." J Bacteriol **187**(6): 2127-2137.
Lu, Y., S. Iyoda, H. Satou, H. Satou, K. Itoh, T. Saitoh and H. Watanabe (2006). "A new immunoglobulin-binding protein, EibG, is responsible for the chain-like adhesion phenotype of locus of enterocyte effacement-negative, shiga toxin-producing Escherichia coli." Infect Immun **74**(10): 5747-5755.
Luck, S. N., S. A. Turner, K. Rajakumar, B. Adler and H. Sakellaris (2004). "Excision of the Shigella resistance locus pathogenicity island in Shigella flexneri is stimulated by a member of a new subgroup of recombination directionality factors." J Bacteriol **186**(16): 5551-5554.
Luo, Q., P. Kumar, T. J. Vickers, A. Sheikh, W. G. Lewis, D. A. Rasko, J. Sistrunk and J. M. Fleckenstein (2014). "Enterotoxigenic Escherichia coli secretes a highly conserved mucin-degrading metalloprotease to effectively engage intestinal epithelial cells." Infect Immun **82**(2): 509-521.
Macarisin, D., J. Patel, G. Bauchan, J. A. Giron and V. K. Sharma (2012). "Role of curli and cellulose expression in adherence of Escherichia coli O157:H7 to spinach leaves." Foodborne Pathog Dis **9**(2): 160-167.
Macnab, R. M. (1987). "Escherichia Coli and Salmonella Typhimurium: Cellular and Molecular Biology " American Society for Microbiology **Vol I** 70-83, 732-759.
Macnab, R. M. (2004). "Type III flagellar protein export and flagellar assembly." Biochim Biophys Acta **1694**(1-3): 207-217.
Maeda, S., M. Ito, T. Ando, Y. Ishimoto, Y. Fujisawa, H. Takahashi, A. Matsuda, A. Sawamura and S. Kato (2006). "Horizontal transfer of nonconjugative plasmids in a colony biofilm of Escherichia coli." FEMS Microbiol Lett **255**(1): 115-120.
Mahajan, A., C. G. Currie, S. Mackie, J. Tree, S. McAteer, I. McKendrick, T. N. McNeilly, A. Roe, R. M. La Ragione, M. J. Woodward, D. L. Gally and D. G. Smith (2009). "An investigation of the expression and adhesin function of H7 flagella in the interaction of Escherichia coli O157 : H7 with bovine intestinal epithelium." Cell Microbiol **11**(1): 121-137.
Mainil, J. G. and G. Daube (2005). "Verotoxigenic Escherichia coli from animals, humans and foods: who's who?" J Appl Microbiol **98**(6): 1332-1344.
Mallick, E. M., M. J. Brady, S. A. Luperchio, V. K. Vanguri, L. Magoun, H. Liu, B. J. Sheppard, J. Mukherjee, A. Donohue-Rolfe, S. Tzipori, J. M. Leong and D. B. Schauer (2012). "Allele- and tir-independent functions of intimin in diverse animal infection models." Front Microbiol **3**: 11.
Mann, E. A. and S. A. Saeed (2012). "Gastrointestinal infection as a trigger for inflammatory bowel disease." Curr Opin Gastroenterol **28**(1): 24-29.
Manwaring, N. P., R. A. Skurray and N. Firth (1999). "Nucleotide sequence of the F plasmid leading region." Plasmid **41**(3): 219-225.
Mao, Y., M. P. Doyle and J. Chen (2006). "Role of colanic acid exopolysaccharide in the survival of enterohaemorrhagic Escherichia coli O157:H7 in simulated gastrointestinal fluids." Lett Appl Microbiol **42**(6): 642-647.
Mariani-Kurkdjian, P. and E. Bingen (2012). "Physiopathologie et virulence des Escherichia coli

producteurs de Shiga-toxines." Réanimation (2012) 21:268-279.
Martinez-Gil, M., K. G. K. Goh, E. Rackaityte, C. Sakamoto, B. Audrain, D. G. Moriel, M. Totsika, J. M. Ghigo, M. A. Schembri and C. Beloin (2017). "YeeJ is an inverse autotransporter

from Escherichia coli that binds to peptidoglycan and promotes biofilm formation." Sci Rep **7**(1): 11326.
Mason, T. G. and G. Richardson (1981). "Escherichia coli and the human gut: some ecological considerations." J Appl Bacteriol **51**(1): 1-16.
Matthysse, A. G., R. Deora, M. Mishra and A. G. Torres (2008). "Polysaccharides cellulose, poly-beta-1,6-n-acetyl-D-glucosamine, and colanic acid are required for optimal binding of Escherichia coli O157:H7 strains to alfalfa sprouts and K-12 strains to plastic but not for binding to epithelial cells." Appl Environ Microbiol **74**(8): 2384-2390.
May, T. and S. Okabe (2008). "Escherichia coli harboring a natural IncF conjugative F plasmid develops complex mature biofilms by stimulating synthesis of colanic acid and Curli." J Bacteriol **190**(22): 7479-7490.
May, T., K. Tsuruta and S. Okabe (2011). "Exposure of conjugative plasmid carrying Escherichia coli biofilms to male-specific bacteriophages." ISME J **5**(4): 771-775.
Mazel, D. and J. Davies (1999). "Gene capture in : Response." Trends Microbiol **7**(3): 95.
McCrate, O. A., X. Zhou, C. Reichhardt and L. Cegelski (2013). "Sum of the parts: composition and architecture of the bacterial extracellular matrix." J Mol Biol **425**(22): 4286-4294.
McNeilly, T. N., S. W. Naylor, A. Mahajan, M. C. Mitchell, S. McAteer, D. Deane, D. G. Smith, J. C. Low, D. L. Gally and J. F. Huntley (2008). "Escherichia coli O157:H7 colonization in cattle following systemic and mucosal immunization with purified H7 flagellin." Infect Immun **76**(6): 2594-2602.
McWilliams, B. D. and A. G. Torres (2014). "EHEC Adhesins." Microbiol Spectr **2**(2): EHEC00032013.
Mellies, J. L. and E. Lorenzen (2014). "Enterohemorrhagic Escherichia coli Virulence Gene Regulation." Microbiol Spectr **2**(4): EHEC-0004-2013.
Mellmann, A., D. Harmsen, C. A. Cummings, E. B. Zentz, S. R. Leopold, A. Rico, K. Prior, R. Szczepanowski, Y. Ji, W. Zhang, S. F. McLaughlin, J. K. Henkhaus, B. Leopold, M. Bielaszewska, R. Prager, P. M. Brzoska, R. L. Moore, S. Guenther, J. M. Rothberg and H. Karch (2011). "Prospective genomic characterization of the German enterohemorrhagic Escherichia coli O104:H4 outbreak by rapid next generation sequencing technology." PLoS One **6**(7): e22751.
Melton-Celsa, A. R. (2014). "Shiga Toxin (Stx) Classification, Structure, and Function." Microbiol Spectr **2**(4): EHEC-0024-2013.
Meng, G., N. Spahich, R. Kenjale, G. Waksman and J. W. St Geme, 3rd (2011). "Crystal structure of the Haemophilus influenzae Hap adhesin reveals an intercellular oligomerization mechanism for bacterial aggregation." EMBO J **30**(18): 3864-3874.
Merkel, V., B. Ohder, M. Bielaszewska, W. Zhang, A. Fruth, C. Menge, E. Borrmann, B. Middendorf, J. Muthing, H. Karch and A. Mellmann (2010). "Distribution and phylogeny of immunoglobulin-binding protein G in Shiga toxin-producing Escherichia coli and its association with adherence phenotypes." Infect Immun **78**(8): 3625-3636.
Mermin, J. H. and P. M. Griffin (1999). "Public health in crisis: outbreaks of Escherichia coli O157:H7 infections in Japan." Am J Epidemiol **150**(8): 797-803; discussion 804-795.
Molin, S. and T. Tolker-Nielsen (2003). "Gene transfer occurs with enhanced efficiency in biofilms and induces enhanced stabilisation of the biofilm structure." Curr Opin Biotechnol **14**(3): 255-261.
Monteiro, B. T., L. C. Campos, M. P. Sircili, M. R. Franzolin, L. F. Bevilacqua, J. P. Nataro and W. P. Elias (2009). "The dispersin-encoding gene (aap) is not restricted to enteroaggregative Escherichia coli." Diagn Microbiol Infect Dis **65**(1): 81-84.
Monteiro, R., V. Ageorges, M. Rojas-Lopez, H. Schmidt, A. Weiss, Y. Bertin, E. Forano, G. Jubelin, I. R. Henderson, V. Livrelli, A. P. Gobert, R. Rosini, M. Soriani and M. Desvaux (2016). "A secretome view of colonisation factors in Shiga toxin-encoding Escherichia coli

(STEC): from enterohaemorrhagic E. coli (EHEC) to related enteropathotypes." FEMS Microbiol Lett **363**(16).
Montero, D. A., J. Velasco, F. Del Canto, J. L. Puente, N. L. Padola, D. A. Rasko, M. Farfan, J. C. Salazar and R. Vidal (2017). "Locus of Adhesion and Autoaggregation (LAA), a pathogenicity island present in emerging Shiga Toxin-producing Escherichia coli strains." Sci Rep **7**(1): 7011.
Moreira, C. G., S. M. Carneiro, J. P. Nataro, L. R. Trabulsi and W. P. Elias (2003). "Role of type I fimbriae in the aggregative adhesion pattern of enteroaggregative Escherichia coli." FEMS Microbiol Lett **226**(1): 79-85.
Moreira, C. G. and V. Sperandio (2016). "The Epinephrine/Norepinephrine/Autoinducer-3 Interkingdom Signaling System in Escherichia coli O157:H7." Adv Exp Med Biol **874**: 247-261.
Moriel, D. G., I. Bertoldi, A. Spagnuolo, S. Marchi, R. Rosini, B. Nesta, I. Pastorello, V. A. Corea, G. Torricelli, E. Cartocci, S. Savino, M. Scarselli, U. Dobrindt, J. Hacker, H. Tettelin, L. J. Tallon, S. Sullivan, L. H. Wieler, C. Ewers, D. Pickard, G. Dougan, M. R. Fontana, R. Rappuoli, M. Pizza and L. Serino (2010). "Identification of protective and broadly conserved vaccine antigens from the genome of extraintestinal pathogenic Escherichia coli." Proc Natl Acad Sci U S A **107**(20): 9072-9077.
Moxley, R. A. and G. R. Acuff (2014). "Peri- and Postharvest Factors in the Control of Shiga Toxin-Producing Escherichia coli in Beef." Microbiol Spectr **2**(6).
Muniesa, M., J. A. Hammerl, S. Hertwig, B. Appel and H. Brussow (2012). "Shiga toxin-producing Escherichia coli O104:H4: a new challenge for microbiology." Appl Environ Microbiol **78**(12): 4065-4073.
Musken, A., M. Bielaszewska, L. Greune, C. H. Schweppe, J. Muthing, H. Schmidt, M. A. Schmidt, H. Karch and W. Zhang (2008). "Anaerobic conditions promote expression of Sfp fimbriae and adherence of sorbitol-fermenting enterohemorrhagic Escherichia coli O157:NM to human intestinal epithelial cells." Appl Environ Microbiol **74**(4): 1087-1093.
Nagy, A., J. Mowery, G. R. Bauchan, L. Wang, L. Nichols-Russell and X. Nou (2015). "Role of Extracellular Structures of Escherichia coli O157:H7 in Initial Attachment to Biotic and Abiotic Surfaces." Appl Environ Microbiol **81**(14): 4720-4727.
Naili, I., B. Baudner, A. Bernalier-Donadille, M. Pizza, M. Desvaux, G. Jubelin and C. Buonsanti (2017). "In vivo characterization of the immune response towards the pathogenic Escherichia coli antigen SslE and modulation of the intestinal microbiota." JOURNAL OF IMMUNOLOGY **198 (1S)**: 147.123.
Naili, I., C. Buonsanti, M. Soriani, M. Desvaux, G. Jubelin, M. Vieru, B. Baudner and U. D'Oro (2016). "In vivo characterization of the immune response towards the pathogenic Escherichia coli antigen SslE and modulation of the intestinal microbiota. ." JOURNAL OF IMMUNOLOGY **196 (1S)**: 215.219.
Nakao, R., M. Ramstedt, S. N. Wai and B. E. Uhlin (2012). "Enhanced biofilm formation by Escherichia coli LPS mutants defective in Hep biosynthesis." PLoS One **7**(12): e51241.
Nataro, J. P. and J. B. Kaper (1998). "Diarrheagenic Escherichia coli." Clin Microbiol Rev **11**(1): 142-201.
Navarro-Garcia, F. and W. P. Elias (2011). "Autotransporters and virulence of enteroaggregative E. coli." Gut Microbes **2**(1): 13-24.
Naylor, S. W., J. C. Low, T. E. Besser, A. Mahajan, G. J. Gunn, M. C. Pearce, I. J. McKendrick, D. G. Smith and D. L. Gally (2003). "Lymphoid follicle-dense mucosa at the terminal rectum is the principal site of colonization of enterohemorrhagic Escherichia coli O157:H7 in the bovine host." Infect Immun **71**(3): 1505-1512.
Nesta, B., G. Spraggon, C. Alteri, D. G. Moriel, R. Rosini, D. Veggi, S. Smith, I. Bertoldi, I. Pastorello, I. Ferlenghi, M. R. Fontana, G. Frankel, H. L. Mobley, R. Rappuoli, M. Pizza, L.

Serino and M. Soriani (2012). "FdeC, a novel broadly conserved Escherichia coli adhesin eliciting protection against urinary tract infections." MBio **3**(2).
Nesta, B., M. Valeri, A. Spagnuolo, R. Rosini, M. Mora, P. Donato, C. J. Alteri, M. Del Vecchio, S. Buccato, A. Pezzicoli, I. Bertoldi, L. Buzzigoli, G. Tuscano, M. Falduto, V. Rippa, Y. Ashhab, G. Bensi, M. R. Fontana, K. L. Seib, H. L. Mobley, M. Pizza, M. Soriani and L. Serino (2014). "SslE elicits functional antibodies that impair in vitro mucinase activity and in vivo colonization by both intestinal and extraintestinal Escherichia coli strains." PLoS Pathog **10**(5): e1004124.
Nicholls, L., T. Grant and R. Robins-Browne (2002). "Identification of a novel genetic locus that is required for in vitro adhesion of a clinical isolate of enterohaemorrhagic Escherichia coli to epithelial cells." Molecular Microbiology **35**(2): 275-288.
Nishi J, S. J., Mizuguchi K et al. (2003). "The export of coat protein from enteroaggregative Escherichia coli by a specific ATP-binding cassette transporter system." J Biol Chem.
Nishikawa, K., M. Watanabe, E. Kita, K. Igai, K. Omata, M. B. Yaffe and Y. Natori (2006). "A multivalent peptide library approach identifies a novel Shiga toxin inhibitor that induces aberrant cellular transport of the toxin." FASEB J **20**(14): 2597-2599.
Nyholm, O., J. Halkilahti, G. Wiklund, U. Okeke, L. Paulin, P. Auvinen, K. Haukka and A. Siitonen (2015). "Comparative Genomics and Characterization of Hybrid Shigatoxigenic and Enterotoxigenic Escherichia coli (STEC/ETEC) Strains." PLoS One **10**(8): e0135936.
Nyholm, O., S. Heinikainen, S. Pelkonen, S. Hallanvuo, K. Haukka and A. Siitonen (2015). "Hybrids of Shigatoxigenic and Enterotoxigenic Escherichia coli (STEC/ETEC) Among Human and Animal Isolates in Finland." Zoonoses Public Health **62**(7): 518-524.
O'Brien, A. D., J. W. Newland, S. F. Miller, R. K. Holmes, H. W. Smith and S. B. Formal (1984). "Shiga-like toxin-converting phages from Escherichia coli strains that cause hemorrhagic colitis or infantile diarrhea." Science **226**(4675): 694-696.
O'Loughlin, E. V. and R. M. Robins-Browne (2001). "Effect of Shiga toxin and Shiga-like toxins on eukaryotic cells." Microbes Infect **3**(6): 493-507.
Obadia, B., S. Lacour, P. Doublet, H. Baubichon-Cortay, A. J. Cozzone and C. Grangeasse (2007). "Influence of tyrosine-kinase Wzc activity on colanic acid production in Escherichia coli K12 cells." J Mol Biol **367**(1): 42-53.
Oberhettinger, P., M. Schutz, J. C. Leo, N. Heinz, J. Berger, I. B. Autenrieth and D. Linke (2012). "Intimin and invasin export their C-terminus to the bacterial cell surface using an inverse mechanism compared to classical autotransport." PLoS One **7**(10): e47069.
Ochman, H. and R. K. Selander (1984). "Standard reference strains of Escherichia coli from natural populations." J Bacteriol **157**(2): 690-693.
Oliver, D. C., G. Huang, E. Nodel, S. Pleasance and R. C. Fernandez (2003). "A conserved region within the Bordetella pertussis autotransporter BrkA is necessary for folding of its passenger domain." Mol Microbiol **47**(5): 1367-1383.
Olsen, A., A. Arnqvist, M. Hammar and S. Normark (1993). "Environmental regulation of curli production in Escherichia coli." Infect Agents Dis **2**(4): 272-274.
Omisakin, F., M. MacRae, I. D. Ogden and N. J. C. Strachan (2003). "Concentration and prevalence of Escherichia coli O157 in cattle feces at slaughter." Applied and Environmental Microbiology **69**(5): 2444-2447.
Oomen, C. J., P. van Ulsen, P. van Gelder, M. Feijen, J. Tommassen and P. Gros (2004). "Structure of the translocator domain of a bacterial autotransporter." EMBO J **23**(6): 1257-1266.
Owen, P., P. Caffrey and L. G. Josefsson (1987). "Identification and partial characterization of a novel bipartite protein antigen associated with the outer membrane of Escherichia coli." J Bacteriol **169**(8): 3770-3777.

Owen, P., M. Meehan, H. de Loughry-Doherty and I. Henderson (1996). "Phase-variable outer membrane proteins in Escherichia coli." FEMS Immunol Med Microbiol **16**(2): 63-76.
Palchevskiy, V. and S. E. Finkel (2006). "Escherichia coli competence gene homologs are essential for competitive fitness and the use of DNA as a nutrient." J Bacteriol **188**(11): 3902-3910.
Paton, A. W., R. Morona and J. C. Paton (2000). "A new biological agent for treatment of Shiga toxigenic Escherichia coli infections and dysentery in humans." Nat Med **6**(3): 265-270.
Paton, A. W., P. Srimanote, M. C. Woodrow and J. C. Paton (2001). "Characterization of Saa, a novel autoagglutinating adhesin produced by locus of enterocyte effacement-negative Shiga-toxigenic Escherichia coli strains that are virulent for humans." Infect Immun **69**(11): 6999-7009.
Pawar, D. M., M. L. Rossman and J. Chen (2005). "Role of curli fimbriae in mediating the cells of enterohaemorrhagic Escherichia coli to attach to abiotic surfaces." J Appl Microbiol **99**(2): 418-425.
Pedersen, K., & Clark, L. (2007). "A review of shiga toxin Eschrichia coli ans Salmonella enterica in cattle and free-ranging birds: potential association and epidemiological links." Human-Wildlife Conflicts: 68-77.
Pereira, A. L., T. N. Silva, A. C. Gomes, A. C. Araujo and L. G. Giugliano (2010). "Diarrhea-associated biofilm formed by enteroaggregative Escherichia coli and aggregative Citrobacter freundii: a consortium mediated by putative F pili." BMC Microbiol **10**: 57.
Perna, N. T., G. Plunkett, 3rd, V. Burland, B. Mau, J. D. Glasner, D. J. Rose, G. F. Mayhew, P. S. Evans, J. Gregor, H. A. Kirkpatrick, G. Posfai, J. Hackett, S. Klink, A. Boutin, Y. Shao, L. Miller, E. J. Grotbeck, N. W. Davis, A. Lim, E. T. Dimalanta, K. D. Potamousis, J. Apodaca, T. S. Anantharaman, J. Lin, G. Yen, D. C. Schwartz, R. A. Welch and F. R. Blattner (2001). "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7." Nature **409**(6819): 529-533.
Peterson, J. H., R. L. Szabady and H. D. Bernstein (2006). "An unusual signal peptide extension inhibits the binding of bacterial presecretory proteins to the signal recognition particle, trigger factor, and the SecYEG complex." J Biol Chem **281**(14): 9038-9048.
Power, M. L., B. C. Ferrari, J. Littlefield-Wyer, D. M. Gordon, M. B. Slade and D. A. Veal (2006). "A naturally occurring novel allele of Escherichia coli outer membrane protein A reduces sensitivity to bacteriophage." Appl Environ Microbiol **72**(12): 7930-7932.
Price, S. B., J. C. Wright, F. J. DeGraves, M. P. Castanie-Cornet and J. W. Foster (2004). "Acid resistance systems required for survival of Escherichia coli O157:H7 in the bovine gastrointestinal tract and in apple cider are different." Appl Environ Microbiol **70**(8): 4792-4799.
Priest, F. G. and M. Barker (2010). "Gram-negative bacteria associated with brewery yeasts: reclassification of Obesumbacterium proteus biogroup 2 as Shimwellia pseudoproteus gen. nov., sp. nov., and transfer of Escherichia blattae to Shimwellia blattae comb. nov." Int J Syst Evol Microbiol **60**(Pt 4): 828-833.
Puttamreddy, S., N. A. Cornick and F. C. Minion (2010). "Genome-wide transposon mutagenesis reveals a role for pO157 genes in biofilm development in Escherichia coli O157:H7 EDL933." Infect Immun **78**(6): 2377-2384.
Raetz, C. R. and C. Whitfield (2002). "Lipopolysaccharide endotoxins." Annu Rev Biochem **71**: 635-700.
Randall, L. L. and S. J. Hardy (2002). "SecB, one small chaperone in the complex milieu of the cell." Cell Mol Life Sci **59**(10): 1617-1623.
Rangel, J. M., P. H. Sparling, C. Crowe, P. M. Griffin and D. L. Swerdlow (2005). "Epidemiology of Escherichia coli O157:H7 outbreaks, United States, 1982-2002." Emerg Infect Dis **11**(4): 603-609.

Rasko, D. A., D. R. Webster, J. W. Sahl, A. Bashir, N. Boisen, F. Scheutz, E. E. Paxinos, R. Sebra, C. S. Chin, D. Iliopoulos, A. Klammer, P. Peluso, L. Lee, A. O. Kislyuk, J. Bullard, A. Kasarskis, S. Wang, J. Eid, D. Rank, J. C. Redman, S. R. Steyert, J. Frimodt-Moller, C. Struve, A. M. Petersen, K. A. Krogfelt, J. P. Nataro, E. E. Schadt and M. K. Waldor (2011). "Origins of the E. coli strain causing an outbreak of hemolytic-uremic syndrome in Germany." N Engl J Med **365**(8): 709-717.

Reading, N. C., D. A. Rasko, A. G. Torres and V. Sperandio (2009). "The two-component system QseEF and the membrane protein QseG link adrenergic and stress sensing to bacterial pathogenesis." Proc Natl Acad Sci U S A **106**(14): 5889-5894.

Reidl, S., A. Lehmann, R. Schiller, A. Salam Khan and U. Dobrindt (2009). "Impact of O-glycosylation on the molecular and cellular adhesion properties of the Escherichia coli autotransporter protein Ag43." Int J Med Microbiol **299**(6): 389-401.

Rendon, M. A., Z. Saldana, A. L. Erdem, V. Monteiro-Neto, A. Vazquez, J. B. Kaper, J. L. Puente and J. A. Giron (2007). "Commensal and pathogenic Escherichia coli use a common pilus adherence factor for epithelial cell colonization." Proc Natl Acad Sci U S A **104**(25): 10637-10642.

Restieri, C., G. Garriss, M. C. Locas and C. M. Dozois (2007). "Autotransporter-encoding sequences are phylogenetically distributed among Escherichia coli clinical isolates and reference strains." Appl Environ Microbiol **73**(5): 1553-1562.

Riley, L. W., R. S. Remis, S. D. Helgerson, H. B. McGee, J. G. Wells, B. R. Davis, R. J. Hebert, E. S. Olcott, L. M. Johnson, N. T. Hargrett, P. A. Blake and M. L. Cohen (1983). "Hemorrhagic colitis associated with a rare Escherichia coli serotype." N Engl J Med **308**(12): 681-685.

Roche, A., J. McFadden and P. Owen (2001). "Antigen 43, the major phase-variable protein of the Escherichia coli outer membrane, can exist as a family of proteins encoded by multiple alleles." Microbiology **147**(Pt 1): 161-169.

Rojas-Lopez, M., M. A. Zorgani, L. A. Kelley, X. Bailly, A. V. Kajava, I. R. Henderson, F. Polticelli, M. Pizza, R. Rosini and M. Desvaux (2017). "Identification of the Autochaperone Domain in the Type Va Secretion System (T5aSS): Prevalent Feature of Autotransporters with a beta-Helical Passenger." Front Microbiol **8**: 2607.

Rosenshine, I., S. Ruschkowski and B. B. Finlay (1996). "Expression of attaching/effacing activity by enteropathogenic Escherichia coli depends on growth phase, temperature, and protein synthesis upon contact with epithelial cells." Infect Immun **64**(3): 966-973.

Ross, B. N., M. Rojas-Lopez, R. J. Cieza, B. D. McWilliams and A. G. Torres (2015). "The Role of Long Polar Fimbriae in Escherichia coli O104:H4 Adhesion and Colonization." PLoS One **10**(10): e0141845.

Rossez, Y., A. Holmes, H. Lodberg-Pedersen, L. Birse, J. Marshall, W. G. Willats, I. K. Toth and N. J. Holden (2014). "Escherichia coli common pilus (ECP) targets arabinosyl residues in plant cell walls to mediate adhesion to fresh produce plants." J Biol Chem **289**(49): 34349-34365.

Rossi, E., A. Cimdins, P. Luthje, A. Brauner, A. Sjoling, P. Landini and U. Romling (2018). ""It's a gut feeling" - Escherichia coli biofilm formation in the gastrointestinal tract environment." Crit Rev Microbiol **44**(1): 1-30.

Rossiter, A. E., D. L. Leyton, K. Tveen-Jensen, D. F. Browning, Y. Sevastsyanovich, T. J. Knowles, K. B. Nichols, A. F. Cunningham, M. Overduin, M. A. Schembri and I. R. Henderson (2011). "The essential beta-barrel assembly machinery complex components BamD and BamA are required for autotransporter biogenesis." J Bacteriol **193**(16): 4250-4253.

Roux, A., C. Beloin and J. M. Ghigo (2005). "Combined inactivation and expression strategy to study gene function under physiological conditions: application to identification of new Escherichia coli adhesins." J Bacteriol **187**(3): 1001-1013.

Saeedi, P., M. Yazdanparast, E. Behzadi, A. H. Salmanian, S. L. Mousavi, S. Nazarian and J. Amani (2017). "A review on strategies for decreasing E. coli O157:H7 risk in animals." Microb Pathog **103**: 186-195.
Sakellaris, H., S. N. Luck, K. Al-Hasani, K. Rajakumar, S. A. Turner and B. Adler (2004). "Regulated site-specific recombination of the she pathogenicity island of Shigella flexneri." Mol Microbiol **52**(5): 1329-1336.
Saldana, Z., J. Xicohtencatl-Cortes, F. Avelino, A. D. Phillips, J. B. Kaper, J. L. Puente and J. A. Giron (2009). "Synergistic role of curli and cellulose in cell adherence and biofilm formation of attaching and effacing Escherichia coli and identification of Fis as a negative regulator of curli." Environ Microbiol **11**(4): 992-1006.
Samadder, P., J. Xicohtencatl-Cortes, Z. Saldana, D. Jordan, P. I. Tarr, J. B. Kaper and J. A. Giron (2009). "The Escherichia coli ycbQRST operon encodes fimbriae with laminin-binding and epithelial cell adherence properties in Shiga-toxigenic E. coli O157:H7." Environ Microbiol **11**(7): 1815-1826.
Sanchez-Torres, V., H. Hu and T. K. Wood (2011). "GGDEF proteins YeaI, YedQ, and YfiN reduce early biofilm formation and swimming motility in Escherichia coli." Appl Microbiol Biotechnol **90**(2): 651-658.
Sandt, C. H. and C. W. Hill (2000). "Four different genes responsible for nonimmune immunoglobulin-binding activities within a single strain of Escherichia coli." Infect Immun **68**(4): 2205-2214.
Sandvig, K., J. Bergan, A. B. Dyve, T. Skotland and M. L. Torgersen (2010). "Endocytosis and retrograde transport of Shiga toxin." Toxicon **56**(7): 1181-1185.
Schembri, M. A., D. W. Ussery, C. Workman, H. Hasman and P. Klemm (2002). "DNA microarray analysis of fim mutations in Escherichia coli." Mol Genet Genomics **267**(6): 721-729.
Scheutz, F. (2014). "Taxonomy Meets Public Health: The Case of Shiga Toxin-Producing Escherichia coli." Microbiol Spectr **2**(3).
Scheutz, F., L. D. Teel, L. Beutin, D. Pierard, G. Buvens, H. Karch, A. Mellmann, A. Caprioli, R. Tozzoli, S. Morabito, N. A. Strockbine, A. R. Melton-Celsa, M. Sanchez, S. Persson and A. D. O'Brien (2012). "Multicenter evaluation of a sequence-based protocol for subtyping Shiga toxins and standardizing Stx nomenclature." J Clin Microbiol **50**(9): 2951-2963.
Scotland, S. M., H. R. Smith, G. A. Willshaw and B. Rowe (1983). "Vero cytotoxin production in strain of Escherichia coli is determined by genes carried on bacteriophage." Lancet **2**(8343): 216.
Serra-Moreno, R., J. Jofre and M. Muniesa (2007). "Insertion site occupancy by stx2 bacteriophages depends on the locus availability of the host strain chromosome." J Bacteriol **189**(18): 6645-6654.
Shaw RK, B. C., Feys B et al. (2008). "Enterohemorrhagic Escherichia coli exploits EspA filaments for attachment to salad leaves." Appl Environ Microbiol 2908-2914.
Sheikh, A., R. Rashu, Y. A. Begum, F. M. Kuhlman, M. A. Ciorba, S. J. Hultgren, F. Qadri and J. M. Fleckenstein (2017). "Highly conserved type 1 pili promote enterotoxigenic E. coli pathogen-host interactions." PLoS Negl Trop Dis **11**(5): e0005586.
Sheikh, J., J. R. Czeczulin, S. Harrington, S. Hicks, I. R. Henderson, C. Le Bouguenec, P. Gounon, A. Phillips and J. P. Nataro (2002). "A novel dispersin protein in enteroaggregative Escherichia coli." J Clin Invest **110**(9): 1329-1337.
Sheng, H., J. Y. Lim, M. K. Watkins, S. A. Minnich and C. J. Hovde (2008). "Characterization of an Escherichia coli O157:H7 O-antigen deletion mutant and effect of the deletion on bacterial persistence in the mouse intestine and colonization at the bovine terminal rectal mucosa." Appl Environ Microbiol **74**(16): 5015-5022.

Sheoran, A. S., S. Chapman-Bonofiglio, B. R. Harvey, J. Mukherjee, G. Georgiou, A. Donohue-Rolfe and S. Tzipori (2005). "Human antibody against shiga toxin 2 administered to piglets after the onset of diarrhea due to Escherichia coli O157:H7 prevents fatal systemic complications." Infect Immun **73**(8): 4607-4613.
Sherlock, O., U. Dobrindt, J. B. Jensen, R. Munk Vejborg and P. Klemm (2006). "Glycosylation of the self-recognizing Escherichia coli Ag43 autotransporter protein." J Bacteriol **188**(5): 1798-1807.
Shrivastava S, M. S. (2008). "Identification and functional characterization of gene components of Type VI Secretion system in bacterial genomes." PloS one.
Smajs, D., H. Pilsl and V. Braun (1997). "Colicin U, a novel colicin produced by Shigella boydii." J Bacteriol **179**(15): 4919-4928.
Smith, S. G., V. Mahon, M. A. Lambert and R. P. Fagan (2007). "A molecular Swiss army knife: OmpA structure, function and expression." FEMS Microbiol Lett **273**(1): 1-11.
Smyth, C. J., M. B. Marron, J. M. Twohig and S. G. Smith (1996). "Fimbrial adhesins: similarities and variations in structure and biogenesis." FEMS Immunol Med Microbiol **16**(2): 127-139.
Solano, C., B. Garcia, J. Valle, C. Berasain, J. M. Ghigo, C. Gamazo and I. Lasa (2002). "Genetic analysis of Salmonella enteritidis biofilm formation: critical role of cellulose." Mol Microbiol **43**(3): 793-808.
Solheim, H. T., C. Sekse, A. M. Urdahl, Y. Wasteson and L. L. Nesse (2013). "Biofilm as an environment for dissemination of stx genes by transduction." Appl Environ Microbiol **79**(3): 896-900.
Sorensen, H. P. and K. K. Mortensen (2005). "Advanced genetic strategies for recombinant protein expression in Escherichia coli." J Biotechnol **115**(2): 113-128.
Spooner, R. A. and J. M. Lord (2012). "How ricin and Shiga toxin reach the cytosol of target cells: retrotranslocation from the endoplasmic reticulum." Curr Top Microbiol Immunol **357**: 19-40.
Stevens, M. P. and G. M. Frankel (2014). "The Locus of Enterocyte Effacement and Associated Virulence Factors of Enterohemorrhagic Escherichia coli." Microbiol Spectr **2**(4): EHEC-0007-2013.
Stevens, M. P., A. J. Roe, I. Vlisidou, P. M. van Diemen, R. M. La Ragione, A. Best, M. J. Woodward, D. L. Gally and T. S. Wallis (2004). "Mutation of toxB and a truncated version of the efa-1 gene in Escherichia coli O157:H7 influences the expression and secretion of locus of enterocyte effacement-encoded proteins but not intestinal colonization in calves or sheep." Infect Immun **72**(9): 5402-5411.
Storz, G. and J. A. Imlay (1999). "Oxidative stress." Curr Opin Microbiol **2**(2): 188-194.
Sugawara, E. and H. Nikaido (1992). "Pore-forming activity of OmpA protein of Escherichia coli." J Biol Chem **267**(4): 2507-2511.
Szabady, R. L., J. H. Peterson, K. M. Skillman and H. D. Bernstein (2005). "An unusual signal peptide facilitates late steps in the biogenesis of a bacterial autotransporter." Proc Natl Acad Sci U S A **102**(1): 221-226.
Szabo, E., A. Skedsmo, A. Sonnevend, K. Al-Dhaheri, L. Emody, A. Usmani and T. Pal (2005). "Curli expression of enterotoxigenic Escherichia coli." Folia Microbiol (Praha) **50**(1): 40-46.
Szczesny, P. and A. Lupas (2008). "Domain annotation of trimeric autotransporter adhesins--daTAA." Bioinformatics **24**(10): 1251-1256.
Takeda, T. (1998). "Strategy to prevent the progression of enterohemorrhagic Escherichia coli O157 infection to hemolytic uremic syndrome." Jpn J Med Sci Biol **51 Suppl**: S124-128.
Tamura, K., R. Sakazaki, Y. Kosako and E. Yoshizaki (1986). "Leclercia-Adecarboxylata Gen-Nov, Comb-Nov, Formerly Known as Escherichia-Adecarboxylata." Current Microbiology **13**(4): 179-184.

Tarr, P. I., S. S. Bilge, J. C. Vary, Jr., S. Jelacic, R. L. Habeeb, T. R. Ward, M. R. Baylor and T. E. Besser (2000). "Iha: a novel Escherichia coli O157:H7 adherence-conferring molecule encoded on a recently acquired chromosomal island of conserved structure." Infect Immun **68**(3): 1400-1407.
Tartaglia, L. A., G. Storz and B. N. Ames (1989). "Identification and molecular analysis of oxyR-regulated promoters important for the bacterial adaptation to oxidative stress." J Mol Biol **210**(4): 709-719.
Tatsuno, I., M. Horie, H. Abe, T. Miki, K. Makino, H. Shinagawa, H. Taguchi, S. Kamiya, T. Hayashi and C. Sasakawa (2001). "toxB gene on pO157 of enterohemorrhagic Escherichia coli O157:H7 is required for full epithelial cell adherence phenotype." Infect Immun **69**(11): 6660-6669.
Tatsuno, I., H. Kimura, A. Okutani, K. Kanamaru, H. Abe, S. Nagai, K. Makino, H. Shinagawa, M. Yoshida, K. Sato, J. Nakamoto, T. Tobe and C. Sasakawa (2000). "Isolation and characterization of mini-Tn5Km2 insertion mutants of enterohemorrhagic Escherichia coli O157:H7 deficient in adherence to Caco-2 cells." Infect Immun **68**(10): 5943-5952.
Tenaillon, O., D. Skurnik, B. Picard and E. Denamur (2010). "The population genetics of commensal Escherichia coli." Nat Rev Microbiol **8**(3): 207-217.
Tesh, V. L. (2010). "Induction of apoptosis by Shiga toxins." Future Microbiol **5**(3): 431-453.
Tesh, V. L., J. A. Burris, J. W. Owens, V. M. Gordon, E. A. Wadolkowski, A. D. O'Brien and J. E. Samuel (1993). "Comparison of the relative toxicities of Shiga-like toxins type I and type II for mice." Infect Immun **61**(8): 3392-3402.
Tilden, J., Jr., W. Young, A. M. McNamara, C. Custer, B. Boesel, M. A. Lambert-Fair, J. Majkowski, D. Vugia, S. B. Werner, J. Hollingsworth and J. G. Morris, Jr. (1996). "A new route of transmission for Escherichia coli: infection from dry fermented salami." Am J Public Health **86**(8): 1142-1145.
Tobe, T., S. A. Beatson, H. Taniguchi, H. Abe, C. M. Bailey, A. Fivian, R. Younis, S. Matthews, O. Marches, G. Frankel, T. Hayashi and M. J. Pallen (2006). "An extensive repertoire of type III secretion effectors in Escherichia coli O157 and the role of lambdoid phages in their dissemination." Proc Natl Acad Sci U S A **103**(40): 14941-14946.
Toma, C., E. Martinez Espinosa, T. Song, E. Miliwebsky, I. Chinen, S. Iyoda, M. Iwanaga and M. Rivas (2004). "Distribution of putative adhesins in different seropathotypes of Shiga toxin-producing Escherichia coli." J Clin Microbiol **42**(11): 4937-4946.
Torres, A. G., C. Jeter, W. Langley and A. G. Matthysse (2005). "Differential binding of Escherichia coli O157:H7 to alfalfa, human epithelial cells, and plastic is mediated by a variety of surface structures." Appl Environ Microbiol **71**(12): 8008-8015.
Torres, A. G. and J. B. Kaper (2003). "Multiple elements controlling adherence of enterohemorrhagic Escherichia coli O157:H7 to HeLa cells." Infect Immun **71**(9): 4985-4995.
Torres, A. G., N. T. Perna, V. Burland, A. Ruknudin, F. R. Blattner and J. B. Kaper (2002). "Characterization of Cah, a calcium-binding and heat-extractable autotransporter protein of enterohaemorrhagic Escherichia coli." Mol Microbiol **45**(4): 951-966.
Torres, A. G., T. M. Slater, S. D. Patel, V. L. Popov and M. M. Arenas-Hernandez (2008). "Contribution of the Ler- and H-NS-regulated long polar fimbriae of Escherichia coli O157:H7 during binding to tissue-cultured cells." Infect Immun **76**(11): 5062-5071.
Totsika, M., T. J. Wells, C. Beloin, J. Valle, L. P. Allsopp, N. P. King, J. M. Ghigo and M. A. Schembri (2012). "Molecular characterization of the EhaG and UpaG trimeric autotransporter proteins from pathogenic Escherichia coli." Appl Environ Microbiol **78**(7): 2179-2189.
Tozzoli, R., L. Grande, V. Michelacci, P. Ranieri, A. Maugliani, A. Caprioli and S. Morabito (2014). "Shiga toxin-converting phages and the emergence of new pathogenic Escherichia coli: a world in motion." Front Cell Infect Microbiol **4**: 80.

Trunk, T., H. S. Khalil and J. C. Leo (2018). "Bacterial autoaggregation." Aims Microbiology **4**(1): 140-164.
Tsai, J. C., M. R. Yen, R. Castillo, D. L. Leyton, I. R. Henderson and M. H. Saier, Jr. (2010). "The bacterial intimins and invasins: a large and novel family of secreted proteins." PLoS One **5**(12): e14403.
Turner, L., R. Zhang, N. C. Darnton and H. C. Berg (2010). "Visualization of Flagella during bacterial Swarming." J Bacteriol **192**(13): 3259-3267.
Turner, S. A., S. N. Luck, H. Sakellaris, K. Rajakumar and B. Adler (2004). "Role of attP in integrase-mediated integration of the Shigella resistance locus pathogenicity island of Shigella flexneri." Antimicrob Agents Chemother **48**(3): 1028-1031.
Tuttle, J., T. Gomez, M. P. Doyle, J. G. Wells, T. Zhao, R. V. Tauxe and P. M. Griffin (1999). "Lessons from a large outbreak of Escherichia coli O157:H7 infections: insights into the infectious dose and method of widespread contamination of hamburger patties." Epidemiol Infect **122**(2): 185-192.
Ude, S., D. L. Arnold, C. D. Moon, T. Timms-Wilson and A. J. Spiers (2006). "Biofilm formation and cellulose expression among diverse environmental Pseudomonas isolates." Environ Microbiol **8**(11): 1997-2011.
UE (2004). "RÈGLEMENT (CE) No 853/2004 DU PARLEMENT EUROPÉEN ET DU CONSEIL du 29 avril 2004 fixant des règles spécifiques d'hygiène applicables aux denrées alimentaires d'origine animale." Journal officiel de l'Union européenne.
Ulett, G. C., J. Valle, C. Beloin, O. Sherlock, J. M. Ghigo and M. A. Schembri (2007). "Functional analysis of antigen 43 in uropathogenic Escherichia coli reveals a role in long-term persistence in the urinary tract." Infect Immun **75**(7): 3233-3244.
Ulett, G. C., R. I. Webb and M. A. Schembri (2006). "Antigen-43-mediated autoaggregation impairs motility in Escherichia coli." Microbiology **152**(Pt 7): 2101-2110.
USDA (2012). "Rules and Regulations: Shiga Toxin-Producing Escherichia coli in Certain Raw Beef Products." ederale Register 77, 31975-31981.
Vakonakis, I. and I. D. Campbell (2007). "Extracellular matrix: from atomic resolution to ultrastructure." Curr Opin Cell Biol **19**(5): 578-583.
van der Woude, M. W. (2006). "Re-examining the role and random nature of phase variation." FEMS Microbiol Lett **254**(2): 190-197.
van der Woude, M. W. (2017). "Epigenetic Phase Variation in Bacterial Pathogens." Epigenetics of Infectious Diseases: 159-173.
van der Woude, M. W. and A. J. Baumler (2004). "Phase and antigenic variation in bacteria." Clin Microbiol Rev **17**(3): 581-611, table of contents.
van der Woude, M. W. and I. R. Henderson (2008). "Regulation and function of Ag43 (flu)." Annu Rev Microbiol **62**: 153-169.
van Diemen, P. M., F. Dziva, M. P. Stevens and T. S. Wallis (2005). "Identification of enterohemorrhagic Escherichia coli O26:H- genes required for intestinal colonization in calves." Infect Immun **73**(3): 1735-1743.
Velarde, J. J. and J. P. Nataro (2004). "Hydrophobic residues of the autotransporter EspP linker domain are important for outer membrane translocation of its passenger." J Biol Chem **279**(30): 31495-31504.
Velarde JJ, V. K., Inman KG et al. (2007). "Solution structure of the novel dispersin protein of enteroaggregative Escherichia coli." Molecular microbiology 1123-1135.
Vikram, A., G. K. Jayaprakasha, R. M. Uckoo and B. S. Patil (2013). "Inhibition of Escherichia coli O157:H7 motility and biofilm by beta-sitosterol glucoside." Biochim Biophys Acta **1830**(11): 5219-5228.

Vo, J. L., G. C. Martinez Ortiz, P. Subedi, S. Keerthikumar, S. Mathivanan, J. J. Paxman and B. Heras (2017). "Autotransporter Adhesins in Escherichia coli Pathogenesis." Proteomics **17**(23-24).
Vogeleer, P., Y. D. Tremblay, G. Jubelin, M. Jacques and J. Harel (2015). "Biofilm-Forming Abilities of Shiga Toxin-Producing Escherichia coli Isolates Associated with Human Infections." Appl Environ Microbiol **82**(5): 1448-1458.
Vogeleer, P., Y. D. Tremblay, A. A. Mafu, M. Jacques and J. Harel (2014). "Life on the outside: role of biofilms in environmental persistence of Shiga-toxin producing Escherichia coli." Front Microbiol **5**: 317.
Wakimoto, N., J. Nishi, J. Sheikh, J. P. Nataro, J. Sarantuya, M. Iwashita, K. Manago, K. Tokuda, M. Yoshinaga and Y. Kawano (2004). "Quantitative biofilm assay using a microtiter plate to screen for enteroaggregative Escherichia coli." Am J Trop Med Hyg **71**(5): 687-690.
Waldron, D. E., P. Owen and C. J. Dorman (2002). "Competitive interaction of the OxyR DNA-binding protein and the Dam methylase at the antigen 43 gene regulatory region in Escherichia coli." Mol Microbiol **44**(2): 509-520.
Walk, S. T., E. W. Alm, D. M. Gordon, J. L. Ram, G. A. Toranzos, J. M. Tiedje and T. S. Whittam (2009). "Cryptic lineages of the genus Escherichia." Appl Environ Microbiol **75**(20): 6534-6544.
Wallecha, A., J. Correnti, V. Munster and M. van der Woude (2003). "Phase variation of Ag43 is independent of the oxidation state of OxyR." J Bacteriol **185**(7): 2203-2209.
Wallecha, A., V. Munster, J. Correnti, T. Chan and M. van der Woude (2002). "Dam- and OxyR-dependent phase variation of agn43: essential elements and evidence for a new role of DNA methylation." J Bacteriol **184**(12): 3338-3347.
Wallecha, A., H. Oreh, M. W. van der Woude and P. L. deHaseth (2014). "Control of gene expression at a bacterial leader RNA, the agn43 gene encoding outer membrane protein Ag43 of Escherichia coli." J Bacteriol **196**(15): 2728-2735.
Wang, G. and M. P. Doyle (1998). "Survival of enterohemorrhagic Escherichia coli O157:H7 in water." J Food Prot **61**(6): 662-667.
Wang, P., Y. Xiong, R. Lan, C. Ye, H. Wang, J. Ren, H. Jing, Y. Wang, Z. Zhou, Z. Cui, H. Zhao, Y. Chen, D. Jin, X. Bai, A. Zhao, Y. Wang, S. Zhang, H. Sun, J. Li, T. Wang, L. Wang and J. Xu (2011). "pO157_Sal, a novel conjugative plasmid detected in outbreak isolates of Escherichia coli O157:H7." J Clin Microbiol **49**(4): 1594-1597.
Wang, X., J. F. Preston, 3rd and T. Romeo (2004). "The pgaABCD locus of Escherichia coli promotes the synthesis of a polysaccharide adhesin required for biofilm formation." J Bacteriol **186**(9): 2724-2734.
Wells, T. J., T. N. McNeilly, M. Totsika, A. Mahajan, D. L. Gally and M. A. Schembri (2009). "The Escherichia coli O157:H7 EhaB autotransporter protein binds to laminin and collagen I and induces a serum IgA response in O157:H7 challenged cattle." Environ Microbiol **11**(7): 1803-1814.
Wells, T. J., O. Sherlock, L. Rivas, A. Mahajan, S. A. Beatson, M. Torpdahl, R. I. Webb, L. P. Allsopp, K. S. Gobius, D. L. Gally and M. A. Schembri (2008). "EhaA is a novel autotransporter protein of enterohemorrhagic Escherichia coli O157:H7 that contributes to adhesion and biofilm formation." Environ Microbiol **10**(3): 589-604.
Wieler, L. H., A. Schwanitz, E. Vieler, B. Busse, H. Steinruck, J. B. Kaper and G. Baljer (1998). "Virulence properties of Shiga toxin-producing Escherichia coli (STEC) strains of serogroup O118, a major group of STEC pathogens in calves." J Clin Microbiol **36**(6): 1604-1607.
Winkler, H. H. and T. H. Wilson (1966). "The role of energy coupling in the transport of beta-galactosides by Escherichia coli." J Biol Chem **241**(10): 2200-2211.
Wurpel, D. J., S. A. Beatson, M. Totsika, N. K. Petty and M. A. Schembri (2013). "Chaperone-usher fimbriae of Escherichia coli." PLoS One **8**(1): e52835.

Xi, C. and J. Wu (2010). "dATP/ATP, a multifunctional nucleotide, stimulates bacterial cell lysis, extracellular DNA release and biofilm development." PLoS One **5**(10): e13355.
Xicohtencatl-Cortes, J., V. Monteiro-Neto, Z. Saldana, M. A. Ledesma, J. L. Puente and J. A. Giron (2009). "The type 4 pili of enterohemorrhagic Escherichia coli O157:H7 are multipurpose structures with pathogenic attributes." J Bacteriol **191**(1): 411-421.
Xicohtencatl-Cortes, J., Z. Saldana, W. Deng, E. Castaneda, E. Freer, P. I. Tarr, B. B. Finlay, J. L. Puente and J. A. Giron (2010). "Bacterial macroscopic rope-like fibers with cytopathic and adhesive properties." J Biol Chem **285**(42): 32336-32342.
Xicohtencatl-Cortes, J., E. Sanchez Chacon, Z. Saldana, E. Freer and J. A. Giron (2009). "Interaction of Escherichia coli O157:H7 with leafy green produce." J Food Prot **72**(7): 1531-1537.
Yamagami, S., M. Motoki, T. Kimura, H. Izumi, T. Takeda, Y. Katsuura and Y. Matsumoto (2001). "Efficacy of postinfection treatment with anti-Shiga toxin (Stx) 2 humanized monoclonal antibody TMA-15 in mice lethally challenged with Stx-producing Escherichia coli." J Infect Dis **184**(6): 738-742.
Yang, J., D. L. Baldi, M. Tauschek, R. A. Strugnell and R. M. Robins-Browne (2007). "Transcriptional regulation of the yghJ-pppA-yghG-gspCDEFGHIJKLM cluster, encoding the type II secretion pathway in enterotoxigenic Escherichia coli." J Bacteriol **189**(1): 142-150.
Yen, M. R., C. R. Peabody, S. M. Partovi, Y. Zhai, Y. H. Tseng and M. H. Saier (2002). "Protein-translocating outer membrane porins of Gram-negative bacteria." Biochim Biophys Acta **1562**(1-2): 6-31.
Yin, X., R. Wheatcroft, J. R. Chambers, B. Liu, J. Zhu and C. L. Gyles (2009). "Contributions of O island 48 to adherence of enterohemorrhagic Escherichia coli O157:H7 to epithelial cells in vitro and in ligated pig ileal loops." Appl Environ Microbiol **75**(18): 5779-5786.
Zhang, X., A. D. McDaniel, L. E. Wolf, G. T. Keusch, M. K. Waldor and D. W. Acheson (2000). "Quinolone antibiotics induce Shiga toxin-encoding bacteriophages, toxin production, and death in mice." J Infect Dis **181**(2): 664-670.
Zhao, B. and W. A. Houry (2010). "Acid stress response in enteropathogenic gammaproteobacteria: an aptitude for survival." Biochem Cell Biol **88**(2): 301-314.
Zhao, H., C. Chen, Y. Xiong, X. Xu, R. Lan, H. Wang, X. Yao, X. Bai, X. Liu, Q. Meng, X. Zhang, H. Sun, A. Zhao, X. Bai, Y. Cheng, Q. Chen, C. Ye and J. Xu (2013). "Global transcriptional and phenotypic analyses of Escherichia coli O157:H7 strain Xuzhou21 and its pO157_Sal cured mutant." PLoS One **8**(5): e65466.
Zhao, T., M. P. Doyle, J. Shere and L. Garber (1995). "Prevalence of enterohemorrhagic Escherichia coli O157:H7 in a survey of dairy herds." Appl Environ Microbiol **61**(4): 1290-1293.
Zude, I., A. Leimbach and U. Dobrindt (2014). "Prevalence of autotransporters in Escherichia coli: what is the impact of phylogeny and pathotype?" Int J Med Microbiol **304**(3-4): 243-256.

LISTE DES COMMUNICATIONS

Communication orales :

Caractérisation de l'Antigène 43 d'*E. coli* dans le processus de colonisation de surface.

V. Ageorges [1], N. Caccia [1], C. Provot [2], F. Chaucheyras-Durand [1,3], E. Peyretaillade [4] et M. Desvaux [1].

[1] Université Clermont Auvergne, INRA, UMR454 MEDiS, 63000 Clermont-Ferrand, France
[2] BioFilm Control, Saint-Beauzire, F-63360, France
[3] Lallemand SAS, 31702 Blagnac Cedex, France.
[4] Université Clermont Auvergne, CNRS, LMGE, 63000 Clermont-Ferrand, France

- *Journées des microbiologistes du pôle Clermontois (France 2018)*
- *Journées de l'école doctorale SVSAE (France 2018)*

Posters :

Characterisation of the molecular determinants involved in the colonisation ability of *Escherichia coli* O157:H7 to the extracellular matrix.

V. Ageorges [1], N. Caccia [1], R. Monteiro [2], C. Chagnot [1], C. Provot [3], F. Chaucheyras-Durand [1,4] et M. Desvaux 1.

[1] Université Clermont Auvergne, INRA, UMR454 MEDiS, 63000 Clermont-Ferrand, France
[2] Institute for Research and Innovation in Health, Host interaction and Response Unit, 4200-135 Porto, Portugal.
[3] BioFilm Control, Saint-Beauzire, F-63360, France
[4] Lallemand SAS, 31702 Blagnac Cedex, France.

- *Congrés FEMS (Espagne 2017)*
- *Congrés du Réseau Nation Biofilm (France 2017)*
- *Congrès Microbiology Socitety (Angleterre 2018)*
- *Congrès VTEC (Italie 2018)*

Characterisation of the Antigen 43 of *Escherichia coli* in the surface colonisation process.

V. Ageorges [1], N. Caccia [1], C. Provot [2], F. Chaucheyras-Durand [1,3], E. Peyretaillade [4] et M. Desvaux [1].

[1] Université Clermont Auvergne, INRA, UMR454 MEDiS, 63000 Clermont-Ferrand, France
[2] BioFilm Control, Saint-Beauzire, F-63360, France
[3] Lallemand SAS, 31702 Blagnac Cedex, France.
[4] Université Clermont Auvergne, CNRS, LMGE, 63000 Clermont-Ferrand, France

- *Symposium BioPhysAdh (France 2018)*

ANNEXES

ANNEXE n°1 : Revue de littérature

Ricardo MONTEIRO, Valentin AGEORGES, Maricarmen ROJAS-LOPEZ, Herbert SCHMIDT, Agnes WEISS, Yolande BERTIN, Evelyne FORANO, Grégory JUBELIN, Ian R. HENDERSON, Valérie LIVRELLI, Alain P. GOBERT, Roberto ROSINI, Marco SORIANI and Mickaël DESVAUX.

A secretome view of colonisation factors in Shiga toxin-encoding *Escherichia coli* (STEC) : from enterohaemorrhagic *E. coli* (EHEC) to related enteropathotypes

Publiée dans FEMS Microbiology Letters, 363, 2016, fnw179

Annexe n°1

FEMS Microbiology Letters, 363, 2016, fnw179

doi: 10.1093/femsle/fnw179
Advance Access Publication Date: 26 July 2016
Minireview

MINIREVIEW – Pathogens & Pathogenicity

A secretome view of colonisation factors in Shiga toxin-encoding *Escherichia coli* (STEC): from enterohaemorrhagic *E. coli* (EHEC) to related enteropathotypes

Ricardo Monteiro[1,2], Valentin Ageorges[1], Maricarmen Rojas-Lopez[1,2], Herbert Schmidt[3], Agnes Weiss[3], Yolande Bertin[1], Evelyne Forano[1], Grégory Jubelin[1], Ian R. Henderson[4], Valérie Livrelli[5], Alain P. Gobert[1,†], Roberto Rosini[2], Marco Soriani[2] and Mickaël Desvaux[1,*]

[1]INRA, UR454 Microbiologie, F-63122 Saint-Genès Champanelle, France, [2]GlaxoSmithKline Vaccines, Via Fiorentina 1, 53100 Siena, Italy, [3]Department of Food Microbiology and Hygiene, Institute of Food Science and Biotechnology, University of Hohenheim, Garbenstrasse 28, 70599 Stuttgart, Germany, [4]Institute of Microbiology and Infection, University of Birmingham, Birmingham B15 2TT, UK and [5]Centre de Recherche en Nutrition Humaine Auvergne, UMR INSERM UdA U1071, USC-INRA 2018, Clermont Université - Université d'Auvergne, Faculté de Pharmacie, CHU Clermont-Ferrand, Service Bactériologie Mycologie Parasitologie, Clermont-Ferrand, France

*Corresponding author: INRA Auvergne-Rhône-Alpes, Clermont-Ferrand Research Centre, UR454 Microbiology, Site of Theix, F-63122 Saint-Genès Champanelle, France. Tel: +33 (0)4-73-62-47-23; Fax: +33 (0)4-73-62-45-81; E-mail: mickael.desvaux@clermont.inra.fr

†Present address: Center for Mucosal Inflammation and Cancer, Vanderbilt University Medical Center, 2215B Garland Avenue, 1015C MRB IV, Nashville, TN 37232-0252, USA.

One sentence summary: The diversity of colonisation factors in the different subgroups of Shiga-toxin encoding *Escherichia coli* and related enteropathotypes is reviewed respective to their cognate protein secretion routes in Gram-negative bacteria.

Editor: David Clarke

Downloaded from http://femsle.oxfordjournals.org/ by guest on January 17, 2017

ABSTRACT

Shiga toxin-encoding *Escherichia coli* (STEC) regroup strains that carry genes encoding Shiga toxin (Stx). Among intestinal pathogenic *E. coli*, enterohaemorrhagic *E. coli* (EHEC) constitute the major subgroup of virulent STEC. EHEC cause serious human disease such as haemorrhagic colitis and haemolytic-uremic syndrome. While EHEC have evolved from enteropathogenic *E. coli*, hybrids with enteroaggregative *E. coli* have recently emerged. Of note, some enteroinvasive *E. coli* also belong to the STEC group. While the LEE (locus of enterocyte effacement) is a key and prominent molecular determinant in the pathogenicity, neither all EHEC nor STEC contain the LEE, suggesting that they possess additional virulence and colonisation factors. Currently, nine protein secretion systems have been described in diderm-lipopolysaccharide bacteria (archetypal Gram-negative) and can be involved in the secretion of extracellular effectors, cell-surface proteins or assembly of cell-surface organelles, such as flagella or pili. In this review, we focus on the secretome of STEC and related enteropathotypes, which are relevant to the colonisation of biotic and abiotic surfaces. Considering the

Received: 11 April 2016; **Accepted:** 18 July 2016
© FEMS 2016. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com

Annexe n°1

wealth of potential protein trafficking mechanisms, the different combinations of colonisation factors and modulation of their expression is further emphasised with regard to the ecophysiology of STEC.

Keywords: secretion systems; cell-surface display; surface organelle; outer membrane proteins; autotransporters; bacterial protein secretion

INTRODUCTION

Escherichia coli is an inhabitant of the intestines of warm-blooded animals, including humans (Tenaillon *et al.* 2010). *Escherichia coli* typically coexists with the host as a commensal with mutual benefit for both organisms (Hooper and Gordon 2001) and is usually shed in faeces (Kaper, Nataro and Mobley 2004). Despite its wide prevalence, it rarely causes disease. However, based on their site of infection, pathogenic strains can be broadly divided into intestinal pathogenic *E. coli* (InPEC) resulting in diarrhoea, and extraintestinal pathogenic *E. coli* responsible for urinary tract infections, sepsis and/or meningitis (Nataro and Kaper 1998). In humans, one of the greatest disease burdens from *E. coli* results from intestinal infections with severity ranging from mild watery diarrhoea and abdominal cramps that proceed to haemorrhagic colitis and severe complications that may lead to death (Karch, Tarr and Bielaszewska 2005). InPEC represents a leading cause for traveler's diarrhoea, an emerging source of diarrhoea in industrialised countries, and a major cause of paediatric disease and morbidity in children in developing countries (Cohen *et al.* 2005). Infection mainly results from consumption of food products exposed to direct and/or indirect faecal contamination.

Escherichia coli genomes consist of conserved regions, the so-called core genome and more flexible regions, which are interspersed (Dobrindt *et al.* 2010). The latter contain frequently foreign DNA such as bacteriophages, genomic and pathogenicity islands (PAIs), which can define new phenotypic characteristics for the host (Dobrindt 2005; Schmidt 2010). Moreover, plasmids contribute to the horizontal gene transfer and the creation of novel combinations of virulence factors leading to new phenotypes and thus the emergence of new pathogens (Kaper, Nataro and Mobley 2004; Nataro *et al.* 2006). Besides, some of these virulence factors are horizontally transferred due to the promiscuity nature of *E. coli*. During evolution, numerous gene exchange events have occurred, resulting in the appearance of new pathotypes that cause new types of disease and outbreaks (Creuzburg *et al.* 2011a,b). This concept of different *E. coli* enteropathotypes was originally developed by Nataro and Kaper (1998) and Kaper, Nataro and Mobley (2004). Based on a combination of clinical, epidemiological and molecular criteria, six InPEC are currently recognised: enterotoxigenic *E. coli* (ETEC), enteropathogenic *E. coli* (EPEC), enterohaemorrhagic *E. coli* (EHEC), enteroaggregative *E. coli* (EAEC), enteroinvasive *E. coli* (EIEC) and diffusely adherent *E. coli* (Qadri *et al.* 2005; Amisano *et al.* 2011).

Shiga toxin-encoding *E. coli* (STEC), or Shiga toxin-producing *E. coli*, comprise a group of bacteria that carry genes encoding Shiga toxin (Stx) and/or have been shown to produce one or more Stx (Nataro and Kaper 1998). The term STEC was originally introduced by O'Brien and colleagues to describe the molecular and functional similarity of the Stx encoded by the genetically closely related *Shigella dysenteriae* type 1 (McDonough and Butterton 1999). In fact, *Shigella* and EIEC form a single enteropathotype exhibiting different virulence levels and causing varying degrees of dysentery (Lan *et al.* 2004; Croxen *et al.* 2013). A large amount of different *stx* subtypes with different designations have been detected in *E. coli* during the last decades, and in a recent large multicenter study, a unified Stx nomenclature has been recommended (Scheutz *et al.* 2012). Notably, some individuals still use older designations such as verotoxins or verocytotoxins (Calderwood *et al.* 1996; Melton-Celsa and O'Brien 2000).

Several subgroups of STEC are defined and among InPEC, EHEC is the most important group. From early definitions, EHEC are pathogenic STEC that are able to cause the typical symptoms from watery diarrhoea, bloody diarrhoea, up to haemorrhagic colitis (HC), haemolytic uremic syndrome (HUS) and/or other extraintestinal sequelae in humans (Tarr, Gordon and Chandler 2005). Little progress has been made in reducing the number of infections associated with these pathogens, and few interventions to reduce the food contamination and the infectious complications of this enteric disease are available (Walker and Black 2010). While the vast majority of EHEC infections are sporadic, they can also lead to major outbreaks worldwide (Karch *et al.* 1999). Of note, ruminants are the natural reservoir of EHEC but they are asymptomatic carriers. There are over 380 distinct serotypes of STEC (Karmali *et al.* 2003; Karmali, Gannon and Sargeant 2010) but only EHEC of serotypes O157:H7, as well as the big six serotypes O26:H11, O45:H2, O103:H2, O111:H8, O121:H19 and O145:H28 are the most frequently associated with human disease (Mora *et al.* 2007). Besides Stx production, some EHEC can form attaching and effacing (A/E) lesions, as demonstrated in human cultured cells (Kelly *et al.* 1987; Nataro and Kaper 1998; Lewis *et al.* 2015), and encoded by a chromosomal PAI known as the locus of enterocyte effacement (LEE) (Perna *et al.* 2001). LEE-positive EHEC and LEE-positive EPEC are sometimes referred to A/E encoding *E. coli* (AEEC) (Frohlicher *et al.* 2008). Interestingly, EHEC O157:H7 would have evolved from EPEC O55:H7 following a series of stepwise acquisitions and losses of genes (Kyle *et al.* 2012). The terms AEEC and STEC rely mainly on the genetic traits of the bacteria (genotype), i.e. harbouring *eae* (enterocyte attaching and effacing) or *stx* respectively, whereas EHEC and EPEC are primarily classified on the clinical symptoms of the patients (phenotype) and in a second line on the molecular traits (Caprioli *et al.* 2005). Recently, LEE-negative STEC and EHEC strains have also been reported (Hauser *et al.* 2013). LEE-negative STEC occur mainly in animals or the environment, and are frequently isolated from food sources. LEE-negative EHEC are more rarely isolated from patients with HUS. However, a LEE-negative *E. coli* O104:H4 caused the large outbreak in Europe in 2011 and appeared as a hybrid of EAEC and EHEC (Brzuszkiewicz *et al.* 2011). These *E. coli* O104:H4 strains have been tentatively named enteroaggregative haemorrhagic *E. coli* (Mellmann *et al.* 2011) and can be categorised as a subgroup of EHEC. While the LEE is a key and prominent molecular determinant in the pathogenicity, neither all EHEC nor STEC possess the LEE, suggesting that they possess additional virulence and colonisation factors (Paton *et al.* 2001).

Basically, pathogenic organisms differentiate from their non-pathogenic counterparts by encoding specific virulence factors secreted to the bacterial cell surface or beyond (Finlay and

Downloaded from http://femsle.oxfordjournals.org/ by guest on January 17, 2017

Annexe n°1

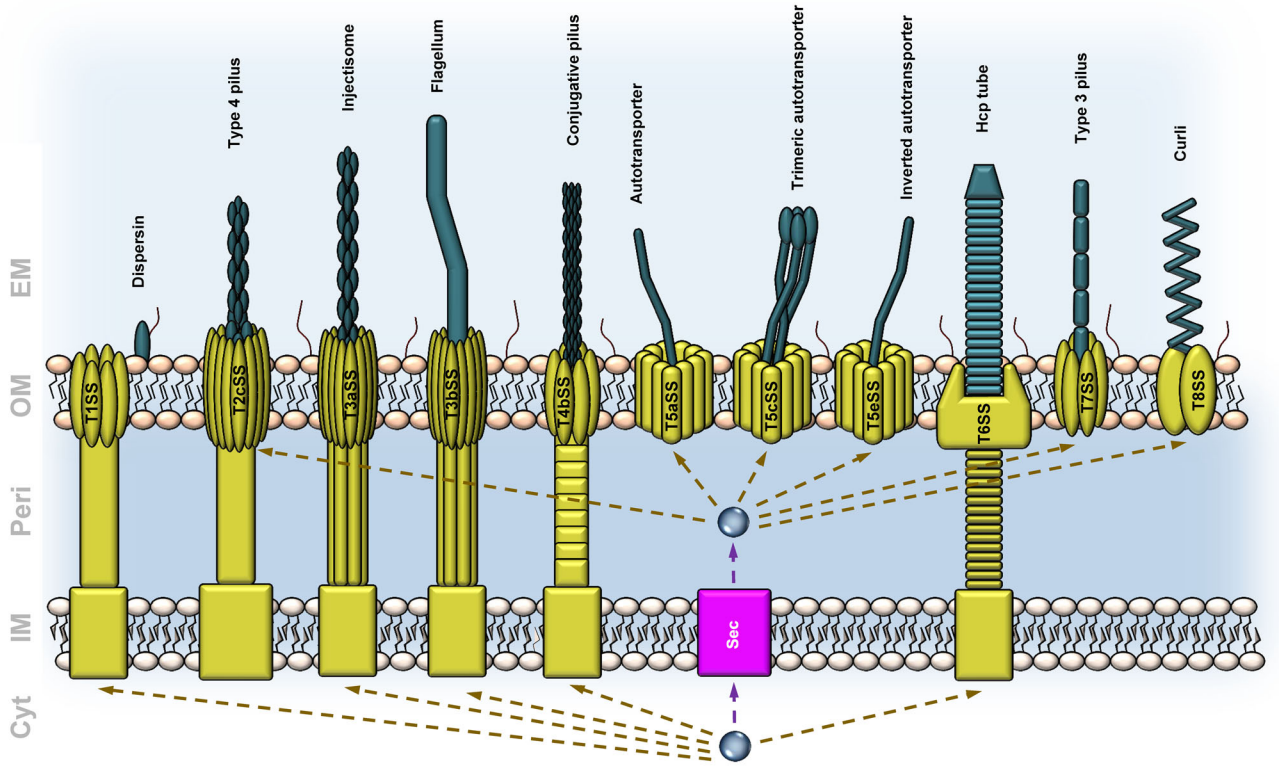

Figure 1. Overview of the complement of the secretome associated with the surface colonisation factors in STEC. Protein export systems are coloured in violet, protein secretion systems are coloured in green and their cognate effectors directly involved in surface colonisation are coloured in blue. For details of the respective secretion system and cognate colonisation factors, refer to the main text. T1SS: type I protein secretion system or the ABC-MFP-TolC heterotrimeric system. T2SS: type II protein secretion system or the SDP, especially the T4P system (T2cSS). T3SS: type III protein secretion system, including the injectisome (T3aSS) and flagellar (T3bSS) systems. T4SS: type IV protein secretion system, especially the conjugative Tra system (T4bSS). T5SS: type V protein secretion system, especially including the autotransporter (T5aSS), trimeric autotransporter (T5cSS) and inverted autotransporter (T5eSS) systems. T6SS: type VI protein secretion system. T7SS: type VII protein secretion system or the CU pathway. T8SS: type VIII protein secretion system or the ENP pathway. Cyt, cytoplasm; IM, inner membrane; Peri, periplasm; OM, outer membrane; EM, extracellular milieu.

Falkow 1997). To date, nine protein secretion systems, numbered from type I to type IX secretion system (T1SS to T9SS), have been uncovered in diderm-LPS (lipopolysaccharide) bacteria (archetypal Gram-negative bacteria) (Chagnot *et al.* 2013). Considering the wealth of systems and effectors, the secretome is a powerful concept to apprehend protein trafficking as a whole in living cells by considering both the secretion routes and their cognate secreted proteins (Tjalsma *et al.* 2000; Antelmann *et al.* 2001). Respective to the framework for STEC including related enteropathotypes such as EHEC, the complement of the secretome associated with the expression of colonisation factors is reviewed here. The colonisation processes are essential not only in human infection but also for bacterial multiplication in animals and contamination of foodstuffs. In order to highlight the molecular mechanisms responsible for their correct expression and subcellular localisation, those colonisation effectors are stated here from a protein secretion system perspective.

COLONISATION FACTORS OF STEC FROM A SECRETOME PERSPECTIVE

Protein secretion is an essential cellular function present in all living cells. In bacteria, it is involved in a huge diversity of functions playing important roles in the bacterial physiology from adaptation and response to environmental cues, adhesion and biofilm formation, to pathogenicity and virulence (Chagnot *et al.* 2013). While secretion refers to the active transport from the interior to the exterior of the cell compartment, secreted proteins have three main possible fates in diderm-LPS bacteria (archetypal Gram-negative bacteria): (i) remain associated with the bacterial outer membrane (OM), (ii) be released into the extracellular milieu or (iii) be injected into a target cell (either a eukaryotic or bacterial cell) (Gerlach and Hensel 2007). Among the nine protein secretion systems in diderm-LPS bacteria secreted proteins (Chagnot *et al.* 2013), T1SS to T8SS could potentially be involved in surface colonisation of STEC (Fig. 1). Here, the secretion systems and their cognate secreted proteins involved in the surface colonisation process, namely adhesion and biofilm formation, are reviewed in STEC (Table 1) including EHEC and EIEC, as well as their related enteropathotypes, namely EPEC and EAEC. In addition, some OM proteins (Omp) assembled by the Bam (β-barrel assembly machinery) complex (Knowles *et al.* 2009), namely IhA (IrgA homologue adhesion) (Tarr *et al.* 2000; Herold *et al.* 2009) or OmpA (Torres and Kaper 2003), can also participates in the colonisation process in STEC.

Type I secretion system: dispersin

The type I secretion system (T1SS) refers to a heterotrimeric complex constituted of an inner membrane ABC exporter, a membrane fusion protein (MFP) and a TolC-like pore-forming Omp (Thomas, Holland and Schmitt 2014). Dispersin Aap (anti-aggregation protein) is a well-known effector secreted via the T1SS and present in some STEC, such as EHEC O104:H4 (Muniesa

Downloaded from http://femsle.oxfordjournals.org/ by guest on January 17, 2017

Annexe n°1

Table 1. Colonisation factors of STEC, including EHEC and related enteropathotypes, according to the protein secretion systems in diderm-LPS bacteria. Please refer to the main text for details.

T1SS[a]	T2SS[b]	T3SS[c]		T4SS	T5SS[d]			T6SS[e]	T7SS[f]	T8SS	Non-classical secretion[g]
	T2cSS	T3aSS	T3bSS	T4bSS	T5aSS	T5cSS	T5eSS				
Aap	HCP	Injectisome Tir	Flagellum	Conjugative pilus	EhaA EhaB EhaD EhaJ Cah EspP EspC PssA Sab	EhaG Saa EibG	Intimin FdeC	Hcp	ECP Lpf ELF SFP F9 AAF	Curli	GAPDH

[a]Aap: dispersin, anti-aggregation protein.
[b]HCP: haemorrhagic coli pilus.
[c]Tir: translocated intimin receptor.
[d]Eha: enterohaemorrhagic *E. coli* autotransporters; Cah: calcium-binding antigen 43 homologue; EspP: extracellular serine protease plasmid encoded; EspC: EPEC-secreted protein C; PssA: protease secreted by STEC; Sab: STEC autotransporter; Saa: STEC autoagglutinating adhesion autotransporter; EibG: *E. coli* immunoglobulin-binding protein G; FdeC: factor adherence *E. coli*.
[e]Hcp: haemolysin-coregulated protein.
[f]ECP: *E. coli* common pilus; Lpf: Long polar fimbria; ELF: *E. coli* YcbQ laminin-binding fimbria; SFP: sorbitol-fermenting fimbriae; AAF: aggregative adherence fimbria.
[g]Non-classical secretion refers to unknown secretion system (i.e. as yet uncovered) but, in the end, it could also correspond to protein trafficking mechanisms unrelated to protein secretion; GAPDH: glyceraldehyde 3-phosphate dehydrogenase.

et al. 2012). From investigations in EAEC, the T1SS is encoded in the enteroaggregative ABC transporter locus (*aat*) present on a large virulence plasmid called pAA (aggregative adherence) (Nishi *et al.* 2003). This system is involved in the secretion of the Aap dispersin at the cell surface following non-covalent binding to LPS (Velarde *et al.* 2007). By preventing hyperaggregation of bacteria and collapse of aggregative adherence fimbriae (AAF), dispersin allows dispersion of bacterial cells along the host intestinal mucosa, thus contributing to the adherence and colonisation of EAEC (Sheikh *et al.* 2002). As it is found in some commensal *E. coli* but is also absent from some EAEC (Monteiro *et al.* 2009), the role of dispersin in pathogenesis is not clear.

In EPEC, the bundle-forming pili (BFP) are responsible for initial bacterial attachment to the intestinal epithelium and formation of compact microcolonies (Scaletsky, Silva and Trabulsi 1984; Giron, Ho and Schoolnik 1991; Stone *et al.* 1996). BFP are encoded on the plasmid pEAF (EPEC adherence factor) and are therefore not systematically present in all EPEC (Giron *et al.* 1993; Croxen *et al.* 2013). In some EAEC, T4P are encoded on conjugative plasmids (thus encoding a T4bSS in addition, see below) and contribute to bacterial adherence to epithelial cells and abiotic surfaces, as well as biofilm formation (Dudley *et al.* 2006a). The presence and prevalence of those alternative T4P, such BFP, in STEC would require further investigations.

Type II subtype c secretion systems: type 4 pili

The T2SS, or secretion-dependent pathway (SDP), allows secretion of some proteins first exported via the Sec or Tat systems (Voulhoux *et al.* 2001). While the T2SS is primarily involved in the secretion of free soluble extracellular proteins, namely through the subfamilies Xcp, i.e. the T2SS of subtype a (T2aSS), and Hxc, i.e. the T2SS of subtype b (T2bSS), it can also be involved in pili assembly (Durand *et al.* 2011). Indeed, the system responsible for the subunit secretion, assembly and biogenesis of the type 4 pili (T4P) is a distant homologue to the paradigm Xcp T2aSS, and actually corresponds to the type II subtype c secretion system (T2cSS) (Ayers, Howell and Burrows 2010; Chagnot *et al.* 2013). In pathogenic *E. coli*, T4P are involved in bacterial virulence, including the colonisation process, especially host cell adherence, biofilm formation, bacterial aggregation and the twitching motility (Craig, Pique and Tainer 2004). In STEC, the T4P named HCP (haemorrhagic coli pilus) is composed of a 19-kDa pilin subunit encoded by the *hcpA* chromosomal gene, also called *ppdD* (prepilin peptidase-dependent) in *E. coli* K-12. In EHEC O157:H7 EDL933, HcpA is directly involved in cell adherence and in the ability to invade human and bovine epithelial cells. HCP also contributes to the biofilm formation due to its ability to bind some extracellular matrix (ECM) proteins, especially fibronectin and laminin (Xicohtencatl-Cortes *et al.* 2009).

Type III secretion systems: injectisome and flagella

The injectisome and the flagellum are assembled by two phylogenetically distinct but homologous type III secretion systems (T3SS) of subtypes a and b, i.e. T3aSS (injectisome system) and the T3bSS (flagellar system), respectively (Desvaux *et al.* 2006; Diepold and Armitage 2015). Among InPEC, the T3aSS is the hallmark of EPEC but it is also present in some but not all EHEC (Galan and Wolf-Watz 2006). The T3aSS is a double-membrane embedded nanomachine and promotes the delivery of bacterial effectors to the cytoplasm or plasma membrane of target eukaryotic cells, where they can modulate or subvert a large variety of host cell mechanisms but also promote bacterial invasion and colonisation (Cornelis 2006). The injectisome formed by the T3aSS is a needle structure leaning on a multiring complex that spans the bacterial envelope and the plasma membrane of the host cell providing a continuous path up to the cytosol of the infected host cell. In EPEC, the T3aSS is directly involved in cell adhesion and pedestal formation resulting in characteristic histopathological A/E lesions (Wong *et al.* 2011). However, it must be stressed that such lesions were never observed from clinical samples after EHEC infection (Kelly *et al.* 1987; Nataro and Kaper 1998; Lewis *et al.* 2015). Nonetheless, A/E lesions are frequently observed *in vitro* from EHEC-infected epithelial cell cultures. The T3aSS structure is encoded by genes located in the

Downloaded from http://femsle.oxfordjournals.org/ by guest on January 17, 2017

Annexe n°1

LEE PAI (McDaniel *et al.* 1995). Among all the proteins secreted by the injectisome T3aSS, Tir (translocated intimin receptor) is the primary molecule associated with intimate bacterial interaction with epithelia and A/E lesions (Donnenberg *et al.* 1993). Tir, encoded by the *espE* gene located in the LEE, is secreted and injected into the host cell by the injectisome T3aSS, and later localised at the plasma membrane of the infected host epithelial cell. This bacterial protein exposed at the host cell surface then acts as a receptor for a direct and specific binding of the bacterial cell-surface exposed intimin. By protein–protein interaction, Tir activates the recruitment and rearrangement of the host-cell cytoskeletal actin, which rearranges and results in the formation of a characteristic pedestal structure (Liu *et al.* 1999; Liu, Magoun and Leong 1999). Furthermore, the injectisome plays a role in the adhesion to plants with a marked tropism for the stomata (Shaw *et al.* 2008; Berger *et al.* 2010). Besides adhesion to the host cell, the T3aSS also participates in invasion capabilities in EIEC (Schroeder and Hilbi 2008; Croxen *et al.* 2013), which is apparently not the case in EHEC or EPEC.

While T3aSS is only present in a subset of InPEC, namely EPEC and some EHEC, flagellum is considered quite ubiquitous across *E. coli*. In *E. coli*, the flagella are peritrichous forming a ponytail when in motion (DiLuzio *et al.* 2005). The different components of the flagellum are secreted by the T3bSS, namely the hook-filament junction protein, the filament-capping protein and flagellin, the major subunit of the filament (Macnab 2004; Altegoer and Bange 2015). Of course, flagella are primarily involved in cell motility, especially swimming and/or swarming in *E. coli* depending on environmental conditions (Turner *et al.* 2010). Besides coordinated movement at surfaces contributing to colonisation, that is swarming, flagella participate in adhesion and invasion by providing motility towards surfaces or target cells (Haiko and Westerlund-Wikstrom 2013). In EHEC, flagella play a role in the adhesion to the epithelial cells but are downregulated after contact with the epithelium and would then just initiate early stages of the adhesion process (Mahajan *et al.* 2009). In EPEC, the flagellum tip protein FliD can mediate adhesion to Caco-2 cells. Following vaccination targeting the flagellin FliC, the colonisation of EHEC in cattle was significantly decreased providing evidence of the importance of flagella in host gut colonisation (McNeilly *et al.* 2008). Most EIEC, though, have no flagellum and, consequently, are non-motile with no H antigen serotype (Croxen *et al.* 2013).

Type IV subtype b secretion system: conjugative pili

Respective to protein secretion, the T4SS can be discriminated into two subtypes. The T4aSS includes effector translocator systems homologous to the prototypical VirB/D4 complex, whereas the type IV subtype b secretion system (T4bSS) comprises the conjugation machines homologous to the prototypical F-episome conjugal transfer (Tra) system (Christie and Vogel 2000; Alvarez-Martinez and Christie 2009; Christie, Whitaker and Gonzalez-Rivera 2014). While no T4aSS has been described so far in *E. coli*, conjugative plasmids can be present (Lawley *et al.* 2003). From what is known of the F episome in *E. coli* K-12 (Ghigo 2001), conjugative plasmids could further contribute to bacterial colonisation. Indeed, conjugative F pili assembled by the T4bSS induce biofilm formation by improving adhesion capability of the bacterial cells. Horizontal transfer of the F episome to siblings within the biofilm further increases the proportion of transconjugant cells and consequently expands the surface colonisation propensity of the bacterial population (Ghigo 2001; May and Okabe 2008; May, Ito and Okabe 2010). Besides pO157 (Lim, Yoon and Hovde 2010), plasmid profiling in different *E. coli* O157:H7 strains indeed revealed the presence of numerous additional plasmids, which highly vary in size and number (Ostroff *et al.* 1989; Paros *et al.* 1993). Among them, pO157-Sal was identified and demonstrated to be a novel conjugative plasmid of the IncI family in STEC (Fratamico *et al.* 2011; Wang *et al.* 2011). This plasmid contains the full set of *tra* genes and thus encodes a T4bSS, but its contribution to sessile development remains to be established.

Type V secretion system: adhesins by numbers

The type V secretion system (T5SS) can be categorised into five subtypes: (i) the autotransporter system (subtype a), (ii) the two-partner secretion pathway (subtype b), (iii) the trimeric autotransporter system (subtype c), (iv) the hybrid autotransporter system (subtype d) and (v) the inverted autotransporter system (subtype e). Broadly speaking, the T5SS refers to protein secretion via an OM pore formed by a β-barrel, the secreted proteins first being exported via the Sec translocon. Except for the T5dSS, all four other subtypes can secrete proteins involved in surface colonisation. Although a T5bSS is present in *E. coli* O157:H7 (Choi, Dawson and Bernstein 2007), the exoprotein OtpA (O157:H7 two-partner protein A) does not display any adhesive properties towards intestinal epithelial cells, and the possible contribution of this secretion system and cognate secreted proteins to surface colonisation remains to be established.

Type V subtype a secretion system (T5aSS): SAAT and SPATE adhesins

Autotransporters are single polypeptides that can drive their own secretion through the OM via a C-terminal translocator forming a β-barrel in the OM allowing translocation of the N-terminal passenger domain (Desvaux, Parham and Henderson 2004). The self-associating autotransporters (SAATs) are cell-surface exposed outer-membrane proteins (OMPs) encompassing (i) AIDA (adhesin diffuse adherence), (ii) TibA (enterotoxigenic invasion locus b protein A) and (iii) Ag43 (antigen 43) autotransporter family (Benz and Schmidt 1989; Klemm, Vejborg and Sherlock 2006). Besides autoaggregation, SAATs are actively involved in bacteria–host interaction and biofilm formation (Sherlock *et al.* 2004). Five autotransporters belonging to the AIDA family have been identified in *E. coli* O157:H7 and are called enterohaemorrhagic *E. coli* autotransporters (Eha) (Wells *et al.* 2008). The EhaA overexpression in *E. coli* K-12 appeared to promote autoaggregation, biofilm formation and adhesion to bovine primary epithelial cells from the terminal rectum (Wells *et al.* 2008). EhaB specifically binds to some ECM proteins, especially laminin and collagen I (Wells *et al.* 2009). Similarly, EhaJ mediates specific adhesion to collagens I, II, III and V as well as fibronectin, fibrinogen and laminin (Easton *et al.* 2011). While EhaD and EhaJ also promote biofilm formation, the possible function of EhaC in the colonisation processes remains to be established (Wells *et al.* 2008). Interestingly enough, glycosylation of EhaJ appeared of importance for biofilm formation but not for binding to the ECM (Easton *et al.* 2011). While no TibA homologue has been reported in STEC, *E. coli* O157:H7 encodes an Ag43 homologue called Cah (calcium-binding antigen 43 homologous) promoting cell autoaggregation but apparently not involved in adhesion to the host cells (Torres *et al.* 2005). The SPATEs (serine protease autotransporters of *Enterobacteriaceae*) constitute another autotransporter family generally released to the extracellular milieu and primarily with proteolytic activity against various

Downloaded from http://femsle.oxfordjournals.org/ by guest on January 17, 2017

substrates but also exhibiting adhesion properties for some of them (Henderson *et al.* 2004). In *E. coli* O157:H7, EspP (extracellular serine protease plasmid-encoded) is directly involved in biofilm formation but also adherence to epithelial cells (Dziva *et al.* 2007; Puttamreddy, Cornick and Minion 2010). Recently, EspP was shown to oligomerise to form megastructural ropes, which possess adhesive and cytopathic activities on host epithelial cells (Puttamreddy, Cornick and Minion 2010; Xicohtencatl-Cortes *et al.* 2010). In EPEC, EspC oligomerises similarly and serves as substratum for bacterial adherence and biofilm formation (Xicohtencatl-Cortes *et al.* 2010). In STEC O26, PssA (protease secreted by STEC) was shown to participate in intestinal colonisation of calves (van Diemen *et al.* 2005). In STEC O113, Sab (STEC autotransporter contributing to biofilm formation) is involved in adherence to abiotic surfaces and epithelial cells (Herold, Paton and Paton 2009). Of note, the high variability in the presence of those different autotransporters in STEC may significantly contribute to differences in colonisation abilities and even modulate virulence. For example, the identification of Sab in LEE-negative STEC O113:H2 contrasts with its absence from a LEE-positive strain collection and would suggest an alternative way to adhere to the host cells for strains defective in their ability to induce A/E lesions (Herold, Paton and Paton 2009).

Type V subtype c secretion system: trimeric autotransporter adhesins

The type V subtype c secretion system (T5cSS) corresponds to autotransporters formed upon homotrimerisation. All trimeric autotransporters characterised to date are exposed to the bacterial cell surface and play a role in adhesion (El Tahir and Skurnik 2001; Linke *et al.* 2006; Szczesny and Lupas 2008). EhaG (EHEC adhesin G) is present in diarrheagenic *E. coli*, from EHEC, ETEC, EPEC, EAEC to EIEC, and is highly prevalent in STEC (Totsika *et al.* 2012). Upon overexpression in *E. coli* K-12, EhaG from *E. coli* O157:H7 was shown to mediate autoaggregation resulting in the formation of strong biofilm. In addition, adhesion to intestinal epithelial cells and specific binding to collagens I, II, III and V as well as to laminin, fibronectin and fibrinogen was reported (Totsika *et al.* 2012). Saa (STEC autoagglutinating adhesin) is involved in adhesion to intestinal epithelial cells and was identified in various LEE-negative STEC (Paton *et al.* 2001). EibG (*E. coli* immunoglobulin-binding protein G) contributes to a chain-like adhesion pattern to human epithelial cells also in LEE-negative STEC (Lu *et al.* 2006; Merkel *et al.* 2010).

Type V subtype e secretion system: intimins

The type V subtype e secretion system (T5eSS) refers to inverted autotransporters in the sense that the translocator is located at the N-terminal instead of the C-terminal region of the monomeric autotransporter (Oberhettinger *et al.* 2012, 2015; Leo *et al.* 2015). Intimin from EPEC and EHEC is a prototypical member of this OMP family. Along with the T3aSS, intimin is encoded by the *eae* gene in the LEE and, as already mentioned above, interacts specifically with Tir resulting in intimate attachment of the bacteria with the host-cell surface, pedestal formation and A/E lesions (Schmidt 2010). While five alleles (α, β, γ, δ and ε) have been reported for *eae* with a total of 27 variants (Adu-Bobie *et al.* 1998; Oswald *et al.* 2000), these intimins appeared functionally interchangeable (Liu *et al.* 1999; Mallick *et al.* 2012). Besides Tir, the binding of intimin to alternative receptors such as β_1 integrins or nucleolin remains unclear (Liu, Magoun and Leong 1999; Leo *et al.* 2015). Nonetheless, intimin can also contribute to intestinal colonisation independently from its ability to bind to Tir (Mallick *et al.* 2012).

While no invasion has been reported to date in STEC (Leo *et al.* 2015), FdeC (factor adherence *E. coli*) was recently uncovered as another representative of the T5eSS and appeared widely distributed in *E. coli*, including InPEC, e.g. STEC and EHEC (Nesta *et al.* 2012; Easton *et al.* 2014). FdeC binds to human epithelial cells and exhibits affinity for some ECM components, notably collagens I, III, V and VI (Nesta *et al.* 2012). Its contribution to the kidney and bladder colonisation was associated with the propensity of UPEC bacterial cells for autoaggregation. In EHEC O26:H11, the expression of FdeC regulated by H-NS (histone-like nucleoid-structuring protein) and occurring at temperatures above 39°C was demonstrated to mediate biofilm formation and could contribute to the colonisation of the terminal rectum of cattle (Easton *et al.* 2014).

Type VI secretion system: aggregation to host cell surface

The type VI secretion system (T6SS) consists of two principal associated complexes with an additional bridging and cytoplasmic elements: a membrane-associated assembly and an assembly similar to the bacteriophage sheath, tube and spike proteins (Leiman *et al.* 2009). In some EHEC, EPEC and EAEC strains, bioinformatic analysis identified more than 10 orthologues of known T6SS components present (Shrivastava and Mande 2008; Lloyd *et al.* 2009; Aschtgen *et al.* 2010; Moriel *et al.* 2010), suggesting that they could be crucial mediators in the aggregation to the host cell surface (Dudley *et al.* 2006b). In EAEC, the correct expression of the T6SS apparatus, including the Hcp (haemolysin-coregulated protein) tube forming the inverted phage tail, appeared critical for biofilm formation (Aschtgen *et al.* 2008). Still, the exact contribution of this secretion system in the colonisation process remains to be established in STEC.

Type VII secretion system: type 3 pili by numbers

In diderm-LPS bacteria, the type VII secretion system (T7SS) corresponds to the chaperone-usher (CU) pathway involved in the secretion and assembly of pili (Desvaux *et al.* 2009). The nomenclature for pili formed via the T7SS is messy, quite confusing and essentially species dependent, e.g. under the general denomination of type 3 pili, the designations of CS (coli surface) pili, P pili, aggregative adherence fimbriae (AafD), adhesive fimbriae on RDEC-1 and diffuse adherence fibrillar adhesin/Dr blood group antigen (Afa/Dr) are also found but all fall under the T7SS umbrella (Zav'yalov *et al.* 2010; Busch and Waksman 2012; Thanassi, Bliska and Christie 2012). While numerous operons encoding such pili remain putative (Hayashi *et al.* 2001, Low *et al.* 2006b, Korea *et al.* 2010), STEC secrete and assemble several pili via the T7SS. The operon encoding ECP (*E. coli* common pilus) is highly prevalent across *E. coli*, including EPEC (Saldana *et al.* 2009). These pili participate in bacterial adhesion to human epithelial cells, intestinal colonisation and biofilm development through interorganelle binding via the EcpA pilin (Rendon *et al.* 2007). In *E. coli* O157:H7, these pili were further demonstrated to target arabinosyl residues in plant cell walls to mediate adhesion to vegetables (Rossez *et al.* 2014). *Escherichia coli* O157:H7 also contains two operons encoding Lpf (long polar fimbriae), namely *lpf1* with six genes (*lpfABCC'DE*) forming LpfA pilus, and *lpf2* containing five genes (*lpfABCDD'*) forming LpfD2 pilus (Doughty *et al.* 2002). Regarding the Lpf major subunits, five different genetic variants have been described for LpfA1 against three for LpfA2

Downloaded from http://femsle.oxfordjournals.org/ by guest on January 17, 2017

Annexe n°1

(Farfan and Torres 2012). Lpf interact specifically with the ECM, namely fibronectin, laminin or collagen IV components. They contribute to the adhesion of EHEC to the intestinal epithelium and can even influence bacterial tropism towards different intestinal tissues (Torres *et al.* 2008; Farfan *et al.* 2011). In EHEC O104:H4, LpfA contributes to bacterial adhesion to polarised intestinal epithelial cells, biofilm formation as well as caecum and large intestine colonisation (Ross *et al.* 2015). In EHEC, ELF (*E. coli* YcbQ laminin-binding fimbriae) were found to bind laminin as well as to adhere to human epithelial cells (Samadder *et al.* 2009). SFP (sorbitol-fermenting fimbriae protein) is not responsible for haemagglutination activity in *E. coli* O157:H7 but also acts as adhesin via the SfpG pilin, involved in adherence to epithelial cells (Musken *et al.* 2008). F9 pili are encoded in O-island 61 and are involved in binding of EHEC to fibronectin and to bovine intestinal epithelial cells (Low *et al.* 2006a). While encoded in numerous STEC (Li, Koch and Cebula 1997), expression of type 1 pili was only reported in a subset of STEC O26 and O118 strains (Enami *et al.* 1999), where they not only participate in the colonisation of the bovine gut by mediating adhesion to intestinal epithelial cells but also enable adhesion to abiotic surfaces (Galfi *et al.* 1998; Cookson, Cooley and Woodward 2002). Actually, the expression of type 1 pili would be subjected to phase variation of *fimA* encoding the major prepilin but the triggering environmental factors remain to be determined (Li, Koch and Cebula 1997; Roe *et al.* 2001).

Like EAEC, EHEC O104:H4 lacks the LEE but encodes typical AAF (aggregative adherence fimbria) (Brzuszkiewicz *et al.* 2011; Mellmann *et al.* 2011). Just like all EPEC are AEEC, all EAEC are aggregative-adherence encoding *E. coli*, i.e. *aaf*$^+$, from which atypical EHEC O104:H4 would have emerged (Brzuszkiewicz *et al.* 2011; Mellmann *et al.* 2011; Tietze *et al.* 2015). In EAEC, AAF are associated with a strong ability to form biofilms on biotic and abiotic surfaces as well as haemagglutination with human erythrocytes. Four variants of AAF have been identified, namely AAF/I to AAF/IV, encoded on virulence plasmids of the pAA family (Boisen *et al.* 2008). AAF/II was further demonstrated to bind fibronectin (Farfan, Inman and Nataro 2008). AAF act in concert with dispersin in the colonisation of the intestinal mucosa, which would result in a highly virulent combination in EHEC O104:H4 when Stx is also present (Boisen *et al.* 2015). While the expression of those different pili is subjected to regulation by various environmental factors (Gonyar and Kendall 2014), their global expression and respective contribution to the colonisation process along the food chain, from the environment, animal reservoirs, food matrices to human infection are far from being understood. Besides, several operons encoding putative T7SS remain to be characterised in STEC.

Type VIII secretion system: curli

The type VIII secretion system (T8SS) corresponds to the extracellular nucleation-precipitation (ENP) pathway involved in the secretion and assembly of peculiar pili, called curli, which are functional amyloid fibres (Desvaux *et al.* 2009; Goyal *et al.* 2014). Curli fibres are extremely adherent and involved in cell aggregation, bacterial adhesion and, ultimately, biofilm development (Hammar, Bian and Normark 1996; Fronzes, Christie and Waksman 2009). *Escherichia coli* O157:H7 expressing curli is more virulent and exhibits higher adherence abilities to eukaryotic cells than non-curliated strains (Uhlich, Keen and Elder 2002). The expression of curli seems to compensate the absence of Lpf (Lloyd *et al.* 2012). It also appeared that rather than affecting initial attachment, curli enhanced sessile development (Ryu and Beuchat 2005).

Non-classical secretion and other protein trafficking mechanisms: glycolytic enzymes

Besides secretion via known protein secretion systems, some proteins devoid of any identifiable signal sequence or cognate secretion system can be identified outside the cytoplasm. In some conditions, it can be reasonably thought those proteins could be secreted by alternative but unknown protein secretion machineries, the so-called non-classical secretion. Besides piggybacking, where a protein like a chaperone associates with a secreted protein and is collaterally transported, the true secretion mechanism is sometimes uncovered years after, such as for the secreted proteins via the most recently uncovered T6SS or T9SS (Chagnot *et al.* 2013). In parallel, mechanisms unrelated to secretion also participate in protein trafficking, such as (i) cell lysis, which can be controlled (allolysis), results from bacteriophage infection or follows the entry into a lytic cycle (resulting in the extracellular release of Stx), or (ii) membrane budding resulting in outer membrane vesicles released from the bacterial cell surface.

Among unexpected extracytoplasmic proteins, glycolytic enzymes are frequently uncovered (Henderson and Martin 2011). Because of their second hidden function revealing when localised outside the cytoplasm, these proteins are qualified of moonlighting. In EHEC and EPEC, the GAPDH (glyceraldehyde 3-phosphate dehydrogenase) present at cell surface binds plasminogen and fibrinogen and remains associated with intestinal epithelial cells upon infection (Egea *et al.* 2007). Depending on the growth conditions, secretion of GAPDH could occur via piggybacking through the T3aSS (Aguilera *et al.* 2012). The exact contribution of GAPDH and other putative moonlighting glycolytic enzymes to colonisation in STEC would deserve further investigations.

CONCLUSION

While the Stx and the LEE are regarded as major virulence factors in classical EHEC, it is clear that the virulence levels of *Escherichia coli* strains encoding Stx are highly variable and that the LEE is not systematically present in all pathogenic strains. Describing the complement of the secretome involved in surface colonisation provides a rational view of the diversity of these effectors respective to their cognate secretion systems. It also highlights that some secretion systems and cognate effectors are clearly demonstrated as involved in the colonisation process of STEC (e.g. Lpf of the T7SS), whereas some other remain putative (e.g. like Hcp of the T6SS) or characterised only in laboratory strain (*E. coli* K12) background (e.g. Eha of the T5aSS) and would thus require further investigations to determine their possible or exact contribution in natural isolates. Considering the wealth of colonisation effectors secreted by various protein trafficking mechanisms, the different possible combinations could modulate the virulence of the STEC strains. Besides, the regulation of the secretome in very diversified ecosystems, both in terms of biotope and biocenose, all along the food chain most certainly contributes to differential protein expression resulting in phenotypic variation. Indeed, these factors can not only contribute and modulate the virulence of EHEC but also have a role in bacterial colonisation along the food chain from asymptomatic animal carriage to contamination of food matrices. Arguably,

Downloaded from http://femsle.oxfordjournals.org/ by guest on January 17, 2017

some colonisation factors can further physically shield others as observed when pili prevent the interaction of autotransporter adhesins with their substratum (Hasman, Chakraborty and Klemm 1999; Ulett, Webb and Schembri 2006; Wells *et al.* 2008). Furthermore, disease incidence and severity of the infection are likely to be functions of a complex interplay between host, pathogen and ecological factors and may not necessarily be explained on the basis of the pathogen's attributes alone. One promising future research direction would be aiming at and considering both (i) the ecophysiology of the STEC along the food chain from their natural environment, animal reservoir, food-processing environment, food matrices to human ingestion (and not just the infection) and (ii) the global expression of virulence and colonisation factors (rather than focusing on one particular system, such as the T3aSS) considering the different regulation levels (pre-transcriptional, transcriptional, post-transcriptional, pre-translational, translational, post-translational, translocational and post-translocational). Such an approach is most certainly the promise to open new perspectives for innovative preventive strategies against zoonotic foodborne pathogens, especially pathogenic STEC, such as EHEC and some EIEC, or related EPEC and EAEC.

ACKNOWLEDGEMENTS

Authors member of the COST (european COoperation in Science and Technology) BacFoodNet (A European network for mitigating bacterial colonisation and persistence on foods and food processing environments; http://www.bacfoodnet.org/; n°FA1202) acknowledge this action for facilitating the collaborative networking that assisted with this study.

FUNDING

This work was supported in part by the Institut National de la Recherche Agronomique (INRA) and EU FP7 (seventh framework programme) Marie Curie Actions ITN (Initiale Training Network) EID (European Industrial Doctorate) DISCo (A multidisciplinary Doctoral Industrial School on novel preventive strategies against *E. Coli* infections; http://www.discoproject.eu/; n° FP7-PEOPLE-607611) project. Ricardo MONTEIRO and Maricarmen ROJAS-LOPEZ are Marie Curie PhD Research Fellows granted by ITN EID DISCo.

Conflict of interest. None declared.

REFERENCES

Adu-Bobie J, Frankel G, Bain C *et al.* Detection of intimins alpha, beta, gamma, and delta, four intimin derivatives expressed by attaching and effacing microbial pathogens. *J Clin Microbiol* 1998;**36**:662–8.

Aguilera L, Ferreira E, Gimenez R *et al.* Secretion of the housekeeping protein glyceraldehyde-3-phosphate dehydrogenase by the LEE-encoded type III secretion system in enteropathogenic *Escherichia coli*. *Int J Biochem Cell B* 2012;**44**:955–62.

Altegoer F, Bange G. Undiscovered regions on the molecular landscape of flagellar assembly. *Curr Opin Microbiol* 2015;**28**:98–105.

Alvarez-Martinez CE, Christie PJ. Biological diversity of prokaryotic Type IV secretion systems. *Microbiol Mol Biol R* 2009;**73**:775–808.

Amisano G, Fornasero S, Migliaretti G *et al.* Diarrheagenic *Escherichia coli* in acute gastroenteritis in infants in North-West Italy. *New Microbiol* 2011;**34**:45–51.

Antelmann H, Tjalsma H, Voigt B *et al.* A proteomic view on genome-based signal peptide predictions. *Genome Res* 2001;**11**:1484–502.

Aschtgen MS, Bernard CS, De Bentzmann S *et al.* SciN is an outer membrane lipoprotein required for type VI secretion in enteroaggregative *Escherichia coli*. *J Bacteriol* 2008;**190**:7523–31.

Aschtgen MS, Gavioli M, Dessen A *et al.* The SciZ protein anchors the enteroaggregative *Escherichia coli* Type VI secretion system to the cell wall. *Mol Microbiol* 2010;**75**:886–99.

Ayers M, Howell PL, Burrows LL. Architecture of the type II secretion and type IV pilus machineries. *Future Microbiol* 2010;**5**:1203–18.

Benz I, Schmidt MA. Cloning and expression of an adhesin (AIDA-I) involved in diffuse adherence of enteropathogenic *Escherichia coli*. *Infect Immun* 1989;**57**:1506–11.

Berger CN, Sodha SV, Shaw RK *et al.* Fresh fruit and vegetables as vehicles for the transmission of human pathogens. *Environ Microbiol* 2010;**12**:2385–97.

Boisen N, Struve C, Scheutz F *et al.* New adhesin of enteroaggregative *Escherichia coli* related to the Afa/Dr/AAF family. *Infect Immun* 2008;**76**:3281–92.

Boisen N, Melton-Celsa AR, Scheutz F *et al.* Shiga toxin 2a and Enteroaggregative *Escherichia coli*—a deadly combination. *Gut Microbes* 2015;**6**:272–8.

Brzuszkiewicz E, Thurmer A, Schuldes J *et al.* Genome sequence analyses of two isolates from the recent *Escherichia coli* outbreak in Germany reveal the emergence of a new pathotype: Entero-Aggregative-Haemorrhagic *Escherichia coli* (EAHEC). *Arch Microbiol* 2011;**193**:883–91.

Busch A, Waksman G. Chaperone-usher pathways: diversity and pilus assembly mechanism. *Philos T Roy Soc B* 2012;**367**:1112–22.

Calderwood S, Acheson DWK, Keusch GT *et al.* Proposed new nomenclature for SLT (VT) family. *ASM News* 1996;**62**:118–9.

Caprioli A, Morabito S, Brugere H *et al.* Enterohaemorrhagic *Escherichia coli*: emerging issues on virulence and modes of transmission. *Vet Res* 2005;**36**:289–311.

Chagnot C, Zorgani MA, Astruc T *et al.* Proteinaceous determinants of surface colonization in bacteria: bacterial adhesion and biofilm formation from a protein secretion perspective. *Front Microbiol* 2013;**4**:303.

Choi PS, Dawson AJ, Bernstein HD. Characterization of a novel two-partner secretion system in *Escherichia coli* O157:H7. *J Bacteriol* 2007;**189**:3452–61.

Christie PJ, Vogel JP. Bacterial type IV secretion: conjugation systems adapted to deliver effector molecules to host cells. *Trends Microbiol* 2000;**8**:354–60.

Christie PJ, Whitaker N, Gonzalez-Rivera C. Mechanism and structure of the bacterial type IV secretion systems. *Biochim Biophys Acta* 2014;**1843**:1578–91.

Cohen MB, Nataro JP, Bernstein DI *et al.* Prevalence of diarrheagenic *Escherichia coli* in acute childhood enteritis: a prospective controlled study. *J Pediatr* 2005;**146**:54–61.

Cookson AL, Cooley WA, Woodward MJ. The role of type 1 and curli fimbriae of Shiga toxin-producing *Escherichia coli* in adherence to abiotic surfaces. *Int J Med Microbiol* 2002;**292**:195–205.

Cornelis GR. The Type III secretion injectisome. *Nat Rev Microbiol* 2006;**4**:811–25.

Craig L, Pique ME, Tainer JA. Type IV pilus structure and bacterial pathogenicity. *Nat Rev Microbiol* 2004;**2**:363–78.

Downloaded from http://femsle.oxfordjournals.org/ by guest on January 17, 2017

Annexe n°1

Creuzburg K, Heeren S, Lis CM *et al*. Genetic background and mobility of variants of the gene *nleA* in attaching and effacing *Escherichia coli*. *Appl Environ Microb* 2011a;**77**:8705–13.

Creuzburg K, Middendorf B, Mellmann A *et al*. Evolutionary analysis and distribution of type III effector genes in pathogenic *Escherichia coli* from human, animal and food sources. *Environ Microbiol* 2011b;**13**:439–52.

Croxen MA, Law RJ, Scholz R *et al*. Recent advances in understanding enteric pathogenic *Escherichia coli*. *Clin Microbiol Rev* 2013;**26**:822–80.

Desvaux M, Hebraud M, Henderson IR *et al*. Type III secretion: what's in a name? *Trends Microbiol* 2006;**14**:157–60.

Desvaux M, Hébraud M, Talon R *et al*. Secretion and subcellular localizations of bacterial proteins: a semantic awareness issue. *Trends Microbiol* 2009;**17**:139–45.

Desvaux M, Parham NJ, Henderson IR. Type V protein secretion: simplicity gone awry? *Curr Issues Mol Biol* 2004;**6**:111–24.

Diepold A, Armitage JP. Type III secretion systems: the bacterial flagellum and the injectisome. *Philos T Roy Soc B* 2015;**370**:20150020.

DiLuzio WR, Turner L, Mayer M *et al*. *Escherichia coli* swim on the right-hand side. *Nature* 2005;**435**:1271–4.

Dobrindt U. (Patho-)Genomics of Escherichia coli. *Int J Med Microbiol* 2005;**295**:357–71.

Dobrindt U, Chowdary MG, Krumbholz G *et al*. Genome dynamics and its impact on evolution of *Escherichia coli*. *Med Microbiol Immunol* 2010;**199**:145–54.

Donnenberg MS, Tacket CO, James SP *et al*. Role of the eaeA gene in experimental enteropathogenic *Escherichia coli* infection. *J Clin Invest* 1993;**92**:1412–7.

Doughty S, Sloan J, Bennett-Wood V *et al*. Identification of a novel fimbrial gene cluster related to long polar fimbriae in locus of enterocyte effacement-negative strains of enterohemorrhagic *Escherichia coli*. *Infect Immun* 2002;**70**:6761–9.

Dudley EG, Abe C, Ghigo JM *et al*. An IncI1 plasmid contributes to the adherence of the atypical enteroaggregative *Escherichia coli* strain C1096 to cultured cells and abiotic surfaces. *Infect Immun* 2006a;**74**:2102–14.

Dudley EG, Thomson NR, Parkhill J *et al*. Proteomic and microarray characterization of the AggR regulon identifies a pheU pathogenicity island in enteroaggregative *Escherichia coli*. *Mol Microbiol* 2006b;**61**:1267–82.

Durand E, Alphonse S, Brochier-Armanet C *et al*. The assembly mode of the pseudopilus: a hallmark to distinguish a novel secretion system subtype. *J Biol Chem* 2011;**286**:24407–16.

Dziva F, Mahajan A, Cameron P *et al*. EspP, a Type V-secreted serine protease of enterohaemorrhagic *Escherichia coli* O157:H7, influences intestinal colonization of calves and adherence to bovine primary intestinal epithelial cells. *FEMS Microbiol Lett* 2007;**271**:258–64.

Easton DM, Allsopp LP, Phan MD *et al*. The intimin-like protein FdeC is regulated by H-NS and temperature in enterohemorrhagic *Escherichia coli*. *Appl Environ Microb* 2014;**80**:7337–47.

Easton DM, Totsika M, Allsopp LP *et al*. Characterization of EhaJ, a new autotransporter protein from enterohemorrhagic and enteropathogenic *Escherichia coli*. *Front Microbiol* 2011;**2**:120.

Egea L, Aguilera L, Gimenez R *et al*. Role of secreted glyceraldehyde-3-phosphate dehydrogenase in the infection mechanism of enterohemorrhagic and enteropathogenic *Escherichia coli*: interaction of the extracellular enzyme with human plasminogen and fibrinogen. *Int J Biochem Cell B* 2007;**39**:1190–203.

El Tahir Y, Skurnik M. YadA, the multifaceted *Yersinia* adhesin. *Int J Med Microbiol* 2001;**291**:209–18.

Enami M, Nakasone N, Honma Y *et al*. Expression of type I pili is abolished in verotoxin-producing *Escherichia coli* O157. *FEMS Microbiol Lett* 1999;**179**:467–72.

Farfan MJ, Cantero L, Vidal R *et al*. Long polar fimbriae of enterohemorrhagic *Escherichia coli* O157:H7 bind to extracellular matrix proteins. *Infect Immun* 2011;**79**:3744–50.

Farfan MJ, Inman KG, Nataro JP. The major pilin subunit of the AAF/II fimbriae from enteroaggregative *Escherichia coli* mediates binding to extracellular matrix proteins. *Infect Immun* 2008;**76**:4378–84.

Farfan MJ, Torres AG. Molecular mechanisms that mediate colonization of Shiga toxin-producing *Escherichia coli* strains. *Infect Immun* 2012;**80**:903–13.

Finlay BB, Falkow S. Common themes in microbial pathogenicity revisited. *Microbiol Mol Biol Rev* 1997;**61**:136–69.

Fratamico PM, Yan X, Caprioli A *et al*. The complete DNA sequence and analysis of the virulence plasmid and of five additional plasmids carried by Shiga toxin-producing *Escherichia coli* O26:H11 strain H30. *Int J Med Microbiol* 2011;**301**:192–203.

Frohlicher E, Krause G, Zweifel C *et al*. Characterization of attaching and effacing *Escherichia coli* (AEEC) isolated from pigs and sheep. *BMC Microbiol* 2008;**8**:144.

Fronzes R, Christie PJ, Waksman G. The structural biology of Type IV secretion systems. *Nat Rev Microbiol* 2009;**7**:703–14.

Galan JE, Wolf-Watz H. Protein delivery into eukaryotic cells by type III secretion machines. *Nature* 2006;**444**:567–73.

Galfi P, Neogrady S, Semjen G *et al*. Attachment of different *Escherichia coli* strains to cultured rumen epithelial cells. *Vet Microbiol* 1998;**61**:191–7.

Gerlach RG, Hensel M. Protein secretion systems and adhesins: the molecular armory of Gram-negative pathogens. *Int J Med Microbiol* 2007;**297**:401–15.

Ghigo JM. Natural conjugative plasmids induce bacterial biofilm development. *Nature* 2001;**412**:442–5.

Giron JA, Donnenberg MS, Martin WC *et al*. Distribution of the bundle-forming pilus structural gene (bfpA) among enteropathogenic *Escherichia coli*. *J Infect Dis* 1993;**168**:1037–41.

Giron JA, Ho AS, Schoolnik GK. An inducible bundle-forming pilus of enteropathogenic *Escherichia coli*. *Science* 1991;**254**:710–3.

Gonyar LA, Kendall MM. Ethanolamine and choline promote expression of putative and characterized fimbriae in enterohemorrhagic *Escherichia coli* O157:H7. *Infect Immun* 2014;**82**:193–201.

Goyal P, Krasteva PV, Van Gerven N *et al*. Structural and mechanistic insights into the bacterial amyloid secretion channel CsgG. *Nature* 2014;**516**:250–3.

Haiko J, Westerlund-Wikstrom B. The role of the bacterial flagellum in adhesion and virulence. *Biology* 2013;**2**:1242–67.

Hammar M, Bian Z, Normark S. Nucleator-dependent intercellular assembly of adhesive curli organelles in *Escherichia coli*. *P Natl Acad Sci USA* 1996;**93**:6562–6.

Hasman H, Chakraborty T, Klemm P. Antigen-43-mediated autoaggregation of *Escherichia coli* is blocked by fimbriation. *J Bacteriol* 1999;**181**:4834–41.

Hauser E, Mellmann A, Semmler T *et al*. Phylogenetic and molecular analysis of food-borne shiga toxin-producing *Escherichia coli*. *Appl Environ Microb* 2013;**79**:2731–40.

Hayashi T, Makino K, Ohnishi M *et al*. Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genomic comparison with a laboratory strain K-12. *DNA Res* 2001;**8**:11–22.

Downloaded from http://femsle.oxfordjournals.org/ by guest on January 17, 2017

Annexe n°1

Henderson B, Martin A. Bacterial virulence in the moonlight: multitasking bacterial moonlighting proteins are virulence determinants in infectious disease. *Infect Immun* 2011;**79**:3476–91.

Henderson IR, Navarro-Garcia F, Desvaux M *et al*. Type V protein secretion pathway: the autotransporter story. *Microbiol Mol Biol R* 2004;**68**:692–744.

Herold S, Paton JC, Paton AW. Sab, a novel autotransporter of locus of enterocyte effacement-negative shiga-toxigenic *Escherichia coli* O113:H21, contributes to adherence and biofilm formation. *Infect Immun* 2009a;**77**:3234–43.

Herold S, Paton JC, Srimanote P *et al*. Differential effects of short-chain fatty acids and iron on expression of iha in Shiga-toxigenic *Escherichia coli*. *Microbiology* 2009b;**155**:3554–63.

Hooper LV, Gordon JI. Commensal host-bacterial relationships in the gut. *Science* 2001;**292**:1115–8.

Kaper JB, Nataro JP, Mobley HL. Pathogenic *Escherichia coli*. *Nat Rev Microbiol* 2004;**2**:123–40.

Karch H, Bielaszewska M, Bitzan M *et al*. Epidemiology and diagnosis of Shiga toxin-producing *Escherichia coli* infections. *Diagn Micr Infec Dis* 1999;**34**:229–43.

Karch H, Tarr PI, Bielaszewska M. Enterohaemorrhagic *Escherichia coli* in human medicine. *Int J Med Microbiol* 2005;**295**:405–18.

Karmali MA, Gannon V, Sargeant JM. Verocytotoxin-producing Escherichia coli (VTEC). *Vet Microbiol* 2010;**140**:360–70.

Karmali MA, Mascarenhas M, Shen S *et al*. Association of genomic O island 122 of *Escherichia coli* EDL 933 with verocytotoxin-producing *Escherichia coli* seropathotypes that are linked to epidemic and/or serious disease. *J Clin Microbiol* 2003;**41**:4930–40.

Kelly JK, Pai CH, Jadusingh IH *et al*. The histopathology of rectosigmoid biopsies from adults with bloody diarrhea due to verotoxin-producing *Escherichia coli*. *Am J Clin Pathol* 1987;**88**:78–82.

Klemm P, Vejborg RM, Sherlock O. Self-associating autotransporters, SAATs: functional and structural similarities. *Int J Med Microbiol* 2006;**296**:187–95.

Knowles TJ, Scott-Tucker A, Overduin M *et al*. Membrane protein architects: the role of the BAM complex in outer membrane protein assembly. *Nat Rev Microbiol* 2009;**7**:206–14.

Korea CG, Badouraly R, Prevost MC *et al*. *Escherichia coli* K-12 possesses multiple cryptic but functional chaperone-usher fimbriae with distinct surface specificities. *Environ Microbiol* 2010;**12**:1957–77.

Kyle JL, Cummings CA, Parker CT *et al*. *Escherichia coli* serotype O55:H7 diversity supports parallel acquisition of bacteriophage at Shiga toxin phage insertion sites during evolution of the O157:H7 lineage. *J Bacteriol* 2012;**194**:1885–96.

Lan R, Alles MC, Donohoe K *et al*. Molecular evolutionary relationships of enteroinvasive *Escherichia coli* and *Shigella* spp. *Infect Immun* 2004;**72**:5080–8.

Lawley TD, Klimke WA, Gubbins MJ *et al*. F factor conjugation is a true Type IV secretion system. *FEMS Microbiol Lett* 2003;**224**:1–15.

Leiman PG, Basler M, Ramagopal UA *et al*. Type VI secretion apparatus and phage tail-associated protein complexes share a common evolutionary origin. *P Natl Acad Sci USA* 2009;**106**:4154–9.

Leo JC, Oberhettinger P, Schutz M *et al*. The inverse autotransporter family: intimin, invasin and related proteins. *Int J Med Microbiol* 2015;**305**:276–82.

Lewis SB, Cook V, Tighe R *et al*. Enterohemorrhagic *Escherichia coli* colonization of human colonic epithelium *in vitro* and *ex vivo*. *Infect Immun* 2015;**83**:942–9.

Li B, Koch WH, Cebula TA. Detection and characterization of the fimA gene of *Escherichia coli* O157:H7. *Mol Cell Probes* 1997;**11**:397–406.

Lim JY, Yoon J, Hovde CJ. A brief overview of *Escherichia coli* O157:H7 and its plasmid O157. *J Microbiol Biotechn* 2010;**20**:5–14.

Linke D, Riess T, Autenrieth IB *et al*. Trimeric autotransporter adhesins: variable structure, common function. *Trends Microbiol* 2006;**14**:264–70.

Liu H, Magoun L, Leong JM. beta1-chain integrins are not essential for intimin-mediated host cell attachment and enteropathogenic *Escherichia coli*-induced actin condensation. *Infect Immun* 1999;**67**:2045–9.

Liu H, Magoun L, Luperchio S *et al*. The Tir-binding region of enterohaemorrhagic *Escherichia coli* intimin is sufficient to trigger actin condensation after bacterial-induced host cell signalling. *Mol Microbiol* 1999;**34**:67–81.

Lloyd AL, Henderson TA, Vigil PD *et al*. Genomic islands of uropathogenic *Escherichia coli* contribute to virulence. *J Bacteriol* 2009;**191**:3469–81.

Lloyd SJ, Ritchie JM, Rojas-Lopez M *et al*. A double, long polar fimbria mutant of *Escherichia coli* O157:H7 expresses curli and exhibits reduced *in vivo* colonization. *Infect Immun* 2012;**80**:914–20.

Low AS, Dziva F, Torres AG *et al*. Cloning, expression, and characterization of fimbrial operon F9 from enterohemorrhagic *Escherichia coli* O157:H7. *Infect Immun* 2006a;**74**:2233–44.

Low AS, Holden N, Rosser T *et al*. Analysis of fimbrial gene clusters and their expression in enterohaemorrhagic *Escherichia coli* O157:H7. *Environ Microbiol* 2006b;**8**:1033–47.

Lu Y, Iyoda S, Satou H *et al*. A new immunoglobulin-binding protein, EibG, is responsible for the chain-like adhesion phenotype of locus of enterocyte effacement-negative, shiga toxin-producing *Escherichia coli*. *Infect Immun* 2006;**74**:5747–55.

Macnab RM. Type III flagellar protein export and flagellar assembly. *Biochim Biophys Acta* 2004;**1694**:207–17.

McDaniel TK, Jarvis KG, Donnenberg MS *et al*. A genetic locus of enterocyte effacement conserved among diverse enterobacterial pathogens. *P Natl Acad Sci USA* 1995;**92**:1664–8.

McDonough MA, Butterton JR. Spontaneous tandem amplification and deletion of the shiga toxin operon in *Shigella dysenteriae* 1. *Mol Microbiol* 1999;**34**:1058–69.

McNeilly TN, Naylor SW, Mahajan A *et al*. *Escherichia coli* O157:H7 colonization in cattle following systemic and mucosal immunization with purified H7 flagellin. *Infect Immun* 2008;**76**:2594–602.

Mahajan A, Currie CG, Mackie S *et al*. An investigation of the expression and adhesin function of H7 flagella in the interaction of *Escherichia coli* O157 : H7 with bovine intestinal epithelium. *Cell Microbiol* 2009;**11**:121–37.

Mallick EM, Brady MJ, Luperchio SA *et al*. Allele- and tir-independent functions of intimin in diverse animal infection models. *Front Microbiol* 2012;**3**:11.

May T, Ito A, Okabe S. Characterization and global gene expression of F- phenocopies during *Escherichia coli* biofilm formation. *Mol Genet Genomics* 2010;**284**:333–42.

May T, Okabe S. *Escherichia coli* harboring a natural IncF conjugative F plasmid develops complex mature biofilms by stimulating synthesis of colanic acid and Curli. *J Bacteriol* 2008;**190**:7479–90.

Mellmann A, Harmsen D, Cummings CA *et al*. Prospective genomic characterization of the German enterohemorrhagic *Escherichia coli* O104:H4 outbreak by rapid next generation sequencing technology. *PLoS One* 2011;**6**:e22751.

Downloaded from http://femsle.oxfordjournals.org/ by guest on January 17, 2017

Annexe n°1

Melton-Celsa AR, O'Brien AD. Shiga toxins of Shigella dysenteriae and Escherichia coli. In: Aktories K, Just I (eds). *Bacterial Protein Toxins*, Vol. **145**. Berlin, Heidelberg: Springer, 2000, 385–406.

Merkel V, Ohder B, Bielaszewska M *et al*. Distribution and phylogeny of immunoglobulin-binding protein G in Shiga toxin-producing *Escherichia coli* and its association with adherence phenotypes. *Infect Immun* 2010;**78**:3625–36.

Monteiro BT, Campos LC, Sircili MP *et al*. The dispersin-encoding gene (aap) is not restricted to enteroaggregative *Escherichia coli*. *Diagn Micro Infec Dis* 2009;**65**:81–4.

Mora A, Blanco M, Blanco JE *et al*. Serotypes, virulence genes and intimin types of Shiga toxin (verocytotoxin)-producing *Escherichia coli* isolates from minced beef in Lugo (Spain) from 1995 through 2003. *BMC Microbiol* 2007;**7**:13.

Moriel DG, Bertoldi I, Spagnuolo A *et al*. Identification of protective and broadly conserved vaccine antigens from the genome of extraintestinal pathogenic *Escherichia coli*. *P Natl Acad Sci USA* 2010;**107**:9072–7.

Muniesa M, Hammerl JA, Hertwig S *et al*. Shiga toxin-producing *Escherichia coli* O104:H4: a new challenge for microbiology. *Appl Environ Microb* 2012;**78**:4065–73.

Musken A, Bielaszewska M, Greune L *et al*. Anaerobic conditions promote expression of Sfp fimbriae and adherence of sorbitol-fermenting enterohemorrhagic *Escherichia coli* O157:NM to human intestinal epithelial cells. *Appl Environ Microb* 2008;**74**:1087–93.

Nataro JP, Kaper JB. Diarrheagenic *Escherichia coli*. *Clin Microbiol Rev* 1998;**11**:142–201.

Nataro JP, Mai V, Johnson J *et al*. Diarrheagenic *Escherichia coli* infection in Baltimore, Maryland, and New Haven, Connecticut. *Clin Infect Dis* 2006;**43**:402–7.

Nesta B, Spraggon G, Alteri C *et al*. FdeC, a novel broadly conserved *Escherichia coli* adhesin eliciting protection against urinary tract infections. *MBio* 2012;**3**:e00010-12.

Nishi J, Sheikh J, Mizuguchi K *et al*. The export of coat protein from enteroaggregative *Escherichia coli* by a specific ATP-binding cassette transporter system. *J Biol Chem* 2003;**278**:45680–9.

Oberhettinger P, Leo JC, Linke D *et al*. The inverse autotransporter intimin exports its passenger domain via a hairpin intermediate. *J Biol Chem* 2015;**290**:1837–49.

Oberhettinger P, Schutz M, Leo JC *et al*. Intimin and invasin export their C-terminus to the bacterial cell surface using an inverse mechanism compared to classical autotransport. *PLoS One* 2012;**7**:e47069.

Ostroff SM, Tarr PI, Neill MA *et al*. Toxin genotypes and plasmid profiles as determinants of systemic sequelae in *Escherichia coli* O157:H7 infections. *J Infect Dis* 1989; **160**:994–8.

Oswald E, Schmidt H, Morabito S *et al*. Typing of intimin genes in human and animal enterohemorrhagic and enteropathogenic *Escherichia coli*: characterization of a new intimin variant. *Infect Immun* 2000;**68**:64–71.

Paros M, Tarr PI, Kim H *et al*. A comparison of human and bovine *Escherichia coli* O157:H7 isolates by toxin genotype, plasmid profile, and bacteriophage lambda-restriction fragment length polymorphism profile. *J Infect Dis* 1993;**168**: 1300–3.

Paton AW, Srimanote P, Woodrow MC *et al*. Characterization of Saa, a novel autoagglutinating adhesin produced by *locus* of enterocyte effacement-negative Shiga-toxigenic *Escherichia coli* strains that are virulent for humans. *Infect Immun* 2001;**69**:6999–7009.

Perna NT, Plunkett G, Burland V *et al*. Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. *Nature* 2001;**409**:529–33.

Puttamreddy S, Cornick NA, Minion FC. Genome-wide transposon mutagenesis reveals a role for pO157 genes in biofilm development in *Escherichia coli* O157:H7 EDL933. *Infect Immun* 2010;**78**:2377–84.

Qadri F, Svennerholm AM, Faruque AS *et al*. Enterotoxigenic *Escherichia coli* in developing countries: epidemiology, microbiology, clinical features, treatment, and prevention. *Clin Microbiol Rev* 2005;**18**:465–83.

Rendon MA, Saldana Z, Erdem AL *et al*. Commensal and pathogenic *Escherichia coli* use a common pilus adherence factor for epithelial cell colonization. *P Natl Acad Sci USA* 2007;**104**:10637–42.

Roe AJ, Currie C, Smith DG *et al*. Analysis of type 1 fimbriae expression in verotoxigenic *Escherichia coli*: a comparison between serotypes O157 and O26. *Microbiology* 2001;**147**:145–52.

Ross BN, Rojas-Lopez M, Cieza RJ *et al*. The role of long polar fimbriae in *Escherichia coli* O104:H4 adhesion and colonization. *PLoS One* 2015;**10**:e0141845.

Rossez Y, Holmes A, Lodberg-Pedersen H *et al*. *Escherichia coli* common pilus (ECP) targets arabinosyl residues in plant cell walls to mediate adhesion to fresh produce plants. *J Biol Chem* 2014;**289**:34349–65.

Ryu JH, Beuchat LR. Biofilm formation by *Escherichia coli* O157:H7 on stainless steel: effect of exopolysaccharide and Curli production on its resistance to chlorine. *Appl Environ Microb* 2005;**71**:247–54.

Saldana Z, Erdem AL, Schuller S *et al*. The *Escherichia coli* common pilus and the bundle-forming pilus act in concert during the formation of localized adherence by enteropathogenic *E. coli*. *J Bacteriol* 2009;**191**:3451–61.

Samadder P, Xicohtencatl-Cortes J, Saldana Z *et al*. The *Escherichia coli ycbQRST* operon encodes fimbriae with laminin-binding and epithelial cell adherence properties in Shiga-toxigenic *E. coli* O157:H7. *Environ Microbiol* 2009;**11**:1815–26.

Scaletsky IC, Silva ML, Trabulsi LR. Distinctive patterns of adherence of enteropathogenic *Escherichia coli* to HeLa cells. *Infect Immun* 1984;**45**:534–6.

Scheutz F, Teel LD, Beutin L *et al*. Multicenter evaluation of a sequence-based protocol for subtyping Shiga toxins and standardizing Stx nomenclature. *J Clin Microbiol* 2012;**50**:2951–63.

Schmidt MA. LEEways: tales of EPEC, ATEC and EHEC. *Cell Microbiol* 2010;**12**:1544–52.

Schroeder GN, Hilbi H. Molecular pathogenesis of *Shigella* spp.: controlling host cell signaling, invasion, and death by type III secretion. *Clin Microbiol Rev* 2008;**21**:134–56.

Shaw RK, Berger CN, Feys B *et al*. Enterohemorrhagic *Escherichia coli* exploits EspA filaments for attachment to salad leaves. *Appl Environ Microb* 2008;**74**:2908–14.

Sheikh J, Czeczulin JR, Harrington S *et al*. A novel dispersin protein in enteroaggregative *Escherichia coli*. *J Clin Invest* 2002;**110**:1329–37.

Sherlock O, Schembri MA, Reisner A *et al*. Novel roles for the AIDA adhesin from diarrheagenic *Escherichia coli*: cell aggregation and biofilm formation. *J Bacteriol* 2004;**186**:8058–65.

Shrivastava S, Mande SS. Identification and functional characterization of gene components of Type VI Secretion system in bacterial genomes. *PLoS One* 2008;**3**:e2955.

Stone KD, Zhang HZ, Carlson LK *et al*. A cluster of fourteen genes from enteropathogenic *Escherichia coli* is sufficient for the biogenesis of a type IV pilus. *Mol Microbiol* 1996;**20**:325–37.

Downloaded from http://femsle.oxfordjournals.org/ by guest on January 17, 2017

Annexe n°1

Szczesny P, Lupas A. Domain annotation of trimeric autotransporter adhesins–daTAA. *Bioinformatics* 2008;**24**: 1251–6.

Tarr PI, Bilge SS, Vary JC *et al*. Iha: a novel *Escherichia coli* O157:H7 adherence-conferring molecule encoded on a recently acquired chromosomal island of conserved structure. *Infect Immun* 2000;**68**:1400–7.

Tarr PI, Gordon CA, Chandler WL. Shiga-toxin-producing *Escherichia coli* and haemolytic uraemic syndrome. *Lancet* 2005;**365**:1073–86.

Tenaillon O, Skurnik D, Picard B *et al*. The population genetics of commensal *Escherichia coli*. *Nat Rev Microbiol* 2010;**8**: 207–17.

Thanassi DG, Bliska JB, Christie PJ. Surface organelles assembled by secretion systems of Gram-negative bacteria: diversity in structure and function. *FEMS Microbiol Rev* 2012;**36**:1046–82.

Thomas S, Holland IB, Schmitt L. The Type 1 secretion pathway - the hemolysin system and beyond. *Biochim Biophys Acta* 2014;**1843**:1629–41.

Tietze E, Dabrowski PW, Prager R *et al*. Comparative genomic analysis of two novel sporadic Shiga toxin-producing *Escherichia coli* O104:H4 strains isolated 2011 in Germany. *PLoS One* 2015;**10**:e0122074.

Tjalsma H, Bolhuis A, Jongbloed JD *et al*. Signal peptide-dependent protein transport in *Bacillus subtilis*: a genome-based survey of the secretome. *Microbiol Mol Biol R* 2000;**64**:515–47.

Torres AG, Kaper JB. Multiple elements controlling adherence of enterohemorrhagic *Escherichia coli* O157:H7 to HeLa cells. *Infect Immun* 2003;**71**:4985–95.

Torres AG, Jeter C, Langley W *et al*. Differential binding of *Escherichia coli* O157:H7 to alfalfa, human epithelial cells, and plastic is mediated by a variety of surface structures. *Appl Environ Microb* 2005;**71**:8008–15.

Torres AG, Slater TM, Patel SD *et al*. Contribution of the Ler- and H-NS-regulated long polar fimbriae of *Escherichia coli* O157:H7 during binding to tissue-cultured cells. *Infect Immun* 2008;**76**:5062–71.

Totsika M, Wells TJ, Beloin C *et al*. Molecular characterization of the EhaG and UpaG trimeric autotransporter proteins from pathogenic *Escherichia coli*. *Appl Environ Microb* 2012;**78**:2179–89.

Turner L, Zhang R, Darnton NC *et al*. Visualization of Flagella during bacterial Swarming. *J Bacteriol* 2010;**192**:3259–67.

Uhlich GA, Keen JE, Elder RO. Variations in the csgD promoter of *Escherichia coli* O157:H7 associated with increased virulence in mice and increased invasion of HEp-2 cells. *Infect Immun* 2002;**70**:395–9.

Ulett GC, Webb RI, Schembri MA. Antigen-43-mediated autoaggregation impairs motility in *Escherichia coli*. *Microbiology* 2006;**152**:2101–10.

van Diemen PM, Dziva F, Stevens MP *et al*. Identification of enterohemorrhagic *Escherichia coli* O26:H- genes required for intestinal colonization in calves. *Infect Immun* 2005;**73**:1735–43.

Velarde JJ, Varney KM, Inman KG *et al*. Solution structure of the novel dispersin protein of enteroaggregative *Escherichia coli*. *Mol Microbiol* 2007;**66**:1123–35.

Voulhoux R, Ball G, Ize B *et al*. Involvement of the twin-arginine translocation system in protein secretion via the type II pathway. *EMBO J* 2001;**20**:6735–41.

Walker CL, Black RE. Diarrhoea morbidity and mortality in older children, adolescents, and adults. *Epidemiol Infect* 2010;**138**:1215–26.

Wang P, Xiong Y, Lan R *et al*. pO157´Sal, a novel conjugative plasmid detected in outbreak isolates of *Escherichia coli* O157:H7. *J Clin Microbiol* 2011;**49**:1594–7.

Wells TJ, McNeilly TN, Totsika M *et al*. The *Escherichia coli* O157:H7 EhaB autotransporter protein binds to laminin and collagen I and induces a serum IgA response in O157:H7 challenged cattle. *Environ Microbiol* 2009;**11**:1803–14.

Wells TJ, Sherlock O, Rivas L *et al*. EhaA is a novel autotransporter protein of enterohemorrhagic *Escherichia coli* O157:H7 that contributes to adhesion and biofilm formation. *Environ Microbiol* 2008;**10**:589–604.

Wong AR, Pearson JS, Bright MD *et al*. Enteropathogenic and enterohaemorrhagic *Escherichia coli*: even more subversive elements. *Mol Microbiol* 2011;**80**:1420–38.

Xicohtencatl-Cortes J, Monteiro-Neto V, Saldana Z *et al*. The Type 4 pili of enterohemorrhagic *Escherichia coli* O157:H7 are multipurpose structures with pathogenic attributes. *J Bacteriol* 2009;**191**:411–21.

Xicohtencatl-Cortes J, Saldana Z, Deng W *et al*. Bacterial macroscopic rope-like fibers with cytopathic and adhesive properties. *J Biol Chem* 2010;**285**:32336–42.

Zav'yalov V, Zavialov A, Zav'yalova G *et al*. Adhesive organelles of Gram-negative pathogens assembled with the classical chaperone/usher machinery: structure and function from a clinical standpoint. *FEMS Microbiol Rev* 2010;**34**:317–78.

Downloaded from http://femsle.oxfordjournals.org/ by guest on January 17, 2017

ANNEXE n°2 : Revue de littérature

Valentin AGEORGES, Ricardo MONTEIRO, Sabine LEROY, Mariagrazia PIZZA, Kaye BURGESS, Christian PROVOT, Frédérique CHAUCHEYRAS-DURAND and Mickaël DESVAUX.

Molecular Determinants of Surface Colonisation in Diarrhoeagenic *Escherichia coli* (DEC): from Adhesion to Biofilm Formation

Pour Publication dans Frontiers in Microbiology

Molecular Determinants of Surface Colonisation in Diarrhoeagenic *Escherichia coli* (DEC): from Adhesion to Biofilm Formation

Valentin AGEORGES [1], Ricardo MONTEIRO [1,2], Sabine LEROY [1], Mariagrazia PIZZA [3], Kaye BURGESS [4], Christian PROVOT [5], Frédérique CHAUCHEYRAS-DURAND [1,6] and Mickaël DESVAUX [1,*]

[1] Université Clermont Auvergne, INRA, UMR454 MEDiS, F-63000 Clermont-Ferrand, France.

[2] Institute for Research and Innovation in Health, Host interaction and Response Unit, Moluclar Microbiology group, 4200-135 Porto, Portugal.

[3] GSK, Via Fiorentina 1, 53100 Siena, Italy.

[4] Teagasc Food Research Centre, Ashtown, Dublin. D15 KN3K

[5] BioFilm Control, Saint-Beauzire, F-63360, France.

[6] Lallemand Animal Nutrition SAS, F-31702 Blagnac Cedex, France.

Runnig title: Surface colonisation factors in diarrhoeagenic *Escherichia coli*

Keywords: Biofilm formation, Pathogenic *Escherichia coli*, Surface colonisation, Adhesion, Surface proteins, Exopolymers.

*Corresponding author: Dr Hab Mickaël DESVAUX-LENÔTRE, INRA (Institut National de la Recherche Agronomique) Auvergne-Rhône-Alpes, Centre de Recherche de Clermont-Ferrand, UMR454 MEDiS (Microbiologie, Environnement Digestif, Santé), Site de Theix, F-63122 Saint-Genès Champanelle, France. Tel.: +33 (0)4 73 62 47 23, Fax: +33 (0)4 73 62 45 81, Email: mickael.desvaux@inra.fr.

Abstract

Escherichia coli is primarily known as a harmless commensal colonising very early the gastrointestinal tract of infants. However, some bacterial strains are of medical importance as they can be responsible for diarrhoea, which can be especially severe in young children. Primarily based on a combination of clinical symptoms as well as serotypical and histological criteria, intestinal pathogenic *E. coli* include six pathotypes of diarrhoeagenic *E. coli* (DEC). Prior to human infection, DEC can be found in natural environments, animal reservoirs, food processing environments and food matrices. From an ecophysiological point of view, DEC thus deal with very different biotopes and biocoenoses all along the food chain. In this context, the wide range of surface molecular determinants acting as colonisation factors (CFs) in DEC is here exhaustively reviewed. In the first instance, CFs can be broadly discriminated into (i) extracellular polysaccharides, (ii) extracellular DNA, and (iii) surface proteins. Surface proteins constitute the most diverse group of CFs broadly discriminated into (i) monomeric CFs, including autotransporter (AT) adhesins, self-associating ATs, inverted ATs, OmpA, heat-resistant agglutinins, dispersin and some moonlighting proteins, (ii) oligomeric CFs, namely the trimeric ATs, to (iii) supramolecular CFs, including flagella and numerous pili, *e.g.* the injectisome, type 4 pili and or conjugative pili. In DEC, the regulation network of these numerous CFs is extremely complex as it occurs at various stages from pre-transcriptional to post-translocational levels.

Introduction

Most recent phylogenetic analyses revealed the *Escherichia* genus is subdivided into eight groups containing three species, namely *Escherichia coli*, *Escherichia fergusonii*, and *Escherichia albertii*, as well as five clades numbered from I to V (Lawrence *et al.* 1991, Walk *et al.* 2009). *E. coli* is undoubtedly the most investigated bacterial species used as a model organism in microbiology. This lipopolysaccharidic diderm bacterium (archetypical Gram-negative bacterium) is primarily known as a harmless commensal of the gastrointestinal tract (GIT) of warm-blooded animals (Mason *et al.* 1981, Chagnot *et al.* 2013). *E. coli* early colonises the GIT of young children and usually represents less than 1% of the human intestinal microbiota (Eckburg *et al.* 2005).

Nevertheless, some *E. coli* species acquired virulence factors that enable them to cause a broad range of human extraintestinal and intestinal infections. On one side extraintestinal pathogenic *E. coli* (ExPEC) mainly comprises the uropathogenic *E. coli* (UPEC), neonatal meningitis *E. coli* (NMEC), necrotoxic *E. coli* (NTEC) and sepsis-associated *E. coli* (SEPEC). On the other side and besides the adherent invasive *E. coli* (AIEC) associated with Crohn's disease (Mann *et al.* 2012), the intestinal pathogenic *E. coli* (InPEC) regroups six pathotypes of diarrhoeagenic *E. coli* (DEC), namely the (i) enterotoxigenic *E. coli* (ETEC), (ii) enteroaggregative *E. coli* (EAEC), (iii) enteropathogenic *E. coli* (EPEC), (iv) enterohemorragic *E. coli* (EHEC), (v) enteroinvasive *E. coli* (EIEC), and (vi) diffusely-adherent *E. coli* (DAEC) (Kaper *et al.* 2004, Croxen *et al.* 2010). The pathogenicity of these strains is well documented and their main virulence factors are also well defined (Croxen *et al.* 2010). Among them, shiga-toxins (Stx) are important virulent factors of EHEC and can be also found in ETEC, EIEC/*Shigella*, EAEC or environmental strains grouped together under the name shiga-toxing producing (or encoding) *E. coli* (STEC) (Nataro *et al.* 1998, EFSA 2015). Despite the high genome plasticity witnessing intensive gene flow, the population structure of *E. coli* remains mostly clonal (Touchon *et al.* 2009) with a clear delineation into seven principal phylogenetic groups (A, B1, B2, C, D, E and F) (Jaureguy *et al.* 2008, Walk *et al.* 2009, Tenaillon *et al.* 2010, Clermont *et al.* 2013, Beghain *et al.* 2018). Commensal *E. coli* strains generally belong to phylogroup A, whereas DEC usually belong to phylogroups A, B1, C, D and E (Jaureguy *et al.* 2008, Croxen *et al.* 2013, Rossi *et al.* 2018). EHEC strains are mostly found in phylogroups B1 and D (with the main serotype O157:H7) but also in E with the serotype O104:H4. EAEC are found within the phylogroup A but also B1, D and to a

smaller extend B2 (Okeke *et al.* 2010, Croxen *et al.* 2013). *Shigella*/EIEC belongs to phylogroup C (Croxen *et al.* 2013); it is no more a matter of debate that all *Shigella* actually belong to the *E. coli* species (Brenner *et al.* 1972, Chaudhuri *et al.* 2012). ETEC can be found in phylogroup A and B1 and at lesser extent in D. This distinct grouping suggests a parallel evolution of the different pathovars on multiple occasions, possibly the intervention of mobile elements enabling the acquisition of specific combinations of virulence factors (Chaudhuri *et al.* 2012, Croxen *et al.* 2013).

DEC can be found all along the food chain (Giaouris *et al.* 2014, Kim *et al.* 2017). They can have various environmental reservoirs, such as ruminants for EHEC, and can mainly be transmitted to humans by the faecal-oral route through consumption of contaminated food or contact with soiled surfaces (Croxen *et al.* 2013). Besides anthropozoonosis, transmission can also occur between humans. In any case, the colonisation of the food chain by DEC is a major issue for the agri-food sectors and in public health. In order to colonise and persist in their habitat, these bacteria can adhere various biotic and abiotic surfaces. When the reversible adhesion to the surface by low energy linkages (e.g. electrostatics, Van der Waals interactions) is overcome, some bacteria can grow at the surface and form biofilm (Azeredo *et al.* 2017). The bacterial biofilm formation is *per se* divided in several steps: (i) initial and irreversible adhesion of bacterial cells to the surface, (ii) bacterial division at the site of adhesion resulting in the formation of microcolonies, (iii) maturation of the biofilm architecture into a three-dimensional structure, and (iv) bacterial dispersion enabling the colonisation of other sites (O'Toole *et al.* 2000, Hall-Stoodley *et al.* 2002). Biofilm can be monospecies but more generally multispecies in natural environment by forming a complex multicellular community, which is often embedded in an exopolymeric matrix (EPM) (Costerton 1995, Costerton *et al.* 1999). It confers to bacterial cells an increased resistance against environmental stress, antibiotics and/or immunological defences of the host. Biofilm formation can thus plays a key role in DEC ecophysiology by enabling colonisation of various environmental niches (soil, water, vegetables, agri-food surfaces…), the asymptomatic and direct colonisation of some hosts, as well as contributing to transmission through the food chain and ultimately human infection (Ahmed *et al.* 2013).

Most information about the colonisation process in *E. coli* is focused on the domesticated laboratory strain K12 that was commonly considered as representative of the *E. coli* species (Beloin *et al.* 2008). However, this notion is biased due to the numerous and very significant genotypic and/or phenotypic differences with commensal and pathogenic

E. coli isolates (Hobman *et al.* 2007). Indeed, *E. coli* K12 has one of the smallest genome compared to other genome-sequenced strains of *E. coli* due to the lost a large variety of genes during its domestication (Lenski 2017). With regards to the selective pressures that shape the genome evolution, *E. coli* K12 had been replicated and studied for long a time under laboratory conditions, far from those encountered in natural environments (Hobman *et al.* 2007).

Commensal and pathogenic *E. coli* are now demonstrated to harbour some colonisation factors (CFs) that are lacking or misregulated in domesticated laboratory strains of *E. coli*. As the interface between the bacterial cell and its surrounding, the molecular surface determinants are key players in the initial adhesion and sessile development processes and this review aims at summarising exhaustively the CFs at play in DEC. The complexity of the regulation network occurring at various stages, from pre-transcriptional to post-translocational levels, is further highlighted.

1. Molecular determinants involved in surface colonisation by DEC

The colonisation processes along the food chain, from the natural environments, such as the soil, plants and animals, to food environments, including industrial processing food chain and food matrices, and ultimately infection or asymptomatic carriage in human, are very complex and involve many molecular determinants. Sessile development at a surface or interface is generally accompanied by the formation of an EPM embedding the bacterial cells in biofilm (Figure 1). These exopolymers can act as glue for adherence of the bacterial cell to the support and shape the architecture of the biofilm (Hobley *et al.* 2015). Furthermore, the EPM provides a protection by shielding the bacteria from desiccation and anti-microbial compounds but also participates to the channelling of nutrients and signalling molecules (Sutherland 2001, Melissa Starkey 2004, Beloin *et al.* 2008). As such, the EPM contribute to the survival strategy and persistence of bacteria in various environmental conditions (Branda *et al.* 2005). Molecular determinants participating to the surface colonisation by DEC can be either closely associated with the bacterial cell surface and form the cell-associated EPM (caEPM) or present in the extracellular milieu, namely the interstitial EPM (iEPM) (Figure 1) (Van Houdt *et al.* 2005).

At a biochemical level, EPM components can be broadly discriminated between (i) extracellular polysaccharides (EPS), (ii) extracellular DNA (eDNA), and (iii) surface proteins. Depending on the different DEC pathovars, these various determinants can be either present or absent (Table 1). Outer membrane vesicles (OMVs) containing DNA, RNA, proteins or virulence molecules can also be released from the outer membrane (OM) of diderm bacteria to communicate with other microorganisms in the environment or with the host (Mayrand *et al.* 1989, Kuehn *et al.* 2005). While OMVs have been reported to be components of the EPM in *E. coli* K12 (Schooling *et al.* 2006), there is no report whatsoever of their presence in biofilm from DEC and/or their contribution to biofilm formation as observed in *Pseudomonas aeruginosa* or *Helicobacter pylori* (Yonezawa *et al.* 2009, Wang *et al.* 2015), therefore this aspect would deserve further investigations in DEC. Of note, poly-γ-glutamate (PGA) can be found as component of the EPM of numerous bacteria, especially parietal monoderm bacteria (archetypical Gram-positive bacterium) and only a few LPS-diderm bacteria, where it can either be released or cell-surface attached to form a capsule (Candela *et al.* 2006, Ogunleye *et al.* 2015, Radchenkova *et al.* 2018) but, to date, this has never been reported in any *E. coli* strain.

1.1. Exopolysaccharides (EPS)

EPS are one of the main component of the EPM in *E. coli* biofilms (Beloin *et al.* 2008). DEC can biosynthesise a variety of EPS, namely the (i) lipopolysaccharide (LPS), (ii) poly-β-1,6-N-acetyl-D-glucosamine (PNAG), (iii) colanic acid, and (iv) cellulose. Because of their intimate association with the bacterial cell surface, several of these EPS can participate to the formation of a so-called capsule. Actually, *E. coli* harbours some serotype-specific polysaccharides, namely lipopolysaccharides (LPS) (O antigen) and capsular polysaccharides (K antigen). *E. coli* capsules are high-molecular weight polysaccharides embedding the bacterium and linked to the cell-surface via covalent attachments which can also be a constituent of the biofilm EPM (Whitfield 2006, Yaron *et al.* 2014). More than 80 capsular antigens are found in *E. coli* and divided into four groups, from G1 to G4 (Whitfield 2006, Yaron *et al.* 2014). DEC (including EPEC, ETEC and EHEC) produce G1 and G4 capsules that share a common assembly system. During an infection, these capsules allow bacteria to be protected from opsonophagocytosis and complement-mediated killing (Whitfield 2006). In EHEC O104:H4, the capsule was shown to play a role in bacterial survival in the environment and in the direct bacterial interaction with plants (Jang *et al.* 2018).

1.1.1. Lipopolysaccharide (LPS)

LPS is a glycolipidic polymer located at the outer leaflet of the outer membrane of diderm-LPS bacteria (Raetz *et al.* 2002). The LPS is formed around a toxic component, the lipid A, and for this reason is also considered as an endotoxin. The LPS is further composed of the core region linked to the lipid A (divided into an inner and an outer part) and the O-antigen that is linked to the outer part of the core region (Raetz *et al.* 2002). Biosynthesis and assembly pathways of LPS have been fully described and involve more than 50 genes encoded in operons or monocistrons scattered on the bacterial chromosome (Sandkvist 2001, Szalo 2006). Structures of lipid A and core region are highly conserved in *E. coli*, whereas the core region has five basic structures, called R1, R2, R3, R4 and K12. In *E. coli* clinical isolates (InPEC or ExPEC), R1 is most prevalent, whereas the K12 core is as expected not detected (Gibb *et al.* 1992, Appelmelk *et al.* 1994). More exactly, R1 is the most prevalent in non-STEC clinical isolates of *E. coli* and R3 is more associated with STEC strains (Gibb *et al.* 1992, Appelmelk *et al.* 1994, Currie *et al.* 1999, Amor *et al.* 2000). More than 170 O-antigens have been identified and consist of 10–25 repeating units containing one to eight sugar residues (Stenutz *et al.* 2006). O-antigen can be absent (smooth LPS) or present (rough LPS) depending on the *E. coli* strain; if the core region is also absent, it is called deep-rough LPS (Hitchcock *et al.* 1986). Smooth strains are the most commonly found in nature, including DEC, whereas the rough phenotype is usually found in laboratory strains (Whitfield *et al.* 1995, Nataro *et al.* 1998). For smooth strains, the LPS length is positively correlated with the force of adhesion (Strauss *et al.* 2009). O-antigen assists adhesion through hydrogen binding (Tomme *et al.* 1996). For example, O-antigen enabled EHEC O157:H7 strain to colonise animal host (Sheng *et al.* 2008). Mutants in LPS biosynthesis have been shown to affect the adhesion of *E. coli* to abiotic surfaces and its biofilm formation ability (Bilge *et al.* 1996, Genevaux *et al.* 1999, Landini *et al.* 2002, Beloin *et al.* 2006). Besides, LPS could promote or inhibit biofilm formation. Indeed, it was shown that alteration of LPS synthesis could impair type 1 pili and colanic acid expression as well as bacterial motility, whereas the reduction in LPS expression could unmask *E. coli* adhesins and thus promote adhesion or biofilm formation like in the EHEC O157:H7 strain (Bilge *et al.* 1996, Beloin *et al.* 2006, Beloin *et al.* 2008).

1.1.2. Poly-N-acetyl glucosamine (PNAG)

PNAG is an EPS attached to the bacterial surface and involved in biofilm formation on abiotic surfaces (Wang *et al.* 2004). The biosynthetic pathway for PNAG is encoded by the *pgaABCD* locus (formerly *ycdSRQP*). Initiation of the PNAG production occurs with the PgaDC, a glycosyl transferase localised on the cytoplasmic side of the inner membrane that uses the UDP-N-acetyl-D-glucosamine as substrate (Wang *et al.* 2004, Itoh *et al.* 2005, Yoshikane Itoh 2008). The PNAG polymer is exported and anchored to the bacterial surface through the *β*-barrel formed by two outer membrane proteins (OMPs), namely PgaA and PgaB. Although PNAG forms a surface capsule and is one of the main components of diverse bacterial biofilms, the *pga* locus was identified in few DEC strains including EHEC and ETEC (Wang *et al.* 2004, Cerca *et al.* 2007, Al Safadi *et al.* 2012, Gonzales-Siles *et al.* 2016, Cimdins *et al.* 2017). In DEC, PNAG would take part in the stabilisation of biofilm architecture (Wang *et al.* 2004, Al Safadi *et al.* 2012). It has been demonstrated to be important for biofilm formation of EHEC on sprouts and tomato roots (Matthysse *et al.* 2008). Besides, *in vivo* expression of *pgaA* during infection by EHEC O104:H4 suggests that biofilm formation is a key step in pathogenesis (Al Safadi *et al.* 2012). PNAG is also expressed by some ETEC strains and often induced by conditions found in the environment (Gonzales-Siles *et al.* 2016).

1.1.3. Colanic acid

Colanic acid is a negatively charged polymer of glucose, galactose, fucose, and glucuronic acid produced by most of *E. coli* strains, including DEC (Obadia *et al.* 2007). The *wca* operon (or *cps*) encoded 19 genes including polymerases involved in colanic acid synthesis from sugar residues (Stevenson *et al.* 1996). Colanic acid actually corresponds to G1 capsule but a significant portion of the colanic acid produced can also be released into the extracellular milieu (Chris Whitfield 1999, Beloin *et al.* 2008, Beloin *et al.* 2008). The exact contribution of colanic acid to biofilm formation is still unclear (Matthysse *et al.* 2008, May *et al.* 2008). Nonetheless, it forms a physical barrier that helps bacteria to survive in the environment with the formation of a protective capsule around the bacterial cell that allows *E. coli* biofilms to resist to osmotic and oxidative stresses as well as to temperature variations (Chris Whitfield 1999, Chen *et al.* 2004). In EHEC O157:H7, it was further shown to play a role in the bacterial survival in simulated GIT fluids (Mao *et al.* 2006). In EAEC, presence of

colanic acid was linked with the formation of large biofilm structures on surface of sprouts (Borgersen *et al.* 2018). In contrast, production of colanic acid could also mask some cell-surface adhesins and consequently impair initial adhesion to surface (Hanna *et al.* 2003, Schembri *et al.* 2004, Beloin *et al.* 2008).

1.1.4. Cellulose

Cellulose is a linear homopolysaccharide composed of D-glucopyranose units linked by β-1→4 glycosidic bonds. While this widespread biopolymer is generally related to plant biology, it is also present in some bacterial species where it plays a role in protection, maturation and structure of the biofilm (Solano *et al.* 2002, Ude *et al.* 2006). In *E. coli*, cellulose biosynthesis-genes are located in two operons, namely *bcsQABZC* and *bcsEFG* (Zogaj *et al.* 2001, Solano *et al.* 2002, Le Quere *et al.* 2009). The cellulose synthase is formed by BcsAB, which catalyses the cellulose biosynthesis from UDP-glucose subunits and forms a transmembrane pore across the inner membrane for cellulose export prior to secretion across the OM via a β-barrel pore formed by BcsC (Keiski *et al.* 2010, Omadjela *et al.* 2013). The role of *bcsEFG* operon is still unclear but its presence was necessary for cellulose production (Solano *et al.* 2002). These genes are found in both commensal and pathogenic *E. coli* strains (Beloin *et al.* 2008). Although cellulose production is essential for biofilm maturation, an over-production was reported to negatively impact immerged biofilm formation and bacterial aggregation possibly by coating and thus masking the adhesive properties of surface proteins such as curli (Gualdi *et al.* 2008). In EHEC O157:H7 and EPEC O127:H6 cellulose production favoured biofilm formation and consequently host colonisation and survival in different environments (Saldana *et al.* 2009). Involvement of cellulose in *E. coli* colonisation of plant materials has also been demonstrated but it would depend on the vegetable considered since its presence seemed dispensable for biofilm formation of *E. coli* O157:H7 to spinach leaves but it was required for bacterial adhesion to alfalfa sprouts (Matthysse *et al.* 2008, Macarisin *et al.* 2012). Expression of these genes in some ETEC strains is often induced at ambient temperatures, low ionic strength and nutrient limitation (Bokranz *et al.* 2005, Szabo *et al.* 2005).

1.2. Extracellular DNA (eDNA)

The importance of eDNA in biofilm maturation has been demonstrated in numerous bacterial species (Muto *et al.* 1986, Kadurugamuwa *et al.* 1995, Steinberger *et al.* 2002), including *E. coli* (Xi *et al.* 2010, Nakao *et al.* 2012). As a component of the EPM, eDNA serves as structural component of the biofilm but can also contribute to cations gradient, participate as a nutrient source, induce antibiotic resistance, and ease horizontal gene transfer (Bockelmann *et al.* 2006, Palchevskiy *et al.* 2006, Sanchez-Torres *et al.* 2011). However, the role of eDNA in DEC strains remains to be eluded. The molecular mechanism explaining the presence of eDNA has been intriguing for some times but would most certainly and essentially result from the release of genomic DNA upon cell lysis, following bacteriophage lytic cycle or bacterial cell apoptosis (Palmen *et al.* 1995, Steinmoen *et al.* 2002, Qin *et al.* 2007). Nonetheless, the lysis of OMVs containing DNA (Kadurugamuwa *et al.* 1996, Whitchurch *et al.* 2002) as well as DNA secretion through the conjugative Type IV, subtype b, secretion system (T4bSS) (Hamilton *et al.* 2005, Chagnot *et al.* 2013) could also contribute to the presence of eDNA. The extent and respective contribution of these different mechanisms to the presence of eDNA would undoubtedly require further investigations, especially in DEC.

1.3. Cell-surface proteins

The cell surface of LPS-diderm bacteria can be decorated with proteins associated with the OM. Proteinaceous determinants found at the bacterial cell surface and acting as CFs can be broadly discriminated into (i) monomeric proteins, (ii) multimeric proteins (Figure 2).

1.3.1. Monomeric protein CFs

In *E. coli*, monomeric protein acting as CFs essentially comprise some autotransporters (ATs), inverted autotransporters (IATs), and some OMPs but also the surface-exposed lipoprotein SslE, Efa-1 (*E. coli* factor adherence 1), dispersin as well as some moonlighting proteins.

1.3.1.1. Autotransporters (ATs)

ATs acting as CFs comprise the autotransporter adhesins (ATAs), the self-associating autotransporters (SAATs), and some serine protease autotransporters from Enterobacteriaceae

(SPATES) (Henderson *et al.* 2004, Rojas-Lopez *et al.* 2017). ATs correspond to the Type V, subtype a, secretion system (T5aSS) (Desvaux *et al.* 2003, Desvaux *et al.* 2004, Leo *et al.* 2012).

1.3.1.1.1. Autotransporter adhesins (ATAs)

ATAs were shown to enable direct adhesion to abiotic supports, e.g. glass, stainless steel or plastic ware, and/or biotic, e.g. mammalian cells or extracellular matrix (ECM) components such as collagens (Vo *et al.* 2017). As such, they can also belong to MSCRAMM (microbial surface components recognizing adhesive matrix molecules) proteins (Chagnot *et al.* 2012).

In the EHEC, several enterohemorrhagic *E. coli* autotransporters (Eha) have been identified (Wells *et al.* 2008). Among them, EhaB was shown to promote bacterial cells binding to laminin and collagen I (Wells *et al.* 2008, Wells *et al.* 2009), whereas EhaJ further adhered strongly to fibronectin, fibrinogen, collagens II, III and V, and laminin (Easton *et al.* 2011). EhaB was also identified in EPEC and ETEC (Zude *et al.* 2014). Immediately adjacent to the *ehaJ*, *egtA* encodes a glycosyltransferase. EhaJ required to be glycosylated to mediate strong biofilm formation but not for adhesion to ECM components (Easton *et al.* 2011). Following genomic analysis, *ehaJ* appeared to be also present in EAEC, EIEC and ETEC where its function is still unknown. In EPEC, its exact function in the colonisation process remain unclear since it did not seem to be required for bacterial adhesion and biofilm formation (Easton *et al.* 2011). While EhaD was shown to mediate biofilm formation, its role in bacterial adhesion has not been determined yet and its contribution to sessile development in DEC would require more in-depth investigations (Wells *et al.* 2008). In the laboratory strain *E. coli* K12, the EhaD homologue YpjA was shown to promote adhesion to glass and polyvinyl chloride (PVC) as well as biofilm formation together with the EhaC homologue YfaL and YcgV (Roux *et al.* 2005). In EHEC, though, EhaC was not shown to promote biofilm formation (Wells *et al.* 2008). Homologue of *ycgV* was genetically identified in several DEC, namely EPEC, ETEC, EAEC and EIEC (Wells *et al.* 2010, Zude *et al.* 2014). Altogether, this information stresses the need for further experimental characterisations of the adhesive functions of Eha considering the diversity of DEC.

Some ATs originally identified in UPEC and acting as adhesins were further identified in DEC, namely UpaB (uropathogenic *E. coli* autotransporter B) and UpaI (Zude *et al.* 2014). From UPEC investigations, these proteins appeared to promote adhesion to a wide range of

ECM components (Allsopp *et al.* 2012, Zude *et al.* 2014), whereas UpaI was further demonstrated to mediate biofilm formation (Zude *et al.* 2014). Although the genes are found in EPEC and STEC, none of them have been functionally characterised in any DEC to date (Zude *et al.* 2014).

Following genomic analysis, AatA (avian pathogenic *E. coli* autotransporter A) appeared to be also present in some DEC strains (Zude *et al.* 2014). In APEC (avian pathogenic *E. coli*), AatA is important for pathogenesis as it enhanced adhesion to chicken fibroblast cells (Dai *et al.* 2010, Li *et al.* 2010, Wang *et al.* 2011). However, its role and contribution in DEC is still unknown.

1.3.1.1.2. Self-associating autotransporters (SAATs)

SAATs are enable to associate one to another resulting in bacterial cell autoaggregation (Klemm 2006). In *E. coli*, the SAATs regroup ATs from the Ag43 (antigen 43), AIDA-I (adhesin involved in diffuse adherence phenotype) and TibA (toxigenic invasion locus b) families (Trunk *et al.* 2018).

Ag43 is probably the SAAT having triggered the most interest so far and where most of the information results from investigations in the laboratory strain *E. coli* K12 (van der Woude *et al.* 2008). Besides autoaggregation, Ag43 was demonstrated to increase biofilm formation on abiotic surfaces (Kjaergaard *et al.* 2000) and adhesion to epithelial cells (Sherlock *et al.* 2006, de Luna *et al.* 2008) but to decrease bacterial motility (Ulett *et al.* 2007). Gene encoding Ag43 was shown to be highly expressed at early stage of biofilm formation (Schembri *et al.* 2003) but not in mature biofilms (Beloin *et al.* 2004). While biofilm formation is favoured by the autoaggregation phenomenon (Chauhan *et al.* 2013), Ag43 was not involved in gut colonisation in *E. coli* K12 (de Luna *et al.* 2008). It is also known that the expression of pili would shield the interaction between Ag43 and thus prevents the autoaggregation (Korea *et al.* 2010). Phylogenetic analysis revealed the *agn43* distribute into two subfamilies, namely subfamily I (SF-I) and SF-II, and is found only among, but not all, *E. coli* (including some *Shigella* spp.) (van der Woude *et al.* 2008). It is suggested the *agn43* is more prevalent in pathogenic *E. coli* strains than in commensal *E. coli* strains (van der Woude *et al.* 2008). It can be detected as a single gene copy, like in *E. coli* K12, or in multiple alleles, like in EHEC O157:H7 EDL933 where two identical copies are found in two different pathogenicity islands, namely the O-island 43 (OI-43) and OI-48 (Torres *et al.* 2002). In UPEC CFT073, the two alleles encode for different Ag43, namely

Ag43a and Ag43b (Ulett *et al.* 2007); while Ag43a was shown to promote a strong autoaggregation phenotype as well as biofilm formation, Ag43b was associated with slower autoaggregation kinetics and lower propension for biofilm formation. Autoaggregation results from the L-shape structure of Ag43, which drives molecular interaction via salt bridge and hydrogen bonds along the β-helix structure in a molecular Velcro-like handshake mechanism (Heras *et al.* 2014). In *E. coli* O157:H7, Ag43 was shown to promote autoaggregation, calcium binding and biofilm formation but was unable to mediate adhesion to epithelial cells (Torres *et al.* 2002). While present in other DEC, such as EPEC, ETEC and EAEC (Zude *et al.* 2014, Vo *et al.* 2017), functional characterisation of Ag43 in these different genomic backgrounds is still awaited.

AIDA-I is involved in the diffuse adherence of DEC strains (Benz *et al.* 1989, Benz *et al.* 1992) and also in bacterial autoaggregation, biofilm formation and adherence to a wide range of human and non-human cells (Benz *et al.* 1989, Sherlock *et al.* 2006). While the function of AIDA-I is quite similar to Ag43, they clearly belong to different protein families (Vo *et al.* 2017). The gene encoding AIDA-I was originally identified in DAEC (Benz *et al.* 1989, Benz *et al.* 1992, Benz *et al.* 1992) but is especially prevalent in ETEC, EPEC and STEC strains from porcine origin (Niewerth *et al.* 2001, Ha *et al.* 2003). The AIDA-I gene (*aidA*) is associated with *aah* encoding a 45-KDa heptosyltransferase (Benz *et al.* 2001). O-glycosylation mediated by Aah (adhesin associated heptosyltransferase) modifies the AIDA-I by addition of 19 heptose residues in average, which enables for examples EPEC to adhere to human cells (Benz *et al.* 1992, Benz *et al.* 2001, Laarmann *et al.* 2003, Schembri *et al.* 2004). In EHEC O157:H7, though, AIDA-I did not play a role in adherence to cultured cells or to pig intestinal epithelial cells (Yin *et al.* 2009). This suggests different subfamilies of AIDA-I could exist as observed for Ag43, which would require further in-depth investigations.

TibA was found to self-aggregate, form biofilm and facilitate colonisation of the intestinal epithelium (Sherlock *et al.* 2005, Cote *et al.* 2011). In ETEC, TibA is encoded by the *tib* operon further encoding the glycosyltransferase TibC (Lindenthal *et al.* 1999). Glycosylation of TibA is important to its function since its unglycosylated form is less stable, cannot oligomerise properly and in turn cannot promote bacterial adhesion to epithelial cells (Cote *et al.* 2013); nonetheless, it can autoaggregate, promote biofilm formation and cell invasion. Interestingly, TibA, AIDA-I and Ag43 were briefly reported to interact one with another resulting in the formation of mixed bacterial aggregates (Klemm 2006). These interesting findings would deserve further in-depth characterisation.

More recently in *E. coli* O157:H7, EhaA was shown to mediate autoaggregation and adhesion to primary epithelial cells derived from the bovine terminal rectum as well as biofilm formation (Wells *et al.* 2008). As such, EhaA can be considered as an additional member of SAAT also found in EAEC, EPEC and ETEC (Vo *et al.* 2017). Similarly, UpaC was reported to promote autoaggregation as well as biofilm formation (Zude *et al.* 2014). UpaC is found in a wide range of InPEC (Zude *et al.* 2014). Of note, some ATAs such as UpaI can further promote autoaggregation at some extents (Zude *et al.* 2014).

1.3.1.1.2. Serine protease autotransporters from enterobacteriaceae (SPATEs)

SPATEs correspond to a subfamily of protease autotransporters that specifically exhibit a serine protease domain in the passenger (IPR034061) (Rojas-Lopez *et al.* 2017). While their primary function is associated with the degradation of various proteins, such as mucin or haemoglobin, they can contribute to bacterial virulence via their cytotoxic effect, and some can even be involved in bacterial colonisation (Dautin 2010).

In EHEC, EspP (extracellular serine protease plasmid-encoded), also known as PssA (protein secreted by Stx-producing *E. coli*), contributes to biofilm formation, bacterial adherence to intestinal epithelial cells, including bovine primary rectal cells, and colonisation of the bovine intestine (Dziva *et al.* 2007, Supraja Puttamreddy 2010, Farfan *et al.* 2012). EspP is plasmid encoded on the pO157 and can be found in diverse STEC isolates (van Diemen *et al.* 2005, Dziva *et al.* 2007, Ruiz-Perez *et al.* 2014). At the bacterial cell surface, EspP passengers self-assemble to form supramolecular structures, called ropes (Xicohtencatl-Cortes *et al.* 2010). Besides cytopathic activities, the EspP ropes displayed strong adhesive properties on host epithelial cells and could further serve as a substratum for bacterial adherence and biofilm formation. Similar observations were also made for EspC from EPEC (Xicohtencatl-Cortes *et al.* 2010).

In EAEC, Pic (protein involved in colonisation) was involved in mucin degradation but also directly in mucin binding (Gutierrez-Jimenez *et al.* 2008, Andrade *et al.* 2017). It thus participated to the intestinal colonisation and could also be involved in bacterium-mucus biofilm (Navarro-Garcia *et al.* 2011). Pic is also expressed by the hybrid EHEC/EAEC *E. coli* O104:H4 but its exact contribution to the colonisation process in this genetic background remains to be ascertained (Henderson *et al.* 1999, Harrington *et al.* 2009, Abreu *et al.* 2015, Abreu *et al.* 2016). Of note, Shmu is a mucinase identical to Pic found in *Shigella* (Rajakumar *et al.* 1997).

1.3.1.2. Inverted autotransporters (IATs)

In IATs, which correspond to the Type V, subtype e, secretion system (T5eSS) T5eSS, the translocator is located in the N-terminal region and the passenger at the C-terminal, which the opposite of the modular organisation found in ATs (Tsai *et al.* 2010, Oberhettinger *et al.* 2012). In DEC, several IATs acting as CFs can be considered, namely intimin, FdeC (Factor adherence of *E. coli*), and YeeJ.

1.3.1.2.1. Intimin

The intimin is the prototypical member of IATs (Leo *et al.* 2015). In EPEC and EHEC, the intimin is encoded by *eae* (for *E. coli* attachment effacement) gene in the locus of enterocyte effacement (LEE) (Nataro *et al.* 1998). This protein interacts specifically with its receptor Tir (translocated intimin receptor) allowing the establishment of the intimate attachment of the bacteria with the host cell, pedestal formation and attaching/effacing lesions (A/E) (Schmidt 2010). Besides, intimin contributes to intestinal colonisation in a Tir-independent manner (Mallick *et al.* 2012). Intimin could also bind to alternative receptors such as β_1 integrins or nucleolin but this remains to be clarified (Liu *et al.* 1999, Leo *et al.* 2015).

1.3.1.2.2. Factor adherence of *E. coli* (FdeC)

FdeC is a widespread IAT in *E. coli* and present in all DEC pathotypes (Nesta *et al.* 2012, Easton *et al.* 2014). In the EHEC O26:H11, FdeC was shown to take part in biofilm formation and potentially in colonisation of the terminal rectum of cattle (Easton *et al.* 2014).

1.3.1.2.3. YeeJ

More recently, the gene encoding YeeJ was reported to be amplified in some DEC, namely EHEC, EPEC, ETEC and EIEC (Martinez-Gil *et al.* 2017). In *E. coli* K12, this IAT was shown to participate to biofilm formation. While YeeJ exists into two distinct variants of different lengths, no functional difference could be detected between them. Still, the contribution of YeeJ to biofilm formation in DEC remains to be established.

1.3.1.3. OMPs

Besides ATs, several OMPs can act as CFs in DEC, namely OmpA, Hra (Heat-resistant agglutinin), and Iha (Iron-regulated protein A homologue adhesin). OMPs are

integrated to the OM via the β-barrel assembly machinery (Bam) complex (Leyton *et al.* 2015, Botos *et al.* 2017, Schiffrin *et al.* 2017).

1.3.1.3.1. Outer membrane protein A (OmpA)

While originally considered as a pore forming protein (Sugawara *et al.* 1992), the OmpA β-barrel offering a channel for the continuous passage of water or solutes remains controversial (Smith *et al.* 2007). Nowadays, OmpA is rather view as a multifaceted protein including functions of adhesin as well as invasin. In EHEC O157:H7, OmpA is involved in adhesion to intestinal epithelial cells (Torres *et al.* 2003, Kudva *et al.* 2015). In this pathotype, OmpA further appeared as the key molecular determinant for bacterial adhesion to plant surfaces, such as alfalfa sprouts (Torres *et al.* 2005). The role of OmpA as invasin was demonstrated in NMEC (Prasadarao *et al.* 1996) but remains to be established in DEC. Interestingly, OmpA could be encoded by at least two different alleles, namely *ompA1* and *ompA2* (Power *et al.* 2006). Many of the interaction properties of OmpA would emanate from protein loops external to the OM and displayed on the bacterial cell surface (Smith *et al.* 2007); in the two alleles, differences in these regions could influence the adhesin and/or invasin properties of the protein. Of note, OmpA further served as a receptor for bacteriophages and bacteriocins (Smajs *et al.* 1997, Power *et al.* 2006). Regarding biofilm formation, the direct contribution of OmpA remains controversial; while OmpA from *E. coli* K12 was shown to bind to abiotic surfaces and to significantly influence biofilm formation (Lower *et al.* 2005, Barrios *et al.* 2006), the role of OmpA in EHEC O157:H7 would be minor and rather a modulator than a contributor to sessile development (Torres *et al.* 2005, Kudva *et al.* 2015). While further investigations in these directions are clearly needed, allelic variation of OmpA should also be more carefully considered to decipher their exact contribution.

1.3.1.3.2. Heat-resistant agglutinin (Hra)

The Hra family are OMPs first described with Hek (haemagglutinin from *E. coli* K1) in NMEC where it was reported to promote autoaggregation, interactions with human erythrocytes and epithelial cells as well as adhesion to and invasion of epithelial cells (Fagan *et al.* 2007). Actually, Hek was originally identified because of its homology with Tia (toxigenic invasion protein A) (Bhargava *et al.* 2009). In ETEC, Tia mediated attachment to intestinal epithelial cells as well as their invasion (Fleckenstein *et al.* 1996, Sjoling *et al.* 2015). It further appeared to bind heparin as well as several mammalian heparan sulfate binding proteins suggesting that ETEC used these ubiquitous cell surface heparan sulfate

proteoglycans as receptors to adhere and invade host epithelial cells (Fleckenstein *et al.* 2002). While *hra1/hek* was originally considered absent from DEC but restricted to UPEC, NMEC and sepsis *E. coli* (Dobrindt *et al.* 2002, Cooke *et al.* 2010), it later clearly appeared that *hra1* and *tia* are common among DEC, especially EAEC but also EPEC (Fleckenstein *et al.* 1996, Mancini *et al.* 2011). In EAEC O42, Hra1 (heat-resistant agglutinin 1) was demonstrated to be responsible for autoaggregation and aggregative adherence as well as biofilm formation (Bhargava *et al.* 2009). While these observations were made upon protein expression in nonadherent and nonpathogenic laboratory *E. coli* strains, an EAEC 042 *hra1* deletion mutant is not deficient in these phenotypes indicating that Hra1 is an accessory colonisation factor in this genetic background (Bhargava *et al.* 2009). In EAEC strain 60A, Hra2 was not involved in autoaggregation or invasion but only in adherence to epithelial cells (Mancini *et al.* 2011); its involvement in bacterial adhesion to abiotic support and biofilm formation remains to be elucidated. The prevalence of *hra2*, however, seems to be very low among DEC. More recently, a novel member of the Hra family was uncovered, in STEC, namely Hes (Hemagglutinin from Shiga toxin-encoding *E. coli*) (Montero *et al.* 2017). Hes was shown to promote autoaggregation and biofilm formation as well as erythrocyte agglutination and adherence to epithelial cells but not invasion. The gene was observed to be prevalently present in LEE-negative STEC but not LEE-positive STEC (Montero *et al.* 2017).

1.3.1.3.3. Iron-regulated protein A homologue adhesin (Iha)

Iha is an adherent-conferring protein homologous to IrgA (iron-regulated protein A) found in *Vibrio cholerae* (Tarr *et al.* 2000). Besides a β-barrel structure enabling membrane anchoring as in any OMP, Iha has externally exposed domains. Rather than localised adherence, Iha conferred diffuse adherence pattern in *E. coli* O157:H7 (Tarr *et al.* 2000). Besides STEC, *iha* was identified in EPEC and UPEC (Szalo *et al.* 2002, Kanamaru *et al.* 2003, Gomes *et al.* 2011). In UPEC, Iha was shown to further act as a catecholate siderophore receptor (Herold *et al.* 2009) and a virulent factor in a mouse model (Johnson *et al.* 2005) but these roles in DEC remain to be established. In EHEC, Iha was demonstrated to be involved in intestinal colonisation and would contribute to pathogenesis by promoting adherence to intestinal epithelium (Yin *et al.* 2009).

1.3.1.4. Secreted and surface-associated lipoprotein of E. coli (SslE)

SslE, formerly known as YghJ (Yang *et al.* 2007, Iguchi *et al.* 2009), was recently described as a novel *E. coli* mucinase thanks to its zinc metallopeptidase motif (Luo *et al.*

2014, Nesta *et al.* 2014). This protein is secreted by a Type II, subtype a, secretion system (T2aSS) but the molecular mechanisms of its maturation as a surface lipoprotein remains unclear. The gene encoding SslE is present in different DEC pathotypes such as EPEC, ETEC and EHEC (Decanio *et al.* 2013). In EPEC, SslE was shown to mediate biofilm formation and intestinal colonisation (Baldi *et al.* 2012). This protein can be divided into two main variants and antibodies raised against variant I (from ExPEC strain IHE3034) are able to inhibit translocation of *E. coli* strains through a mucin-based matrix. Besides, immunisation of animals with SslE I significantly reduces gut colonisation by strains from differents pathotypes expressing SslE II (Nesta *et al.* 2014). These observations make SslE a key actor in *E. coli* colonisation of mucosal surface and could serve as a component for a protective vaccine against DEC (Naili *et al.* 2016, Naili *et al.* 2017).

1.3.1.5. E. coli factor adherence 1 (Efa-1)

Efa-1 also known as LifA (lymphostatin A) is present in EPEC and some non-O157 EHEC strains (such as O5:H– and O111:H–) and known to inhibit the proliferation of mitogen-activated lymphocytes as well as the synthesis of proinflammatory cytokines and gamma interferon and to take in the colonisation of the bovine intestine (Klapproth *et al.* 2000, Stevens *et al.* 2002, Abu-Median *et al.* 2006, Deacon *et al.* 2010). In EHEC O157 strains, ToxB is homologous to Efa-1 and appeared to contribute to adherence to cultured epithelial intestinal cells (Tatsuno *et al.* 2001). However, no lymphostatin-like activity has been associated with this protein and it is not involved in intestinal colonisation in animal models (Stevens *et al.* 2004, Abu-Median *et al.* 2006). While Efa-1 has an extracytoplasmic domain and is presumably cell-surface exposed (Nicholls *et al.* 2002), the molecular mechanisms at play for its secretion and cell-surface display remain unknown.

1.3.1.6. Dispersin

Dispersin is an anti-aggregation protein (Aap) involved in the spreading of bacterial cells along the host intestinal mucosa (Sheikh *et al.* 2002). Actually, this protein contributes to adherence and colonisation of EAEC by preventing hyper-aggregation and collapse of AAF (aggregative adherence fimbriae). Dispersin is present at the bacterial cell-surface via binding to LPS in a non-covalent manner after secretion through a Type I secretion system (T1SS) (Velarde JJ 2007). This secretion system and cognate secreted protein are encoded in the *aat* (aggregative ABC transporter) locus located in the aggregative adherence plasmid (pAA) of

some EAEC (Nishi J 2003). Dispersin is also present in some STEC strains such as O104:H4 as well as commensal strains (Monteiro *et al.* 2009, Muniesa *et al.* 2012).

1.3.1.7. Moonlighting proteins

At the bacterial cell surface of *E. coli*, some unexpected proteins have been reported. Indeed, some proteins primarily well known to be localised in the cytoplasm could be identified. Among these unexpected cell surface proteins, most of the glycolytic enzymes are frequently uncovered (Henderson *et al.* 2011). These so-called moonlighting proteins were demonstrated to exhibit second function at the bacterial cell-surface completely unrelated to their primary function in the cytoplasm (Khan *et al.* 2014). As a common glycolytic enzyme found at the bacterial cell surface, GAPDH (glyceraldehyde 3-phosphate dehydrogenase) was demonstrated to bind plasminogen and fibrinogen in EHEC and EPEC (Egea *et al.* 2007). Besides, GAPDH is clearly involved in adhesion to intestinal epithelial cells upon infection. As a common theme for moonlighting proteins present at the bacterial cell surface, these proteins lack a N-terminal signal peptide for translocation across the CM and the protein secretion systems enabling their translocation across the OM are often unknown, which is covered by the generic term of non-classical protein secretion. For GAPDH, though, it was strongly suggested to occur via piggybacking through the Type III, subtype a, secretion system (T3aSS) (Aguilera *et al.* 2012). While it is also known that enolase can also be extracellularly located in *E. coli* (Boel *et al.* 2004), its contribution to bacterial adhesion remains to be determined. The elongation factor Tu (EF-Tu) is also found at the bacterial cell surface and reported as involved in bacterial aggregation (Amimanan *et al.* 2017). In DEC, the contribution of putative moonlighting glycolytic enzymes and other moonlighting proteins to the colonisation process would deserve more intensive investigations.

1.3.2. Multimeric protein complex

Multimeric protein complexes acting as CFs can be discriminated between (i) homooligomeric proteins, namely some trimeric autotransporter adhesins (TAAs), and (ii) cell-surface supramolecular structures, including flagella, and numerous pili.

1.3.2.1. Trimeric autotransporter adhesins (TAAs)

TAAs are characterised by the presence of a short translocator domain, which is functional upon homotrimeric assembly and corresponds to the Type V, subtype c, secretion

system (T5cSS) (Cotter *et al.* 2005, Leo *et al.* 2012). In DEC, TAAs include UpaG (UPEC autotransporter G), Eib (*E. coli* immunoglobulin-binding protein), Sab (STEC-autotransporter mediating biofilm formation) and Saa (STEC autoagglutining adhesin).

1.3.2.1.1. UPEC autotransporter G (UpaG)

While UpaG was originally identified in UPEC, it was further identified in EAEC 042 strain (Zude *et al.* 2014). UpaG is involved in autoaggregation, biofilm formation, adhesion to fibronectin, and laminin, as well as human epithelial cells (Valle *et al.* 2008). In EHEC, EhaG (EHEC autotransporter G) is a positional orthologue of UpaG, which is also involved in autoaggregation, biofilm formation, adhesion to laminin, fibronectin and collagens I, II, II and IV as well as some epithelial cells (Valle *et al.* 2008, Totsika *et al.* 2012, Zude *et al.* 2014). Gene encoding EhaG was further identified in a wide range of DEC, including EPEC, including EPEC, EIEC, ETEC and EAEC (Zude *et al.* 2014).

1.3.2.1.2. *E. coli* immunoglobulin-binding protein (Eib)

Eibs were originally characterised for their ability to bind immunoglobulin fraction, especially to the Fc fraction of IgA and IgG (Sandt *et al.* 2000, Sandt *et al.* 2001, Leo *et al.* 2009); up to 7 different Eibs have been identified to date, namely EibA, B, C, D, E, F and G. In LEE-negative STEC O91, it further appeared that EibG is involved in adherence to bovine and human epithelial cells, with peculiar adhesion pattern called chain-like adhesion (CLA) corresponding to the formation of long chain cell aggregates (Lu *et al.* 2006). The gene encoding EibG would distribute into 21 different alleles clustered into three *eibG* subtypes, namely *eibG*-α, -β, and -γ (Merkel *et al.* 2010). While EibG-α and EibG-β are responsible for a typical CLA phenotype, EibG-γ induces adherence in much shorter cell chains and smaller cell aggregates, corresponding to an atypical CLA. EibD was further shown to promote autoaggregation and biofilm formation on abiotic surfaces (Leo *et al.* 2011). Considering their structural similarity, other Eibs were suggested to have similar biological functions but experimental confirmation is still awaited to ascertain this. Eib genes are found in some STEC strains as well as some *E. coli* commensal strains (Lu *et al.* 2006).

1.3.2.1.3. STEC-autotransporter mediating biofilm formation (Sab)

Sab contributes to the diffusive adherence to human epithelial cells and biofilm formation to abiotic surfaces (Paton *et al.* 2001, Jenkins *et al.* 2003, Herold *et al.* 2009, Farfan *et al.* 2012). Sab is especially present in LEE-negative STEC.

1.3.2.1.3. STEC autoagglutining adhesin (Saa)

Saa promotes adhesion to HEp-2 cells in semilocalised adherence pattern (Paton *et al.* 2001, Jenkins *et al.* 2003, Herold *et al.* 2009, Farfan *et al.* 2012). Saa is essentially present in STEC, including some LEE-negative EHEC strains.

1.3.2.2. Cell-surface supramolecular structures

Flagella and pili are organelles resulting from the supramolecular assembly of different protein subunits to form heteromultimeric protein complexes on the bacterial cell-surface.

1.3.2.2.1. Flagella

Flagellar components are secreted and assembled via the Type III, subtype b, secretion system (T3bSS) and more than fifty genes divided in three hierarchical classes are involved in the flagellar apparatus formation (Young *et al.* 1999, Chilcott *et al.* 2000). The main component of the flagellum filament is the flagellin, which has considerable diversity in ultrastructure and is actually responsible for the H-antigen variability (H1 to H56) (Zhou *et al.* 2015). In *E. coli*, the flagellation is peritrichous but the sites of cell surface localisation and the number of flagella (typically around 6-10) are considered random (Macnab 1987). Nonetheless, it must be stressed that when swimming, the flagella in motion coalesce into an undulating bundle forming one rigid helical ponytail about 14 nm in diameter and 10 µm long that appears as polarly localised in *E. coli* (Bray 2001). A swimming bacterial cell has a run-and-tumble behaviour, where it progresses linearly (run) and then changing abruptly in direction (tumble), but also slow-random-walk behaviour, where it moves at relatively low speed (Qu *et al.* 2018). The surface approach is an important step towards initial bacterial adhesion and subsequent sessile development. Active motility involving the flagella allows the bacterial cells to overcome repulsive electrostatic and hydrodynamic forces at the adhesion site (Donlan 2002).

Besides swimming, flagella can participate in an alternative type of motility called swarming where bacterial cells move and spread at surface (Kaiser 2007). Swarming directly participates to the surface colonisation process and is associated with the expression of an alternative system, the lateral flagella (Merino *et al.* 2006). In EAEC O42, the Flag-2 locus encodes such a system (Ren *et al.* 2005). However, a mutation frameshift has likely inactivated this system in this strain. Nonetheless, the Flag-2 cluster appeared to be present in

about 20 % of *E. coli* strains from the ECOR collection (Ren *et al.* 2005). In the environmental strain *E. coli* SMS-3-5, although the Flag-2 gene cluster is complete and intact, swarming motility could not be observed (Fricke *et al.* 2008); to date, the functionality and contribution of this system to the colonisation processes in *E. coli* remains to be elucidated. In the absence of flagella, *E. coli* is not so efficient at surface colonisation but is still considered a temperate swarmer enable to swarm over surface with rheology corresponding to 0.5%-0.8% agar (against ≥1.5% for robust swarmers) (Partridge *et al.* 2013).

Besides motility, flagella can directly act as adhesins as shown in EPEC where they are involved in adhesion to epithelial cells (Giron *et al.* 2002, Cleary *et al.* 2004). In EAEC, flagella contribute to adhesion to plant leaves (Berger *et al.* 2009). In EHEC, the flagellin FliC favours initial attachment, adhesion to epithelial cells and biofilm formation on abiotic surfaces as well as spinach leaves (McNeilly *et al.* 2008, Mahajan *et al.* 2009, Xicohtencatl-Cortes *et al.* 2009, Vikram *et al.* 2013, Nagy *et al.* 2015). In ETEC, flagella take part to bacterial adhesion to salad leaves and intestinal epithelial cells as well as biofilm formation (Shaw *et al.* 2011, Duan *et al.* 2012, Zhou *et al.* 2013, Zhou *et al.* 2014). Interestingly, in this species, flagella can also mediate indirect adhesion through EtpA (ETEC two-partner secretion protein A), a protein secreted by a T5bSS (two-partner secretion system), which bridges the flagella with host cell receptors thus allowing bacterial cell attachment to some epithelial cells and mucin-expressing regions in mouse small intestine (Fleckenstein *et al.* 2006, Roy *et al.* 2009). In both EHEC and ETEC, flagella would also contribute to bacterial invasion of intestinal epithelial cells but the molecular mechanisms at work remains to be clarified (Luck *et al.* 2006, Rogers *et al.* 2012, Duan *et al.* 2013). In EHEC and EPEC, the adhesion of H6 and H7 flagella to intestinal epithelium and epithelial cells was suggested to occur though mucins (Giron *et al.* 2002, Mahajan *et al.* 2009) as reported for H1 flagella from the probiotic *E. coli* Nissle 1917 (Troge *et al.* 2012); this aspect undoubtedly deserves further in-depth investigations in DEC for confirmation. While different flagellin variants have been shown to be involved in direct binding to host cells, such as H1 and H19 flagella in ETEC (Duan *et al.* 2012, Duan *et al.* 2013), systematic analysis of the colonisation properties of all the different H-antigens in *E. coli* has not been investigated as yet. In DEC, the contribution of flagella as motility factor over adhesion factor in the colonisation processes, has not been clearly resolved as yet, especially regarding bacterial adhesion and biofilm formation to biotic and abiotic surfaces (Wood *et al.* 2006). Of note, while EIEC are generally considered as

nonmotile (Nataro *et al.* 1998), no data are available as yet concerning the involvement of flagella in the colonisation process in DAEC (Servin 2014).

1.3.2.2.2. Pili

Pili are clearly key actors during the irreversible attachment of bacteria to surfaces, which is characterised by a stronger and longer interaction coupled with a decrease of bacterial motility (Pruss *et al.* 2006). These pili can be secreted and assembled by different protein secretion systems, namely the Type II, subtype c (T2cSS), Type III, subtype a (T3aSS), Type IV, subtype b (T4bSS), Type VII (T7SS) or Type VIII (T8SS) secretion systems (Figure 2). Of note, this protein secretion nomenclature is restricted to the LPS-diderm bacteria (Desvaux *et al.* 2009). As such the T7SS in this type of bacterial cell envelope corresponds to the chaperone-usher pathway, which is the main pathway responsible for the secretion of a wealth of pili in *E. coli* (Wurpel *et al.* 2013).

1.3.2.2.2.1. The injectisome

The injectisome is a bacterial molecular syringe secreted, assembled and corresponding to the T3aSS (Desvaux *et al.* 2006, Galan *et al.* 2018). In DEC, this peculiar pilus is encoded by genes located in the LEE pathogenicity island (McDaniel TK 1995), a landmark for all EPEC but it is also present in some EHEC strains (namely the LEE-positive over the LEE-negative strains), such as *E. coli* O157:H7, and EIEC (including *Shigella* spp.) (Hueck 1998, Galan *et al.* 2006, Coburn *et al.* 2007). Basically, Tir (translocated intimin receptor) is encoded by the *tir* gene located in the LEE and is injected in the host cell by the injectisome (Hueck 1998). This protein is then exposed at the host cell surface and serves as the receptor for the intimin enabling intimate bacterial interaction with the intestinal epithelium (Donnenberg *et al.* 1993, Liu *et al.* 1999). In EPEC, the injectisome is involved in cell adhesion and pedestal formation that occur during attaching and effacing (A/E) lesions upon actin rearrangement in the infected eukaryotic cell (Wong *et al.* 2011). Of note, while A/E lesions are observed *in vitro* from infected epithelial cell cultures or colonic epithelium with LEE-positive EHEC (Lewis *et al.* 2015), this kind of lesions are never observed from clinical samples of EHEC infections (Nataro *et al.* 1998); a clear explanation of this fact is still matter of controversy in the field but would undoubtedly deserve further investigations in confronting lab experiments with clinical reality (Lewis *et al.* 2015). Besides infection of mammalian cells, the injectisome is involved in adhesion to plants with a marked tropism for the stomata (Schroeder GN 2008, Shaw RK 2008, Berger CN 2010, Croxen *et al.* 2013).

EspA, the main component of the filament in the injectisome is directly involved in adhesion as well as in biofilm formation in EPEC (Knutton *et al.* 1998, Moreira *et al.* 2006). In EIEC, the injectisome takes part in the invasion capabilities (Hueck 1998).

1.3.2.2.2.2. Type 4 pili (T4P)

T4P are secreted by the T2cSS (Ramer *et al.* 2002, Chagnot *et al.* 2013). T4P have been demonstrated to play a role in several *E. coli* pathotypes, including host cell adherence and bacterial aggregation (Craig *et al.* 2004). In EPEC, T4P are also known as BFP (bundle-forming pili) and their subunits assemble in a helical manner to form polymeric fibers and can further interact to create higher-order bundles or tangled aggregates (Giltner *et al.* 2012, Melville *et al.* 2013). These T4P are involved in the colonisation of the GIT and contribute to bacterial virulence (Bieber *et al.* 1998, Tacket *et al.* 1998). BFP are encoded by the *bfp* operon comprising 14 genes, including *bfpA*, which encodes the major repeating subunit of the pilus fiber (Ramer *et al.* 1996, Sohel *et al.* 1996). In EHEC strains, the T4P are called HCP (haemorrhagic *E. coli* pili). Inactivation of the *hcpA* gene in EHEC O157:H7 reduced adherence to human and bovine epithelial cells. HCP were also able to bind to fibronectin and laminin, to agglutinate rabbit red blood cells, to mediate biofilm formation and to promote twitching motility (Xicohtencatl-Cortes *et al.* 2009). Peculiar T4P called longus pili have been reported in ETEC (Giron *et al.* 1994). The N-terminal part of the major subunit LngA of these longus pili is homologous with BfpA of EPEC, CofA subunit of CFA/III (Colonisation factor antigen) of ETEC and TCP (the toxin-coregulated pilin) of *V. cholerae* (Giron *et al.* 1995, Gomez-Duarte *et al.* 1995). Longus pili were involved in colonisation of the human gut (Clavijo *et al.* 2010, Mazariego-Espinosa *et al.* 2010), in bacterium-bacterium interaction and resistance to antimicrobial agents most certainly as a result of biofilm formation (Clavijo *et al.* 2010).

1.3.2.2.2.3. Conjugative pili (CP)

Classically, the genes encoding for F-plasmid transfer are encoded in the *tra* operon located in the conjugative F plasmid (Manwaring *et al.* 1999). In *E. coli*, T4bSS comprises the conjugation machines homologous to the Tra system and assembles CP (Lawley *et al.* 2003). CP are responsible for nucleoprotein transfer between a donor bacterial cell (harbouring the F plasmid) and a recipient bacterial cell (Lawley *et al.* 2003). Bacterial conjugation is a well-known process enabling horizontal transfer of virulence genes or adherence factors (Manwaring *et al.* 1999, Mazel *et al.* 1999, Llosa *et al.* 2002, Sorensen *et al.* 2005). Gene

transfer is especially promoted in biofilm where physical contact between sessile donor and recipient cells is favoured (Lebaron *et al.* 1997, Hausner *et al.* 1999, Dionisio *et al.* 2002, Molin *et al.* 2003, Maeda *et al.* 2006). Besides transfer of genetic material, CP can be directly involved in bacterial adhesion (Beloin *et al.* 2008, May *et al.* 2008, May *et al.* 2011). In biofilm, this can be further amplified as cells carrying a conjugative F plasmid promote the establishment of F pili mating pairs and consequently induce adhesion and biofilm formation to abiotic surfaces to poor biofilm former cell. EAEC strains expressing F pili were demonstrated to improve mixed biofilm formation (Pereira *et al.* 2010). In EAEC C1096, pili encoded in the conjugative plasmid Incl1 further contributed to adherence to abiotic surfaces and epithelial cells (Dudley *et al.* 2006). In EHEC O157:H7 Xuzhou, a novel conjugative plasmid called pO157-Sal encodes a complete set of genes for the T4bSS, but its involvement in the colonisation process has not been questioned as yet (Wang *et al.* 2011, Zhao *et al.* 2013).

1.3.2.2.2.4. Type 1 pili (T1P)

T1P are certainly the most investigated pili secreted and assembled via a T7SS (Capitani *et al.* 2006). The expression of T1P (also called Type 1 fimbriae) is induced at least during the initial adhesion (Schembri *et al.* 2003, Beloin *et al.* 2004, Reisner *et al.* 2014). Besides adhesion to abiotic surfaces (Harris *et al.* 1990, Pratt *et al.* 1998, Cookson *et al.* 2002, Orndorff *et al.* 2004, Reisner *et al.* 2014), they are involved in early and late stages of biofilm formation (Schembri *et al.* 2003, Beloin *et al.* 2004). T1P also have a role in the formation of SIgA (secretory IgA) mediated biofilm of the microbiota within the gut (Bollinger *et al.* 2003, Orndorff *et al.* 2004, Bollinger *et al.* 2006). While FimA (fimbrillin A) constitutes the pilus rod, FimH is at the apex of the pilus tip. FimH is the key adhesin component in T1P as it can link to mannose residues of some receptors on eukaryotic cells (Kaper *et al.* 2004, Duncan *et al.* 2005) but also have nonspecific binding activity at abiotic surfaces (Pratt *et al.* 1998, Beloin *et al.* 2008). The absence of the FimH adhesin was shown to hinder biofilm formation by preventing cell-to-surface and cell-to-cell contacts (Danese *et al.* 2000). In *E. coli*, different *fimH* alleles have reported as conferring distinct colonisation abilities and thus different role in biofilm formation (Martinez *et al.* 2000, Weissman *et al.* 2006, Khan *et al.* 2007). It was shown that contact between T1P and abiotic surface would alter the composition of the OM and change some physicochemical properties of the bacterial surface, which in turn influence adhesion (Otto *et al.* 2001, Orndorff *et al.* 2004). While *E. coli* K12 and UPEC NU14 gather the majority of the investigations about the T1P, their implication in bacterial

adhesion and/or biofilm formation was further demonstrated in EPEC, EAEC, ETEC and STEC strains (Elliott *et al.* 1997, Cookson *et al.* 2002, Moreira *et al.* 2003, Sheikh *et al.* 2017). T1P are encoded by the *fimBEAICDGHF* gene cluster, which is quite widespread in *E. coli* in both commensal and pathogenic isolates (Sauer *et al.* 2000, Kaper *et al.* 2004, Wurpel *et al.* 2013). While present in EHEC O157:H7 (Abraham *et al.* 1988, Li *et al.* 1997, Roe *et al.* 2001, McWilliams *et al.* 2014), evidence of their expression and contribution to the colonisation process are still awaited.

Genes encoding the F1C pili are present in approximately 7 % of *E. coli* faecal isolates (Werneburg *et al.* 2018). F1C pili have been characterised in UPEC strains where they are encoded in the *foc* (fimbriae of serotype 1C) operon homologous to the *fim* locus (Klemm *et al.* 1994). In UPEC, F1C pili are involved in adherence to the bladder and kidney cells as well as in biofilm formation (Werneburg *et al.* 2018). Their prevalence and contribution to the colonisation process in DEC remain to be investigated.

1.3.2.2.2.5. CS31A pili

The CS31A (coli surface associated 31a antigen) play a key role in the virulence of septicemic *E. coli* and ETEC as well as some EPEC and DAEC (Girardeau *et al.* 1988, Contrepois *et al.* 1989, Jallat *et al.* 1994, Adams *et al.* 1997). Because of their thin structure as well as their close and packed association to the bacterial cell surface, CS31A was early described as capsule-like or even nonfimbrial antigens (Bertin *et al.* 1993, Mechin *et al.* 1996) before being clearly identified as thin capsular pili secreted and assembled by a chaperone-usher pathway (T7SS) (Thanassi *et al.* 1998). These pili are synthesised from the *clp* operon located on high-molecular-weight self-transmissible R plasmid, called p31A (Martin *et al.* 1991, Jallat *et al.* 1994, Martin 1996). CS31A are considered homologous to the K88/F4 (*fae* operon) and F41 pili but with some functional dissimilarities, *e.g.* since they do show haemagglutinin activity unlike CS31A (Girardeau *et al.* 1991). In ETEC, F4 pili was shown to allow bacterial adhesion to F4-specific receptors present on brush borders of villous enterocytes thus promoting the colonisation of the small intestine (Snoeck *et al.* 2008). The locus for diffuse adherence (*ldaCDEFGHI*) (Scaletsky *et al.* 2005) from EPEC is homologue to the K88 *fae* and ETEC CS31A *clp* operons. LdaH mediates diffuse adherence on Hep-2 cells. The LdaH encoding gene was also found in STEC strains but no functional characterisation has been as yet reported in this species (Scaletsky *et al.* 2005).

1.3.2.2.2.6. Aggregative adherence fimbriae (AAF)

AAF actually belong to the Afa/Dr (afimbrial adhesin/decay-accelerating factor receptor) hemagglutinin family together with F1845 pili (Nowicki *et al.* 1990, Le Bouguenec *et al.* 2006). In DAEC and EIEC, Afa and Dr hemagglutinins recognise the Dr blood group antigen (Nowicki *et al.* 1990). Among the five genes encoded in the *afa* clusters, *afaB*, *afaC* and *afaE* are required for mannose-resistant hemagglutination (MRHA) (Servin 2005). The Dr hemagglutinin is encoded by the *draABCDE* operon, where *draA*, *draB*, *draC*, and *draD* encode accessory proteins and *draE* encodes the adhesin part (Nowicki *et al.* 1987, Servin 2005). It further and specifically binds collagen IV (Nowicki *et al.* 1988). Afa and Dr hemagglutinins can link to decay-accelerating factor (DAF) and to carcinoembryonic antigen-related cellular adhesion molecules (CEACAMs) (Nowicki *et al.* 1988, Westerlund *et al.* 1989, Berger *et al.* 2004). While some members of the Afa/Dr family were believed not to form pili as they could not be observed by electron microscopy examination, it is now clear they are secreted as AAF and F1845 by T7SS, to form pili of various architecture depending on the pilin subunits secreted and assembled by a T7SS (Anderson *et al.* 2004, Pettigrew *et al.* 2004). In EAEC, colonisation of the gut occurs through aggregative adherence (AA) due to AAF, which bind to ECM proteins such as fibronectin, laminin and collagen IV (Farfan *et al.* 2008, Berry *et al.* 2014) and then promote biofilm formation (Hicks *et al.* 1996, Wakimoto *et al.* 2004). To date, five AAF (AAF/I to AAF/V) have been identified, all encoded by the pAA of EAEC. The main subunits of these fimbriae are AggA, AafA, Agg3A, Agg4a and Agg5a, respectively (Nataro *et al.* 1992, Czeczulin *et al.* 1997, Boisen *et al.* 2008, Jonsson *et al.* 2015). F1845 pili are responsible for diffuse adherence to epithelial cells of the gut and are encoded by the *daaABCDE* operon (Bilge *et al.* 1989, Bilge *et al.* 1993). In DAEC and EIEC, F1845 pili are involved in gut colonisation (Servin 2005). Another hypothetical Dr-related pilin called HdaA (HUS-associated diffuse adherence) also appeared to confer the capacity to cause the AA phenotype to EAEC (Boisen *et al.* 2008).

1.3.2.2.2.7. Colonisation factor antigens (CFA)

In ETEC, colonisation factor antigens (CFA), also called coli surface antigens (CS), form pili that take part in adhesion to the small intestine and are critical for the virulence (Gaastra *et al.* 1996). CFA/I, CFA/II (CS1, 2 and 3) and CFA/IV (CS4, 5 and 6) are the most virulent (Sjoberg *et al.* 1988, Knutton *et al.* 1989, Taniguchi *et al.* 1995, Gaastra *et al.* 1996, Svennerholm *et al.* 2012) but CS12, 14, 17, 18, 19, 20 and 31 can also bind to intestinal cells (Werneburg *et al.* 2018). CFA/CS are encoded by operons; taking CFA/I as an example, it is

encoded by the *cfaABCE* operon, where *cfaB* encodes the main subunit, *cfaE* the distal subunit, *cfaA* a chaperone and *cfaC* the usher involved in pilin transport across the OM (Jordi *et al.* 1992). Cell adhesion is enable by CfaB through its ability to bind glycosphingolipid (Jansson *et al.* 2006).

1.3.2.2.2.8. F9 pili

F9 pili of EHEC O157:H7 were shown to promote colonisation of very young calves (Low *et al.* 2006). When expressed in *E. coli* K12, these pili were involved in the colonisation of epithelial bovine cells and binding to fibronectin (Low *et al.* 2006). In contrast, in the EHEC O157:H7 background, expression of F9 pili reduced adhesion to explant tissue from the bovine extraintestinal tract, probably because of physical hindrance of type III secretion-dependent attachment (Low *et al.* 2006). These pili are short but are able to form longer bundles (Low *et al.* 2006). They are encoded by the F9 gene cluster, a six genes operon located in the pathogenicity island O161 (Low *et al.* 2006, Wurpel *et al.* 2013). This operon was further identified in EPEC as well as EAEC (Wurpel *et al.* 2013). F9 pili are secreted and assembled by a T7SS (Wurpel *et al.* 2013).

1.3.2.2.2.9. E. coli YcbQ laminin-binding fimbriae (ELF)

In EHEC O157:H7, it was shown that ELF bind laminin and are involved in adherence to epithelial cells in human, cow and pig (Samadder *et al.* 2009). ELF form peritrichous flexible fine fibres and are encoded by the *elfADCG*, originally called *ycbQRST* operon previously identified in UPEC and some commensal *E. coli* strains (Spurbeck *et al.* 2011). This operon is actually homologous to the F17 pili biogenesis genes found in ETEC, assembled and secreted by a T7SS (Lintermans *et al.* 1988, Lintermans *et al.* 1991, Bertin *et al.* 1996, Bertin *et al.* 2000). More generally, ELF are further homologous to 20K, K99 and G pili found in various pathogenic *E. coli* (Guinee *et al.* 1976, Contrepois *et al.* 1983). These pili were shown to mediate binding to intestinal mucosal cells, especially to N-acetyl-D-glucosamine-containing receptors (Bertin *et al.* 1996). Composition of the pili and the sequence of the tip-adhesin differ between the strains and could explain phenotypic divergence associated with expression of this family of pili in different *E. coli* (Korea *et al.* 2010).

1.3.2.2.2.10. Long polar fimbriae (LPF)

LPF are encoded by two operons *lpf1* and *lpf2* located on the pathogenicity islands O141 and O154 in EHEC O157:H7, respectively (Perna *et al.* 2001). LPF are also present in other DEC, e.g. LEE-negative EHEC, EPEC, rabbit-specific EPEC, EAEC and ETEC, as well as in several commensal strains (Doughty *et al.* 2002, Wurpel *et al.* 2013). They share homology with the LPF of *Salmonella enterica* serovar Typhimurium involved in adherence to Peyer's patches and M cells in the human gut (Baumler *et al.* 1995, Baumler *et al.* 1996). The *lpf1* operon is composed of five genes with *lpfA* encoding the main pilus subunit, *lpfD* and *lpfE* encoding minor subunits, whereas *lpfB* and *lpfC* encoding the chaperone and usher, respectively (Doughty *et al.* 2002, Torres *et al.* 2004). The *lpf2* operon also contains five genes with a duplication of *lpfD* called *lpfD'* but with no *lpfE* paralogue (Torres *et al.* 2004). In *E. coli* O157:H7, it has been proposed that LPF2 is expressed in early stages whereas LPF1 is expressed in late stages of growth (Torres *et al.* 2004). LPF are secreted and assembled by a T7SS and can bind fibronectin, laminin, collagen IV as well as follicule-associated epithelium (FAE) of Peyer's patches in human (Fitzhenry *et al.* 2006, Farfan *et al.* 2012, McWilliams *et al.* 2014). Besides, expression of *lpf2* is increased under condition to those similar for biofilm formation (Torres *et al.* 2007). Recently, it has been demonstrated that STEC isolates positive for *lpf2* formed significantly more biofilm than *lpf2*-negatives isolates (Vogeleer *et al.* 2015). In EPEC, LPF was shown to contribute to the early stages of colonisation of rabbits and severity of diarrhoea (Newton *et al.* 2004).

1.3.2.2.2.11. E. coli common pilus (ECP)

In EHEC, ECP (previously called Mat for meningitis-associated temperature dependent pilus) provides adherence to HEp-2, HeLa and HT-29 cells and allows interaction between bacterial cells (Rendon *et al.* 2007). Secreted and assembled by a T7SS, ECP expression is increased under environmental conditions that are experienced in the GIT, *e.g.* low oxygen and high CO_2 concentrations (Rendon *et al.* 2007). However, its role seems to be secondary in the colonisation of human or bovine gut (Tatsuno *et al.* 2000, Dziva *et al.* 2004). The *ecp* operon was identified in numerous commensal and pathogenic *E. coli*, including DEC (Rendon *et al.* 2007).

1.3.2.2.2.12. Sorbitol-fermenting frimbriae protein (SFP)

In EHEC, expression of SFP pili is induced in anaerobic conditions and leads to an increased adherence to Caco-2 and HCT-8 cells with a mannose-resistance hemagglutination

phenotype (Brunder *et al.* 2001, Musken *et al.* 2008, Bielaszewska 2009). These pili are encoded by the *sfpABDCDJG* operon harboured in the virulence plasmid pSFO157 (Brunder *et al.* 2006). SFP pili are secreted and assembled by a T7SS (Brunder *et al.* 2001). Besides *E. coli* O157, *sfp* was identified in other EHEC serotype, such as 0165 (Bielaszewska 2009), but it prevalence among STEC in general is probably quite low (Toma *et al.* 2004). Distribution of the *sfp* operon in other DEC has never been questioned as yet.

1.3.2.2.2.13. Curli

Curli are thin aggregative pili generally considered as one of the major proteinaceous component of *E. coli* biofilm matrix (Smyth *et al.* 1996, Stathopoulos *et al.* 2000, Kostakioti *et al.* 2005, Evans *et al.* 2014). These peculiar pili are secreted and assembled by the T8SS or extracellular-nucleation-pathway (ENP). Curli are helical filamentous amyloid fibers that facilitate cell-surface and cell-cell interactions and promote biofilm formation (Olsen *et al.* 1993, Cookson *et al.* 2002, Szabo *et al.* 2005, Beloin *et al.* 2008, McCrate *et al.* 2013). In EHEC O157:H7, curli are associated with cellulose production, adherence to spinach leaves and Hep-2 cells as well as abiotic surfaces (Kim *et al.* 2004, Pawar *et al.* 2005, Macarisin *et al.* 2012). In ETEC, curli facilitate adherence to plastic surfaces (Szabo *et al.* 2005). Although curli was originally thought not be expressed by EPEC (Ben Nasr *et al.* 1996), some strains were later reported to synthetise curli playing a role in bacterial adhesion and biofilm formation in condition closed to human or bovine host (Saldana *et al.* 2009). However, curli do not seem to be required for biofilm formation and/or adhesion of EAEC strains (Sheikh *et al.* 2001, Berger *et al.* 2009, Pereira *et al.* 2010). In *Shigella spp*. and EIEC, CsgD and curli expression is often inactivated (Sakellaris *et al.* 2000). Two operons are involved in curli production, (i) the *csgBAC* operon, encoding the structural components of curli (CsgA and CsgB) and an accessory protein (CsgC), and (ii) the *csgDEFG* operon, encoding a transcriptional regulator (CsgD) and the secretion machinery for transport across the OM (CsgE-G) (Arnqvist *et al.* 1994, Hammar *et al.* 1995, Beloin *et al.* 2008). Recently, CsgC and CsgE were demonstrated to highly inhibit CsgA aggregation and CsgE was shown to prevent pellicle biofilm formation when added exogenously (Andersson *et al.* 2013, Evans *et al.* 2015).

1.3.2.2.2.14. Haemolysin-coregulated protein (Hcp)

In EAEC, the Hcp tube formed by the Type VI secretion system (T6SS) was suggested of importance for biofilm formation (Aschtgen *et al.* 2008). More than ten

orthologues of the T6SS components have been identified in some EHEC and EPEC strains. This system could also contribute to bacterial aggregation at the host cell surface (Dudley *et al.* 2006, Shrivastava S 2008, Lloyd *et al.* 2009, Aschtgen *et al.* 2010, Moriel *et al.* 2010). Further investigations to determine the exact role and molecular mechanisms involved the colonisation processes by the Hcp and T6SS are definitely required in DEC.

2. The different regulation levels for CFs expression

As a general trend, the expression of genes encoded on genomes into proteins are regulated from pre-transcriptional, transcriptional, post-transcriptional, translational to post-translational levels, as well as at translocational and post-translocational levels, which are especially relevant and important for molecular determinants expressed at bacterial cell surface (Figure 3). With the rise of omic approaches, however, some basic bacterial physiology concepts are sometimes overlooked and gene/protein expression is very often considered as limited to regulatory networks involving transcriptional repressors or activators. Because this is primarily proteins that matter for the functions, it must be kept in mind that the relationship between mRNA and protein abundances only very partially correlates; mRNA levels are just an excellent proxy for the presence of a protein but is not directly proportionate with the increase or decrease folds of protein expression and even less with its activity when we consider an enzyme for instance (Vogel *et al.* 2012). Using key examples from different CFs presented here above and involved in bacterial adhesion and biofilm formation in DEC, these different regulatory levels were aimed at being highlighted.

2.1. Regulations at pre-transcriptional level: the phase variation

Prior to transcription, some regulatory mechanisms can already be at work at the DNA level, i.e. the phase variation. Four main mechanisms of phase variation can be considered, (i) DNA inversion, (ii) slipped-strand mispairing, (iii) DNA methylation, and (iv) DNA deletion (Henderson *et al.* 1999). As a commonality, all these regulatory mechanisms primarily occur at the stage of DNA replication and a large majority of genes regulated by phase variations are actually bacterial cell surface molecular determinants (Owen *et al.* 1996, Holden *et al.* 2004).

In *E. coli*, T1P are well-known to be subjected to phase variation following DNA inversion (Blomfield 2001). The expression of the *fim* operon is under the control of the *fim*

promoter, which is located within the *fimS*-invertible element (Abraham *et al.* 1985, Wright *et al.* 2007). The orientation of the promoter determines the ON or OFF phase and then induces the expression of upstream genes or not. Two tyrosine recombinases, FimB and FimE, are known to control the orientation of the *fimS*-invertible region. FimB predominantly switches *fim* operon transcription from OFF to ON, while FimE mediates ON to OFF phase switching (Klemm 1986, Gally *et al.* 1996, Hannan *et al.* 2008). The relevance of this regulation in DEC has not been questioned as yet.

Slipped-strand mispairing occurs in the course of DNA replication in repetitive DNA regions, which can be positioned either upstream of a CDS (coding DNA sequence) and then influence the transcription, such as the promoter efficiency, or within a CDS and then affect the translational reading frame resulting in a mutation frameshift (Henderson *et al.* 1999). In *E. coli*, phase variation resulting from strand-slippage has never been reported as yet, nonetheless, there is no molecular mechanistic constraint for it to occur in *E. coli* (Torres-Cruz *et al.* 2003).

Phase variation resulting from DNA methylation truly corresponds to a bacterial epigenetic mechanism (Henderson *et al.* 1999). The Ag43 is probably one of most investigated surface protein subjected to such a regulatory mechanisms (van der Woude *et al.* 2008). This epigenetic regulation involves two proteins, the DNA adenine methylase Dam and the transcriptional regulator OxyR (van der Woude *et al.* 2008). When Dam has methylated the GATC sites present in the operator region in the course of DNA replication, then the repressor OxyR cannot bind and the transcription by the RNA polymerase occurs and Ag43 is finally expressed (ON phase); however, if OxyR binds the GATC sites before they are methylated by Dam, there is no transcription and no Ag43 expression (OFF phase). Besides Ag43, several pili secreted and assembled by the T7SS were reported to be subjected to such an epigenetic regulation in *E. coli* (Henderson *et al.* 1999, Blomfield 2001). The *pap* (pyelonephritis-associated pilus) operon in UPEC is considered as a paradigm where the Dam methylation of a proximal GATC site in the operator region prevent binding of the repressor Lrp (leucine-responsive regulatory protein), consequently the *papBA* operon is transcribed and the pili expressed (ON phase); at the opposite, in the absence of methylation at the proximal GATC, Lrp can bind to the operator, repress the transcription and ultimately prevent pili formation (OFF phase). Besides, this repression can be lifted when Lrp binds to another site called distal GATC. Among DEC, CS31A pili are subjected to this same regulation mechanism (Crost *et al.* 2003, Graveline *et al.* 2014).

As a general trend, phase variation due to DNA deletion is irreversible due to the loss of the genetic element bearing the gene of interest. In *E. coli*, DNA deletion is at work for unilateral flagellar phase variation as reported in the H3, H47 and H17 strains (Zhou *et al.* 2015). While most flagellins are encoded by *fliC* in *E. coli*, H3 and H47 are encoded by *flkA* and H17 by *flnA*. For H3 and H47, their production results from the expression of *flkAB* operon, where the transcriptional regulator FlkB represses *fliC* (Feng *et al.* 2008). Upon excision of the *flk* region from the chromosome, *flkAB* is irreversibly deleted, the repression of *fliC* is released and the FliC flagellin is produced. Similarly, H17 strain can irreversibly switch flagellar antigens to H4 (Ratiner Iu 1967). It appeared this flagellar phase variation was caused by excision of *flnA* (Liu *et al.* 2012). When *flnA* is present in the chromosome, the translation of FliC H4 is inhibited and only FlnA H17 is produced; once *flnA* is excised, the repression of the *fliC* is released and only the FliC H4 is produced. The ~35 kb DNA deletion region containing *flnA* gene is excised as a covalently closed extrachromosomal circular form. While some DNA deletion can occur through homologous recombination (Henderson *et al.* 1999), flagellar phase variation in *E. coli* would be mediated by non-homologous recombination via integrase of the tyrosine recombinase family (Feng *et al.* 2008). The flagellar phase variation mechanisms in some other *E. coli* H variants and especially in DEC remain to be defined.

2.2. Regulations at transcriptional level: regulators and effectors

Regulation at the transcriptional level is the most well-known level of gene regulation and unfortunately the only one really considered most of the time to apprehend protein expression *in fine*. Transcriptional regulators can be of two types, either repressor or activator. It must be stressed, this is wrong to assume a repressor will systematically repress transcription or an activator will activate transcription. In fact, a second important partner to the process must be considered, that is the effector, which can be of two types, either an inducer or a co-repressor. In the end, four possibilities for the regulation at transcriptional level can be discriminated, (i) positive control of an inducible gene, where an activator is activated by an inducer, (ii) positive control of a repressible gene, where an activator is inactivated by an inhibitor, (iii) negative control of an inducible gene, where a repressor is inactivated by an inducer, or (iv) negative control of a repressible gene, where a repressor is activated by a co-repressor. Besides, a so-called repressor can act as an activator for some genes and vice versa. In other words, the up-expression or down-expression of a regulator is

not sufficient to know what kind of regulation is taking place without knowing the nature and level of the inducer. Bacteria can sense and response to environmental cues thanks to a large range of two-component signal transduction systems where a sensor activates a transcriptional regulator, which further represses or activates gene expression (Hoch 2000, Zschiedrich *et al.* 2016). Some of these systems actually participate to cell-to-cell communication (CTCC) via signal molecule called auto-inducer (AI) (Bassler 2002). Of note, quorum sensing (QS) is only one of the different function for CTCC, which specifically refers to the sensing of the cell density (quorum); QS must not be considered synonymous to CTCC because some sensing can be unrelated to a QS *sensu stricto* but to diffusion sensing, confinement or efficiency sensing for instance (Redfield 2002, Platt *et al.* 2010, West *et al.* 2012). Nonetheless, they all belong to bacterial CTCC system in the end. This semantic issue is especially of importance in biofilm formation since by definition bacteria cells are at a high density following sessile development and therefore the notion of QS makes no or few sense. Transcriptional regulators of virulence and CFs are subjected to intense and extensive research and scientific literature in DEC (Beloin *et al.* 2008, Tobe 2008, Pruss 2017, Rossi *et al.* 2018). For these reasons only some key examples will be provided to make the point of differentiating the regulations at these different levels.

At transcriptional level, PNAG production is regulated by NhaR, a transcriptional regulator of the LysR family, which activates the transcription of the *pgaABCD* operon by binding to two sites near the -35 region of the promoter [55]. Among EPS, the production of colanic acid is consistently upregulated within biofilms by the RcsA transcriptional activator (Matthysse *et al.* 2008, May *et al.* 2008). The transcription of the *wca* operon is regulated by the *rcsABCF* locus that encodes a two-component system (Gervais *et al.* 1992, Ebel *et al.* 1999, Beloin *et al.* 2008). However, the signal sensed by the RcsC sensor kinase remains unknown (Whitfield *et al.* 1999, Oropeza *et al.* 2015). The H-NS is known to act as transcriptional repressor in bacteria, a so-called bacterial transcriptional silencing to underline the reminiscence with the eukaryotic silencing by histones (Landick *et al.* 2015, Grainger 2016). While RcsA is present at a low amount in the cell, it was found to be partially due to transcriptional silencing by H-NS (Sledjeski *et al.* 1995). Cellulose synthesis is under control of the CsgD transcriptional regulator (Romling *et al.* 2000, Zorraquino *et al.* 2013). Interestingly in EIEC, *csgD* expression is often inactivated (Sakellaris *et al.* 2000), suggesting that biofilm formation interferes with pathogenesis and making them poor biofilm former.

While no specific transcriptional regulators could be identified for expression of AIDA-I, it was shown that transcription of both genes was enhanced in the absence of H-NS and RfaH transcriptional regulators [155]. Similarly, the transcription of *ehaG* is negatively regulated by H-NS (Totsika *et al.* 2012). While FdeC is transcriptionally regulated by H-NS (Easton *et al.* 2014), *yeeJ* transcription is increased in absence of the mRNA regulator PNPase (Martinez-Gil *et al.* 2017).

CS31A synthesis is dramatically reduced in media containing alanine or leucine, suggesting they can play a role as effectors (Crost *et al.* 2003). Actually, the ON/OFF switch is locked in the OFF phase by alanine, whereas leucine repressed transcription but without affecting the switch frequency. Analysis of *clp* expression indicated that alanine and leucine could repress the *clp* transcription by a methylation-independent mechanism but also by either promoting methylation or methylation protection of GATC-II and GATC-I respectively, which increased the methylation pattern characteristic of repressed cells. Furthermore, alanine prevented the AfaF-dependent methylation protection and thus the appearance of cells in ON phase. Additional regulatory proteins, including ClpB, cAMP, receptor protein (CRP) and H-NS, also play important roles in the expression of the operons of the *pap* family but regulations by phase variation have been reported from one *E. coli* strain to another (Blomfield *et al.* 2007). For the T4P in EPEC, the expression of the *bfp* operon is controlled by the BfpT (also called PerA) transcriptional regulator, a member of the AraC family, encoded on the enteroadherence factor plasmid (Tobe *et al.* 1992, Gomez-Duarte *et al.* 1995). The expression of CFA/I is positively regulated by CfaR, whereas for the expression of CFA/II, CS1 and CS2 is positively regulated by the *rns* gene product (a homogue to *cfaR* with 96% identity) (Caron *et al.* 1989, Caron *et al.* 1990, Savelkoul *et al.* 1990). The expression of AAF is induced by the transcriptional activator AggR (an homolog of AraC) located on pAA too (Nataro *et al.* 1994); YafK and Fis were also reported to regulate AAF/II transcription (Sheikh *et al.* 2001). From a transcriptional regulatory point of view, *lpf1* is repressed by H-NS and activated by Ler in response to different environmental conditions (Torres *et al.* 2007, Rojas-Lopez *et al.* 2011). The *lpf2* transcription appeared to be activated by Fur (Torres *et al.* 2007). Regulation of curli biogenesis is complex and involves several two-component systems, such as EnvZ/OmpR, CpxA/CpxR or CpxR/H-NS/RstA/IHF/OmpR (Vidal *et al.* 1998, Prigent-Combaret *et al.* 2000, Prigent-Combaret *et al.* 2001, Beloin *et al.* 2008, Ogasawara *et al.* 2010, Laverty *et al.* 2014). In EPEC, Fis (factor for inversion stimulation) has been identified as a negative transcriptional regulator of *csgA* expression (Saldana *et al.*

2009). Their expression can be trigger by a marge range of environmental signals such as the temperature, osmolarity or redox potential (Olsen *et al.* 1993, Prigent-Combaret *et al.* 1999, Gerstel *et al.* 2001, Evans *et al.* 2014).

The transcriptional regulatory control of the locus of enterocyte effacement (LEE) encoding the injectisome is undoubtedly one of the most extensively investigated in DEC, namely EPEC and EHEC (Schmidt 2010, Stevens *et al.* 2014, Franzin *et al.* 2015). Regarding knowledge about the complex regulation network by specific, global and phage encoded regulators as well as environmental signals such as nutrient sources or metabolic products from the host or microbiota that can affect the transcription of the LEE-encoded genes readers can refer to specific and recent reviews (Connolly *et al.* 2015, Furniss *et al.* 2018, Platenkamp *et al.* 2018, Turner *et al.* 2018).

2.3. Regulations at post-transcriptional level

At least three main regulation mechanisms can take place post-transcriptionally, (i) the stability of mRNA, which can be quantified by determining its half-life, (ii) the riboswitch, where molecule such as a metabolite can change the fold of an mRNA with the formation of a termination hairpin that stops the on-going transcription by the RNA polymerase, or (iii) the attenuation based on the formation of terminator/anti-terminator loops, which couple or uncouple the transcription by the RNA polymerase with the translation of the mRNA. Post-transcriptional regulations are important regulatory mechanisms that are generally overlooked and underestimated, most certainly because they cannot be easily investigated and estimated by transcriptomic analysis on its own (Vogel *et al.* 2012).

Recently, it was shown that the expression level of *agn43* is also controlled by antitermination of transcription and translation initiation in the leader mRNA (Wallecha *et al.* 2014). Among EPS determinants, PNAG production is regulated by the RNA-binding protein CsrA (carbon storage regulatory protein A) post-transcriptionally (Boles *et al.* 2011, Wang *et al.* 2017). Actually, CsrA that binds cooperatively to the *pgaA* mRNA and competes with for recognition with the 30S ribosomal subunit. By binding to sites located in the mRNA leader, CsrA would further destabilise the *pgaA* transcript.

Pili produced by the *pap* operon appeared to be regulated post-transcriptionally as a result of differential mRNA stability (Baga *et al.* 1988). The *papBA* transcript appeared to be

processed and the resulting mRNA encoding the major pilin subunit accumulates. The difference in abundance of the two mRNA species was readily explained by differences in their half-life. In *E. coli*, RNA degradation occurs via the degradosome thanks to the combination of endoribonuclease and exoribonuclease activities (Burger *et al.* 2011, Bandyra *et al.* 2013).

2.4. Regulations at translational level

While the attenuation collaterally affects the translation, three main mechanisms can be directly involved in the regulation of the translation, (i) the anti-sense RNAs (including the small RNAs), which hybridises with mRNA and thus block the binding of the ribosome, (ii) the riboregulation, where a ligand changes the fold of the mRNA, which consequently prevents the binding of the ribosome, and (iii) the translational efficiency depending on the codon usage.

PNAG synthesis is also regulated by two small RNAs CsrB and CsrC, which actually sequester CsrA and thus activate the translation of the *pgaABCD* transcript (Liu *et al.* 1997, Weilbacher *et al.* 2003). For colanic acid production, the low level of expression from the *rcsA* promoter by H-NS transcriptional silencing is alleviated by the DsrA small RNA (Sledjeski *et al.* 1995).

In *E. coli*, OmpA protein is expressed to very high levels, is growth rate dependent and is a paradigm for riboregulation (Lugtenberg *et al.* 1976, Koebnik *et al.* 2000). Actually, the mRNA half-life increases proportionally with the bacterial growth rate (Nilsson *et al.* 1984). While a specific region of the transcript is targeted by the RNaseE (endoribonuclease E), binding of the ribosome induces conformational changes that mitigate the mRNA degradation (Emory *et al.* 1990, Emory *et al.* 1992, Hansen *et al.* 1994). As an antagonist, Hfq can bind the transcript to decrease its stability and thus inducing RNA decay (Nilsson *et al.* 1984, Vytvytska *et al.* 2000). Actually, Hfq facilitates the binding of a small RNA called MicA in the vicinity of the ribosome-binding site thus preventing ribosomal recruitment (Udekwu *et al.* 2005).

2.5. Regulations at post-translational level

Regulations at post-translational level concentrate the most diverse range of molecular mechanisms and are hierarchically the most important (Figure 3). In metabolic pathways, regulations at post-translational levels are key mechanisms, especially with the modulation of the enzymatic activity, which can be influenced by physical parameters (pH, temperature, ionic force, redox, etc…), inducers and inhibitors (irreversible or reversible: competitive, non-competitive, incompetitive or mixed inhibition); retro-inhibition and pro-activation can further occur and even involve allosteric enzymes. Protein activity can be further mitigated by numerous post-translational modifications, namely (i) proteolytic cleavage, and (ii) chemical modifications such as disulphide bond, phosphorylation, acetylation, methylation, adenylation or uridylation. Post-translational regulations also include the protein folding, association/dissociation of homo- and heteromers, the degradation of proteins following the N-terminal rule by the ClpAP proteolytic complex, which influence the protein half-life, as well as the protein translocation to a final subcellular location. Indeed, the maturation of a protein can further occur at translocational and post-translocational levels.

At 37°C, the decrease production of colanic acid results from the degradation of the RcsA transcriptional activator by the Lon protease (Ebel *et al.* 1999). This post-translational regulation alleviates the *wca* transcription and explains the low amount of RcsA in cell (Sailer *et al.* 2003). As a two-component system, the RcsA regulator is activated by the transfer of a phosphate group from the RcsC sensor, which is per se another post-translational regulation level (Desai *et al.* 2017). For cellulose biosynthesis, the catalytic activity of BcsA-B complex using UDP-glucose as substrate is allosterically controlled by cyclic-di-GMP (c-di-GMP) on the PilZ domain of the catalytic BcsA [74]. Furthermore, the diguanylate cyclase AdrA regulates the c-di-GMP production (Romling *et al.* 2000, Zorraquino *et al.* 2013). C-di-GMP is an ubiquitous second messenger produced by diguanylate cyclase and degraded by phosphodiesterases. This molecule controls the motility and virulence of planktonic cells as well as cell adhesion and persistence of multicellular communities [41, 75, 76].

As an autotransporter, Ag43 exhibits a signal peptide driving the preprotein to the Sec export system for translocation across the CM before being cleaved off after translocation into the periplasm. In the periplasm, several chaperones can participate to the folding prior to the translocation across the OM through a cooperative mechanism involving the translocation assembly (TAM) and β-barrel assemby (BAM) machineries [128]. Besides, the passenger of

Ag43 is glycosylated, which stabilises its conformation (Sherlock *et al.* 2006). These different post-translational, translocational and post-translocational levels truly participate to the regulation of the expression of this surface protein. While glycosylation is not that important for the functions of Ag43 (Reidl *et al.* 2009), in TibA it is necessary for autoaggregation, adhesion to epithelial cells and biofilm formation (Cote *et al.* 2013).

Conclusion and perspectives

Reviewing the extremely various cell-surface molecular determinants that can participate to the surface colonisation process in DEC, from bacterial adhesion to biofilm formation, the wealth of CFs at play was highlighted. While some of these molecular determinants still remain to be better characterised, their interplay in surface colonisation must also be carefully considered and kept in mind. The flagella as force-generating cell-surface organelles were demonstrated to be important for biofilm formation (Hobley *et al.* 2015), but expression of strong adherence factors could replace motility in early stages of biofilm formation (Pratt *et al.* 1998, Donlan 2002). For instance, curli or conjugative plasmid can compensate motility for initial adhesion and biofilm development (Prigent-Combaret *et al.* 2000, Ghigo 2001, Reisner *et al.* 2003, Beloin *et al.* 2008). Although flagella expression is repressed during the switch from planktonic to sessile lifestyle to reduce the motility capacity of the bacteria, these surface organelles have a structural and architectural role in the EPM (Hung *et al.* 2013, Serra *et al.* 2013). While expression of flagellar genes are repressed, genes involved in biosynthesis of the EPM components are activated during the biofilm maturation step (Guttenplan *et al.* 2013).

In EAEC, optimal colonisation results from interactions through multiple adhesion factors, including AafA, and where the Aap (antiaggregation protein) dispersin masks the Hra1 aggregative factor (Blanton *et al.* 2018). Besides autoaggregation, Ag43 was reported to co-aggregate with AIDA-I and TibA in an *E. coli* K12 background (Klemm 2006, Trunk *et al.* 2018). Still in *E. coli* K12, T1P and capsule polysaccharide appeared to block the autoaggregation mediated by Ag43 by physically shielding intercellular Ag43-Ag43 interaction (Hasman *et al.* 1999, Schembri *et al.* 2004), whereas, in turn, the autoaggregation overrides bacterial motility (Ulett *et al.* 2006). CP can functionally replace other adhesion factors including Ag43 or some other pili (Reisner *et al.* 2003). However, this interplay has never been reported as yet in any DEC.

As a general trend, the average number of pili types appears lower in commensal than in pathogenic *E. coli* (Spurbeck *et al.* 2011). Many regulators are involved in the regulation of the expression of T1P, such as Lsr, CRP, H-NS, themselves regulated by other transcriptional regulators (Kawula *et al.* 1991, van der Woude *et al.* 1992, Kostakioti *et al.* 2009, Muller *et al.* 2009, Hadjifrangiskou *et al.* 2011). In some ExPEC, T1P expression appeared to be further modulated and influenced by OmpA (Teng *et al.* 2006). The inversion the promoter-containing element upstream of the *fim* operon is predominantly in OFF phase in *ompA* mutants. At the opposite, *ompX* mutant upregulates the production of T1P together with the increase of exopolysaccharide production as well as the decrease of the motility of bacteria (Otto *et al.* 2004). This suggests the level and composition of OMPs in the OM may represent a signal for physiological adaptation of bacteria to adhere to the surface, which research direction that has not been explored as yet in DEC. In NMEC, OmpA would act together with Hek in invasion of epithelial cells (Smith *et al.* 2007, Fagan *et al.* 2008).

While conjugative plasmids could take part to biofilm formation (Ghigo 2001), plasmids in general can encode numerous colonisation factors as shown in ETEC or EAEC (Amabile-Cuevas *et al.* 1996, Mainil *et al.* 1998, Ghigo 2001, Molin *et al.* 2003, Kaper *et al.* 2004, Wuertz *et al.* 2004, Beloin *et al.* 2008, Ong *et al.* 2009). While they could confer initial adhesion capacity and modulate the biofilm architecture (Ghigo 2001, Wuertz *et al.* 2004), the genetic mobility of this extrachromosomal gene pool and contribution to biofilm formation remain poorly investigated in DEC (Dudley *et al.* 2006).

Regulatory network for the production of colonisation factors is often depicted as restricted to the transcriptional level. However, this review stressed that the range of regulation levels is much broader and even more complex (Figure 3). It can indeed occur at pre-transcriptional, post-transcriptional, translational and/or post-translation levels as exemplified here above, e.g. for PNAG, colanic acid, curli or Ag43. With such a variety of regulation levels comes the question about their hierarchy in the regulatory network. As a general trend, it is important to stress and keep in mind that the primary regulation level is post-translational and not transcriptional as sometimes wrongly assumed. Indeed, whenever DNA replication, RNA polymerisation or protein synthesis occur, enzymes are basically and primarily required for these regulation processes; any abrupt changes in the environmental conditions, such as some physico-chemical parameters (e.g. pH, temperature, redox potential), will have a first and direct effect on the enzyme activity before the cell could have even time to change its transcription profile.

For the CFs, the interplay taking place at the other different regulation levels is extremely complex and their hierarchy is still extremely difficult to establish at a global scale. Besides, some regulation mechanisms in the expression of CFs in DEC have not been fully considered, e.g. attenuation, riboswitch or translational efficiency, but their involvement cannot be excluded. As already mentioned, non-coding small RNAs can establish antisense interactions with mRNA but they can also antagonising RNA-binding protein (Valverde *et al.* 2008). It is also known that the expression of the small RNAs can be itself controlled by two-component systems. As molecular cell-surface determinants, the CFs in DEC need to be translocated across a LPS-diderm bacterial cell envelope to be functional and active, which further involve translocational and post-translocational regulation levels that should not be overlooked in a regulatory network. In the end, our view of the regulatory network for the production of CFs in *E. coli* remains quite partial, scattered and still far from an integrated view of all regulation mechanisms. Besides, findings from investigations using domesticated strains of *E. coli* must be interpreted with caution and reinvestigation in DEC genetic backgrounds can be wise (Hobman *et al.* 2007). This is the promise of new discoveries in the field in the years to come.

In DEC, CFs have often been pondered for their contribution to bacterial virulence and thus investigated in conditions related to human infection (Nataro *et al.* 1998, Kaper *et al.* 2004, Rossi *et al.* 2018). In addition to human, the GIT of a wide range of animals is also harbouring some *E. coli* strains, both commensal and pathogenic (Escobar-Paramo *et al.* 2006, Croxen *et al.* 2013, Smati *et al.* 2015, Torres 2017). Following shedding from these animal reservoirs, *E. coli* is also found in the environment. Outside the host, the range of extraintestinal environmental conditions that can be encountered by this species is wide from soil, water to plant, as well as food matrices and food processing sites (van Elsas *et al.* 2011, Giaouris *et al.* 2014, Jang *et al.* 2017). As foodborne zoonotic pathogens, understanding the ecophysiology of DEC necessitate to consider its lifestyle outside the human host. As for virulence factors, which are nowadays rather replaced by the concept of coincidental by-products of commensalism (Le Gall *et al.* 2007, Diard *et al.* 2010, Leimbach *et al.* 2013) or niche factors (Hill 2012), the notion of CFs must be place in a context much broader than the colonisation of the GIT since they can also participate to the colonisation of other environmental niches. Being minded focus on the physiopathology and GIT environment, this could actually bias and divert the view from the true relevance of the wide diversity of CFs in *E. coli*.

Besides their involvement in the binding to mucin in the animal intestine, HCP and ECP together with flagella can further participate in the formation of stable biofilms on abiotic surface (Farfan *et al.* 2012). While the injectisome was primarily considered as involved in the physiopathology of EPEC and EHEC, it was also reported as clearly involved in the adhesion to stomata of salad leaves (Shaw RK 2008, Berger CN 2010). A similar observation was made for flagella in ETEC (Shaw *et al.* 2011) or AAF and flagella in EAEC, where they were reported to be required for epidermis and stomatal binding to salad leaves, respectively (Berger *et al.* 2009). In human and animal cutaneous pathogens *Erysipelothrix rhusiopathiae*, the RspA (rhusiopathiae surface protein A) and RspB surface proteins were demonstrated to specifically bind several ECM components, namely fibronectin, collagens I and IV, but also polystyrene shedding light on the ecophysiology of this microorganism through its binding ability to adhere to both biotic and abiotic surfaces (Shimoji *et al.* 2003). The well-correlated biosynthesis of cellulose and curli was proposed as a mechanism for bacterial survival against various environmental stresses (Gualdi *et al.* 2008).

In *Pseudomonas aeruginosa*, T4P were primarily regarded as involved in the attachment of epithelial cells in the course of an infection but later appeared to also bind abiotic surfaces such as polyvinyl chloride, polystyrene and stainless steel (Giltner *et al.* 2006). For the later, it even appeared to exhibit an extremely much higher affinity towards steel than mucosal epithelial surface, which question the relevance of T4P in environmental over infection conditions in the course of evolution (Yu *et al.* 2007, Burgess *et al.* 2014). These aspects have not been reported or questioned as yet in DEC but are particularly relevant considering the presence of T4P and ECM-binding proteins, especially some ATs, in the various *E. coli* enteropathovars.

All-in-all, the physiology of DEC should not only be considered downstream, i.e. at the stage of human infection, but also upstream, i.e. along the food chain. Investigating the ecophysiology of the DEC respective to the various biotopes and biocoenoses encountered in different ecosystems from natural environments, animal reservoirs, food matrices, food-processing environments, to human ingestion should shed new light on the relevance and contribution of the CFs of this species.

Acknowledgments

This work was supported in part by INRA (Institut National de la Recherche Agronomique), by EU FP7 (European Union 7th Framework Programme) Marie Curie Actions ITN (Initiale Training Network) EID (European Industrial Doctorate) DIS*Co* (A multidisciplinary doctoral industrial school on novel preventive strategies against *E. coli* infections) project (n°FP7-PEOPLE-607611) and by the RA-FEDER-Bpifrance (Région Auvergne-Fonds Européen de Développement Régional-Banque Publique d'Investissement des entrepreneurs français) FRI (Fonds Régional Innovation) Cluster IRP (Institut de Recherche Pharmabiotique) CoMBa (Colonisation de la matrice par les bactéries) project (n°AV0003483 and n°DOS0019690/00). VA is a PhD research fellow granted by the RA-FEDER-Bpifrance FRI IRP CoMBa. RM was a Marie Curie PhD research fellow granted by the ITN EID DIS*Co*. All authors declare that the work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Authors contribution statement

VA and MD wrote the first overall draft of the manuscript and draw the original pictures; RM, SL, MP, CP, and FCD wrote sections of the manuscript. MD contributed to conceptualise the overarching aims and had management as well as coordination responsibility for the execution of the work. MD, MP, CP and FCD contributed to the acquisition of the financial supports and resources leading to this publication. All authors contributed to the critical revision of the manuscript, read and approved the submitted version.

Conflict of interest statement

MP is permanent employee of GSK. CP is permanent employee of BioFilm Control. FCD is permanent employee of Lallemand. These authors declare that their respective company provided support in the form of salaries. All other authors declare no competing interests.

References

Abraham, S. N., J. P. Babu, C. S. Giampapa, D. L. Hasty, W. A. Simpson and E. H. Beachey (1985). "Protection against Escherichia coli-induced urinary tract infections with hybridoma antibodies directed against type 1 fimbriae or complementary D-mannose receptors." Infect Immun **48**(3): 625-628.

Abraham, S. N., D. Sun, J. B. Dale and E. H. Beachey (1988). "Conservation of the D-mannose-adhesion protein among type 1 fimbriated members of the family Enterobacteriaceae." Nature **336**(6200): 682-684.

Abreu, A. G., C. M. Abe, K. O. Nunes, C. T. Moraes, L. Chavez-Duenas, F. Navarro-Garcia, A. S. Barbosa, R. M. Piazza and W. P. Elias (2016). "The serine protease Pic as a virulence factor of atypical enteropathogenic Escherichia coli." Gut Microbes **7**(2): 115-125.

Abreu, A. G., T. R. Fraga, A. P. Granados Martinez, M. Y. Kondo, M. A. Juliano, L. Juliano, F. Navarro-Garcia, L. Isaac, A. S. Barbosa and W. P. Elias (2015). "The Serine Protease Pic From Enteroaggregative Escherichia coli Mediates Immune Evasion by the Direct Cleavage of Complement Proteins." J Infect Dis **212**(1): 106-115.

Abu-Median, A. B., P. M. van Diemen, F. Dziva, I. Vlisidou, T. S. Wallis and M. P. Stevens (2006). "Functional analysis of lymphostatin homologues in enterohaemorrhagic Escherichia coli." FEMS Microbiol Lett **258**(1): 43-49.

Adams, L. M., C. P. Simmons, L. Rezmann, R. A. Strugnell and R. M. Robins-Browne (1997). "Identification and characterization of a K88- and CS31A-like operon of a rabbit enteropathogenic Escherichia coli strain which encodes fimbriae involved in the colonization of rabbit intestine." Infect Immun **65**(12): 5222-5230.

Aguilera, L., E. Ferreira, R. Gimenez, F. J. Fernandez, M. Taules, J. Aguilar, M. C. Vega, J. Badia and L. Baldoma (2012). "Secretion of the housekeeping protein glyceraldehyde-3-phosphate dehydrogenase by the LEE-encoded type III secretion system in enteropathogenic Escherichia coli." Int J Biochem Cell Biol **44**(6): 955-962.

Ahmed, D., M. S. Islam, Y. A. Begum, A. Janzon, F. Qadri and A. Sjoling (2013). "Presence of enterotoxigenic Escherichia coli in biofilms formed in water containers in poor households coincides with epidemic seasons in Dhaka." J Appl Microbiol **114**(4): 1223-1229.

Al Safadi, R., G. S. Abu-Ali, R. E. Sloup, J. T. Rudrik, C. M. Waters, K. A. Eaton and S. D. Manning (2012). "Correlation between in vivo biofilm formation and virulence gene expression in Escherichia coli O104:H4." PLoS One **7**(7): e41628.

Allsopp, L. P., C. Beloin, G. C. Ulett, J. Valle, M. Totsika, O. Sherlock, J. M. Ghigo and M. A. Schembri (2012). "Molecular characterization of UpaB and UpaC, two new autotransporter proteins of uropathogenic Escherichia coli CFT073." Infect Immun **80**(1): 321-332.

Amabile-Cuevas, C. F. and M. E. Chicurel (1996). "A possible role for plasmids in mediating the cell-cell proximity required for gene flux." J Theor Biol **181**(3): 237-243.

Amimanan, P., R. Tavichakorntrakool, K. Fong-Ngern, P. Sribenjalux, A. Lulitanond, V. Prasongwatana, C. Wongkham, P. Boonsiri, J. Umka Welbat and V. Thongboonkerd (2017). "Elongation factor Tu on Escherichia coli isolated from urine of kidney stone patients promotes calcium oxalate crystal growth and aggregation." Sci Rep **7**(1): 2953.

Amor, K., D. E. Heinrichs, E. Frirdich, K. Ziebell, R. P. Johnson and C. Whitfield (2000). "Distribution of core oligosaccharide types in lipopolysaccharides from Escherichia coli." Infect Immun **68**(3): 1116-1124.

Anderson, K. L., J. Billington, D. Pettigrew, E. Cota, P. Simpson, P. Roversi, H. A. Chen, P. Urvil, L. du Merle, P. N. Barlow, M. E. Medof, R. A. Smith, B. Nowicki, C. Le

Bouguenec, S. M. Lea and S. Matthews (2004). "An atomic resolution model for assembly, architecture, and function of the Dr adhesins." Mol Cell **15**(4): 647-657.

Andersson, E. K., C. Bengtsson, M. L. Evans, E. Chorell, M. Sellstedt, A. E. Lindgren, D. A. Hufnagel, M. Bhattacharya, P. M. Tessier, P. Wittung-Stafshede, F. Almqvist and M. R. Chapman (2013). "Modulation of curli assembly and pellicle biofilm formation by chemical and protein chaperones." Chem Biol **20**(10): 1245-1254.

Andrade, F. B., A. G. Abreu, K. O. Nunes, T. A. T. Gomes, R. M. F. Piazza and W. P. Elias (2017). "Distribution of serine protease autotransporters of Enterobacteriaceae in typical and atypical enteroaggregative Escherichia coli." Infect Genet Evol **50**: 83-86.

Appelmelk, B. J., Y. Q. An, T. A. Hekker, L. G. Thijs, D. M. MacLaren and J. de Graaf (1994). "Frequencies of lipopolysaccharide core types in Escherichia coli strains from bacteraemic patients." Microbiology **140 (Pt 5)**: 1119-1124.

Arnqvist, A., A. Olsen and S. Normark (1994). "Sigma S-dependent growth-phase induction of the csgBA promoter in Escherichia coli can be achieved in vivo by sigma 70 in the absence of the nucleoid-associated protein H-NS." Mol Microbiol **13**(6): 1021-1032.

Aschtgen, M. S., C. S. Bernard, S. De Bentzmann, R. Lloubes and E. Cascales (2008). "SciN is an outer membrane lipoprotein required for type VI secretion in enteroaggregative Escherichia coli." J Bacteriol **190**(22): 7523-7531.

Aschtgen, M. S., M. Gavioli, A. Dessen, R. Lloubes and E. Cascales (2010). "The SciZ protein anchors the enteroaggregative Escherichia coli Type VI secretion system to the cell wall." Mol Microbiol **75**(4): 886-899.

Azeredo, J., N. F. Azevedo, R. Briandet, N. Cerca, T. Coenye, A. R. Costa, M. Desvaux, G. Di Bonaventura, M. Hebraud, Z. Jaglic, M. Kacaniova, S. Knochel, A. Lourenco, F. Mergulhao, R. L. Meyer, G. Nychas, M. Simoes, O. Tresse and C. Sternberg (2017). "Critical review on biofilm methods." Crit Rev Microbiol **43**(3): 313-351.

Baga, M., M. Goransson, S. Normark and B. E. Uhlin (1988). "Processed mRNA with differential stability in the regulation of E. coli pilin gene expression." Cell **52**(2): 197-206.

Baldi, D. L., E. E. Higginson, D. M. Hocking, J. Praszkier, R. Cavaliere, C. E. James, V. Bennett-Wood, K. I. Azzopardi, L. Turnbull, T. Lithgow, R. M. Robins-Browne, C. B. Whitchurch and M. Tauschek (2012). "The type II secretion system and its ubiquitous lipoprotein substrate, SslE, are required for biofilm formation and virulence of enteropathogenic Escherichia coli." Infect Immun **80**(6): 2042-2052.

Bandyra, K. J., M. Bouvier, A. J. Carpousis and B. F. Luisi (2013). "The social fabric of the RNA degradosome." Biochim Biophys Acta **1829**(6-7): 514-522.

Barrios, A. F., R. Zuo, D. Ren and T. K. Wood (2006). "Hha, YbaJ, and OmpA regulate Escherichia coli K12 biofilm formation and conjugation plasmids abolish motility." Biotechnol Bioeng **93**(1): 188-200.

Bassler, B. L. (2002). "Small talk. Cell-to-cell communication in bacteria." Cell **109**(4): 421-424.

Baumler, A. J. and F. Heffron (1995). "Identification and sequence analysis of lpfABCDE, a putative fimbrial operon of Salmonella typhimurium." J Bacteriol **177**(8): 2087-2097.

Baumler, A. J., R. M. Tsolis and F. Heffron (1996). "Contribution of fimbrial operons to attachment to and invasion of epithelial cell lines by Salmonella typhimurium." Infect Immun **64**(5): 1862-1865.

Beghain, J., A. Bridier-Nahmias, H. Le Nagard, E. Denamur and O. Clermont (2018). "ClermonTyping: an easy-to-use and accurate in silico method for Escherichia genus strain phylotyping." Microb Genom.

Beloin, C., A. Houry, M. Froment, J. M. Ghigo and N. Henry (2008). "A short-time scale colloidal system reveals early bacterial adhesion dynamics." PLoS Biol **6**(7): e167.

Beloin, C., K. Michaelis, K. Lindner, P. Landini, J. Hacker, J. M. Ghigo and U. Dobrindt (2006). "The transcriptional antiterminator RfaH represses biofilm formation in Escherichia coli." J Bacteriol **188**(4): 1316-1331.

Beloin, C., A. Roux and J. M. Ghigo (2008). "Escherichia coli biofilms." Curr Top Microbiol Immunol **322**: 249-289.

Beloin, C., J. Valle, P. Latour-Lambert, P. Faure, M. Kzreminski, D. Balestrino, J. A. Haagensen, S. Molin, G. Prensier, B. Arbeille and J. M. Ghigo (2004). "Global impact of mature biofilm lifestyle on Escherichia coli K-12 gene expression." Mol Microbiol **51**(3): 659-674.

Ben Nasr, A., A. Olsen, U. Sjobring, W. Muller-Esterl and L. Bjorck (1996). "Assembly of human contact phase proteins and release of bradykinin at the surface of curli-expressing Escherichia coli." Mol Microbiol **20**(5): 927-935.

Benz, I. and M. A. Schmidt (1989). "Cloning and expression of an adhesin (AIDA-I) involved in diffuse adherence of enteropathogenic Escherichia coli." Infect Immun **57**(5): 1506-1511.

Benz, I. and M. A. Schmidt (1992). "AIDA-I, the adhesin involved in diffuse adherence of the diarrhoeagenic Escherichia coli strain 2787 (O126:H27), is synthesized via a precursor molecule." Mol Microbiol **6**(11): 1539-1546.

Benz, I. and M. A. Schmidt (1992). "Isolation and serologic characterization of AIDA-I, the adhesin mediating the diffuse adherence phenotype of the diarrhea-associated Escherichia coli strain 2787 (O126:H27)." Infect Immun **60**(1): 13-18.

Benz, I. and M. A. Schmidt (2001). "Glycosylation with heptose residues mediated by the aah gene product is essential for adherence of the AIDA-I adhesin." Mol Microbiol **40**(6): 1403-1413.

Berger, C. N., O. Billker, T. F. Meyer, A. L. Servin and I. Kansau (2004). "Differential recognition of members of the carcinoembryonic antigen family by Afa/Dr adhesins of diffusely adhering Escherichia coli (Afa/Dr DAEC)." Mol Microbiol **52**(4): 963-983.

Berger, C. N., R. K. Shaw, F. Ruiz-Perez, J. P. Nataro, I. R. Henderson, M. J. Pallen and G. Frankel (2009). "Interaction of enteroaggregative *Escherichia coli* with salad leaves." Environ Microbiol Rep **1**(4): 234-239.

Berger CN, S. S., Shaw RK et al. (2010). "Fresh fruit and vegetables as vehicles for the transmission of human pathogens." Environmental microbiology: 2385-2397.

Berman, H. M., T. Battistuz, T. N. Bhat, W. F. Bluhm, P. E. Bourne, K. Burkhardt, Z. Feng, G. L. Gilliland, L. Iype, S. Jain, P. Fagan, J. Marvin, D. Padilla, V. Ravichandran, B. Schneider, N. Thanki, H. Weissig, J. D. Westbrook and C. Zardecki (2002). "The Protein Data Bank." Acta Crystallogr D Biol Crystallogr **58**(Pt 6 No 1): 899-907.

Berry, A. A., Y. Yang, N. Pakharukova, J. A. Garnett, W. C. Lee, E. Cota, J. Marchant, S. Roy, M. Tuittila, B. Liu, K. G. Inman, F. Ruiz-Perez, I. Mandomando, J. P. Nataro, A. V. Zavialov and S. Matthews (2014). "Structural insight into host recognition by aggregative adherence fimbriae of enteroaggregative Escherichia coli." PLoS Pathog **10**(9): e1004404.

Bertin, Y., J. P. Girardeau, A. Darfeuille-Michaud and M. Contrepois (1996). "Characterization of 20K fimbria, a new adhesin of septicemic and diarrhea-associated Escherichia coli strains, that belongs to a family of adhesins with N-acetyl-D-glucosamine recognition." Infect Immun **64**(1): 332-342.

Bertin, Y., J. P. Girardeau, A. Darfeuille-Michaud and C. Martin (2000). "Epidemiological study of pap genes among diarrheagenic or septicemic Escherichia coli strains producing CS31A and F17 adhesins and characterization of Pap(31A) fimbriae." J Clin Microbiol **38**(4): 1502-1509.

Bertin, Y., J. P. Girardeau, M. Der Vartanian and C. Martin (1993). "The ClpE protein involved in biogenesis of the CS31A capsule-like antigen is a member of a periplasmic chaperone family in gram-negative bacteria." FEMS Microbiol Lett **108**(1): 59-67.

Bhargava, S., B. B. Johnson, J. Hwang, T. A. Harris, A. S. George, A. Muir, J. Dorff and I. N. Okeke (2009). "Heat-resistant agglutinin 1 is an accessory enteroaggregative Escherichia coli colonization factor." J Bacteriol **191**(15): 4934-4942.

Bieber, D., S. W. Ramer, C. Y. Wu, W. J. Murray, T. Tobe, R. Fernandez and G. K. Schoolnik (1998). "Type IV pili, transient bacterial aggregates, and virulence of enteropathogenic Escherichia coli." Science **280**(5372): 2114-2118.

Bielaszewska, M., R. Prager, L. Vandivinit, A. Müsken, A. Mellmann, N. J. Holt, P. I. Tarr, H. Karch, and W. Zhang. (2009). "Detection and Characterization of the Fimbrial sfp Cluster in Enterohemorrhagic Escherichia coli O165:H25/NM Isolates from Humans and Cattle." Appl. Environ. Microbiol.

Bilge, S. S., J. M. Apostol, Jr., M. A. Aldape and S. L. Moseley (1993). "mRNA processing independent of RNase III and RNase E in the expression of the F1845 fimbrial adhesin of Escherichia coli." Proc Natl Acad Sci U S A **90**(4): 1455-1459.

Bilge, S. S., C. R. Clausen, W. Lau and S. L. Moseley (1989). "Molecular characterization of a fimbrial adhesin, F1845, mediating diffuse adherence of diarrhea-associated Escherichia coli to HEp-2 cells." J Bacteriol **171**(8): 4281-4289.

Bilge, S. S., J. C. Vary, Jr., S. F. Dowell and P. I. Tarr (1996). "Role of the Escherichia coli O157:H7 O side chain in adherence and analysis of an rfb locus." Infect Immun **64**(11): 4795-4801.

Blanton, L. V., L. T. Wang, J. Hofmann, J. DuBow, A. Lafrance, S. Kwak, L. Bowers, M. A. Levine, C. O. Hale, P. M. Meneely and I. N. Okeke (2018). "Aggregative Adherence and Intestinal Colonization by Enteroaggregative Escherichia coli Are Produced by Interactions among Multiple Surface Factors." mSphere **3**(2).

Blomfield, I. and M. van der Woude (2007). "Regulation of Fimbrial Expression." EcoSal Plus **2**(2).

Blomfield, I. C. (2001). "The regulation of pap and type 1 fimbriation in Escherichia coli." Adv Microb Physiol **45**: 1-49.

Bockelmann, U., A. Janke, R. Kuhn, T. R. Neu, J. Wecke, J. R. Lawrence and U. Szewzyk (2006). "Bacterial extracellular DNA forming a defined network-like structure." FEMS Microbiol Lett **262**(1): 31-38.

Boel, G., V. Pichereau, I. Mijakovic, A. Maze, S. Poncet, S. Gillet, J. C. Giard, A. Hartke, Y. Auffray and J. Deutscher (2004). "Is 2-phosphoglycerate-dependent automodification of bacterial enolases implicated in their export?" J Mol Biol **337**(2): 485-496.

Boisen, N., C. Struve, F. Scheutz, K. A. Krogfelt and J. P. Nataro (2008). "New adhesin of enteroaggregative Escherichia coli related to the Afa/Dr/AAF family." Infect Immun **76**(7): 3281-3292.

Bokranz, W., X. Wang, H. Tschape and U. Romling (2005). "Expression of cellulose and curli fimbriae by Escherichia coli isolated from the gastrointestinal tract." J Med Microbiol **54**(Pt 12): 1171-1182.

Boles, B. R. and A. R. Horswill (2011). "Staphylococcal biofilm disassembly." Trends Microbiol **19**(9): 449-455.

Bollinger, R. R., M. L. Everett, D. Palestrant, S. D. Love, S. S. Lin and W. Parker (2003). "Human secretory immunoglobulin A may contribute to biofilm formation in the gut." Immunology **109**(4): 580-587.

Bollinger, R. R., M. L. Everett, S. D. Wahl, Y. H. Lee, P. E. Orndorff and W. Parker (2006). "Secretory IgA and mucin-mediated biofilm formation by environmental strains of Escherichia coli: role of type 1 pili." Mol Immunol **43**(4): 378-387.

Borgersen, Q., D. T. Bolick, G. L. Kolling, M. Aijuka, F. Ruiz-Perez, R. L. Guerrant, J. P. Nataro and A. E. Santiago (2018). "Abundant production of Exopolysaccharide by EAEC strains enhances the formation of bacterial biofilms in contaminated sprouts." Gut Microbes: 0.

Botos, I., N. Noinaj and S. K. Buchanan (2017). "Insertion of proteins and lipopolysaccharide into the bacterial outer membrane." Philos Trans R Soc Lond B Biol Sci **372**(1726).

Branda, S. S., S. Vik, L. Friedman and R. Kolter (2005). "Biofilms: the matrix revisited." Trends Microbiol **13**(1): 20-26.

Bray, D. (2001). "Cell movements: from molecules to motility." Second Edition. Taylor & Francis Group, Garland Publishing, New York, USA.

Brenner, D. J., G. R. Fanning, F. J. Skerman and S. Falkow (1972). "Polynucleotide sequence divergence among strains of Escherichia coli and closely related organisms." J Bacteriol **109**(3): 953-965.

Brunder, W., H. Karch and H. Schmidt (2006). "Complete sequence of the large virulence plasmid pSFO157 of the sorbitol-fermenting enterohemorrhagic Escherichia coli O157:H- strain 3072/96." Int J Med Microbiol **296**(7): 467-474.

Brunder, W., A. S. Khan, J. Hacker and H. Karch (2001). "Novel type of fimbriae encoded by the large plasmid of sorbitol-fermenting enterohemorrhagic Escherichia coli O157:H(-)." Infect Immun **69**(7): 4447-4457.

Burger, A., C. Whiteley and A. Boshoff (2011). "Current perspectives of the Escherichia coli RNA degradosome." Biotechnol Lett **33**(12): 2337-2350.

Burgess, C., M. Desvaux and H. Olmez (2014). "1st Conference of BacFoodNet: mitigating bacterial colonisation in the food chain: bacterial adhesion, biocide resistance and microbial safety of fresh produce." Res Microbiol **165**(4): 305-310.

Candela, T. and A. Fouet (2006). "Poly-gamma-glutamate in bacteria." Mol Microbiol **60**(5): 1091-1098.

Capitani, G., O. Eidam, R. Glockshuber and M. G. Grutter (2006). "Structural and functional insights into the assembly of type 1 pili from Escherichia coli." Microbes Infect **8**(8): 2284-2290.

Caron, F. and E. Meyer (1989). "Molecular basis of surface antigen variation in paramecia." Annu Rev Microbiol **43**: 23-42.

Caron, J. and J. R. Scott (1990). "A rns-like regulatory gene for colonization factor antigen I (CFA/I) that controls expression of CFA/I pilin." Infect Immun **58**(4): 874-878.

Cerca, N., T. Maira-Litran, K. K. Jefferson, M. Grout, D. A. Goldmann and G. B. Pier (2007). "Protection against Escherichia coli infection by antibody to the Staphylococcus aureus poly-N-acetylglucosamine surface polysaccharide." Proc Natl Acad Sci U S A **104**(18): 7528-7533.

Chagnot, C., A. Listrat, T. Astruc and M. Desvaux (2012). "Bacterial adhesion to animal tissues: protein determinants for recognition of extracellular matrix components." Cell Microbiol **14**(11): 1687-1696.

Chagnot, C., M. A. Zorgani, T. Astruc and M. Desvaux (2013). "Proteinaceous determinants of surface colonization in bacteria: bacterial adhesion and biofilm formation from a protein secretion perspective." Front Microbiol **4**: 303.

Chaudhuri, R. R. and I. R. Henderson (2012). "The evolution of the Escherichia coli phylogeny." Infect Genet Evol **12**(2): 214-226.

Chauhan, A., C. Sakamoto, J. M. Ghigo and C. Beloin (2013). "Did I pick the right colony? Pitfalls in the study of regulation of the phase variable antigen 43 adhesin." PLoS One **8**(9): e73568.

Chen, J., S. M. Lee and Y. Mao (2004). "Protective effect of exopolysaccharide colanic acid of Escherichia coli O157:H7 to osmotic and oxidative stress." Int J Food Microbiol **93**(3): 281-286.

Chilcott, G. S. and K. T. Hughes (2000). "Coupling of flagellar gene expression to flagellar assembly in Salmonella enterica serovar typhimurium and Escherichia coli." Microbiol Mol Biol Rev **64**(4): 694-708.

Chris Whitfield, I. S. R. (1999). "Structure, assembly and regulation of expression of capsules in Escherichia coli." Molecular Microbiology.

Cimdins, A., P. Luthje, F. Li, I. Ahmad, A. Brauner and U. Romling (2017). "Draft Genome Sequences of Semiconstitutive Red, Dry, and Rough Biofilm-Forming Commensal and Uropathogenic Escherichia coli Isolates." Genome Announc **5**(4).

Clavijo, A. P., J. Bai and O. G. Gomez-Duarte (2010). "The Longus type IV pilus of enterotoxigenic Escherichia coli (ETEC) mediates bacterial self-aggregation and protection from antimicrobial agents." Microb Pathog **48**(6): 230-238.

Cleary, J., L. C. Lai, R. K. Shaw, A. Straatman-Iwanowska, M. S. Donnenberg, G. Frankel and S. Knutton (2004). "Enteropathogenic Escherichia coli (EPEC) adhesion to intestinal epithelial cells: role of bundle-forming pili (BFP), EspA filaments and intimin." Microbiology **150**(Pt 3): 527-538.

Clermont, O., J. K. Christenson, E. Denamur and D. M. Gordon (2013). "The Clermont Escherichia coli phylo-typing method revisited: improvement of specificity and detection of new phylo-groups." Environ Microbiol Rep **5**(1): 58-65.

Coburn, B., I. Sekirov and B. B. Finlay (2007). "Type III secretion systems and disease." Clin Microbiol Rev **20**(4): 535-549.

Connolly, J. P., B. B. Finlay and A. J. Roe (2015). "From ingestion to colonization: the influence of the host environment on regulation of the LEE encoded type III secretion system in enterohaemorrhagic Escherichia coli." Front Microbiol **6**: 568.

Contrepois, M., J. M. Fairbrother, Y. K. Kaura and J. P. Girardeau (1989). "Prevalence of CS31A and F165 surface antigens in Escherichia coli isolates from animals in France, Canada and India." FEMS Microbiol Lett **50**(3): 319-323.

Contrepois, M., J. P. Girardeau, P. Gouet and M. Der Vartanian (1983). "Expression of K99 pilus of E. coli." Ann Rech Vet **14**(4): 400-407.

Cooke, N. M., S. G. Smith, M. Kelleher and T. R. Rogers (2010). "Major differences exist in frequencies of virulence factors and multidrug resistance between community and nosocomial Escherichia coli bloodstream isolates." J Clin Microbiol **48**(4): 1099-1104.

Cookson, A. L., W. A. Cooley and M. J. Woodward (2002). "The role of type 1 and curli fimbriae of Shiga toxin-producing Escherichia coli in adherence to abiotic surfaces." Int J Med Microbiol **292**(3-4): 195-205.

Costerton, J. W. (1995). "Overview of microbial biofilms." J Ind Microbiol **15**(3): 137-140.

Costerton, J. W., P. S. Stewart and E. P. Greenberg (1999). "Bacterial biofilms: a common cause of persistent infections." Science **284**(5418): 1318-1322.

Cote, J. P., M. E. Charbonneau and M. Mourez (2013). "Glycosylation of the Escherichia coli TibA self-associating autotransporter influences the conformation and the functionality of the protein." PLoS One **8**(11): e80739.

Cote, J. P. and M. Mourez (2011). "Structure-function analysis of the TibA self-associating autotransporter reveals a modular organization." Infect Immun **79**(5): 1826-1832.

Cotter, S. E., N. K. Surana and J. W. St Geme, 3rd (2005). "Trimeric autotransporters: a distinct subfamily of autotransporter proteins." Trends Microbiol **13**(5): 199-205.

Craig, L., M. E. Pique and J. A. Tainer (2004). "Type IV pilus structure and bacterial pathogenicity." Nat Rev Microbiol **2**(5): 363-378.

Crost, C., A. Garrivier, J. Harel and C. Martin (2003). "Leucine-responsive regulatory protein-mediated repression of clp (encoding CS31A) expression by L-leucine and L-alanine in Escherichia coli." J Bacteriol **185**(6): 1886-1894.

Croxen, M. A. and B. B. Finlay (2010). "Molecular mechanisms of Escherichia coli pathogenicity." Nat Rev Microbiol **8**(1): 26-38.

Croxen, M. A., R. J. Law, R. Scholz, K. M. Keeney, M. Wlodarska and B. B. Finlay (2013). "Recent advances in understanding enteric pathogenic Escherichia coli." Clin Microbiol Rev **26**(4): 822-880.

Currie, C. G. and I. R. Poxton (1999). "The lipopolysaccharide core type of Escherichia coli O157:H7 and other non-O157 verotoxin-producing E. coli." FEMS Immunol Med Microbiol **24**(1): 57-62.

Czeczulin, J. R., S. Balepur, S. Hicks, A. Phillips, R. Hall, M. H. Kothary, F. Navarro-Garcia and J. P. Nataro (1997). "Aggregative adherence fimbria II, a second fimbrial antigen mediating aggregative adherence in enteroaggregative Escherichia coli." Infect Immun **65**(10): 4135-4145.

Dai, J., S. Wang, D. Guerlebeck, C. Laturnus, S. Guenther, Z. Shi, C. Lu and C. Ewers (2010). "Suppression subtractive hybridization identifies an autotransporter adhesin gene of E. coli IMT5155 specifically associated with avian pathogenic Escherichia coli (APEC)." BMC Microbiol **10**: 236.

Danese, P. N., L. A. Pratt, S. L. Dove and R. Kolter (2000). "The outer membrane protein, antigen 43, mediates cell-to-cell interactions within Escherichia coli biofilms." Mol Microbiol **37**(2): 424-432.

Dautin, N. (2010). "Serine protease autotransporters of enterobacteriaceae (SPATEs): biogenesis and function." Toxins (Basel) **2**(6): 1179-1206.

de Luna, M., A. Scott-Tucker, M. Desvaux, P. Ferguson, N. P. Morin, E. G. Dudley, S. Turner, J. P. Nataro, P. Owen and I. R. Henderson (2008). "The Escherichia coli biofilm-promoting protein Antigen 43 does not contribute to intestinal colonization." FEMS Microbiol Lett **284**(2): 237-246.

Deacon, V., F. Dziva, P. M. van Diemen, G. Frankel and M. P. Stevens (2010). "Efa-1/LifA mediates intestinal colonization of calves by enterohaemorrhagic Escherichia coli O26 : H- in a manner independent of glycosyltransferase and cysteine protease motifs or effects on type III secretion." Microbiology **156**(Pt 8): 2527-2536.

Decanio, M. S., R. Landick and R. J. Haft (2013). "The non-pathogenic Escherichia coli strain W secretes SslE via the virulence-associated type II secretion system beta." BMC Microbiol **13**: 130.

Desai, S. K. and L. J. Kenney (2017). "To approximately P or Not to approximately P? Non-canonical activation by two-component response regulators." Mol Microbiol **103**(2): 203-213.

Desvaux, M., M. Hebraud, I. R. Henderson and M. J. Pallen (2006). "Type III secretion: what's in a name?" Trends Microbiol **14**(4): 157-160.

Desvaux, M., M. Hébraud, R. Talon and I. R. Henderson (2009). "Secretion and subcellular localizations of bacterial proteins: a semantic awareness issue." Trends Microbiol **17**(4): 139-145.

Desvaux, M., N. J. Parham and I. R. Henderson (2003). "Type-V secretion system in gram-negative bacteria." Biofutur(237): 34-+.

Desvaux, M., N. J. Parham and I. R. Henderson (2004). "The autotransporter secretion system." Res Microbiol **155**(2): 53-60.

Diard, M., L. Garry, M. Selva, T. Mosser, E. Denamur and I. Matic (2010). "Pathogenicity-associated islands in extraintestinal pathogenic Escherichia coli are fitness elements involved in intestinal colonization." J Bacteriol **192**(19): 4885-4893.

Dionisio, F., I. Matic, M. Radman, O. R. Rodrigues and F. Taddei (2002). "Plasmids spread very fast in heterogeneous bacterial communities." Genetics **162**(4): 1525-1532.

Dobrindt, U., G. Blum-Oehler, G. Nagy, G. Schneider, A. Johann, G. Gottschalk and J. Hacker (2002). "Genetic structure and distribution of four pathogenicity islands (PAI I(536) to PAI IV(536)) of uropathogenic Escherichia coli strain 536." Infect Immun **70**(11): 6365-6372.

Donlan, R. M. (2002). "Biofilms: microbial life on surfaces." Emerg Infect Dis **8**(9): 881-890.

Donnenberg, M. S., S. Tzipori, M. L. McKee, A. D. O'Brien, J. Alroy and J. B. Kaper (1993). "The role of the eae gene of enterohemorrhagic Escherichia coli in intimate attachment in vitro and in a porcine model." J Clin Invest **92**(3): 1418-1424.

Doughty, S., J. Sloan, V. Bennett-Wood, M. Robertson, R. M. Robins-Browne and E. L. Hartland (2002). "Identification of a novel fimbrial gene cluster related to long polar fimbriae in locus of enterocyte effacement-negative strains of enterohemorrhagic Escherichia coli." Infect Immun **70**(12): 6761-6769.

Duan, Q., M. Zhou, H. Liang, X. Zhu, Z. Guo, Y. Li, P. R. Hardwidge and G. Zhu (2013). "Contribution of flagellin subunit FliC to piglet epithelial cells invasion by F18ab E. coli." Vet Microbiol **166**(1-2): 220-224.

Duan, Q., M. Zhou, X. Zhu, W. Bao, S. Wu, X. Ruan, W. Zhang, Y. Yang, J. Zhu and G. Zhu (2012). "The flagella of F18ab Escherichia coli is a virulence factor that contributes to infection in a IPEC-J2 cell model in vitro." Vet Microbiol **160**(1-2): 132-140.

Duan, Q., M. Zhou, X. Zhu, Y. Yang, J. Zhu, W. Bao, S. Wu, X. Ruan, W. Zhang and G. Zhu (2013). "Flagella from F18+Escherichia coli play a role in adhesion to pig epithelial cell lines." Microb Pathog **55**: 32-38.

Dudley, E. G., C. Abe, J. M. Ghigo, P. Latour-Lambert, J. C. Hormazabal and J. P. Nataro (2006). "An IncI1 plasmid contributes to the adherence of the atypical enteroaggregative Escherichia coli strain C1096 to cultured cells and abiotic surfaces." Infect Immun **74**(4): 2102-2114.

Dudley, E. G., N. R. Thomson, J. Parkhill, N. P. Morin and J. P. Nataro (2006). "Proteomic and microarray characterization of the AggR regulon identifies a pheU pathogenicity island in enteroaggregative Escherichia coli." Mol Microbiol **61**(5): 1267-1282.

Duncan, M. J., E. L. Mann, M. S. Cohen, I. Ofek, N. Sharon and S. N. Abraham (2005). "The distinct binding specificities exhibited by enterobacterial type 1 fimbriae are determined by their fimbrial shafts." J Biol Chem **280**(45): 37707-37716.

Dziva, F., A. Mahajan, P. Cameron, C. Currie, I. J. McKendrick, T. S. Wallis, D. G. Smith and M. P. Stevens (2007). "EspP, a Type V-secreted serine protease of enterohaemorrhagic Escherichia coli O157:H7, influences intestinal colonization of calves and adherence to bovine primary intestinal epithelial cells." FEMS Microbiol Lett **271**(2): 258-264.

Dziva, F., P. M. van Diemen, M. P. Stevens, A. J. Smith and T. S. Wallis (2004). "Identification of Escherichia coli O157 : H7 genes influencing colonization of the bovine gastrointestinal tract using signature-tagged mutagenesis." Microbiology **150**(Pt 11): 3631-3645.

Easton, D. M., L. P. Allsopp, M. D. Phan, D. G. Moriel, G. K. Goh, S. A. Beatson, T. J. Mahony, R. N. Cobbold and M. A. Schembri (2014). "The Intimin-Like Protein FdeC Is Regulated by H-NS and Temperature in Enterohemorrhagic Escherichia coli." Appl Environ Microbiol **80**(23): 7337-7347.

Easton, D. M., M. Totsika, L. P. Allsopp, M. D. Phan, A. Idris, D. J. Wurpel, O. Sherlock, B. Zhang, C. Venturini, S. A. Beatson, T. J. Mahony, R. N. Cobbold and M. A. Schembri (2011). "Characterization of EhaJ, a New Autotransporter Protein from Enterohemorrhagic and Enteropathogenic Escherichia coli." Front Microbiol **2**: 120.

Ebel, W. and J. E. Trempy (1999). "Escherichia coli RcsA, a positive activator of colanic acid capsular polysaccharide synthesis, functions To activate its own expression." J Bacteriol **181**(2): 577-584.

Eckburg, P. B., E. M. Bik, C. N. Bernstein, E. Purdom, L. Dethlefsen, M. Sargent, S. R. Gill, K. E. Nelson and D. A. Relman (2005). "Diversity of the human intestinal microbial flora." Science **308**(5728): 1635-1638.

EFSA (2015). "Public health risks associated with Enteroaggregative Escherichia coli (EAEC) as a food-borne pathogen." EFSA Journal 2015;13(12):4330.

Egea, L., L. Aguilera, R. Gimenez, M. A. Sorolla, J. Aguilar, J. Badia and L. Baldoma (2007). "Role of secreted glyceraldehyde-3-phosphate dehydrogenase in the infection mechanism of enterohemorrhagic and enteropathogenic Escherichia coli: interaction of the extracellular enzyme with human plasminogen and fibrinogen." Int J Biochem Cell Biol **39**(6): 1190-1203.

Elliott, S. J. and J. B. Kaper (1997). "Role of type 1 fimbriae in EPEC infections." Microb Pathog **23**(2): 113-118.

Emory, S. A. and J. G. Belasco (1990). "The ompA 5' untranslated RNA segment functions in Escherichia coli as a growth-rate-regulated mRNA stabilizer whose activity is unrelated to translational efficiency." J Bacteriol **172**(8): 4472-4481.

Emory, S. A., P. Bouvet and J. G. Belasco (1992). "A 5'-terminal stem-loop structure can stabilize mRNA in Escherichia coli." Genes Dev **6**(1): 135-148.

Escobar-Paramo, P., A. Le Menac'h, T. Le Gall, C. Amorin, S. Gouriou, B. Picard, D. Skurnik and E. Denamur (2006). "Identification of forces shaping the commensal Escherichia coli genetic structure by comparing animal and human isolates." Environ Microbiol **8**(11): 1975-1984.

Evans, M. L. and M. R. Chapman (2014). "Curli biogenesis: order out of disorder." Biochim Biophys Acta **1843**(8): 1551-1558.

Evans, M. L., E. Chorell, J. D. Taylor, J. Aden, A. Gotheson, F. Li, M. Koch, L. Sefer, S. J. Matthews, P. Wittung-Stafshede, F. Almqvist and M. R. Chapman (2015). "The bacterial curli system possesses a potent and selective inhibitor of amyloid formation." Mol Cell **57**(3): 445-455.

Fagan, R. P., M. A. Lambert and S. G. Smith (2008). "The hek outer membrane protein of Escherichia coli strain RS218 binds to proteoglycan and utilizes a single extracellular loop for adherence, invasion, and autoaggregation." Infect Immun **76**(3): 1135-1142.

Fagan, R. P. and S. G. Smith (2007). "The Hek outer membrane protein of Escherichia coli is an auto-aggregating adhesin and invasin." FEMS Microbiol Lett **269**(2): 248-255.

Farfan, M. J., K. G. Inman and J. P. Nataro (2008). "The major pilin subunit of the AAF/II fimbriae from enteroaggregative Escherichia coli mediates binding to extracellular matrix proteins." Infect Immun **76**(10): 4378-4384.

Farfan, M. J. and A. G. Torres (2012). "Molecular mechanisms that mediate colonization of Shiga toxin-producing Escherichia coli strains." Infect Immun **80**(3): 903-913.

Feng, L., B. Liu, Y. Liu, Y. A. Ratiner, B. Hu, D. Li, X. Zong, W. Xiong and L. Wang (2008). "A genomic islet mediates flagellar phase variation in Escherichia coli strains carrying the flagellin-specifying locus flk." J Bacteriol **190**(13): 4470-4477.

Fitzhenry, R., S. Dahan, A. G. Torres, Y. Chong, R. Heuschkel, S. H. Murch, M. Thomson, J. B. Kaper, G. Frankel and A. D. Phillips (2006). "Long polar fimbriae and tissue tropism in Escherichia coli O157:H7." Microbes Infect **8**(7): 1741-1749.

Fleckenstein, J. M., J. T. Holland and D. L. Hasty (2002). "Interaction of an uuter membrane protein of enterotoxigenic Escherichia coli with cell surface heparan sulfate proteoglycans." Infect Immun **70**(3): 1530-1537.

Fleckenstein, J. M., D. J. Kopecko, R. L. Warren and E. A. Elsinghorst (1996). "Molecular characterization of the tia invasion locus from enterotoxigenic Escherichia coli." Infect Immun **64**(6): 2256-2265.

Fleckenstein, J. M., K. Roy, J. F. Fischer and M. Burkitt (2006). "Identification of a two-partner secretion locus of enterotoxigenic Escherichia coli." Infect Immun **74**(4): 2245-2258.

Franzin, F. M. and M. P. Sircili (2015). "Locus of enterocyte effacement: a pathogenicity island involved in the virulence of enteropathogenic and enterohemorragic Escherichia coli subjected to a complex network of gene regulation." Biomed Res Int **2015**: 534738.

Fricke, W. F., M. S. Wright, A. H. Lindell, D. M. Harkins, C. Baker-Austin, J. Ravel and R. Stepanauskas (2008). "Insights into the environmental resistance gene pool from the genome sequence of the multidrug-resistant environmental isolate Escherichia coli SMS-3-5." J Bacteriol **190**(20): 6779-6794.

Furniss, R. C. D. and A. Clements (2018). "Regulation of the Locus of Enterocyte Effacement in Attaching and Effacing Pathogens." J Bacteriol **200**(2).

Gaastra, W. and A. M. Svennerholm (1996). "Colonization factors of human enterotoxigenic Escherichia coli (ETEC)." Trends Microbiol **4**(11): 444-452.

Galan, J. E. and G. Waksman (2018). "Protein-Injection Machines in Bacteria." Cell **172**(6): 1306-1318.

Galan, J. E. and H. Wolf-Watz (2006). "Protein delivery into eukaryotic cells by type III secretion machines." Nature **444**(7119): 567-573.

Gally, D. L., J. Leathart and I. C. Blomfield (1996). "Interaction of FimB and FimE with the fim switch that controls the phase variation of type 1 fimbriae in Escherichia coli K-12." Mol Microbiol **21**(4): 725-738.

Genevaux, P., P. Bauda, M. S. DuBow and B. Oudega (1999). "Identification of Tn10 insertions in the rfaG, rfaP, and galU genes involved in lipopolysaccharide core biosynthesis that affect Escherichia coli adhesion." Arch Microbiol **172**(1): 1-8.

Gerstel, U. and U. Romling (2001). "Oxygen tension and nutrient starvation are major signals that regulate agfD promoter activity and expression of the multicellular morphotype in Salmonella typhimurium." Environ Microbiol **3**(10): 638-648.

Gervais, F. G. and G. R. Drapeau (1992). "Identification, cloning, and characterization of rcsF, a new regulator gene for exopolysaccharide synthesis that suppresses the division mutation ftsZ84 in Escherichia coli K-12." J Bacteriol **174**(24): 8016-8022.

Ghigo, J. M. (2001). "Natural conjugative plasmids induce bacterial biofilm development." Nature **412**(6845): 442-445.

Giaouris, E., E. Heir, M. Hébraud, N. Chorianopoulos, S. Langsrud, T. Moretro, O. Habimana, M. Desvaux, S. Renier and G. J. Nychas (2014). "Attachment and biofilm formation by foodborne bacteria in meat processing environments: Causes, implications, role of bacterial interactions and control by alternative novel methods." Meat Sci **97**(3): 298-309.

Gibb, A. P., G. R. Barclay, I. R. Poxton and F. di Padova (1992). "Frequencies of lipopolysaccharide core types among clinical isolates of Escherichia coli defined with monoclonal antibodies." J Infect Dis **166**(5): 1051-1057.

Giltner, C. L., Y. Nguyen and L. L. Burrows (2012). "Type IV pilin proteins: versatile molecular modules." Microbiol Mol Biol Rev **76**(4): 740-772.

Giltner, C. L., E. J. van Schaik, G. F. Audette, D. Kao, R. S. Hodges, D. J. Hassett and R. T. Irvin (2006). "The Pseudomonas aeruginosa type IV pilin receptor binding domain functions as an adhesin for both biotic and abiotic surfaces." Mol Microbiol **59**(4): 1083-1096.

Girardeau, J. P., Y. Bertin, C. Martin, M. Der Vartanian and C. Boeuf (1991). "Sequence analysis of the clpG gene, which codes for surface antigen CS31A subunit: evidence of an evolutionary relationship between CS31A, K88, and F41 subunit genes." J Bacteriol **173**(23): 7673-7683.

Girardeau, J. P., M. Der Vartanian, J. L. Ollier and M. Contrepois (1988). "CS31A, a new K88-related fimbrial antigen on bovine enterotoxigenic and septicemic Escherichia coli strains." Infect Immun **56**(8): 2180-2188.

Giron, J. A., M. M. Levine and J. B. Kaper (1994). "Longus: a long pilus ultrastructure produced by human enterotoxigenic Escherichia coli." Mol Microbiol **12**(1): 71-82.

Giron, J. A., A. G. Torres, E. Freer and J. B. Kaper (2002). "The flagella of enteropathogenic Escherichia coli mediate adherence to epithelial cells." Mol Microbiol **44**(2): 361-379.

Giron, J. A., G. I. Viboud, V. Sperandio, O. G. Gomez-Duarte, D. R. Maneval, M. J. Albert, M. M. Levine and J. B. Kaper (1995). "Prevalence and association of the longus pilus structural gene (lngA) with colonization factor antigens, enterotoxin types, and serotypes of enterotoxigenic Escherichia coli." Infect Immun **63**(10): 4195-4198.

Gomes, T. A., R. T. Hernandes, A. G. Torres, F. A. Salvador, B. E. Guth, T. M. Vaz, K. Irino, R. M. Silva and M. A. Vieira (2011). "Adhesin-encoding genes from shiga toxin-producing Escherichia coli are more prevalent in atypical than in typical enteropathogenic E. coli." J Clin Microbiol **49**(9): 3334-3337.

Gomez-Duarte, O. G. and J. B. Kaper (1995). "A plasmid-encoded regulatory region activates chromosomal eaeA expression in enteropathogenic Escherichia coli." Infect Immun **63**(5): 1767-1776.

Gonzales-Siles, L. and A. Sjoling (2016). "The different ecological niches of enterotoxigenic Escherichia coli." Environ Microbiol **18**(3): 741-751.

Grainger, D. C. (2016). "Structure and function of bacterial H-NS protein." Biochem Soc Trans **44**(6): 1561-1569.

Graveline, R., P. Garneau, C. Martin, M. Mourez, M. A. Hancock, R. Lavoie and J. Harel (2014). "Leucine-responsive regulatory protein Lrp and PapI homologues influence phase variation of CS31A fimbriae." J Bacteriol **196**(16): 2944-2953.

Gualdi, L., L. Tagliabue, S. Bertagnoli, T. Ierano, C. De Castro and P. Landini (2008). "Cellulose modulates biofilm formation by counteracting curli-mediated colonization of solid surfaces in Escherichia coli." Microbiology **154**(Pt 7): 2017-2024.

Guinee, P. A., W. H. Jansen and C. M. Agterberg (1976). "Detection of the K99 antigen by means of agglutination and immunoelectrophoresis in Escherichia coli isolates from calves and its correlation with entertoxigenicity." Infect Immun **13**(5): 1369-1377.

Gutierrez-Jimenez, J., I. Arciniega and F. Navarro-Garcia (2008). "The serine protease motif of Pic mediates a dose-dependent mucolytic activity after binding to sugar constituents of the mucin substrate." Microb Pathog **45**(2): 115-123.

Guttenplan, S. B. and D. B. Kearns (2013). "Regulation of flagellar motility during biofilm formation." FEMS Microbiol Rev **37**(6): 849-871.

Ha, S. K., C. Choi and C. Chae (2003). "Prevalence of a gene encoding adhesin involved in diffuse adherence among Escherichia coli isolates in pigs with postweaning diarrhea or edema disease." J Vet Diagn Invest **15**(4): 378-381.

Hadjifrangiskou, M., M. Kostakioti, S. L. Chen, J. P. Henderson, S. E. Greene and S. J. Hultgren (2011). "A central metabolic circuit controlled by QseC in pathogenic Escherichia coli." Mol Microbiol **80**(6): 1516-1529.

Hall-Stoodley, L. and P. Stoodley (2002). "Developmental regulation of microbial biofilms." Curr Opin Biotechnol **13**(3): 228-233.

Hamilton, H. L., N. M. Dominguez, K. J. Schwartz, K. T. Hackett and J. P. Dillard (2005). "Neisseria gonorrhoeae secretes chromosomal DNA via a novel type IV secretion system." Mol Microbiol **55**(6): 1704-1721.

Hammar, M., A. Arnqvist, Z. Bian, A. Olsen and S. Normark (1995). "Expression of two csg operons is required for production of fibronectin- and congo red-binding curli polymers in Escherichia coli K-12." Mol Microbiol **18**(4): 661-670.

Hanna, A., M. Berg, V. Stout and A. Razatos (2003). "Role of capsular colanic acid in adhesion of uropathogenic Escherichia coli." Appl Environ Microbiol **69**(8): 4474-4481.

Hannan, A., S. Saleem, S. Chaudhary, M. Barkaat and M. U. Arshad (2008). "Anti bacterial activity of Nigella sativa against clinical isolates of methicillin resistant Staphylococcus aureus." J Ayub Med Coll Abbottabad **20**(3): 72-74.

Hansen, M. J., L. H. Chen, M. L. Fejzo and J. G. Belasco (1994). "The ompA 5' untranslated region impedes a major pathway for mRNA degradation in Escherichia coli." Mol Microbiol **12**(5): 707-716.

Harrington, S. M., J. Sheikh, I. R. Henderson, F. Ruiz-Perez, P. S. Cohen and J. P. Nataro (2009). "The Pic protease of enteroaggregative Escherichia coli promotes intestinal colonization and growth in the presence of mucin." Infect Immun **77**(6): 2465-2473.

Harris, S. L., D. A. Elliott, M. C. Blake, L. M. Must, M. Messenger and P. E. Orndorff (1990). "Isolation and characterization of mutants with lesions affecting pellicle formation and erythrocyte agglutination by type 1 piliated Escherichia coli." J Bacteriol **172**(11): 6411-6418.

Hasman, H., T. Chakraborty and P. Klemm (1999). "Antigen-43-mediated autoaggregation of Escherichia coli is blocked by fimbriation." J Bacteriol **181**(16): 4834-4841.

Hausner, M. and S. Wuertz (1999). "High rates of conjugation in bacterial biofilms as determined by quantitative in situ analysis." Appl Environ Microbiol **65**(8): 3710-3713.

Henderson, B. and A. Martin (2011). "Bacterial virulence in the moonlight: multitasking bacterial moonlighting proteins are virulence determinants in infectious disease." Infect Immun **79**(9): 3476-3491.

Henderson, I. R., J. Czeczulin, C. Eslava, F. Noriega and J. P. Nataro (1999). "Characterization of pic, a secreted protease of Shigella flexneri and enteroaggregative Escherichia coli." Infect Immun **67**(11): 5587-5596.

Henderson, I. R., F. Navarro-Garcia, M. Desvaux, R. C. Fernandez and D. Ala'Aldeen (2004). "Type V protein secretion pathway: the autotransporter story." Microbiol Mol Biol Rev **68**(4): 692-744.

Henderson, I. R., P. Owen and J. P. Nataro (1999). "Molecular switches--the ON and OFF of bacterial phase variation." Mol Microbiol **33**(5): 919-932.

Heras, B., M. Totsika, K. M. Peters, J. J. Paxman, C. L. Gee, R. J. Jarrott, M. A. Perugini, A. E. Whitten and M. A. Schembri (2014). "The antigen 43 structure reveals a molecular Velcro-like mechanism of autotransporter-mediated bacterial clumping." Proc Natl Acad Sci U S A **111**(1): 457-462.

Herold, S., J. C. Paton and A. W. Paton (2009). "Sab, a novel autotransporter of locus of enterocyte effacement-negative shiga-toxigenic Escherichia coli O113:H21, contributes to adherence and biofilm formation." Infect Immun **77**(8): 3234-3243.

Herold, S., J. C. Paton, P. Srimanote and A. W. Paton (2009). "Differential effects of short-chain fatty acids and iron on expression of iha in Shiga-toxigenic Escherichia coli." Microbiology **155**(Pt 11): 3554-3563.

Hicks, S., D. C. Candy and A. D. Phillips (1996). "Adhesion of enteroaggregative Escherichia coli to formalin-fixed intestinal and ureteric epithelia from children." J Med Microbiol **44**(5): 362-371.

Hill, C. (2012). "Virulence or niche factors: what's in a name?" J Bacteriol **194**(21): 5725-5727.

Hitchcock, P. J., L. Leive, P. H. Makela, E. T. Rietschel, W. Strittmatter and D. C. Morrison (1986). "Lipopolysaccharide nomenclature--past, present, and future." J Bacteriol **166**(3): 699-705.

Hobley, L., C. Harkins, C. E. MacPhee and N. R. Stanley-Wall (2015). "Giving structure to the biofilm matrix: an overview of individual strategies and emerging common themes." FEMS Microbiol Rev **39**(5): 649-669.

Hobman, J. L., C. W. Penn and M. J. Pallen (2007). "Laboratory strains of Escherichia coli: model citizens or deceitful delinquents growing old disgracefully?" Mol Microbiol **64**(4): 881-885.

Hoch, J. A. (2000). "Two-component and phosphorelay signal transduction." Curr Opin Microbiol **3**(2): 165-170.

Holden, N. J. and D. L. Gally (2004). "Switches, cross-talk and memory in Escherichia coli adherence." J Med Microbiol **53**(Pt 7): 585-593.

Hueck, C. J. (1998). "Type III protein secretion systems in bacterial pathogens of animals and plants." Microbiol Mol Biol Rev **62**(2): 379-433.

Hung, C., Y. Zhou, J. S. Pinkner, K. W. Dodson, J. R. Crowley, J. Heuser, M. R. Chapman, M. Hadjifrangiskou, J. P. Henderson and S. J. Hultgren (2013). "Escherichia coli biofilms have an organized and complex extracellular matrix structure." MBio **4**(5): e00645-00613.

Iguchi, A., N. R. Thomson, Y. Ogura, D. Saunders, T. Ooka, I. R. Henderson, D. Harris, M. Asadulghani, K. Kurokawa, P. Dean, B. Kenny, M. A. Quail, S. Thurston, G. Dougan, T. Hayashi, J. Parkhill and G. Frankel (2009). "Complete genome sequence and comparative genome analysis of enteropathogenic Escherichia coli O127:H6 strain E2348/69." J Bacteriol **191**(1): 347-354.

Itoh, Y., X. Wang, B. J. Hinnebusch, J. F. Preston, 3rd and T. Romeo (2005). "Depolymerization of beta-1,6-N-acetyl-D-glucosamine disrupts the integrity of diverse bacterial biofilms." J Bacteriol **187**(1): 382-387.

Jallat, C., A. Darfeuille-Michaud, J. P. Girardeau, C. Rich and B. Joly (1994). "Self-transmissible R plasmids encoding CS31A among human Escherichia coli strains isolated from diarrheal stools." Infect Immun **62**(7): 2865-2873.

Jang, H. and K. R. Matthews (2018). "Influence of surface polysaccharides of Escherichia coli O157:H7 on plant defense response and survival of the human enteric pathogen on Arabidopsis thaliana and lettuce (Lactuca sativa)." Food Microbiol **70**: 254-261.

Jang, J., H. G. Hur, M. J. Sadowsky, M. N. Byappanahalli, T. Yan and S. Ishii (2017). "Environmental Escherichia coli: ecology and public health implications-a review." J Appl Microbiol **123**(3): 570-581.

Jansson, L., J. Tobias, M. Lebens, A. M. Svennerholm and S. Teneberg (2006). "The major subunit, CfaB, of colonization factor antigen i from enterotoxigenic Escherichia coli is a glycosphingolipid binding protein." Infect Immun **74**(6): 3488-3497.

Jaureguy, F., L. Landraud, V. Passet, L. Diancourt, E. Frapy, G. Guigon, E. Carbonnelle, O. Lortholary, O. Clermont, E. Denamur, B. Picard, X. Nassif and S. Brisse (2008). "Phylogenetic and genomic diversity of human bacteremic Escherichia coli strains." BMC Genomics **9**: 560.

Jenkins, C., N. T. Perry, T. Cheasty, D. J. Shaw, G. Frankel, G. Dougan, G. J. Gunn, H. R. Smith, A. W. Paton and J. C. Paton (2003). "Distribution of the saa gene in strains of Shiga toxin-producing Escherichia coli of human and bovine origins." J Clin Microbiol **41**(4): 1775-1778.

Johnson, J. R., S. Jelacic, L. M. Schoening, C. Clabots, N. Shaikh, H. L. Mobley and P. I. Tarr (2005). "The IrgA homologue adhesin Iha is an Escherichia coli virulence factor in murine urinary tract infection." Infect Immun **73**(2): 965-971.

Jonsson, R., C. Struve, N. Boisen, R. V. Mateiu, A. E. Santiago, H. Jenssen, J. P. Nataro and K. A. Krogfelt (2015). "Novel aggregative adherence fimbria variant of enteroaggregative Escherichia coli." Infect Immun **83**(4): 1396-1405.

Jordi, B. J., G. A. Willshaw, B. A. van der Zeijst and W. Gaastra (1992). "The complete nucleotide sequence of region 1 of the CFA/I fimbrial operon of human enterotoxigenic Escherichia coli." DNA Seq **2**(4): 257-263.

Kadurugamuwa, J. L. and T. J. Beveridge (1995). "Virulence factors are released from Pseudomonas aeruginosa in association with membrane vesicles during normal growth and exposure to gentamicin: a novel mechanism of enzyme secretion." J Bacteriol **177**(14): 3998-4008.

Kadurugamuwa, J. L. and T. J. Beveridge (1996). "Bacteriolytic effect of membrane vesicles from Pseudomonas aeruginosa on other bacteria including pathogens: conceptually new antibiotics." J Bacteriol **178**(10): 2767-2774.

Kaiser, D. (2007). "Bacterial swarming: a re-examination of cell-movement patterns." Curr Biol **17**(14): R561-570.

Kanamaru, S., H. Kurazono, S. Ishitoya, A. Terai, T. Habuchi, M. Nakano, O. Ogawa and S. Yamamoto (2003). "Distribution and genetic association of putative uropathogenic virulence factors iroN, iha, kpsMT, ompT and usp in Escherichia coli isolated from urinary tract infections in Japan." J Urol **170**(6 Pt 1): 2490-2493.

Kaper, J. B., J. P. Nataro and H. L. Mobley (2004). "Pathogenic Escherichia coli." Nat Rev Microbiol **2**(2): 123-140.

Kawula, T. H. and P. E. Orndorff (1991). "Rapid site-specific DNA inversion in Escherichia coli mutants lacking the histonelike protein H-NS." J Bacteriol **173**(13): 4116-4123.

Keiski, C. L., M. Harwich, S. Jain, A. M. Neculai, P. Yip, H. Robinson, J. C. Whitney, L. Riley, L. L. Burrows, D. E. Ohman and P. L. Howell (2010). "AlgK is a TPR-containing protein and the periplasmic component of a novel exopolysaccharide secretin." Structure **18**(2): 265-273.

Khan, I., Y. Chen, T. Dong, X. Hong, R. Takeuchi, H. Mori and D. Kihara (2014). "Genome-scale identification and characterization of moonlighting proteins." Biol Direct **9**: 30.

Khan, I. F., M. Vajpayee, V. V. Prasad and P. Seth (2007). "Genetic diversity of HIV type 1 subtype C env gene sequences from India." AIDS Res Hum Retroviruses **23**(7): 934-940.

Kim, N. H., T. J. Cho and M. S. Rhee (2017). "Current Interventions for Controlling Pathogenic Escherichia coli." Adv Appl Microbiol **100**: 1-47.

Kim, S. H. and Y. H. Kim (2004). "Escherichia coli O157:H7 adherence to HEp-2 cells is implicated with curli expression and outer membrane integrity." J Vet Sci **5**(2): 119-124.

Kjaergaard, K., M. A. Schembri, C. Ramos, S. Molin and P. Klemm (2000). "Antigen 43 facilitates formation of multispecies biofilms." Environ Microbiol **2**(6): 695-702.

Klapproth, J. M., I. C. Scaletsky, B. P. McNamara, L. C. Lai, C. Malstrom, S. P. James and M. S. Donnenberg (2000). "A large toxin from pathogenic Escherichia coli strains that inhibits lymphocyte activation." Infect Immun **68**(4): 2148-2155.

Klemm, P. (1986). "Two regulatory fim genes, fimB and fimE, control the phase variation of type 1 fimbriae in Escherichia coli." EMBO J **5**(6): 1389-1393.

Klemm, P., G. Christiansen, B. Kreft, R. Marre and H. Bergmans (1994). "Reciprocal exchange of minor components of type 1 and F1C fimbriae results in hybrid organelles with changed receptor specificities." J Bacteriol **176**(8): 2227-2234.

Klemm, P., Vejborg, R., Sherlock, O. (2006). "Self-associating autotransporters, SAATs: Functional and structural similarities." International Journal of Medical Microbiology: 187–195.

Knutton, S., M. M. McConnell, B. Rowe and A. S. McNeish (1989). "Adhesion and ultrastructural properties of human enterotoxigenic Escherichia coli producing colonization factor antigens III and IV." Infect Immun **57**(11): 3364-3371.

Knutton, S., I. Rosenshine, M. J. Pallen, I. Nisan, B. C. Neves, C. Bain, C. Wolff, G. Dougan and G. Frankel (1998). "A novel EspA-associated surface organelle of enteropathogenic Escherichia coli involved in protein translocation into epithelial cells." EMBO J **17**(8): 2166-2176.

Koebnik, R., K. P. Locher and P. Van Gelder (2000). "Structure and function of bacterial outer membrane proteins: barrels in a nutshell." Mol Microbiol **37**(2): 239-253.

Korea, C. G., R. Badouraly, M. C. Prevost, J. M. Ghigo and C. Beloin (2010). "Escherichia coli K-12 possesses multiple cryptic but functional chaperone-usher fimbriae with distinct surface specificities." Environ Microbiol **12**(7): 1957-1977.

Kostakioti, M., M. Hadjifrangiskou, J. S. Pinkner and S. J. Hultgren (2009). "QseC-mediated dephosphorylation of QseB is required for expression of genes associated with virulence in uropathogenic Escherichia coli." Mol Microbiol **73**(6): 1020-1031.

Kostakioti, M., C. L. Newman, D. G. Thanassi and C. Stathopoulos (2005). "Mechanisms of protein export across the bacterial outer membrane." J Bacteriol **187**(13): 4306-4314.

Kudva, I. T., B. Krastins, A. G. Torres, R. W. Griffin, H. Sheng, D. A. Sarracino, C. J. Hovde, S. B. Calderwood and M. John (2015). "The Escherichia coli O157:H7 cattle immunoproteome includes outer membrane protein A (OmpA), a modulator of adherence to bovine rectoanal junction squamous epithelial (RSE) cells." Proteomics **15**(11): 1829-1842.

Kuehn, M. J. and N. C. Kesty (2005). "Bacterial outer membrane vesicles and the host-pathogen interaction." Genes Dev **19**(22): 2645-2655.

Laarmann, S. and M. A. Schmidt (2003). "The Escherichia coli AIDA autotransporter adhesin recognizes an integral membrane glycoprotein as receptor." Microbiology **149**(Pt 7): 1871-1882.

Landick, R., J. T. Wade and D. C. Grainger (2015). "H-NS and RNA polymerase: a love-hate relationship?" Curr Opin Microbiol **24**: 53-59.

Landini, P. and A. J. Zehnder (2002). "The global regulatory hns gene negatively affects adhesion to solid surfaces by anaerobically grown Escherichia coli by modulating expression of flagellar genes and lipopolysaccharide production." J Bacteriol **184**(6): 1522-1529.

Laverty, G., S. P. Gorman and B. F. Gilmore (2014). "Biomolecular Mechanisms of Pseudomonas aeruginosa and Escherichia coli Biofilm Formation." Pathogens **3**(3): 596-632.

Lawley, T. D., W. A. Klimke, M. J. Gubbins and L. S. Frost (2003). "F factor conjugation is a true type IV secretion system." FEMS Microbiol Lett **224**(1): 1-15.

Lawrence, J. G. and D. L. Hartl (1991). "Unusual codon bias occurring within insertion sequences in Escherichia coli." Genetica **84**(1): 23-29.

Lawson, C. L. (2010). "Unified data resource for cryo-EM." Methods Enzymol **483**: 73-90.

Le Bouguenec, C. and A. L. Servin (2006). "Diffusely adherent Escherichia coli strains expressing Afa/Dr adhesins (Afa/Dr DAEC): hitherto unrecognized pathogens." FEMS Microbiol Lett **256**(2): 185-194.

Le Gall, T., O. Clermont, S. Gouriou, B. Picard, X. Nassif, E. Denamur and O. Tenaillon (2007). "Extraintestinal virulence is a coincidental by-product of commensalism in B2 phylogenetic group Escherichia coli strains." Mol Biol Evol **24**(11): 2373-2384.

Le Quere, B. and J. M. Ghigo (2009). "BcsQ is an essential component of the Escherichia coli cellulose biosynthesis apparatus that localizes at the bacterial cell pole." Mol Microbiol **72**(3): 724-740.

Lebaron, P., P. Bauda, M. C. Lett, Y. Duval-Iflah, P. Simonet, E. Jacq, N. Frank, B. Roux, B. Baleux, G. Faurie, J. C. Hubert, P. Normand, D. Prieur, S. Schmitt and J. C. Block (1997). "Recombinant plasmid mobilization between E. coli strains in seven sterile microcosms." Can J Microbiol **43**(6): 534-540.

Leimbach, A., J. Hacker and U. Dobrindt (2013). "E. coli as an all-rounder: the thin line between commensalism and pathogenicity." Curr Top Microbiol Immunol **358**: 3-32.

Lenski, R. E. (2017). "Experimental evolution and the dynamics of adaptation and genome evolution in microbial populations." ISME J **11**(10): 2181-2194.

Leo, J. C. and A. Goldman (2009). "The immunoglobulin-binding Eib proteins from Escherichia coli are receptors for IgG Fc." Mol Immunol **46**(8-9): 1860-1866.

Leo, J. C., I. Grin and D. Linke (2012). "Type V secretion: mechanism(s) of autotransport through the bacterial outer membrane." Philosophical Transactions of the Royal Society B-Biological Sciences **367**(1592): 1088-1101.

Leo, J. C., I. Grin and D. Linke (2012). "Type V secretion: mechanism(s) of autotransport through the bacterial outer membrane." Philos Trans R Soc Lond B Biol Sci **367**(1592): 1088-1101.

Leo, J. C., A. Lyskowski, K. Hattula, M. D. Hartmann, H. Schwarz, S. J. Butcher, D. Linke, A. N. Lupas and A. Goldman (2011). "The structure of E. coli IgG-binding protein D suggests a general model for bending and binding in trimeric autotransporter adhesins." Structure **19**(7): 1021-1030.

Leo, J. C., P. Oberhettinger, M. Schutz and D. Linke (2015). "The inverse autotransporter family: intimin, invasin and related proteins." Int J Med Microbiol **305**(2): 276-282.

Lewis, S. B., V. Cook, R. Tighe and S. Schuller (2015). "Enterohemorrhagic Escherichia coli colonization of human colonic epithelium in vitro and ex vivo." Infect Immun **83**(3): 942-949.

Leyton, D. L., M. J. Belousoff and T. Lithgow (2015). "The β-barrel assembly machinery complex." Methods Mol Biol **1329**: 1-16.

Li, B., W. H. Koch and T. A. Cebula (1997). "Detection and characterization of the fimA gene of Escherichia coli O157:H7." Mol Cell Probes **11**(6): 397-406.

Li, G., Y. Feng, S. Kariyawasam, K. A. Tivendale, Y. Wannemuehler, F. Zhou, C. M. Logue, C. L. Miller and L. K. Nolan (2010). "AatA is a novel autotransporter and virulence factor of avian pathogenic Escherichia coli." Infect Immun **78**(3): 898-906.

Lindenthal, C. and E. A. Elsinghorst (1999). "Identification of a glycoprotein produced by enterotoxigenic Escherichia coli." Infect Immun **67**(8): 4084-4091.

Lintermans, P., P. Pohl, F. Deboeck, A. Bertels, C. Schlicker, J. Vandekerckhove, J. Van Damme, M. Van Montagu and H. De Greve (1988). "Isolation and nucleotide sequence of the F17-A gene encoding the structural protein of the F17 fimbriae in bovine enterotoxigenic Escherichia coli." Infect Immun **56**(6): 1475-1484.

Lintermans, P. F., A. Bertels, C. Schlicker, F. Deboeck, G. Charlier, P. Pohl, M. Norgren, S. Normark, M. van Montagu and H. De Greve (1991). "Identification, characterization, and nucleotide sequence of the F17-G gene, which determines receptor binding of Escherichia coli F17 fimbriae." J Bacteriol **173**(11): 3366-3373.

Liu, B., B. Hu, Z. Zhou, D. Guo, X. Guo, P. Ding, L. Feng and L. Wang (2012). "A novel non-homologous recombination-mediated mechanism for Escherichia coli unilateral flagellar phase variation." Nucleic Acids Res **40**(10): 4530-4538.

Liu, H., L. Magoun and J. M. Leong (1999). "beta1-chain integrins are not essential for intimin-mediated host cell attachment and enteropathogenic Escherichia coli-induced actin condensation." Infect Immun **67**(4): 2045-2049.

Liu, H., L. Magoun, S. Luperchio, D. B. Schauer and J. M. Leong (1999). "The Tir-binding region of enterohaemorrhagic Escherichia coli intimin is sufficient to trigger actin condensation after bacterial-induced host cell signalling." Mol Microbiol **34**(1): 67-81.

Liu, M. Y., G. Gui, B. Wei, J. F. Preston, 3rd, L. Oakford, U. Yuksel, D. P. Giedroc and T. Romeo (1997). "The RNA molecule CsrB binds to the global regulatory protein CsrA and antagonizes its activity in Escherichia coli." J Biol Chem **272**(28): 17502-17510.

Llosa, M., F. X. Gomis-Ruth, M. Coll and F. de la Cruz Fd (2002). "Bacterial conjugation: a two-step mechanism for DNA transport." Mol Microbiol **45**(1): 1-8.

Lloyd, A. L., T. A. Henderson, P. D. Vigil and H. L. Mobley (2009). "Genomic islands of uropathogenic Escherichia coli contribute to virulence." J Bacteriol **191**(11): 3469-3481.

Low, A. S., F. Dziva, A. G. Torres, J. L. Martinez, T. Rosser, S. Naylor, K. Spears, N. Holden, A. Mahajan, J. Findlay, J. Sales, D. G. Smith, J. C. Low, M. P. Stevens and D. L. Gally (2006). "Cloning, expression, and characterization of fimbrial operon F9 from enterohemorrhagic Escherichia coli O157:H7." Infect Immun **74**(4): 2233-2244.

Lower, B. H., R. Yongsunthon, F. P. Vellano, 3rd and S. K. Lower (2005). "Simultaneous force and fluorescence measurements of a protein that forms a bond between a living bacterium and a solid surface." J Bacteriol **187**(6): 2127-2137.

Lu, Y., S. Iyoda, H. Satou, H. Satou, K. Itoh, T. Saitoh and H. Watanabe (2006). "A new immunoglobulin-binding protein, EibG, is responsible for the chain-like adhesion phenotype of locus of enterocyte effacement-negative, shiga toxin-producing Escherichia coli." Infect Immun **74**(10): 5747-5755.

Luck, S. N., L. Badea, V. Bennett-Wood, R. Robins-Browne and E. L. Hartland (2006). "Contribution of FliC to epithelial cell invasion by enterohemorrhagic Escherichia coli O113:H21." Infect Immun **74**(12): 6999-7004.

Lugtenberg, B., R. Peters, H. Bernheimer and W. Berendsen (1976). "Influence of cultural conditions and mutations on the composition of the outer membrane proteins of Escherichia coli." Mol Gen Genet **147**(3): 251-262.

Luo, Q., P. Kumar, T. J. Vickers, A. Sheikh, W. G. Lewis, D. A. Rasko, J. Sistrunk and J. M. Fleckenstein (2014). "Enterotoxigenic Escherichia coli secretes a highly conserved mucin-degrading metalloprotease to effectively engage intestinal epithelial cells." Infect Immun **82**(2): 509-521.

Macarisin, D., J. Patel, G. Bauchan, J. A. Giron and V. K. Sharma (2012). "Role of curli and cellulose expression in adherence of Escherichia coli O157:H7 to spinach leaves." Foodborne Pathog Dis **9**(2): 160-167.

Macnab, R. M. (1987). "Escherichia Coli and Salmonella Typhimurium: Cellular and Molecular Biology " American Society for Microbiology **Vol I** 70-83, 732-759.

Maeda, S., M. Ito, T. Ando, Y. Ishimoto, Y. Fujisawa, H. Takahashi, A. Matsuda, A. Sawamura and S. Kato (2006). "Horizontal transfer of nonconjugative plasmids in a colony biofilm of Escherichia coli." FEMS Microbiol Lett **255**(1): 115-120.

Mahajan, A., C. G. Currie, S. Mackie, J. Tree, S. McAteer, I. McKendrick, T. N. McNeilly, A. Roe, R. M. La Ragione, M. J. Woodward, D. L. Gally and D. G. Smith (2009). "An investigation of the expression and adhesin function of H7 flagella in the interaction of Escherichia coli O157 : H7 with bovine intestinal epithelium." Cell Microbiol **11**(1): 121-137.

Mainil, J. G., G. Daube, E. Jacquemin, P. Pohl and A. Kaeckenbeeck (1998). "Virulence plasmids of enterotoxigenic Escherichia coli isolates from piglets." Vet Microbiol **62**(4): 291-301.

Mallick, E. M., M. J. Brady, S. A. Luperchio, V. K. Vanguri, L. Magoun, H. Liu, B. J. Sheppard, J. Mukherjee, A. Donohue-Rolfe, S. Tzipori, J. M. Leong and D. B. Schauer (2012). "Allele- and tir-independent functions of intimin in diverse animal infection models." Front Microbiol **3**: 11.

Mancini, J., B. Weckselblatt, Y. K. Chung, J. C. Durante, S. Andelman, J. Glaubman, J. D. Dorff, S. Bhargava, R. S. Lijek, K. P. Unger and I. N. Okeke (2011). "The heat-resistant agglutinin family includes a novel adhesin from enteroaggregative Escherichia coli strain 60A." J Bacteriol **193**(18): 4813-4820.

Mann, E. A. and S. A. Saeed (2012). "Gastrointestinal infection as a trigger for inflammatory bowel disease." Curr Opin Gastroenterol **28**(1): 24-29.

Manwaring, N. P., R. A. Skurray and N. Firth (1999). "Nucleotide sequence of the F plasmid leading region." Plasmid **41**(3): 219-225.

Mao, Y., M. P. Doyle and J. Chen (2006). "Role of colanic acid exopolysaccharide in the survival of enterohaemorrhagic Escherichia coli O157:H7 in simulated gastrointestinal fluids." Lett Appl Microbiol **42**(6): 642-647.

Martin, C. (1996). "The clp (CS31A) operon is negatively controlled by Lrp, ClpB, and L-alanine at the transcriptional level." Mol Microbiol **21**(2): 281-292.

Martin, C., C. Boeuf and F. Bousquet (1991). "Escherichia coli CS31A fimbriae: molecular cloning, expression and homology with the K88 determinant." Microb Pathog **10**(6): 429-442.

Martinez, E., M. A. Garcia-Viejo, J. L. Blanco, L. Bianchi, E. Buira, I. Conget, R. Casamitjana, J. Mallolas and J. M. Gatell (2000). "Impact of switching from human immunodeficiency virus type 1 protease inhibitors to efavirenz in successfully treated adults with lipodystrophy." Clin Infect Dis **31**(5): 1266-1273.

Martinez-Gil, M., K. G. K. Goh, E. Rackaityte, C. Sakamoto, B. Audrain, D. G. Moriel, M. Totsika, J. M. Ghigo, M. A. Schembri and C. Beloin (2017). "YeeJ is an inverse autotransporter from Escherichia coli that binds to peptidoglycan and promotes biofilm formation." Sci Rep **7**(1): 11326.

Mason, T. G. and G. Richardson (1981). "Escherichia coli and the human gut: some ecological considerations." J Appl Bacteriol **51**(1): 1-16.

Matthysse, A. G., R. Deora, M. Mishra and A. G. Torres (2008). "Polysaccharides cellulose, poly-beta-1,6-n-acetyl-D-glucosamine, and colanic acid are required for optimal binding of Escherichia coli O157:H7 strains to alfalfa sprouts and K-12 strains to plastic but not for binding to epithelial cells." Appl Environ Microbiol **74**(8): 2384-2390.

May, T. and S. Okabe (2008). "Escherichia coli harboring a natural IncF conjugative F plasmid develops complex mature biofilms by stimulating synthesis of colanic acid and Curli." J Bacteriol **190**(22): 7479-7490.

May, T., K. Tsuruta and S. Okabe (2011). "Exposure of conjugative plasmid carrying Escherichia coli biofilms to male-specific bacteriophages." ISME J **5**(4): 771-775.

Mayrand, D. and D. Grenier (1989). "Biological activities of outer membrane vesicles." Can J Microbiol **35**(6): 607-613.

Mazariego-Espinosa, K., A. Cruz, M. A. Ledesma, S. A. Ochoa and J. Xicohtencatl-Cortes (2010). "Longus, a type IV pilus of enterotoxigenic Escherichia coli, is involved in adherence to intestinal epithelial cells." J Bacteriol **192**(11): 2791-2800.

Mazel, D. and J. Davies (1999). "Antibiotic resistance in microbes." Cell Mol Life Sci **56**(9-10): 742-754.

McCrate, O. A., X. Zhou, C. Reichhardt and L. Cegelski (2013). "Sum of the parts: composition and architecture of the bacterial extracellular matrix." J Mol Biol **425**(22): 4286-4294.

McDaniel TK, J. K., Donnenberg MS, Kaper JB. (1995). "A genetic locus of enterocyte effacement conserved among diverse enterobacterial pathogens." Proceedings of the National Academy of Sciences of the United States of America 1664-1668.

McNeilly, T. N., S. W. Naylor, A. Mahajan, M. C. Mitchell, S. McAteer, D. Deane, D. G. Smith, J. C. Low, D. L. Gally and J. F. Huntley (2008). "Escherichia coli O157:H7 colonization in cattle following systemic and mucosal immunization with purified H7 flagellin." Infect Immun **76**(6): 2594-2602.

McWilliams, B. D. and A. G. Torres (2014). "EHEC Adhesins." Microbiol Spectr **2**(2): EHEC00032013.

Mechin, M. C., E. Rousset and J. P. Girardeau (1996). "Identification of surface-exposed linear B-cell epitopes of the nonfimbrial adhesin CS31A of Escherichia coli by using overlapping peptides and antipeptide antibodies." Infect Immun **64**(9): 3555-3564.

Melissa Starkey, M. R. P., Kimberly A. Gray, Sung Il Chang (2004). "A Sticky Business: the Extracellular Polymeric Substance Matrix of Bacterial Biofilms." Ghannoum M, O'Toole G (ed), Microbial Biofilms. : 74-191.

Melville, S. and L. Craig (2013). "Type IV pili in Gram-positive bacteria." Microbiol Mol Biol Rev **77**(3): 323-341.

Merino, S., J. G. Shaw and J. M. Tomas (2006). "Bacterial lateral flagella: an inducible flagella system." FEMS Microbiol Lett **263**(2): 127-135.

Merkel, V., B. Ohder, M. Bielaszewska, W. Zhang, A. Fruth, C. Menge, E. Borrmann, B. Middendorf, J. Muthing, H. Karch and A. Mellmann (2010). "Distribution and phylogeny of immunoglobulin-binding protein G in Shiga toxin-producing Escherichia coli and its association with adherence phenotypes." Infect Immun **78**(8): 3625-3636.

Molin, S. and T. Tolker-Nielsen (2003). "Gene transfer occurs with enhanced efficiency in biofilms and induces enhanced stabilisation of the biofilm structure." Curr Opin Biotechnol **14**(3): 255-261.

Monteiro, B. T., L. C. Campos, M. P. Sircili, M. R. Franzolin, L. F. Bevilacqua, J. P. Nataro and W. P. Elias (2009). "The dispersin-encoding gene (aap) is not restricted to enteroaggregative Escherichia coli." Diagn Microbiol Infect Dis **65**(1): 81-84.

Montero, D. A., J. Velasco, F. Del Canto, J. L. Puente, N. L. Padola, D. A. Rasko, M. Farfan, J. C. Salazar and R. Vidal (2017). "Locus of Adhesion and Autoaggregation (LAA), a pathogenicity island present in emerging Shiga Toxin-producing Escherichia coli strains." Sci Rep **7**(1): 7011.

Moreira, C. G., S. M. Carneiro, J. P. Nataro, L. R. Trabulsi and W. P. Elias (2003). "Role of type I fimbriae in the aggregative adhesion pattern of enteroaggregative Escherichia coli." FEMS Microbiol Lett **226**(1): 79-85.

Moreira, C. G., K. Palmer, M. Whiteley, M. P. Sircili, L. R. Trabulsi, A. F. Castro and V. Sperandio (2006). "Bundle-forming pili and EspA are involved in biofilm formation by enteropathogenic Escherichia coli." J Bacteriol **188**(11): 3952-3961.

Moriel, D. G., I. Bertoldi, A. Spagnuolo, S. Marchi, R. Rosini, B. Nesta, I. Pastorello, V. A. Corea, G. Torricelli, E. Cartocci, S. Savino, M. Scarselli, U. Dobrindt, J. Hacker, H. Tettelin, L. J. Tallon, S. Sullivan, L. H. Wieler, C. Ewers, D. Pickard, G. Dougan, M. R. Fontana, R. Rappuoli, M. Pizza and L. Serino (2010). "Identification of protective and broadly conserved vaccine antigens from the genome of extraintestinal pathogenic Escherichia coli." Proc Natl Acad Sci U S A **107**(20): 9072-9077.

Muller, C. M., A. Aberg, J. Straseviciene, L. Emody, B. E. Uhlin and C. Balsalobre (2009). "Type 1 fimbriae, a colonization factor of uropathogenic Escherichia coli, are controlled by the metabolic sensor CRP-cAMP." PLoS Pathog **5**(2): e1000303.

Muniesa, M., J. A. Hammerl, S. Hertwig, B. Appel and H. Brussow (2012). "Shiga toxin-producing Escherichia coli O104:H4: a new challenge for microbiology." Appl Environ Microbiol **78**(12): 4065-4073.

Musken, A., M. Bielaszewska, L. Greune, C. H. Schweppe, J. Muthing, H. Schmidt, M. A. Schmidt, H. Karch and W. Zhang (2008). "Anaerobic conditions promote expression of Sfp fimbriae and adherence of sorbitol-fermenting enterohemorrhagic Escherichia coli O157:NM to human intestinal epithelial cells." Appl Environ Microbiol **74**(4): 1087-1093.

Muto, Y. and S. Goto (1986). "Transformation by extracellular DNA produced by Pseudomonas aeruginosa." Microbiol Immunol **30**(7): 621-628.

Nagy, A., J. Mowery, G. R. Bauchan, L. Wang, L. Nichols-Russell and X. Nou (2015). "Role of Extracellular Structures of Escherichia coli O157:H7 in Initial Attachment to Biotic and Abiotic Surfaces." Appl Environ Microbiol **81**(14): 4720-4727.

Naili, I., B. Baudner, A. Bernalier-Donadille, M. Pizza, M. Desvaux, G. Jubelin and C. Buonsanti (2017). "In vivo characterization of the immune response towards the pathogenic Escherichia coli antigen SslE and modulation of the intestinal microbiota." JOURNAL OF IMMUNOLOGY **198 (1S)**: 147.123.

Naili, I., C. Buonsanti, M. Soriani, M. Desvaux, G. Jubelin, M. Vieru, B. Baudner and U. D'Oro (2016). "In vivo characterization of the immune response towards the pathogenic Escherichia coli antigen SslE and modulation of the intestinal microbiota. ." JOURNAL OF IMMUNOLOGY **196 (1S)**: 215.219.

Nakao, R., M. Ramstedt, S. N. Wai and B. E. Uhlin (2012). "Enhanced biofilm formation by Escherichia coli LPS mutants defective in Hep biosynthesis." PLoS One **7**(12): e51241.

Nataro, J. P., Y. Deng, D. R. Maneval, A. L. German, W. C. Martin and M. M. Levine (1992). "Aggregative adherence fimbriae I of enteroaggregative Escherichia coli mediate adherence to HEp-2 cells and hemagglutination of human erythrocytes." Infect Immun **60**(6): 2297-2304.

Nataro, J. P. and J. B. Kaper (1998). "Diarrheagenic Escherichia coli." Clin Microbiol Rev **11**(1): 142-201.

Nataro, J. P., D. Yikang, D. Yingkang and K. Walker (1994). "AggR, a transcriptional activator of aggregative adherence fimbria I expression in enteroaggregative Escherichia coli." J Bacteriol **176**(15): 4691-4699.

Navarro-Garcia, F. and W. P. Elias (2011). "Autotransporters and virulence of enteroaggregative E. coli." Gut Microbes **2**(1): 13-24.

Nesta, B., G. Spraggon, C. Alteri, D. G. Moriel, R. Rosini, D. Veggi, S. Smith, I. Bertoldi, I. Pastorello, I. Ferlenghi, M. R. Fontana, G. Frankel, H. L. Mobley, R. Rappuoli, M. Pizza, L. Serino and M. Soriani (2012). "FdeC, a novel broadly conserved Escherichia coli adhesin eliciting protection against urinary tract infections." MBio **3**(2).

Nesta, B., M. Valeri, A. Spagnuolo, R. Rosini, M. Mora, P. Donato, C. J. Alteri, M. Del Vecchio, S. Buccato, A. Pezzicoli, I. Bertoldi, L. Buzzigoli, G. Tuscano, M. Falduto, V. Rippa, Y. Ashhab, G. Bensi, M. R. Fontana, K. L. Seib, H. L. Mobley, M. Pizza, M. Soriani and L. Serino (2014). "SslE elicits functional antibodies that impair in vitro mucinase activity and in vivo colonization by both intestinal and extraintestinal Escherichia coli strains." PLoS Pathog **10**(5): e1004124.

Newton, H. J., J. Sloan, V. Bennett-Wood, L. M. Adams, R. M. Robins-Browne and E. L. Hartland (2004). "Contribution of long polar fimbriae to the virulence of rabbit-specific enteropathogenic Escherichia coli." Infect Immun **72**(3): 1230-1239.

Nicholls, L., T. Grant and R. Robins-Browne (2002). "Identification of a novel genetic locus that is required for in vitro adhesion of a clinical isolate of enterohaemorrhagic Escherichia coli to epithelial cells." Molecular Microbiology **35**(2): 275-288.

Niewerth, U., A. Frey, T. Voss, C. Le Bouguenec, G. Baljer, S. Franke and M. A. Schmidt (2001). "The AIDA autotransporter system is associated with F18 and stx2e in Escherichia coli isolates from pigs diagnosed with edema disease and postweaning diarrhea." Clin Diagn Lab Immunol **8**(1): 143-149.

Nilsson, G., J. G. Belasco, S. N. Cohen and A. von Gabain (1984). "Growth-rate dependent regulation of mRNA stability in Escherichia coli." Nature **312**(5989): 75-77.

Nishi J, S. J., Mizuguchi K et al. (2003). "The export of coat protein from enteroaggregative Escherichia coli by a specific ATP-binding cassette transporter system." J Biol Chem.

Nowicki, B., J. P. Barrish, T. Korhonen, R. A. Hull and S. I. Hull (1987). "Molecular cloning of the Escherichia coli O75X adhesin." Infect Immun **55**(12): 3168-3173.

Nowicki, B., A. Labigne, S. Moseley, R. Hull, S. Hull and J. Moulds (1990). "The Dr hemagglutinin, afimbrial adhesins AFA-I and AFA-III, and F1845 fimbriae of uropathogenic and diarrhea-associated Escherichia coli belong to a family of hemagglutinins with Dr receptor recognition." Infect Immun **58**(1): 279-281.

Nowicki, B., J. Moulds, R. Hull and S. Hull (1988). "A hemagglutinin of uropathogenic Escherichia coli recognizes the Dr blood group antigen." Infect Immun **56**(5): 1057-1060.

O'Toole, G., H. B. Kaplan and R. Kolter (2000). "Biofilm formation as microbial development." Annu Rev Microbiol **54**: 49-79.

Obadia, B., S. Lacour, P. Doublet, H. Baubichon-Cortay, A. J. Cozzone and C. Grangeasse (2007). "Influence of tyrosine-kinase Wzc activity on colanic acid production in Escherichia coli K12 cells." J Mol Biol **367**(1): 42-53.

Oberhettinger, P., M. Schutz, J. C. Leo, N. Heinz, J. Berger, I. B. Autenrieth and D. Linke (2012). "Intimin and invasin export their C-terminus to the bacterial cell surface using an inverse mechanism compared to classical autotransport." PLoS One **7**(10): e47069.

Ogasawara, H., K. Yamada, A. Kori, K. Yamamoto and A. Ishihama (2010). "Regulation of the Escherichia coli csgD promoter: interplay between five transcription factors." Microbiology **156**(Pt 8): 2470-2483.

Ogunleye, A., A. Bhat, V. U. Irorere, D. Hill, C. Williams and I. Radecka (2015). "Poly-gamma-glutamic acid: production, properties and applications." Microbiology **161**(Pt 1): 1-17.

Okeke, I. N., F. Wallace-Gadsden, H. R. Simons, N. Matthews, A. S. Labar, J. Hwang and J. Wain (2010). "Multi-locus sequence typing of enteroaggregative Escherichia coli isolates from Nigerian children uncovers multiple lineages." PLoS One **5**(11): e14093.

Olsen, A., A. Arnqvist, M. Hammar and S. Normark (1993). "Environmental regulation of curli production in Escherichia coli." Infect Agents Dis **2**(4): 272-274.

Omadjela, O., A. Narahari, J. Strumillo, H. Melida, O. Mazur, V. Bulone and J. Zimmer (2013). "BcsA and BcsB form the catalytically active core of bacterial cellulose synthase sufficient for in vitro cellulose synthesis." Proc Natl Acad Sci U S A **110**(44): 17856-17861.

Ong, C. L., S. A. Beatson, A. G. McEwan and M. A. Schembri (2009). "Conjugative plasmid transfer and adhesion dynamics in an Escherichia coli biofilm." Appl Environ Microbiol **75**(21): 6783-6791.

Orndorff, P. E., A. Devapali, S. Palestrant, A. Wyse, M. L. Everett, R. R. Bollinger and W. Parker (2004). "Immunoglobulin-mediated agglutination of and biofilm formation by Escherichia coli K-12 require the type 1 pilus fiber." Infect Immun **72**(4): 1929-1938.

Oropeza, R., R. Salgado-Bravo and E. Calva (2015). "Deletion analysis of RcsC reveals a novel signalling pathway controlling poly-N-acetylglucosamine synthesis and biofilm formation in Escherichia coli." Microbiology **161**(Pt 4): 903-913.

Otto, K. and M. Hermansson (2004). "Inactivation of ompX causes increased interactions of type 1 fimbriated Escherichia coli with abiotic surfaces." J Bacteriol **186**(1): 226-234.

Otto, K., J. Norbeck, T. Larsson, K. A. Karlsson and M. Hermansson (2001). "Adhesion of type 1-fimbriated Escherichia coli to abiotic surfaces leads to altered composition of outer membrane proteins." J Bacteriol **183**(8): 2445-2453.

Owen, P., M. Meehan, H. de Loughry-Doherty and I. Henderson (1996). "Phase-variable outer membrane proteins in Escherichia coli." FEMS Immunol Med Microbiol **16**(2): 63-76.

Palchevskiy, V. and S. E. Finkel (2006). "Escherichia coli competence gene homologs are essential for competitive fitness and the use of DNA as a nutrient." J Bacteriol **188**(11): 3902-3910.

Palmen, R. and K. J. Hellingwerf (1995). "Acinetobacter calcoaceticus liberates chromosomal DNA during induction of competence by cell lysis." Curr Microbiol **30**(1): 7-10.

Partridge, J. D. and R. M. Harshey (2013). "Swarming: flexible roaming plans." J Bacteriol **195**(5): 909-918.

Paton, A. W., P. Srimanote, M. C. Woodrow and J. C. Paton (2001). "Characterization of Saa, a novel autoagglutinating adhesin produced by locus of enterocyte effacement-negative Shiga-toxigenic Escherichia coli strains that are virulent for humans." Infect Immun **69**(11): 6999-7009.

Pawar, D. M., M. L. Rossman and J. Chen (2005). "Role of curli fimbriae in mediating the cells of enterohaemorrhagic Escherichia coli to attach to abiotic surfaces." J Appl Microbiol **99**(2): 418-425.

Pereira, A. L., T. N. Silva, A. C. Gomes, A. C. Araujo and L. G. Giugliano (2010). "Diarrhea-associated biofilm formed by enteroaggregative Escherichia coli and aggregative Citrobacter freundii: a consortium mediated by putative F pili." BMC Microbiol **10**: 57.

Perna, N. T., G. Plunkett, 3rd, V. Burland, B. Mau, J. D. Glasner, D. J. Rose, G. F. Mayhew, P. S. Evans, J. Gregor, H. A. Kirkpatrick, G. Posfai, J. Hackett, S. Klink, A. Boutin, Y. Shao, L. Miller, E. J. Grotbeck, N. W. Davis, A. Lim, E. T. Dimalanta, K. D. Potamousis, J. Apodaca, T. S. Anantharaman, J. Lin, G. Yen, D. C. Schwartz, R. A. Welch and F. R. Blattner (2001). "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7." Nature **409**(6819): 529-533.

Pettigrew, D., K. L. Anderson, J. Billington, E. Cota, P. Simpson, P. Urvil, F. Rabuzin, P. Roversi, B. Nowicki, L. du Merle, C. Le Bouguenec, S. Matthews and S. M. Lea (2004). "High resolution studies of the Afa/Dr adhesin DraE and its interaction with chloramphenicol." J Biol Chem **279**(45): 46851-46857.

Platenkamp, A. and J. L. Mellies (2018). "Environment Controls LEE Regulation in Enteropathogenic Escherichia coli." Front Microbiol **9**: 1694.

Platt, T. G. and C. Fuqua (2010). "What's in a name? The semantics of quorum sensing." Trends Microbiol **18**(9): 383-387.

Power, M. L., B. C. Ferrari, J. Littlefield-Wyer, D. M. Gordon, M. B. Slade and D. A. Veal (2006). "A naturally occurring novel allele of Escherichia coli outer membrane protein A reduces sensitivity to bacteriophage." Appl Environ Microbiol **72**(12): 7930-7932.

Prasadarao, N. V., C. A. Wass, J. N. Weiser, M. F. Stins, S. H. Huang and K. S. Kim (1996). "Outer membrane protein A of Escherichia coli contributes to invasion of brain microvascular endothelial cells." Infect Immun **64**(1): 146-153.

Pratt, L. A. and R. Kolter (1998). "Genetic analysis of Escherichia coli biofilm formation: roles of flagella, motility, chemotaxis and type I pili." Mol Microbiol **30**(2): 285-293.

Prigent-Combaret, C., E. Brombacher, O. Vidal, A. Ambert, P. Lejeune, P. Landini and C. Dorel (2001). "Complex regulatory network controls initial adhesion and biofilm formation in Escherichia coli via regulation of the csgD gene." J Bacteriol **183**(24): 7213-7223.

Prigent-Combaret, C., G. Prensier, T. T. Le Thi, O. Vidal, P. Lejeune and C. Dorel (2000). "Developmental pathway for biofilm formation in curli-producing Escherichia coli strains: role of flagella, curli and colanic acid." Environ Microbiol **2**(4): 450-464.

Prigent-Combaret, C., O. Vidal, C. Dorel and P. Lejeune (1999). "Abiotic surface sensing and biofilm-dependent regulation of gene expression in Escherichia coli." J Bacteriol **181**(19): 5993-6002.

Pruss, B. M. (2017). "Involvement of Two-Component Signaling on Bacterial Motility and Biofilm Development." J Bacteriol **199**(18).

Pruss, B. M., C. Besemann, A. Denton and A. J. Wolfe (2006). "A complex transcription network controls the early stages of biofilm development by Escherichia coli." J Bacteriol **188**(11): 3731-3739.

Qin, Z., Y. Ou, L. Yang, Y. Zhu, T. Tolker-Nielsen, S. Molin and D. Qu (2007). "Role of autolysin-mediated DNA release in biofilm formation of Staphylococcus epidermidis." Microbiology **153**(Pt 7): 2083-2092.

Qu, Z., F. Z. Temel, R. Henderikx and K. S. Breuer (2018). "Changes in the flagellar bundling time account for variations in swimming behavior of flagellated bacteria in viscous media." Proc Natl Acad Sci U S A **115**(8): 1707-1712.

Radchenkova, N., I. Boyadzhieva, N. Atanasova, A. Poli, I. Finore, P. Di Donato, B. Nicolaus, I. Panchev, M. Kuncheva and M. Kambourova (2018). "Extracellular polymer substance synthesized by a halophilic bacterium Chromohalobacter canadensis 28." Appl Microbiol Biotechnol **102**(11): 4937-4949.

Raetz, C. R. and C. Whitfield (2002). "Lipopolysaccharide endotoxins." Annu Rev Biochem **71**: 635-700.

Rajakumar, K., C. Sasakawa and B. Adler (1997). "Use of a novel approach, termed island probing, identifies the Shigella flexneri she pathogenicity island which encodes a homolog of the immunoglobulin A protease-like family of proteins." Infect Immun **65**(11): 4606-4614.

Ramer, S. W., D. Bieber and G. K. Schoolnik (1996). "BfpB, an outer membrane lipoprotein required for the biogenesis of bundle-forming pili in enteropathogenic Escherichia coli." J Bacteriol **178**(22): 6555-6563.

Ramer, S. W., G. K. Schoolnik, C. Y. Wu, J. Hwang, S. A. Schmidt and D. Bieber (2002). "The type IV pilus assembly complex: biogenic interactions among the bundle-forming pilus proteins of enteropathogenic Escherichia coli." J Bacteriol **184**(13): 3457-3465.

Ratiner Iu, A. (1967). "[Mutation of E. coli with regard to the H-antigen. I. Isolation of H-antigen mutants from test H-strains of Escherichia cultures]." Zh Mikrobiol Epidemiol Immunobiol **44**(10): 23-29.

Redfield, R. J. (2002). "Is quorum sensing a side effect of diffusion sensing?" Trends Microbiol **10**(8): 365-370.

Reidl, S., A. Lehmann, R. Schiller, A. Salam Khan and U. Dobrindt (2009). "Impact of O-glycosylation on the molecular and cellular adhesion properties of the Escherichia coli autotransporter protein Ag43." Int J Med Microbiol **299**(6): 389-401.

Reisner, A., J. A. Haagensen, M. A. Schembri, E. L. Zechner and S. Molin (2003). "Development and maturation of Escherichia coli K-12 biofilms." Mol Microbiol **48**(4): 933-946.

Reisner, A., M. Maierl, M. Jorger, R. Krause, D. Berger, A. Haid, D. Tesic and E. L. Zechner (2014). "Type 1 fimbriae contribute to catheter-associated urinary tract infections caused by Escherichia coli." J Bacteriol **196**(5): 931-939.

Ren, C. P., S. A. Beatson, J. Parkhill and M. J. Pallen (2005). "The Flag-2 locus, an ancestral gene cluster, is potentially associated with a novel flagellar system from Escherichia coli." J Bacteriol **187**(4): 1430-1440.

Rendon, M. A., Z. Saldana, A. L. Erdem, V. Monteiro-Neto, A. Vazquez, J. B. Kaper, J. L. Puente and J. A. Giron (2007). "Commensal and pathogenic Escherichia coli use a common pilus adherence factor for epithelial cell colonization." Proc Natl Acad Sci U S A **104**(25): 10637-10642.

Roe, A. J., C. Currie, D. G. Smith and D. L. Gally (2001). "Analysis of type 1 fimbriae expression in verotoxigenic Escherichia coli: a comparison between serotypes O157 and O26." Microbiology **147**(Pt 1): 145-152.

Rogers, T. J., C. M. Thorpe, A. W. Paton and J. C. Paton (2012). "Role of lipid rafts and flagellin in invasion of colonic epithelial cells by Shiga-toxigenic Escherichia coli O113:H21." Infect Immun **80**(8): 2858-2867.

Rojas-Lopez, M., M. M. Arenas-Hernandez, A. Medrano-Lopez, C. F. Martinez de la Pena, J. L. Puente, Y. Martinez-Laguna and A. G. Torres (2011). "Regulatory control of the Escherichia coli O157:H7 lpf1 operon by H-NS and Ler." J Bacteriol **193**(7): 1622-1632.

Rojas-Lopez, M., M. A. Zorgani, L. A. Kelley, X. Bailly, A. V. Kajava, I. R. Henderson, F. Polticelli, M. Pizza, R. Rosini and M. Desvaux (2017). "Identification of the Autochaperone Domain in the Type Va Secretion System (T5aSS): Prevalent Feature of Autotransporters with a beta-Helical Passenger." Front Microbiol **8**: 2607.

Romling, U., M. Rohde, A. Olsen, S. Normark and J. Reinkoster (2000). "AgfD, the checkpoint of multicellular and aggregative behaviour in Salmonella typhimurium regulates at least two independent pathways." Mol Microbiol **36**(1): 10-23.

Rose, P. W., A. Prlic, A. Altunkaya, C. Bi, A. R. Bradley, C. H. Christie, L. D. Costanzo, J. M. Duarte, S. Dutta, Z. Feng, R. K. Green, D. S. Goodsell, B. Hudson, T. Kalro, R. Lowe, E. Peisach, C. Randle, A. S. Rose, C. Shao, Y. P. Tao, Y. Valasatava, M. Voigt, J. D. Westbrook, J. Woo, H. Yang, J. Y. Young, C. Zardecki, H. M. Berman and S. K. Burley (2017). "The RCSB protein data bank: integrative view of protein, gene and 3D structural information." Nucleic Acids Res **45**(D1): D271-D281.

Rossi, E., A. Cimdins, P. Luthje, A. Brauner, A. Sjoling, P. Landini and U. Romling (2018). ""It's a gut feeling" - Escherichia coli biofilm formation in the gastrointestinal tract environment." Crit Rev Microbiol **44**(1): 1-30.

Roux, A., C. Beloin and J. M. Ghigo (2005). "Combined inactivation and expression strategy to study gene function under physiological conditions: application to identification of new Escherichia coli adhesins." J Bacteriol **187**(3): 1001-1013.

Roy, K., G. M. Hilliard, D. J. Hamilton, J. Luo, M. M. Ostmann and J. M. Fleckenstein (2009). "Enterotoxigenic Escherichia coli EtpA mediates adhesion between flagella and host cells." Nature **457**(7229): 594-598.

Ruiz-Perez, F. and J. P. Nataro (2014). "Bacterial serine proteases secreted by the autotransporter pathway: classification, specificity, and role in virulence." Cell Mol Life Sci **71**(5): 745-770.

Sailer, F. C., B. M. Meberg and K. D. Young (2003). "beta-Lactam induction of colanic acid gene expression in Escherichia coli." FEMS Microbiol Lett **226**(2): 245-249.

Sakellaris, H., N. K. Hannink, K. Rajakumar, D. Bulach, M. Hunt, C. Sasakawa and B. Adler (2000). "Curli loci of Shigella spp." Infect Immun **68**(6): 3780-3783.

Saldana, Z., J. Xicohtencatl-Cortes, F. Avelino, A. D. Phillips, J. B. Kaper, J. L. Puente and J. A. Giron (2009). "Synergistic role of curli and cellulose in cell adherence and biofilm formation of attaching and effacing Escherichia coli and identification of Fis as a negative regulator of curli." Environ Microbiol **11**(4): 992-1006.

Samadder, P., J. Xicohtencatl-Cortes, Z. Saldana, D. Jordan, P. I. Tarr, J. B. Kaper and J. A. Giron (2009). "The Escherichia coli ycbQRST operon encodes fimbriae with laminin-binding and epithelial cell adherence properties in Shiga-toxigenic E. coli O157:H7." Environ Microbiol **11**(7): 1815-1826.

Sanchez-Torres, V., H. Hu and T. K. Wood (2011). "GGDEF proteins YeaI, YedQ, and YfiN reduce early biofilm formation and swimming motility in Escherichia coli." Appl Microbiol Biotechnol **90**(2): 651-658.

Sandkvist, M. (2001). "Biology of type II secretion." Mol Microbiol **40**(2): 271-283.

Sandt, C. H. and C. W. Hill (2000). "Four different genes responsible for nonimmune immunoglobulin-binding activities within a single strain of Escherichia coli." Infect Immun **68**(4): 2205-2214.

Sandt, C. H. and C. W. Hill (2001). "Nonimmune binding of human immunoglobulin A (IgA) and IgG Fc by distinct sequence segments of the EibF cell surface protein of Escherichia coli." Infect Immun **69**(12): 7293-7303.

Sauer, F. G., M. A. Mulvey, J. D. Schilling, J. J. Martinez and S. J. Hultgren (2000). "Bacterial pili: molecular mechanisms of pathogenesis." Curr Opin Microbiol **3**(1): 65-72.

Savelkoul, P. H., G. A. Willshaw, M. M. McConnell, H. R. Smith, A. M. Hamers, B. A. van der Zeijst and W. Gaastra (1990). "Expression of CFA/I fimbriae is positively regulated." Microb Pathog **8**(2): 91-99.

Scaletsky, I. C., J. Michalski, A. G. Torres, M. V. Dulguer and J. B. Kaper (2005). "Identification and characterization of the locus for diffuse adherence, which encodes a novel afimbrial adhesin found in atypical enteropathogenic Escherichia coli." Infect Immun **73**(8): 4753-4765.

Schembri, M. A., D. Dalsgaard and P. Klemm (2004). "Capsule shields the function of short bacterial adhesins." J Bacteriol **186**(5): 1249-1257.

Schembri, M. A., K. Kjaergaard and P. Klemm (2003). "Global gene expression in Escherichia coli biofilms." Mol Microbiol **48**(1): 253-267.

Schiffrin, B., D. J. Brockwell and S. E. Radford (2017). "Outer membrane protein folding from an energy landscape perspective." BMC Biol **15**(1): 123.

Schmidt, M. A. (2010). "LEEways: tales of EPEC, ATEC and EHEC." Cell Microbiol **12**(11): 1544-1552.

Schooling, S. R. and T. J. Beveridge (2006). "Membrane vesicles: an overlooked component of the matrices of biofilms." J Bacteriol **188**(16): 5945-5957.

Schroeder GN, H. H. (2008). "Molecular pathogenesis of Shigella spp.: controlling host cell signaling, invasion, and death by type III secretion." Clinical microbiology reviews 134-156.

Serra, D. O., A. M. Richter and R. Hengge (2013). "Cellulose as an architectural element in spatially structured Escherichia coli biofilms." J Bacteriol **195**(24): 5540-5554.

Servin, A. L. (2005). "Pathogenesis of Afa/Dr diffusely adhering Escherichia coli." Clin Microbiol Rev **18**(2): 264-292.

Servin, A. L. (2014). "Pathogenesis of human diffusely adhering Escherichia coli expressing Afa/Dr adhesins (Afa/Dr DAEC): current insights and future challenges." Clin Microbiol Rev **27**(4): 823-869.

Shaw RK, B. C., Feys B et al. (2008). "Enterohemorrhagic Escherichia coli exploits EspA filaments for attachment to salad leaves." Appl Environ Microbiol 2908-2914.

Shaw, R. K., C. N. Berger, M. J. Pallen, A. Sjoling and G. Frankel (2011). "Flagella mediate attachment of enterotoxigenic Escherichia coli to fresh salad leaves." Environ Microbiol Rep **3**(1): 112-117.

Sheikh, A., R. Rashu, Y. A. Begum, F. M. Kuhlman, M. A. Ciorba, S. J. Hultgren, F. Qadri and J. M. Fleckenstein (2017). "Highly conserved type 1 pili promote enterotoxigenic E. coli pathogen-host interactions." PLoS Negl Trop Dis **11**(5): e0005586.

Sheikh, J., J. R. Czeczulin, S. Harrington, S. Hicks, I. R. Henderson, C. Le Bouguenec, P. Gounon, A. Phillips and J. P. Nataro (2002). "A novel dispersin protein in enteroaggregative Escherichia coli." J Clin Invest **110**(9): 1329-1337.

Sheikh, J., S. Hicks, M. Dall'Agnol, A. D. Phillips and J. P. Nataro (2001). "Roles for Fis and YafK in biofilm formation by enteroaggregative Escherichia coli." Mol Microbiol **41**(5): 983-997.

Sheng, H., J. Y. Lim, M. K. Watkins, S. A. Minnich and C. J. Hovde (2008). "Characterization of an Escherichia coli O157:H7 O-antigen deletion mutant and effect of the deletion on bacterial persistence in the mouse intestine and colonization at the bovine terminal rectal mucosa." Appl Environ Microbiol **74**(16): 5015-5022.

Sherlock, O., U. Dobrindt, J. B. Jensen, R. Munk Vejborg and P. Klemm (2006). "Glycosylation of the self-recognizing Escherichia coli Ag43 autotransporter protein." J Bacteriol **188**(5): 1798-1807.

Sherlock, O., R. M. Vejborg and P. Klemm (2005). "The TibA adhesin/invasin from enterotoxigenic Escherichia coli is self recognizing and induces bacterial aggregation and biofilm formation." Infect Immun **73**(4): 1954-1963.

Shimoji, Y., Y. Ogawa, M. Osaki, H. Kabeya, S. Maruyama, T. Mikami and T. Sekizaki (2003). "Adhesive surface proteins of Erysipelothrix rhusiopathiae bind to polystyrene, fibronectin, and type I and IV collagens." J Bacteriol **185**(9): 2739-2748.

Shrivastava S, M. S. (2008). "Identification and functional characterization of gene components of Type VI Secretion system in bacterial genomes." PloS one.

Sjoberg, P. O., M. Lindahl, J. Porath and T. Wadstrom (1988). "Purification and characterization of CS2, a sialic acid-specific haemagglutinin of enterotoxigenic Escherichia coli." Biochem J **255**(1): 105-111.

Sjoling, A., A. von Mentzer and A. M. Svennerholm (2015). "Implications of enterotoxigenic Escherichia coli genomics for vaccine development." Expert Rev Vaccines **14**(4): 551-560.

Sledjeski, D. and S. Gottesman (1995). "A small RNA acts as an antisilencer of the H-NS-silenced rcsA gene of Escherichia coli." Proc Natl Acad Sci U S A **92**(6): 2003-2007.

Smajs, D., H. Pilsl and V. Braun (1997). "Colicin U, a novel colicin produced by Shigella boydii." J Bacteriol **179**(15): 4919-4928.

Smati, M., O. Clermont, A. Bleibtreu, F. Fourreau, A. David, A. S. Daubie, C. Hignard, O. Loison, B. Picard and E. Denamur (2015). "Quantitative analysis of commensal Escherichia coli populations reveals host-specific enterotypes at the intra-species level." Microbiologyopen **4**(4): 604-615.

Smith, S. G., V. Mahon, M. A. Lambert and R. P. Fagan (2007). "A molecular Swiss army knife: OmpA structure, function and expression." FEMS Microbiol Lett **273**(1): 1-11.

Smyth, C. J., M. B. Marron, J. M. Twohig and S. G. Smith (1996). "Fimbrial adhesins: similarities and variations in structure and biogenesis." FEMS Immunol Med Microbiol **16**(2): 127-139.

Snoeck, V., W. Van den Broeck, V. De Colvenaer, F. Verdonck, B. Goddeeris and E. Cox (2008). "Transcytosis of F4 fimbriae by villous and dome epithelia in F4-receptor positive pigs supports importance of receptor-dependent endocytosis in oral immunization strategies." Vet Immunol Immunopathol **124**(1-2): 29-40.

Sohel, I., J. L. Puente, S. W. Ramer, D. Bieber, C. Y. Wu and G. K. Schoolnik (1996). "Enteropathogenic Escherichia coli: identification of a gene cluster coding for bundle-forming pilus morphogenesis." J Bacteriol **178**(9): 2613-2628.

Solano, C., B. Garcia, J. Valle, C. Berasain, J. M. Ghigo, C. Gamazo and I. Lasa (2002). "Genetic analysis of Salmonella enteritidis biofilm formation: critical role of cellulose." Mol Microbiol **43**(3): 793-808.

Sorensen, H. P. and K. K. Mortensen (2005). "Advanced genetic strategies for recombinant protein expression in Escherichia coli." J Biotechnol **115**(2): 113-128.

Spurbeck, R. R., A. E. Stapleton, J. R. Johnson, S. T. Walk, T. M. Hooton and H. L. Mobley (2011). "Fimbrial profiles predict virulence of uropathogenic Escherichia coli strains: contribution of ygi and yad fimbriae." Infect Immun **79**(12): 4753-4763.

Stathopoulos, C., D. R. Hendrixson, D. G. Thanassi, S. J. Hultgren, J. W. St Geme, 3rd and R. Curtiss, 3rd (2000). "Secretion of virulence determinants by the general secretory pathway in gram-negative pathogens: an evolving story." Microbes Infect **2**(9): 1061-1072.

Steinberger, R. E., A. R. Allen, H. G. Hansa and P. A. Holden (2002). "Elongation correlates with nutrient deprivation in Pseudomonas aeruginosa-unsaturates biofilms." Microb Ecol **43**(4): 416-423.

Steinmoen, H., E. Knutsen and L. S. Havarstein (2002). "Induction of natural competence in Streptococcus pneumoniae triggers lysis and DNA release from a subfraction of the cell population." Proc Natl Acad Sci U S A **99**(11): 7681-7686.

Stenutz, R., A. Weintraub and G. Widmalm (2006). "The structures of Escherichia coli O-polysaccharide antigens." FEMS Microbiol Rev **30**(3): 382-403.

Stevens, M. P. and G. M. Frankel (2014). "The Locus of Enterocyte Effacement and Associated Virulence Factors of Enterohemorrhagic Escherichia coli." Microbiol Spectr **2**(4): EHEC-0007-2013.

Stevens, M. P., A. J. Roe, I. Vlisidou, P. M. van Diemen, R. M. La Ragione, A. Best, M. J. Woodward, D. L. Gally and T. S. Wallis (2004). "Mutation of toxB and a truncated version of the efa-1 gene in Escherichia coli O157:H7 influences the expression and secretion of locus of enterocyte effacement-encoded proteins but not intestinal colonization in calves or sheep." Infect Immun **72**(9): 5402-5411.

Stevens, M. P., P. M. van Diemen, G. Frankel, A. D. Phillips and T. S. Wallis (2002). "Efa1 influences colonization of the bovine intestine by shiga toxin-producing Escherichia coli serotypes O5 and O111." Infect Immun **70**(9): 5158-5166.

Stevenson, G., K. Andrianopoulos, M. Hobbs and P. R. Reeves (1996). "Organization of the Escherichia coli K-12 gene cluster responsible for production of the extracellular polysaccharide colanic acid." J Bacteriol **178**(16): 4885-4893.

Strauss, J., N. A. Burnham and T. A. Camesano (2009). "Atomic force microscopy study of the role of LPS O-antigen on adhesion of E. coli." J Mol Recognit **22**(5): 347-355.

Sugawara, E. and H. Nikaido (1992). "Pore-forming activity of OmpA protein of Escherichia coli." J Biol Chem **267**(4): 2507-2511.

Supraja Puttamreddy, N. A. C. a. F. C. M. (2010). "Genome-Wide Transposon Mutagenesis Reveals a Role for pO157 Genes in Biofilm Development in Escherichia coli O157:H7 EDL93." American Society for Microbiology.

Sutherland, I. W. (2001). "The biofilm matrix--an immobilized but dynamic microbial environment." Trends Microbiol **9**(5): 222-227.

Svennerholm, A. M. and A. Lundgren (2012). "Recent progress toward an enterotoxigenic Escherichia coli vaccine." Expert Rev Vaccines **11**(4): 495-507.

Szabo, E., A. Skedsmo, A. Sonnevend, K. Al-Dhaheri, L. Emody, A. Usmani and T. Pal (2005). "Curli expression of enterotoxigenic Escherichia coli." Folia Microbiol (Praha) **50**(1): 40-46.

Szalo, I. M., F. Goffaux, V. Pirson, D. Pierard, H. Ball and J. Mainil (2002). "Presence in bovine enteropathogenic (EPEC) and enterohaemorrhagic (EHEC) Escherichia coli of genes encoding for putative adhesins of human EHEC strains." Res Microbiol **153**(10): 653-658.

Szalo, I. M. T., Bernard; Mainil, Jacques (2006). "Le lipopolysaccharide d'Escherichia coli : structure, biosynthèse et rôles." Annales de Médecine Vétérinaire.

Tacket, C. O., R. K. Taylor, G. Losonsky, Y. Lim, J. P. Nataro, J. B. Kaper and M. M. Levine (1998). "Investigation of the roles of toxin-coregulated pili and mannose-sensitive

hemagglutinin pili in the pathogenesis of Vibrio cholerae O139 infection." Infect Immun **66**(2): 692-695.

Taniguchi, T., Y. Fujino, K. Yamamoto, T. Miwatani and T. Honda (1995). "Sequencing of the gene encoding the major pilin of pilus colonization factor antigen III (CFA/III) of human enterotoxigenic Escherichia coli and evidence that CFA/III is related to type IV pili." Infect Immun **63**(2): 724-728.

Tarr, P. I., S. S. Bilge, J. C. Vary, Jr., S. Jelacic, R. L. Habeeb, T. R. Ward, M. R. Baylor and T. E. Besser (2000). "Iha: a novel Escherichia coli O157:H7 adherence-conferring molecule encoded on a recently acquired chromosomal island of conserved structure." Infect Immun **68**(3): 1400-1407.

Tatsuno, I., M. Horie, H. Abe, T. Miki, K. Makino, H. Shinagawa, H. Taguchi, S. Kamiya, T. Hayashi and C. Sasakawa (2001). "toxB gene on pO157 of enterohemorrhagic Escherichia coli O157:H7 is required for full epithelial cell adherence phenotype." Infect Immun **69**(11): 6660-6669.

Tatsuno, I., H. Kimura, A. Okutani, K. Kanamaru, H. Abe, S. Nagai, K. Makino, H. Shinagawa, M. Yoshida, K. Sato, J. Nakamoto, T. Tobe and C. Sasakawa (2000). "Isolation and characterization of mini-Tn5Km2 insertion mutants of enterohemorrhagic Escherichia coli O157:H7 deficient in adherence to Caco-2 cells." Infect Immun **68**(10): 5943-5952.

Tenaillon, O., D. Skurnik, B. Picard and E. Denamur (2010). "The population genetics of commensal Escherichia coli." Nat Rev Microbiol **8**(3): 207-217.

Teng, C. H., Y. Xie, S. Shin, F. Di Cello, M. Paul-Satyaseela, M. Cai and K. S. Kim (2006). "Effects of ompA deletion on expression of type 1 fimbriae in Escherichia coli K1 strain RS218 and on the association of E. coli with human brain microvascular endothelial cells." Infect Immun **74**(10): 5609-5616.

Thanassi, D. G., E. T. Saulino and S. J. Hultgren (1998). "The chaperone/usher pathway: a major terminal branch of the general secretory pathway." Curr Opin Microbiol **1**(2): 223-231.

Tobe, T. (2008). "The roles of two-component systems in virulence of pathogenic Escherichia coli and Shigella spp." Adv Exp Med Biol **631**: 189-199.

Tobe, T., C. Sasakawa, N. Okada, Y. Honma and M. Yoshikawa (1992). "vacB, a novel chromosomal gene required for expression of virulence genes on the large plasmid of Shigella flexneri." J Bacteriol **174**(20): 6359-6367.

Toma, C., E. Martinez Espinosa, T. Song, E. Miliwebsky, I. Chinen, S. Iyoda, M. Iwanaga and M. Rivas (2004). "Distribution of putative adhesins in different seropathotypes of Shiga toxin-producing Escherichia coli." J Clin Microbiol **42**(11): 4937-4946.

Tomme, P., A. L. Creagh, D. G. Kilburn and C. A. Haynes (1996). "Interaction of polysaccharides with the N-terminal cellulose-binding domain of Cellulomonas fimi CenC. 1. Binding specificity and calorimetric analysis." Biochemistry **35**(44): 13885-13894.

Torres, A. G. (2017). "Escherichia coli diseases in Latin America-a 'One Health' multidisciplinary approach." Pathog Dis **75**(2).

Torres, A. G., C. Jeter, W. Langley and A. G. Matthysse (2005). "Differential binding of Escherichia coli O157:H7 to alfalfa, human epithelial cells, and plastic is mediated by a variety of surface structures." Appl Environ Microbiol **71**(12): 8008-8015.

Torres, A. G., K. J. Kanack, C. B. Tutt, V. Popov and J. B. Kaper (2004). "Characterization of the second long polar (LP) fimbriae of Escherichia coli O157:H7 and distribution of LP fimbriae in other pathogenic E. coli strains." FEMS Microbiol Lett **238**(2): 333-344.

Torres, A. G. and J. B. Kaper (2003). "Multiple elements controlling adherence of enterohemorrhagic Escherichia coli O157:H7 to HeLa cells." Infect Immun **71**(9): 4985-4995.

Torres, A. G., L. Milflores-Flores, J. G. Garcia-Gallegos, S. D. Patel, A. Best, R. M. La Ragione, Y. Martinez-Laguna and M. J. Woodward (2007). "Environmental regulation and colonization attributes of the long polar fimbriae (LPF) of Escherichia coli O157:H7." Int J Med Microbiol **297**(3): 177-185.

Torres, A. G., N. T. Perna, V. Burland, A. Ruknudin, F. R. Blattner and J. B. Kaper (2002). "Characterization of Cah, a calcium-binding and heat-extractable autotransporter protein of enterohaemorrhagic Escherichia coli." Mol Microbiol **45**(4): 951-966.

Torres-Cruz, J. and M. W. van der Woude (2003). "Slipped-strand mispairing can function as a phase variation mechanism in Escherichia coli." J Bacteriol **185**(23): 6990-6994.

Totsika, M., T. J. Wells, C. Beloin, J. Valle, L. P. Allsopp, N. P. King, J. M. Ghigo and M. A. Schembri (2012). "Molecular characterization of the EhaG and UpaG trimeric autotransporter proteins from pathogenic Escherichia coli." Appl Environ Microbiol **78**(7): 2179-2189.

Touchon, M., C. Hoede, O. Tenaillon, V. Barbe, S. Baeriswyl, P. Bidet, E. Bingen, S. Bonacorsi, C. Bouchier, O. Bouvet, A. Calteau, H. Chiapello, O. Clermont, S. Cruveiller, A. Danchin, M. Diard, C. Dossat, M. E. Karoui, E. Frapy, L. Garry, J. M. Ghigo, A. M. Gilles, J. Johnson, C. Le Bouguenec, M. Lescat, S. Mangenot, V. Martinez-Jehanne, I. Matic, X. Nassif, S. Oztas, M. A. Petit, C. Pichon, Z. Rouy, C. S. Ruf, D. Schneider, J. Tourret, B. Vacherie, D. Vallenet, C. Medigue, E. P. Rocha and E. Denamur (2009). "Organised genome dynamics in the Escherichia coli species results in highly diverse adaptive paths." PLoS Genet **5**(1): e1000344.

Troge, A., W. Scheppach, B. O. Schroeder, S. A. Rund, K. Heuner, J. Wehkamp, E. F. Stange and T. A. Oelschlaeger (2012). "More than a marine propeller--the flagellum of the probiotic Escherichia coli strain Nissle 1917 is the major adhesin mediating binding to human mucus." Int J Med Microbiol **302**(7-8): 304-314.

Trunk, T., H. S. Khalil and J. C. Leo (2018). "Bacterial autoaggregation." Aims Microbiology **4**(1): 140-164.

Tsai, J. C., M. R. Yen, R. Castillo, D. L. Leyton, I. R. Henderson and M. H. Saier, Jr. (2010). "The bacterial intimins and invasins: a large and novel family of secreted proteins." PLoS One **5**(12): e14403.

Turner, N. C. A., J. P. R. Connolly and A. J. Roe (2018). "Control freaks-signals and cues governing the regulation of virulence in attaching and effacing pathogens." Biochem Soc Trans.

Ude, S., D. L. Arnold, C. D. Moon, T. Timms-Wilson and A. J. Spiers (2006). "Biofilm formation and cellulose expression among diverse environmental Pseudomonas isolates." Environ Microbiol **8**(11): 1997-2011.

Udekwu, K. I., F. Darfeuille, J. Vogel, J. Reimegard, E. Holmqvist and E. G. Wagner (2005). "Hfq-dependent regulation of OmpA synthesis is mediated by an antisense RNA." Genes Dev **19**(19): 2355-2366.

Ulett, G. C., A. N. Mabbett, K. C. Fung, R. I. Webb and M. A. Schembri (2007). "The role of F9 fimbriae of uropathogenic Escherichia coli in biofilm formation." Microbiology **153**(Pt 7): 2321-2331.

Ulett, G. C., J. Valle, C. Beloin, O. Sherlock, J. M. Ghigo and M. A. Schembri (2007). "Functional analysis of antigen 43 in uropathogenic Escherichia coli reveals a role in long-term persistence in the urinary tract." Infect Immun **75**(7): 3233-3244.

Ulett, G. C., R. I. Webb and M. A. Schembri (2006). "Antigen-43-mediated autoaggregation impairs motility in Escherichia coli." Microbiology **152**(Pt 7): 2101-2110.

Valle, J., A. N. Mabbett, G. C. Ulett, A. Toledo-Arana, K. Wecker, M. Totsika, M. A. Schembri, J. M. Ghigo and C. Beloin (2008). "UpaG, a new member of the trimeric

autotransporter family of adhesins in uropathogenic Escherichia coli." J Bacteriol **190**(12): 4147-4161.

Valverde, C. and D. Haas (2008). "Small RNAs controlled by two-component systems." Adv Exp Med Biol **631**: 54-79.

van der Woude, M. W., B. A. Braaten and D. A. Low (1992). "Evidence for global regulatory control of pilus expression in Escherichia coli by Lrp and DNA methylation: model building based on analysis of pap." Mol Microbiol **6**(17): 2429-2435.

van der Woude, M. W. and I. R. Henderson (2008). "Regulation and function of Ag43 (flu)." Annu Rev Microbiol **62**: 153-169.

van Diemen, P. M., F. Dziva, M. P. Stevens and T. S. Wallis (2005). "Identification of enterohemorrhagic Escherichia coli O26:H- genes required for intestinal colonization in calves." Infect Immun **73**(3): 1735-1743.

van Elsas, J. D., A. V. Semenov, R. Costa and J. T. Trevors (2011). "Survival of Escherichia coli in the environment: fundamental and public health aspects." ISME J **5**(2): 173-183.

Van Houdt, R. and C. W. Michiels (2005). "Role of bacterial cell surface structures in Escherichia coli biofilm formation." Res Microbiol **156**(5-6): 626-633.

Velarde JJ, V. K., Inman KG et al. (2007). "Solution structure of the novel dispersin protein of enteroaggregative Escherichia coli." Molecular microbiology 1123-1135.

Vidal, O., R. Longin, C. Prigent-Combaret, C. Dorel, M. Hooreman and P. Lejeune (1998). "Isolation of an Escherichia coli K-12 mutant strain able to form biofilms on inert surfaces: involvement of a new ompR allele that increases curli expression." J Bacteriol **180**(9): 2442-2449.

Vikram, A., G. K. Jayaprakasha, R. M. Uckoo and B. S. Patil (2013). "Inhibition of Escherichia coli O157:H7 motility and biofilm by beta-sitosterol glucoside." Biochim Biophys Acta **1830**(11): 5219-5228.

Vo, J. L., G. C. Martinez Ortiz, P. Subedi, S. Keerthikumar, S. Mathivanan, J. J. Paxman and B. Heras (2017). "Autotransporter Adhesins in Escherichia coli Pathogenesis." Proteomics **17**(23-24).

Vogel, C. and E. M. Marcotte (2012). "Insights into the regulation of protein abundance from proteomic and transcriptomic analyses." Nat Rev Genet **13**(4): 227-232.

Vogeleer, P., Y. D. Tremblay, G. Jubelin, M. Jacques and J. Harel (2015). "Biofilm-Forming Abilities of Shiga Toxin-Producing Escherichia coli Isolates Associated with Human Infections." Appl Environ Microbiol **82**(5): 1448-1458.

Vytvytska, O., I. Moll, V. R. Kaberdin, A. von Gabain and U. Blasi (2000). "Hfq (HF1) stimulates ompA mRNA decay by interfering with ribosome binding." Genes Dev **14**(9): 1109-1118.

Wakimoto, N., J. Nishi, J. Sheikh, J. P. Nataro, J. Sarantuya, M. Iwashita, K. Manago, K. Tokuda, M. Yoshinaga and Y. Kawano (2004). "Quantitative biofilm assay using a microtiter plate to screen for enteroaggregative Escherichia coli." Am J Trop Med Hyg **71**(5): 687-690.

Walk, S. T., E. W. Alm, D. M. Gordon, J. L. Ram, G. A. Toranzos, J. M. Tiedje and T. S. Whittam (2009). "Cryptic lineages of the genus Escherichia." Appl Environ Microbiol **75**(20): 6534-6544.

Wallecha, A., H. Oreh, M. W. van der Woude and P. L. deHaseth (2014). "Control of gene expression at a bacterial leader RNA, the agn43 gene encoding outer membrane protein Ag43 of Escherichia coli." J Bacteriol **196**(15): 2728-2735.

Wang, P., Y. Xiong, R. Lan, C. Ye, H. Wang, J. Ren, H. Jing, Y. Wang, Z. Zhou, Z. Cui, H. Zhao, Y. Chen, D. Jin, X. Bai, A. Zhao, Y. Wang, S. Zhang, H. Sun, J. Li, T. Wang, L. Wang and J. Xu (2011). "pO157_Sal, a novel conjugative plasmid detected in outbreak isolates of Escherichia coli O157:H7." J Clin Microbiol **49**(4): 1594-1597.

Wang, S., Y. Xia, J. Dai, Z. Shi, Y. Kou, H. Li, Y. Bao and C. Lu (2011). "Novel roles for autotransporter adhesin AatA of avian pathogenic Escherichia coli: colonization during infection and cell aggregation." FEMS Immunol Med Microbiol **63**(3): 328-338.

Wang, S., F. Yang and B. Yang (2017). "Global effect of CsrA on gene expression in enterohemorrhagic Escherichia coli O157:H7." Res Microbiol **168**(8): 700-709.

Wang, W., W. Chanda and M. Zhong (2015). "The relationship between biofilm and outer membrane vesicles: a novel therapy overview." FEMS Microbiol Lett **362**(15): fnv117.

Wang, X., J. F. Preston, 3rd and T. Romeo (2004). "The pgaABCD locus of Escherichia coli promotes the synthesis of a polysaccharide adhesin required for biofilm formation." J Bacteriol **186**(9): 2724-2734.

Weilbacher, T., K. Suzuki, A. K. Dubey, X. Wang, S. Gudapaty, I. Morozov, C. S. Baker, D. Georgellis, P. Babitzke and T. Romeo (2003). "A novel sRNA component of the carbon storage regulatory system of Escherichia coli." Mol Microbiol **48**(3): 657-670.

Weissman, S. J., S. Chattopadhyay, P. Aprikian, M. Obata-Yasuoka, Y. Yarova-Yarovaya, A. Stapleton, W. Ba-Thein, D. Dykhuizen, J. R. Johnson and E. V. Sokurenko (2006). "Clonal analysis reveals high rate of structural mutations in fimbrial adhesins of extraintestinal pathogenic Escherichia coli." Mol Microbiol **59**(3): 975-988.

Wells, T. J., T. N. McNeilly, M. Totsika, A. Mahajan, D. L. Gally and M. A. Schembri (2009). "The Escherichia coli O157:H7 EhaB autotransporter protein binds to laminin and collagen I and induces a serum IgA response in O157:H7 challenged cattle." Environ Microbiol **11**(7): 1803-1814.

Wells, T. J., O. Sherlock, L. Rivas, A. Mahajan, S. A. Beatson, M. Torpdahl, R. I. Webb, L. P. Allsopp, K. S. Gobius, D. L. Gally and M. A. Schembri (2008). "EhaA is a novel autotransporter protein of enterohemorrhagic Escherichia coli O157:H7 that contributes to adhesion and biofilm formation." Environ Microbiol **10**(3): 589-604.

Wells, T. J., M. Totsika and M. A. Schembri (2010). "Autotransporters of Escherichia coli: a sequence-based characterization." Microbiology **156**(Pt 8): 2459-2469.

Werneburg, G. T. and D. G. Thanassi (2018). "Pili Assembled by the Chaperone/Usher Pathway in Escherichia coli and Salmonella." EcoSal Plus **8**(1).

West, S. A., K. Winzer, A. Gardner and S. P. Diggle (2012). "Quorum sensing and the confusion about diffusion." Trends Microbiol **20**(12): 586-594.

Westerlund, B., P. Kuusela, J. Risteli, L. Risteli, T. Vartio, H. Rauvala, R. Virkola and T. K. Korhonen (1989). "The O75X adhesin of uropathogenic Escherichia coli is a type IV collagen-binding protein." Mol Microbiol **3**(3): 329-337.

Whitchurch, C. B., T. Tolker-Nielsen, P. C. Ragas and J. S. Mattick (2002). "Extracellular DNA required for bacterial biofilm formation." Science **295**(5559): 1487.

Whitfield, C. (2006). "Biosynthesis and assembly of capsular polysaccharides in Escherichia coli." Annu Rev Biochem **75**: 39-68.

Whitfield, C. and W. J. Keenleyside (1995). "Regulation of expression of group IA capsular polysaccharides in Escherichia coli and related extracellular polysaccharides in other bacteria." J Ind Microbiol **15**(4): 361-371.

Whitfield, C. and I. S. Roberts (1999). "Structure, assembly and regulation of expression of capsules in Escherichia coli." Mol Microbiol **31**(5): 1307-1319.

Wong, A. R., J. S. Pearson, M. D. Bright, D. Munera, K. S. Robinson, S. F. Lee, G. Frankel and E. L. Hartland (2011). "Enteropathogenic and enterohaemorrhagic Escherichia coli: even more subversive elements." Mol Microbiol **80**(6): 1420-1438.

Wood, T. K., A. F. Gonzalez Barrios, M. Herzberg and J. Lee (2006). "Motility influences biofilm architecture in Escherichia coli." Appl Microbiol Biotechnol **72**(2): 361-367.

Wright, K. J., P. C. Seed and S. J. Hultgren (2007). "Development of intracellular bacterial communities of uropathogenic Escherichia coli depends on type 1 pili." Cell Microbiol **9**(9): 2230-2241.

Wuertz, S., S. Okabe and M. Hausner (2004). "Microbial communities and their interactions in biofilm systems: an overview." Water Sci Technol **49**(11-12): 327-336.

Wurpel, D. J., S. A. Beatson, M. Totsika, N. K. Petty and M. A. Schembri (2013). "Chaperone-usher fimbriae of Escherichia coli." PLoS One **8**(1): e52835.

Xi, C. and J. Wu (2010). "dATP/ATP, a multifunctional nucleotide, stimulates bacterial cell lysis, extracellular DNA release and biofilm development." PLoS One **5**(10): e13355.

Xicohtencatl-Cortes, J., V. Monteiro-Neto, Z. Saldana, M. A. Ledesma, J. L. Puente and J. A. Giron (2009). "The type 4 pili of enterohemorrhagic Escherichia coli O157:H7 are multipurpose structures with pathogenic attributes." J Bacteriol **191**(1): 411-421.

Xicohtencatl-Cortes, J., Z. Saldana, W. Deng, E. Castaneda, E. Freer, P. I. Tarr, B. B. Finlay, J. L. Puente and J. A. Giron (2010). "Bacterial macroscopic rope-like fibers with cytopathic and adhesive properties." J Biol Chem **285**(42): 32336-32342.

Xicohtencatl-Cortes, J., E. Sanchez Chacon, Z. Saldana, E. Freer and J. A. Giron (2009). "Interaction of Escherichia coli O157:H7 with leafy green produce." J Food Prot **72**(7): 1531-1537.

Yang, J., D. L. Baldi, M. Tauschek, R. A. Strugnell and R. M. Robins-Browne (2007). "Transcriptional regulation of the yghJ-pppA-yghG-gspCDEFGHIJKLM cluster, encoding the type II secretion pathway in enterotoxigenic Escherichia coli." J Bacteriol **189**(1): 142-150.

Yaron, S. and U. Romling (2014). "Biofilm formation by enteric pathogens and its role in plant colonization and persistence." Microb Biotechnol **7**(6): 496-516.

Yin, X., R. Wheatcroft, J. R. Chambers, B. Liu, J. Zhu and C. L. Gyles (2009). "Contributions of O island 48 to adherence of enterohemorrhagic Escherichia coli O157:H7 to epithelial cells in vitro and in ligated pig ileal loops." Appl Environ Microbiol **75**(18): 5779-5786.

Yonezawa, H., T. Osaki, S. Kurata, M. Fukuda, H. Kawakami, K. Ochiai, T. Hanawa and S. Kamiya (2009). "Outer membrane vesicles of Helicobacter pylori TK1402 are involved in biofilm formation." BMC Microbiol **9**: 197.

Yoshikane Itoh, J. D. R., Carlos Goller, Archana Pannuri, Jeannette Taylor, Jeffrey Meisner, Terry J. Beveridge, James F. Preston III, and Tony Romeo (2008). "Roles of pgaABCD Genes in Synthesis, Modification, and Export of the Escherichia coli Biofilm Adhesin Poly-β-1,6-N-Acetyl-D-Glucosamine." American Society for Microbiology.

Young, G. M., D. H. Schmiel and V. L. Miller (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**(11): 6456-6461.

Yu, B., C. L. Giltner, E. J. van Schaik, D. L. Bautista, R. S. Hodges, G. F. Audette, D. Y. Li and T. Randall (2007). "A Novel Biometallic Interface: High Affinity Tip-AssociatedBinding by Pilin-Derived Protein Nanotubes." Journal of Bionanoscience **1**: 73-83.

Zhao, H., C. Chen, Y. Xiong, X. Xu, R. Lan, H. Wang, X. Yao, X. Bai, X. Liu, Q. Meng, X. Zhang, H. Sun, A. Zhao, X. Bai, Y. Cheng, Q. Chen, C. Ye and J. Xu (2013). "Global transcriptional and phenotypic analyses of Escherichia coli O157:H7 strain Xuzhou21 and its pO157_Sal cured mutant." PLoS One **8**(5): e65466.

Zhou, M., Q. Duan, X. Zhu, Z. Guo, Y. Li, P. R. Hardwidge and G. Zhu (2013). "Both flagella and F4 fimbriae from F4ac+ enterotoxigenic Escherichia coli contribute to attachment to IPEC-J2 cells in vitro." Vet Res **44**: 30.

Zhou, M., Z. Guo, Y. Yang, Q. Duan, Q. Zhang, F. Yao, J. Zhu, X. Zhang, P. R. Hardwidge and G. Zhu (2014). "Flagellin and F4 fimbriae have opposite effects on biofilm formation

and quorum sensing in F4ac+ enterotoxigenic Escherichia coli." Vet Microbiol **168**(1): 148-153.

Zhou, M., Y. Yang, P. Chen, H. Hu, P. R. Hardwidge and G. Zhu (2015). "More than a locomotive organelle: flagella in Escherichia coli." Appl Microbiol Biotechnol **99**(21): 8883-8890.

Zogaj, X., M. Nimtz, M. Rohde, W. Bokranz and U. Romling (2001). "The multicellular morphotypes of Salmonella typhimurium and Escherichia coli produce cellulose as the second component of the extracellular matrix." Mol Microbiol **39**(6): 1452-1463.

Zorraquino, V., B. Garcia, C. Latasa, M. Echeverz, A. Toledo-Arana, J. Valle, I. Lasa and C. Solano (2013). "Coordinated cyclic-di-GMP repression of Salmonella motility through YcgR and cellulose." J Bacteriol **195**(3): 417-428.

Zschiedrich, C. P., V. Keidel and H. Szurmant (2016). "Molecular Mechanisms of Two-Component Signal Transduction." J Mol Biol **428**(19): 3752-3775.

Zude, I., A. Leimbach and U. Dobrindt (2014). "Prevalence of autotransporters in Escherichia coli: what is the impact of phylogeny and pathotype?" Int J Med Microbiol **304**(3-4): 243-256.

Table

Table 1: **Molecular determinants involved in surface colonisation in DEC**.

Colonisation factor	Role	DEC pathotype
Exopolysaccharide (EPS)		
LPS	Adhesion to abiotic surfaces, biofilm formation, intestinal colonisation	EHEC, all other DEC*
PNAG	Biofilm formation on biotic and abiotic surfaces	EHEC, ETEC*
Colanic acid	Biofilm formation	EHEC, EAEC
Cellulose	Biofilm formation	EHEC, EPEC, ETEC*
Extracellular DNA (eDNA)	Biofilm formation	DEC*
Cell-surface protein		
Monomeric protein		
EhaB	Binding to laminin and collagen I	EHEC, EPEC*, ETEC*
EhaJ	Binding to fibronectin, fibrinogen, laminin, collagens II, III and IV, biofilm formation	EHEC, EPEC*, ETEC*, EAEC*, EIEC*
EhaD/YpjA	Adhesion to abiotic surfaces, biofilm formation	EHEC, EAEC*
EhaC/YcgV/YfaL	Adhesion to abiotic surfaces, biofilm formation	EHEC*, EPEC*, ETEC*, EAEC*, EIEC*
UpaB	Adhesion to ECM proteins	EPEC*, EHEC*

UpaI	Adhesion to ECM proteins, biofilm formation, autoaggregation	EHEC*, EPEC*
AatA	Adhesion to chicken fibroblast cells	EAEC*
Ag43	Autoaggregation, adhesion to intestinal epithelial cells, biofilm formation	EHEC, EPEC*, ETEC*, EAEC*
AIDA-I	Autoaggregation, adhesion to intestinal epithelial cells, biofilm formation	EPEC, DAEC, STEC*, ETEC*
TibA	Autoaggregation, colonisation of intestinal epithelium, biofilm formation	ETEC
EhaA	Autoaggregation, adhesion to intestinal epithelial cells, biofilm formation	EHEC, EPEC*, ETEC*, EAEC*
UpaC	Autoaggregation, biofilm formation	EHEC*, EPEC*, ETEC*, EAEC*, EIEC*
EspP	Adhesion to intestinal epithelial cells, biofilm formation	EHEC, EPEC*
EspC	Adhesion to intestinal epithelial cells, biofilm formation	EPEC
Pic/Shmu	Binding to mucin, intestinal colonisation, biofilm formation	EAEC, EHEC*, EIEC
Intimin	A/E lesions, intestinal colonisation	EHEC, EPEC
FdeC	Colonisation of terminal rectum of cattle, biofilm formation	EHEC, all other DEC*
YeeJ	Biofilm formation	EHEC*, EPEC*, ETEC*, EIEC*
OmpA	Adhesion to intestinal epithelial cells and plant surfaces	EHEC
Hra1	Autoaggregation, aggregative adhesion, biofilm formation	EAEC, EPEC*
Hra2	Adhesion to intestinal epithelial cells	EAEC

Tia	Adhesion/invasion to intestinal epithelial cells	ETEC, EPEC*, EAEC*
Hes	Autoaggregation, erythrocyte agglutination, adhesion to intestinal epithelial cells, biofilm formation	EHEC
Iha	Diffuse adherence pattern, intestinal colonisation	EHEC, EPEC*
Dispersin	Intestinal colonisation	EAEC, EHEC*
SslE	Binding and degradation of mucin, biofilm formation, intestinal colonisation	EPEC, EHEC*, ETEC*
Efa-1	Intestinal colonisation, adhesion to epithelial cells	EHEC, EPEC
GAPDH	Binding to plasminogen and fibrinogen	EHEC, EPEC, all other DEC*
EF-Tu	Bacterial aggregation	DEC*
Multimeric protein complex		
UpaG/EhaG	Autoaggregation, adhesion to intestinal epithelial cells, laminin, fibronectin and collagens I to IV, biofilm formation	EHEC, EPEC*, ETEC*, EAEC*, EIEC*
Eib	Binding to immunoglobulin fraction of the Fc fraction of IgA and IgG, Chain-like adhesion, autoaggregation, biofilm formation	EHEC
Sab	Diffuse adherence to intestinal epithelial cells, biofilm formation to abiotic surfaces	EHEC
Saa	Adhesion to intestinal epithelial cells	EHEC
Flagella	Active motility, adhesion to intestinal eptihelial cells and plant leaves, biofilm formation	EHEC, EPEC, ETEC, EAEC
Injectisome	Adhesion to mammalian and plant cells, pedestal formation during attaching and effacing lesions, biofilm formation	EHEC, EPEC, EIEC

T4P	Bacterial aggregation, adhesion to intestinal epithelial cells, fibronectin and laminin, biofilm formation, twitching motility	EHEC, EPEC, ETEC
CP	Bacterial conjugation, adhesion and biofilm formation to abiotic surfaces	EAEC, EHEC*
T1P	Adhesion to abiotic surfaces and eukaryotic cells, biofilm formation	EHEC, EPEC, ETEC, EAEC
AAF	Adhesion to intestinal epithelial cells, fibronectin, laminin, collagen IV, decay accelerating factor and carcinoembryonic antigen-related cell adhesion molecules , biofilm formation	EAEC, EIEC, DAEC, ETEC*, EHEC*
CFA	Intestinal adhesion	ETEC
F9	Adhesion to intestinal epithelial cells and fibronectin	EHEC, EPEC*, EAEC*
ELF	Adhesion to intestinal epithelial cells and laminin	EHEC, ETEC
CS31A	Intestinal colonisation	EPEC, ETEC, DAEC, EHEC*
LPF	Binding to fibronectin, laminin, collagen IV and follicle-associated epithelium of Peyer's patches, biofilm formation	EHEC, STEC, EPEC, ETEC*, EAEC*
ECP	Autoaggregation, adherence to intestinal epithelial cells	EHEC, EPEC*, ETEC*, EAEC*, EIEC*
SFP	Adhesion to intestinal epithelial cells	EHEC
Curli	Adhesion intestinal epithelial cells, plant cells and abiotic surfaces, cell-cell interactions, biofilm formation	EHEC, EPEC, ETEC
Hcp	Bacterial aggregation, biofilm formation	EHEC, EPEC, EAEC

* Protein encoding-genes are present but they have not been functionally characterised as yet.

Figure legends

Figure 1: Schematic representation of the exopolymeric matrix (EPM) in biofilm of DEC. By analogy with the extracellular matrix (ECM) in mammalian tissue, the EPM in bacterial biofilm can be further discriminated between (i) the EPM closely associated with the bacterial cells, i.e. the cell-associated EPM (caEPM) (purple shade background), and (ii) the interstitial EPM (iEPM) (white background). Molecular determinants of the caEPM are essentially attached, anchored or linked to the bacterial cell surface. Besides cell-surface proteinaceous determinants including monomeric proteins (not depicted on the picture) and supramolecular protein structures, such the flagella and pili, molecular components of caEPM further comprise some extracellular polysaccharides (EPS), namely some lipopolysaccharides (LPS) as well as poly-β-1,6-N-acetyl-D-glucosamine (PNAG) and colanic acid, which both form a capsule. Together with colanic acid that can be released from the bacterial cell surface, cellulose can compose the EPS part of the iEPM. Besides, extracellular DNA (eDNA) and some exoproteins (not depicted on the picture) can also constitute the iEPM in *E. coli* biofilm.

Figure 2: Schematic representation of the cell-surface proteinaceous determinants acting as colonisation factors (CFs) in DEC. Monomeric proteins are depicted in shades of red, whereas multimeric protein complexes are depicted in shades of blue. Whenever possible, molecular structures where obtained from the protein databank (PDB) (Berman *et al.* 2002, Rose *et al.* 2017) or the electron microscopy databank (EMBD) (Lawson 2010). Regarding autotransporters (ATs), no structure for autotransporter adhesins (ATAs) is currently available but Ag43 (PDB: 4KH3) is provided as a representative of a self-associating autotransporter (SAAT) and EspP (PDB: 3SLI, 3SZE) as a representative of serine protease autotransporter of *Enterobacteriaceae* (SPATE). Intimin (PDB: 3NCW, 4E1S) is given as representative of an inverted autotransporter (IAT). Proteins secreted across the OM by the T5SS are first exported via the Sec translocase (SecYEG-DF/SecA) (PDB: 2AKH, 3AQO, 5XAM) across the inner membrane (IM). Dispersin (PDB 2JVU) is secreted via T1SS (PDB: 5066). Besides ATs, all OMPs including the Hra, OmpA (PDB: 2GE4) and Iha are first exported via Sec before being integrated to the OM via the β-barrel assembly machinery (BAM) complex (BamABCDE) (PDB: 5LJO). The surface-associated lipoprotein of *E. coli* (SslE) is secreted by a T2aSS (EMDB: 1763, PDB: 3CIO, 3OSS, 4KSR, 2W7V, 2BH1) after Sec export. Like for moonlighting proteins here represented by the GAPDH (PDB: 5ZA0), the secretion mechanisms of Efa-1 remain unknown. EibD (PDB: 2XQH) is provided as a representative of trimeric autotransporters (TAAs). The injectisome is secreted and assembled by the T3aSS (EMDB: 1875). The flagellum (EMDB: 1132, 1873; PDB: 1IO1) is secreted and assembled by the T3bSS (EMDB: 1887). The T4P (EMDB: 0070) is secreted and assembled by the T2bSS. The conjugative pilus (CP) (PDB: 5LEG) is secreted and assembled by the T4bSS (EMDB: 2567). The T1P (EMDB: 3222), CS31A, AAF (PDB: IUT2, 2XQ), CFA (EMDB: 1952), F9 pilus, ELF, LPF (PDB: 5AFO), ECP (PDB: 3QS3) and SFP are all secreted and assembled by T7SS (PDB: 4J3O) after Sec export. The curli are secreted and assembled by the T8SS (EMDB: 2750). Hcp form a tube, which is displayed extracellularly upon triggering of the T6SS (EMDB: 2524; PDB: 4HKH, 3RX9, 4JIV).

Figure 3: Regulation levels and control mechanisms for the expression of genes encoding colonisation factors in DEC. At least five regulation levels can be considered in bacteria. While the sequential steps and events for gene/protein expression flow from RNA transcription to protein translation (as depicted by blue arrows), the hierarchy of control mechanisms and regulations levels (as depicted by shades of red) is at the opposite.

Figures

Figure 1

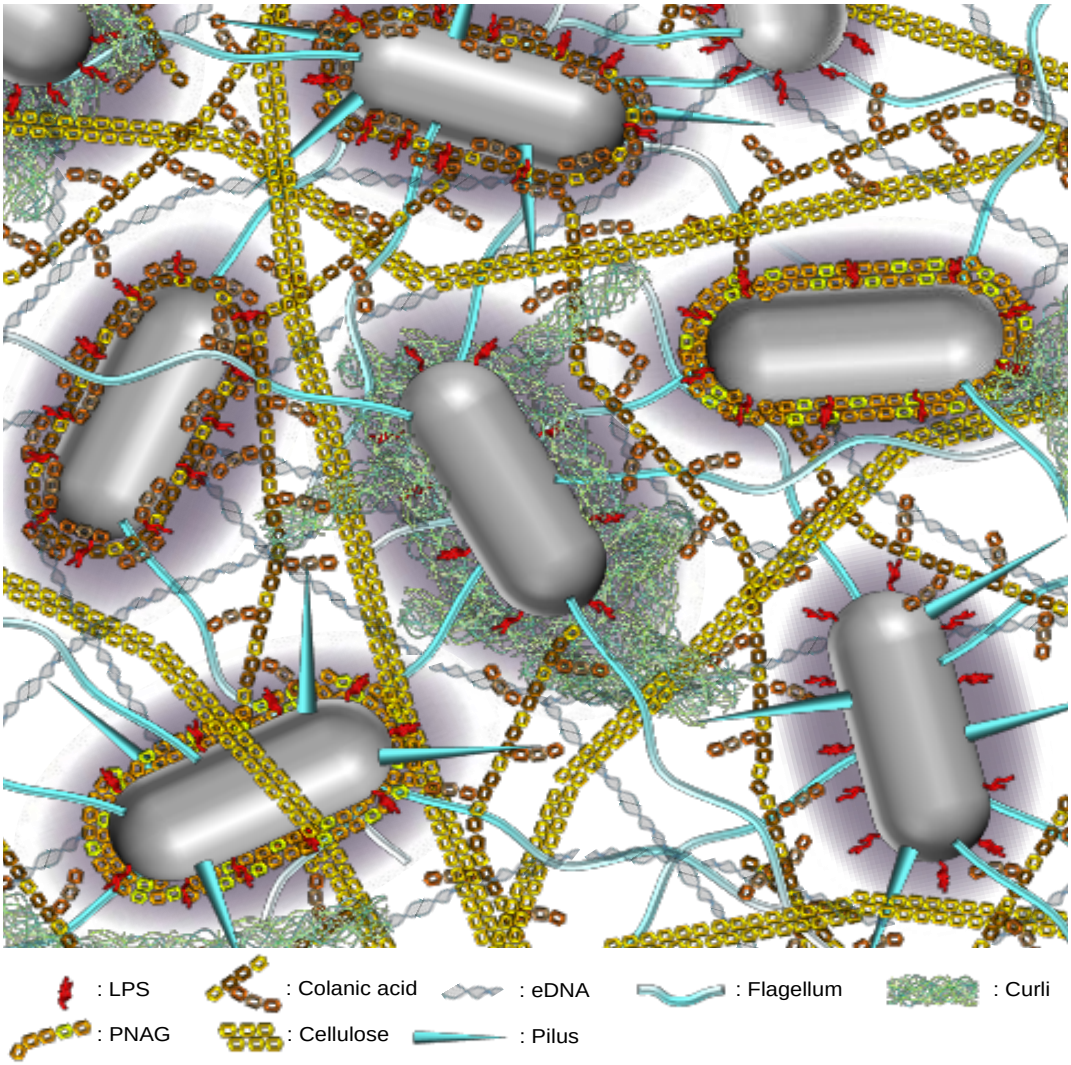

Figure 2

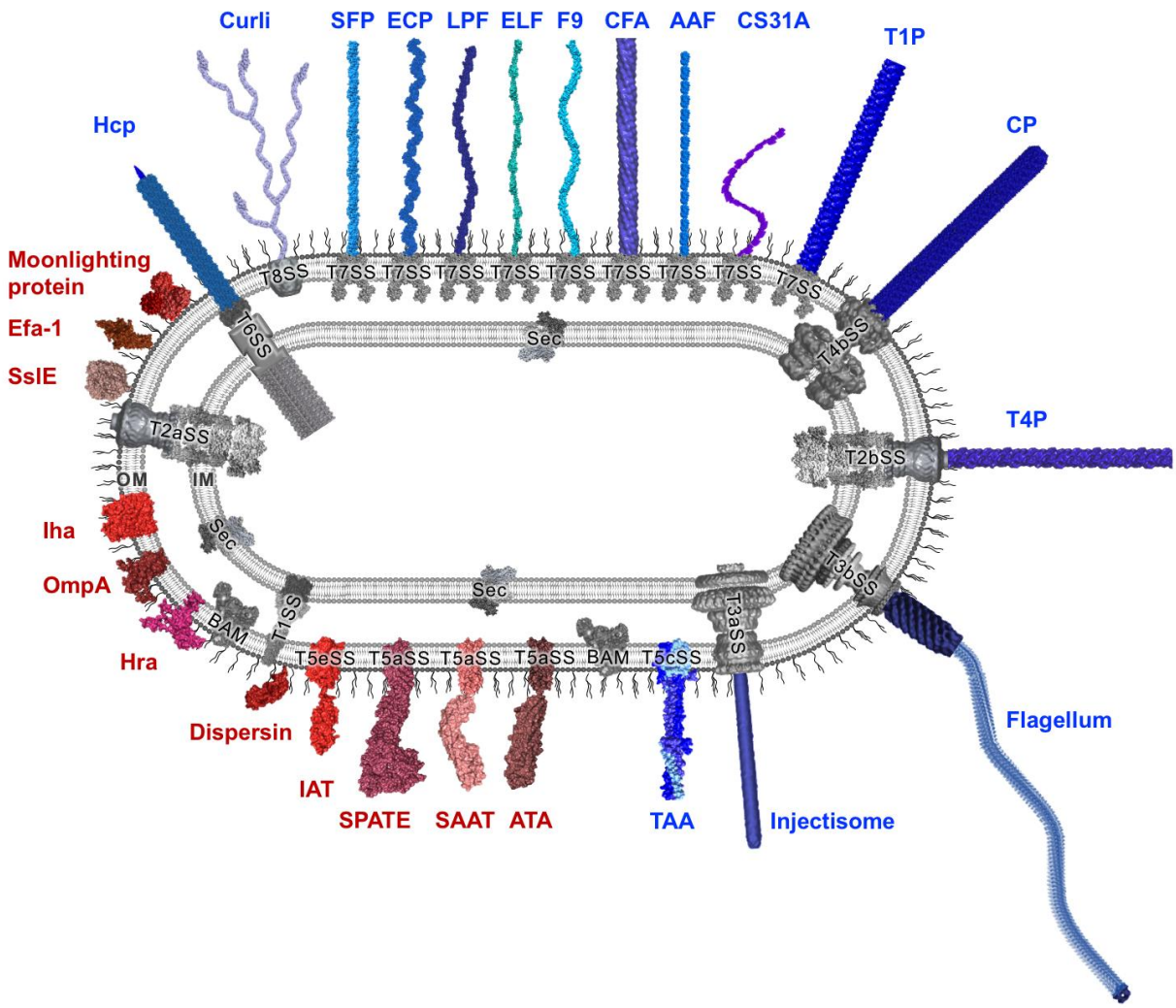
Curli
SFP
ECP
LPF
ELF
F9
CFA
AAF
CS31A
T1P
Hcp
CP
Moonlighting protein
Efa-1
SsIE
T8SS
T7SS
T6SS
Sec
T4bSS
T2aSS
T2bSS
T4P
OM
IM
Iha
OmpA
Hra
BAM
T1SS
T5eSS
T5aSS
T3aSS
T3bSS
T5cSS
Dispersin
IAT
SPATE
SAAT
ATA
TAA
Injectisome
Flagellum

Figure 3

Regulation levels	Control mechanisms	Examples
Pre-transcriptional	**DNA replication (phase variation):**	
	-DNA inversion	***fim* operon (T1P)**
	-Slipped-strand mispairing	
	-DNA methylation	***flu* (Ag43), *clp* operon (CS31A)**
	-DNA deletion	***fliC* (flagellin)**
Transcriptional	**Rate of mRNA synthesis (regulators / effectors):**	***pga* operon (PNAG), *wca* operon (colonic acid)**
	-Positive control of an inducible gene	***cfa* operon (CFA)**
	-Positive control of a repressible gene	
	-Negative control of an inducible gene	
	-Negative control of a repressible gene	***clp* operon (CS31A)**
Post-transcriptional	**Modulate transcripts for translation initiation:**	***flu* (Ag43), *pga* operon (PNAG)**
	-Stability of the mRNA (half-life)	***pap* (CU pili)**
	-Riboswitch	
	-Attenuation	
Translational	**Rate of protein synthesis:**	
	-Anti-sense RNA (incl. small RNA)	**PNAG**
	-Riboregulation	**OmpA**
	-Codon usage (translation rate)	
Post-translational	**Modulate protein activity:**	**Ag43**
	-Physical parameters (pH, etc...)	
	-Inducers and inhibitors (allostery)	**Cellulose**
	-Proteolytic cleavage	
	-Chemical modifications (glycosylation, etc..)	**TibA**
	-Protein folding (chaperones)	
	-Association/dissociation multimers	
	-Stability of the protein (half-life)	**Colonic acid**
	-Translocation and final subcellular localisation	

ANNEXE n°3 : Informations supplémentaires de l'article n°2

Valentin AGEORGES, Marion SCHIAVONE, Grégory JUBELIN, Nelly CACCIA, Philippe RUIZ, Xavier BAILLY, Etienne DAGUE, Sabine LEROY, Jason PAXMAN, Begoña HERAS, Frédérique CHAUCHEYRAS-DURAND, Amanda E. ROSSITER, Ian R. HENDERSON and Mickaël DESVAUX.

Differential homotypic and heterotypic interactions of antigen 43 (Ag43) variants in autotransporter-mediated bacterial autoaggregation

Annexe n°3

SUPPLEMENTARY INFORMATION

Differential homotypic and heterotypic interactions of antigen 43 (Ag43) variants in autotransporter-mediated bacterial autoaggregation

Valentin AGEORGES [1], Marion SCHIAVONE [2,3], Grégory JUBELIN [1], Nelly CACCIA [1], Philippe RUIZ [1], Ingrid CHAFSEY [1], Xavier BAILLY [5], Etienne DAGUE [4], Sabine LEROY [1], Jason PAXMAN [6], Begoña HERAS [6], Frédérique CHAUCHEYRAS-DURAND [1,2], Amanda E. ROSSITER [7], Ian R. HENDERSON [7,8] and Mickaël DESVAUX [1,*]

[1] Université Clermont Auvergne, INRA, UMR454 MEDiS, 63000 Clermont-Ferrand, France

[2] Lallemand Animal Nutrition, 31702 Blagnac Cedex, France.

[3] LISBP, Université de Toulouse, CNRS UMR5504, INRA UMR792, INSA, 31077 Toulouse Cedex, France.

[4] LAAS-CNRS, Université de Toulouse, 31077 Toulouse Cedex, France.

[5] INRA UR346 Epidémiologie Animale, 63122 Saint Genès Champanelle, France.

[6] Department of Biochemistry and Genetics, La Trobe Institute for Molecular Science, La Trobe University, Bundoora, Victoria 3086, Australia.

[7] Institute of Microbiology and Infection, University of Birmingham, B152TT Birmingham, United Kingdom.

[8] Institute for Molecular Biosciences, University of Queensland, St. Lucia, Queensland 4067, Australia.

Table S1: Dataset of 106 Ag43 passengers. Provided as an independent Excel file.

Table S2: Complete list of bacterial strains and plasmids used in this study.

Table S3: Quality check of the Ag43 structural models.

Figure S1: Phylogenetic network trees of Ag43 passengers as well as SL, EJ and BL subdomains. Provided as independent scalable Nexus files. Network trees were generated using SplitsTree including the 106-entries dataset sequences without applying bootstrapping and confidence threshold and A: Tree of the Ag43 passengers. B: Tree of the SL subdomain. C: Tree of the EJ subdomain. D: Tree of the BL. The scale bars represent the number of substitutions per site.

Figure S2: Structure models of Ag43 passengers belonging to C1, C2, C3 and C4. Provided as independent PDB files. Three-dimensional structures of Ag43 presented in Figure 2B were modelled with Phyre, checked and modified with COOT.

Figure S3: Flow cytometry analysis of fluorescent-tagged bacterial cell aggregates. A: SRUD pBAD was used to define events corresponding to individual single cells (red gate) on a SSC-A over FSC-A representation. B: SRUD Ag43 $^{C1\text{-}PT}$ were used to determine events corresponding to bacterial cell aggregates (pink gate) versus individual single cells (red gate) on a SSC-A over FSC-A representation. C: SSC-A over FSC-A representation using *E. coli* SRUD-Green pBAD and SRUD-Blue pBAD indicating the presence of individual single cells (red gate) and the absence of co-interaction as defined above with SRUD pBAD (A). D: Focusing on the gate corresponding to the individual single cells as defined above (C), fluorescent FITC-H over PB450-H representation using SRUD-Green pBAD and SRUD-Blue pBAD allowed discriminating the individual single cells of SRUD-Green pBAD (green gate) and of SRUD-Blue pBAD (blue gate); black dots would correspond to event where two green and blue cells passed simultaneously (representing less than 5% of total events). E: SSC-A over FSC-A representation using *E. coli* SRUD-Green Ag43 $^{C1\text{-}PT}$ and SRUD-Blue Ag43 $^{C1\text{-}PT}$ indicating the formation of bacterial cell aggregates (pink gate) as defined above with

SRUD Ag43 $^{C1-PT}$ (B). F: Focusing on the gate corresponding to the bacterial cell aggregates defined above (E), fluorescent FITC-H over PB450-H representation using SRUD-Green Ag43 $^{C1-PT}$ and SRUD-Blue Ag43 $^{C1-PT}$ allowed to discriminate the aggregates formed by SRUD-Green Ag43 $^{C1-PT}$ (green gate), the aggregates formed by SRUD-Blue Ag43 $^{C1-PT}$ (blue gate) and the mixed-aggregates formed by SRUD-Green Ag43 $^{C1-PT}$/ SRUD-Blue Ag43 $^{C1-PT}$ (orange gate) here indicative of homotypic interactions (referring to interactions between identical Ag43).

Figure S4: Flow cytometry analyses of co-cultures of nonaggregative with aggregative fluorescent-tagged bacterial cells. Flow cytometry analyses were performed by mixing nonaggregative (NA) SRUD-Green pBAD and SRUD-Blue expressing different Ag43 variants, namely Ag43 $^{C1-PT}$ (C1), Ag43 $^{C2-PT}$ (C2), Ag43 $^{C3-PT}$ (C3) or Ag43 $^{C4-PT}$ (C4). A: SSC-A over FSC-A representation indicated the formation of bacterial cell aggregates (pink cloud). Events corresponding to single cells (red cloud) and bacterial cell aggregates were selected using gates previously defined in control experiments (Figure S3). B: Focusing on the gate corresponding to the bacterial cell aggregates, fluorescent FITC-H over PB450-H representation indicated the presence of aggregates only for cells expressing Ag43 variants (blue cloud).

Figure S5: Immunodetection and subcellular localisation of Ag43 by Western blot analyses. A: Expression level of Ag43 from whole cell extracts (1 µg of total protein per lane); Lane 1: *E. coli* SRUD pBAD, Lane 2: *E. coli* SRUD Ag43 $^{C1-PT}$, Lane 3: *E. coli* SRUD Ag43 $^{C2-PT}$, Lane 4: *E. coli* SRUD Ag43 $^{C3-PT}$, Lane 5: *E. coli* SRUD Ag43 $^{C4-PT}$. Subcellular localisation of Ag43 $^{C1-PT}$ (B), Ag43 $^{C2-PT}$ (C), Ag43 $^{C3-PT}$ (D) and Ag43 $^{C4-PT}$ (E) in *E. coli* SRUD determined through immunodetection in different cell compartments, *i.e.* whole cell (lane 1), OM (lane 2), IM (lane 3) and soluble (lane 4) fractions. MW: molecular weight marker (KDa).

Figure S6: Measurement of adhesion forces for homotypic interactions of bacterial cells expressing different Ag43 variants by AFM. Analyses were performed using SRUD expressing the

Ag43 from the different classes, namely Ag43 $^{C1\text{-}PT}$ (C1), Ag43 $^{C2\text{-}PT}$ (C2), Ag43 $^{C3\text{-}PT}$ (C3) or Ag43 $^{C4\text{-}PT}$ (C4). n.a.: non adhesive events.

Figure S7: Measurement of adhesion forces for heterotypic interactions of bacterial cells expressing different Ag43 variants by AFM. Analyses were performed by mixing SRUD expressing Ag43 from different classes, respectively Ag43 $^{C1\text{-}PT}$ (C1), Ag43 $^{C2\text{-}PT}$ (C2), Ag43 $^{C3\text{-}PT}$ (C3) or Ag43 $^{C4\text{-}PT}$ (C4). n.a.: non adhesive events.

Table S2: Complete list of bacterial strains and plasmids used in this study

Bacterial strain/plasmid	Relevant genotype and/or phenotype	References [a]
E. coli strains		
SRUD	*E. coli* K12 M C1655 Δ*fim* Δ*flu*::*kan*	1
SRUD pBAD	Empty pBAD/Myc-HisA vector in *E. coli* SRUD	This study
SRUD Ag43 $^{C1\text{-}UPI00003B25EE}$	pBAD-Ag43 $^{C1\text{-}UPI00003B25EE}$ in *E. coli* SRUD	This study
SRUD Ag43 $^{C1\text{-}UPI000224229A}$	pBAD-Ag43 $^{C1\text{-}UPI000224229A}$ in *E. coli* SRUD	This study
SRUD Ag43 $^{C1\text{-}UPI0001F8D949}$	pBAD-Ag43 $^{C1\text{-}UPI0001F8D949}$ in *E. coli* SRUD	This study
SRUD Ag43 $^{C2\text{-}UPI00000D07E5}$	pBAD-Ag43 $^{C2\text{-}UPI00000D07E5}$ in *E. coli* SRUD	This study
SRUD Ag43 $^{C2\text{-}UPI0001F68B55}$	pBAD-Ag43 $^{C2\text{-}UPI0001F68B55}$ in *E. coli* SRUD	This study
SRUD Ag43 $^{C2\text{-}UPI00005F0EEA}$	pBAD-Ag43 $^{C2\text{-}UPI00005F0EEA}$ in *E. coli* SRUD	This study
SRUD Ag43 $^{C3\text{-}UPI00000E4A66}$	pBAD-Ag43 $^{C3\text{-}UPI00000E4A66}$ in *E. coli* SRUD	This study
SRUD Ag43 $^{C3\text{-}UPI0001EB6ED2}$	pBAD-Ag43 $^{C3\text{-}UPI0001EB6ED2}$ in *E. coli* SRUD	This study
SRUD Ag43 $^{C3\text{-}UPI00000E49D2}$	pBAD-Ag43 $^{C3\text{-}UPI00000E49D2}$ in *E. coli* SRUD	This study
SRUD Ag43 $^{C3\text{-}UPI000037640F}$	pBAD- Ag43 $^{C3\text{-}UPI000037640F}$ in *E. coli* SRUD	This study
SRUD Ag43 $^{C3\text{-}UPI0001888EFC}$	pBAD- Ag43 $^{C3\text{-}UPI0001888EFC}$ in *E. coli* SRUD	This study
SRUD Ag43 $^{C3\text{-}UPI000000CD3C}$	pBAD- Ag43 $^{C3\text{-}UPI000000CD3C}$ in *E. coli* SRUD	This study
SRUD Ag43 $^{C4\text{-}UPI0001E6D993}$	pBAD-Ag43 $^{C4\text{-}UPI0001E6D993}$ in *E. coli* SRUD	This study
SRUD Ag43 $^{C4\text{-}UPI00092D946C}$	pBAD-Ag43 $^{C4\text{-}UPI00092D946C}$ in *E. coli* SRUD	This study
SRUD Ag43 $^{C4\text{-}UPI00092DE8CA}$	pBAD-Ag43 $^{C4\text{-}UPI00092DE8CA}$ in *E. coli* SRUD	This study
SRUD-Green pBAD	pVaAgGreen in *E. coli* SRUD pBAD	This study

SRUD-Red pBAD	pVaAgRed in *E. coli* SRUD pBAD	This study
SRUD-Blue pBAD	pVaAgBlue in *E. coli* SRUD pBAD	This study
SRUD-Green Ag43 $^{C1\text{-}UPI00003B25EE}$	pVaAgGreen in *E. coli* SRUD pBAD-Ag43 $^{C1\text{-}UPI00003B25EE}$	This study
SRUD-Red Ag43 $^{C1\text{-}UPI00003B25EE}$	pVaAgRed in *E. coli* SRUD pBAD-Ag43 $^{C1\text{-}UPI00003B25EE}$	This study
SRUD-Blue Ag43 $^{C1\text{-}UPI00003B25EE}$	pVaAgBlue in *E. coli* SRUD pBAD-Ag43 $^{C1\text{-}UPI00003B25EE}$	This study
SRUD-Green Ag43 $^{C2\text{-}UPI00000D07E5}$	pVaAgGreen in *E. coli* SRUD pBAD-Ag43 $^{C2\text{-}UPI00000D07E5}$	This study
SRUD-Red Ag43 $^{C2\text{-}UPI00000D07E5}$	pVaAgRed in *E. coli* SRUD pBAD-Ag43 $^{C2\text{-}UPI00000D07E5}$	This study
SRUD-Blue Ag43 $^{C2\text{-}UPI00000D07E5}$	pVaAgBlue in *E. coli* SRUD pBAD-Ag43 $^{C2\text{-}UPI00000D07E5}$	This study
SRUD-Green Ag43 $^{C3\text{-}UPI00000E49D2}$	pVaAgGreen in *E. coli* SRUD pBAD-Ag43 $^{C3\text{-}UPI00000E49D2}$	This study
SRUD-Red Ag43 $^{C3\text{-}UPI00000E49D2}$	pVaAgRed in *E. coli* SRUD pBAD-Ag43 $^{C3\text{-}UPI00000E49D2}$	This study
SRUD-Blue Ag43 $^{C3\text{-}UPI00000E49D2}$	pVaAgBlue in *E. coli* SRUD pBAD-Ag43 $^{C3\text{-}UPI00000E49D2}$	This study
SRUD-Green Ag43 $^{C4\text{-}UPI0001E6D993}$	pVaAgGreen in *E. coli* SRUD pBAD-Ag43 $^{C4\text{-}UPI0001E6D993}$	This study
SRUD-Red Ag43 $^{C4\text{-}UPI0001E6D993}$	pVaAgRed in *E. coli* SRUD pBAD-Ag43 $^{C4\text{-}UPI0001E6D993}$	This study
SRUD-Blue Ag43 $^{C4\text{-}UPI0001E6D993}$	pVaAgBlue in *E. coli* SRUD pBAD-Ag43 $^{C4\text{-}UPI0001E6D993}$	This study
Plasmids		
pBAD/Myc-HisA	Protein expression vector under the control of the inducible *ara*BAD promoter (P_{BAD}), Amp^R	Thermo Fisher Scientific
pBAD-Ag43 $^{C1\text{-}UPI00003B25EE}$	Gene encoding Ag43 $^{C1\text{-}UPI00003B25EE}$ in XhoI-HindIII digested pBAD/Myc-HisA	This study
pBAD-Ag43 $^{C1\text{-}UPI000224229A}$	Gene encoding Ag43 $^{C1\text{-}UPI000224229A}$ in XhoI-HindIII digested pBAD/Myc-HisA	This study
pBAD-Ag43 $^{C1\text{-}UPI0001F8D949}$	Gene encoding Ag43 $^{C1\text{-}UPI0001F8D949}$ in XhoI-HindIII digested pBAD/Myc-HisA	This study
pBAD-Ag43 $^{C2\text{-}UPI00000D07E5}$	Gene encoding Ag43 $^{C2\text{-}UPI00000D07E5}$ in XhoI-HindIII digested pBAD/Myc-HisA	This study
pBAD-Ag43 $^{C2\text{-}UPI0001F68B55}$	Gene encoding Ag43 $^{C2\text{-}UPI0001F68B55}$ in XhoI-HindIII digested pBAD/Myc-HisA	This study

pBAD-Ag43 $^{C2\text{-}UPI00005F0EEA}$	Gene encoding Ag43 $^{C2\text{-}UPI00005F0EEA}$ in XhoI-HindIII digested pBAD/Myc-HisA	This study
pBAD-Ag43 $^{C3\text{-}UPI00000E49D2}$	Gene encoding Ag43 $^{C3\text{-}UPI00000E49D2}$ in XhoI-HindIII digested pBAD/Myc-HisA	This study
pBAD-Ag43 $^{C3\text{-}UPI00000E4A66}$	Gene encoding Ag43 $^{C3\text{-}UPI00000E4A66}$ in XhoI-HindIII digested pBAD/Myc-HisA	This study
pBAD-Ag43 $^{C3\text{-}UPI000037640F}$	Gene encoding Ag43 $^{C3\text{-}UPI000037640F}$ in XhoI-HindIII digested pBAD/Myc-HisA	This study
pBAD-Ag43 $^{C3\text{-}UPI0001EB6ED2}$	Gene encoding Ag43 $^{C3\text{-}UPI0001EB6ED2}$ in XhoI-HindIII digested pBAD/Myc-HisA	This study
pBAD-Ag43 $^{C3\text{-}UPI0001888EFC}$	Gene encoding Ag43 $^{C3\text{-}UPI0001888EFC}$ in XhoI-HindIII digested pBAD/Myc-HisA	This study
pBAD-Ag43 $^{C3\text{-}UPI000000CD3C}$	Gene encoding Ag43 $^{C3\text{-}UPI000000CD3C}$ in XhoI-HindIII digested pBAD/Myc-HisA	This study
pBAD-Ag43 $^{C4\text{-}UPI0001E6D993}$	Gene encoding Ag43 $^{C4\text{-}UPI0001E6D993}$ in XhoI-HindIII digested pBAD/Myc-HisA	This study
pBAD-Ag43 $^{C4\text{-}UPI00092D946C}$	Gene encoding Ag43 $^{C4\text{-}UPI00092D946C}$ in XhoI-HindIII digested pBAD/Myc-HisA	This study
pBAD-Ag43 $^{C4\text{-}UPI00092DE8CA}$	Gene encoding Ag43 $^{C4\text{-}UPI00092DE8CA}$ in XhoI-HindIII digested pBAD/Myc-HisA	This study
pACYC184	Cloning vector with a p15A replication origin, Cm^R, Tc^R	2
pVaAgGreen	$P_{BBa\text{-}J23119}$::RBS^{pBAD}::*mEmerald* in AvaI-Bsu36I digested pACYC184	This study
pVaAgBlue	$P_{BBa\text{-}J23119}$::RBS^{pBAD}::*mTagBFP2* in AvaI-Bsu36I digested pACYC184	This study
pVaAgRed	$P_{BBa\text{-}J23119}$::RBS^{pBAD}::*mRuby2* in AvaI-Tth111I digested pACYC184	This study

[a] 1. Reidl, S., Lehmann, A., Schiller, R., Salam Khan, A. & Dobrindt, U. Impact of O-glycosylation on the molecular and cellular adhesion properties of the *Escherichia coli* autotransporter protein Ag43. *Int. J. Med. Microbiol*. **6**, 389-401 (2009).

2. Chang, A.C. & Cohen, S.N. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. *J. Bacteriol*. **3**, 1141-1156 (1978).

Table S3: Quality check of the Ag43 structural models.

Models [a]	Generously allowed (%)	Disallowed (%)
Ag43 C1-UPI00003B25EE (PT)	0.5	0.0
Ag43 C1-UPI000224229A	0.2	0.0
Ag43 C1-UPI000F8D949	0.7	0.0
Ag43 C2-UPI00000D07E5 (PT)	1.2	0.0
Ag43 C2-UPI0001F68B55	0.6	0.3
Ag43 C2-UPI00005F0EEA	1.2	0.0
Ag43 C3-UPI00000E49D2 (PT)	0.5	0.0
Ag43 C3-UPI000IEB6ED2	0.7	0.0
Ag43 C3-UP1000037640F	0.2	0.0
Ag43 C4-UPI0001E6D933 (PT)	1.2	0.0
Ag43 C4-UP100092D946C	1.5	0.3
Ag43 C4-UP100092DE8CA	2.7	0.0

[a] The quality of the Ag43 homolog models was assessed using ProCheck. PDB files of these Ag43 structural models are provided as supplementary materials. PT: prototypical variant.

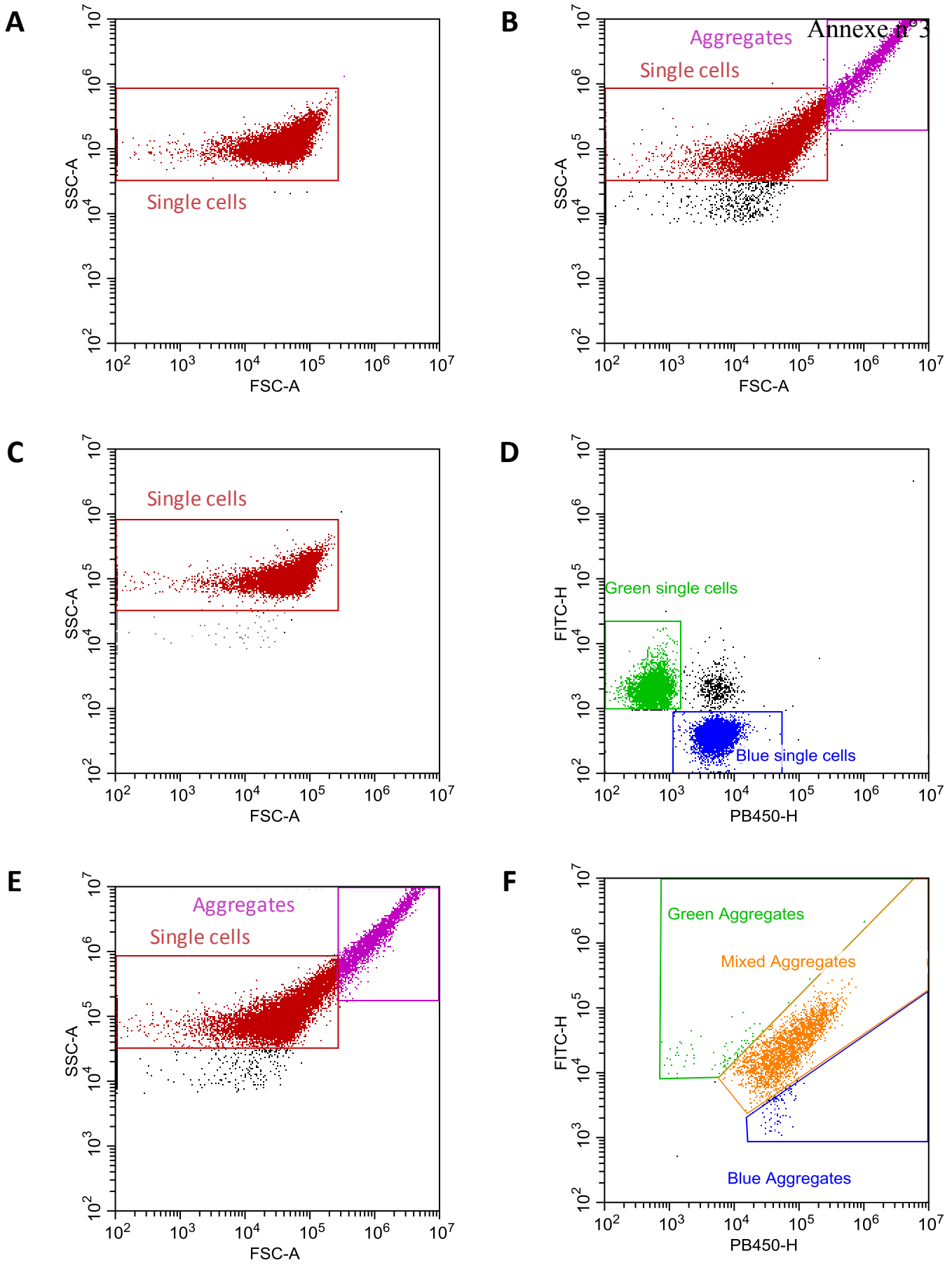
A
Single cells
SSC-A
FSC-A
B
Aggregates
Single cells
SSC-A
FSC-A
C
Single cells
SSC-A
FSC-A
D
Green single cells
Blue single cells
FITC-H
PB450-H
E
Aggregates
Single cells
SSC-A
FSC-A
F
Green Aggregates
Mixed Aggregates
Blue Aggregates
FITC-H
PB450-H

Figure S3

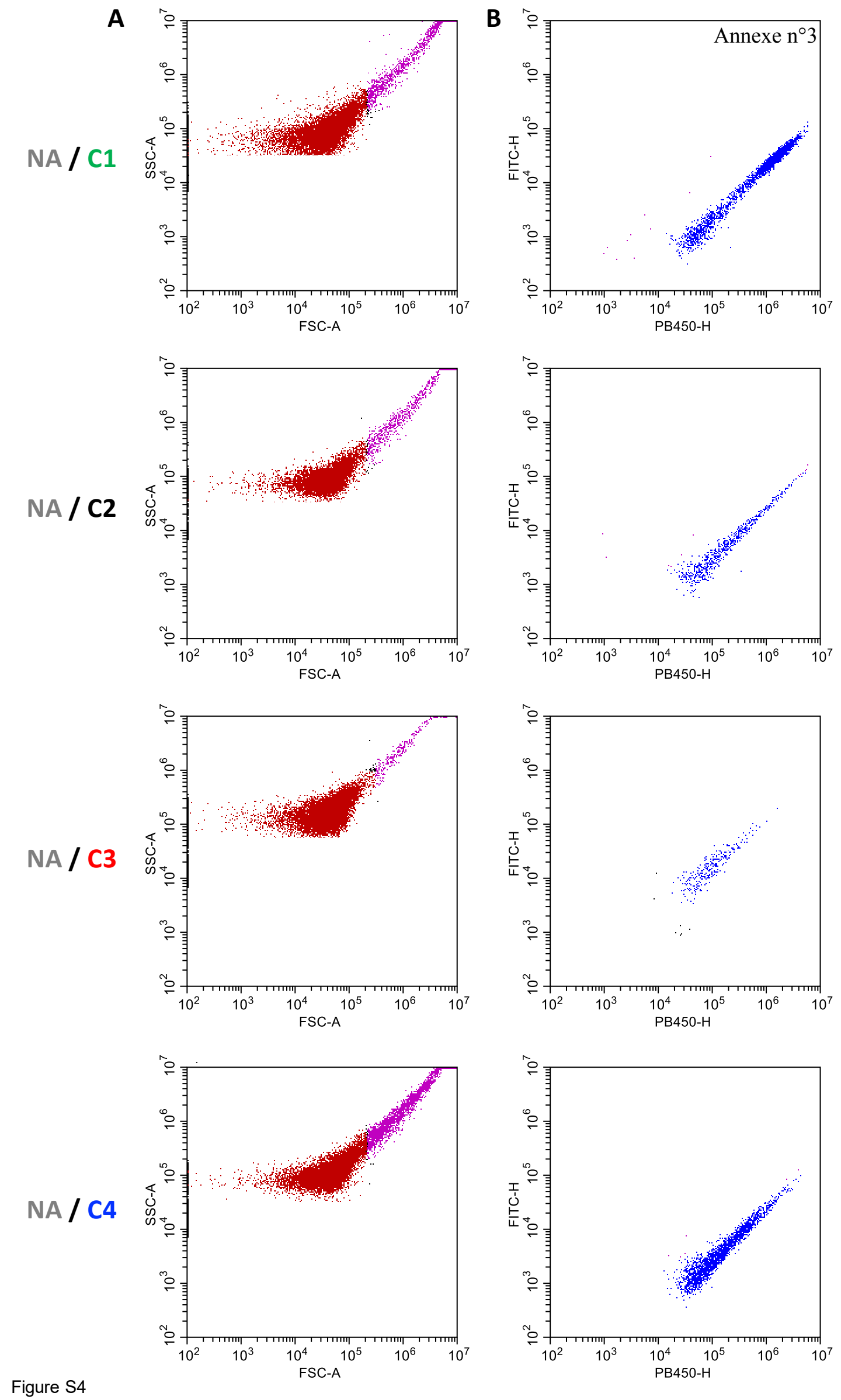

Figure S4

Annexe n°3

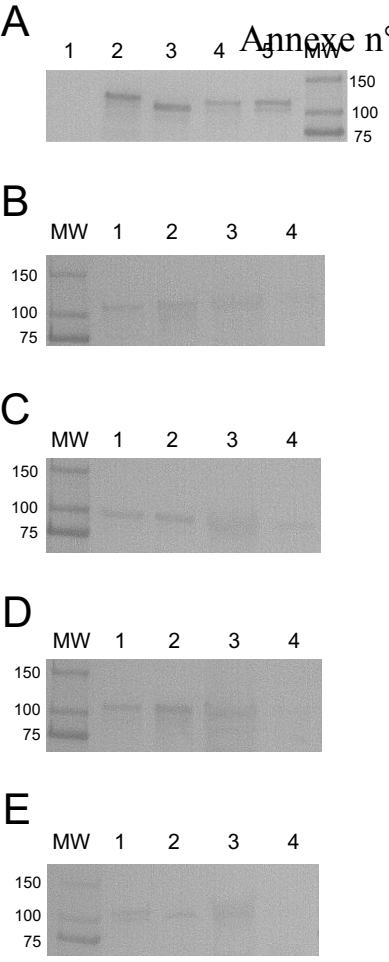

Figure S5

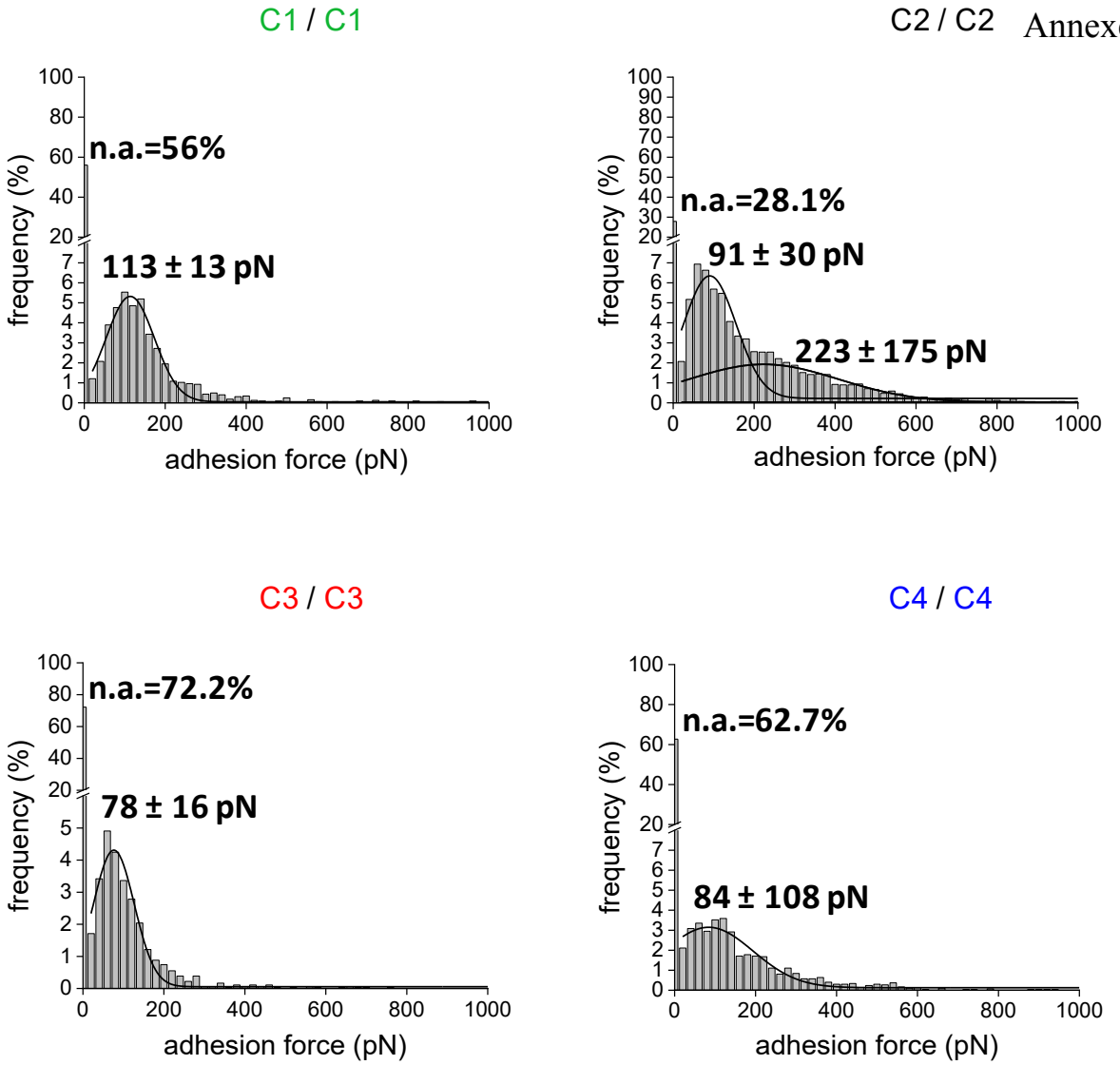
C1 / C1
n.a.=56%
113 ± 13 pN
frequency (%)
adhesion force (pN)
C2 / C2
n.a.=28.1%
91 ± 30 pN
223 ± 175 pN
frequency (%)
adhesion force (pN)
C3 / C3
n.a.=72.2%
78 ± 16 pN
frequency (%)
adhesion force (pN)
C4 / C4
n.a.=62.7%
84 ± 108 pN
frequency (%)
adhesion force (pN)

Figure S6

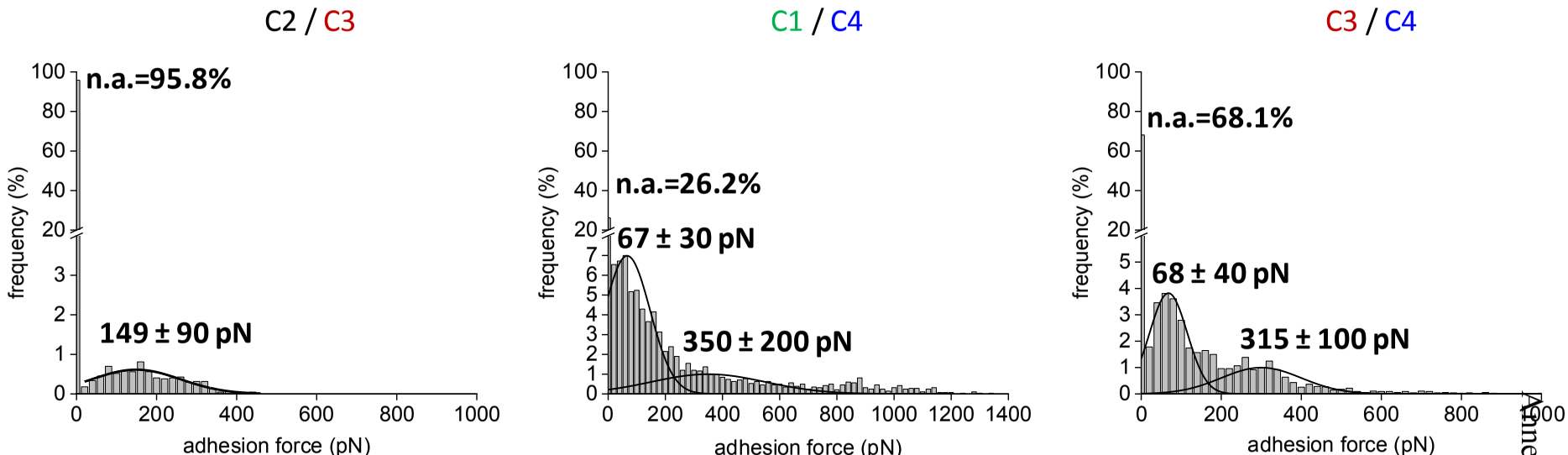

Figure S7

Annexe n°3

ANNEXE n°4 : Portfolio de compétences de l'EDSVSAE

Annexe n°4

UNIVERSITÉ Clermont Auvergne

PORTFOLIO DE COMPETENCES

Ecole Doctorale 65
Sciences de la Vie, Santé, Agronomie, Environnement

Nom, Prénoms	**Ageorges Valentin**
Sujet de recherche doctorale :	**Caractérisation du rôle de l'Antigène 43 dans le processus de colonisation d'*Escherichia coli* O157:H7**
Date de début de la thèse :	**1er février 2016**
Type de financement :	**Bourse région (Fond Régional d'Investissement Auvergne - cluster Institut de Recherche Pharmabiotique)**
Direction de thèse : • Directeur • Co-Directeur • Co-Encadrant	• **Directeur : Mickaël Desvaux** • **Co-encadrants : Christian Provot; Frédérique Chaucheyras-Durand**
Unité de recherche :	**UMR 454 MEDIS (Microbiologie Environnement DIgestif et Santé) – INRA Auvergne Rhône-Alpes, Site de Theix**

1. Formation doctorale

1.1. Modules disciplinaires

Lister les 4 modules effectués (intitulé, descriptif succinct, date de validation)	- **Formation au travail en laboratoire confiné de catégorie 3** : 10 jours au Teagasc d'Aschtown (Irlande) pour voir le fonctionnement et les méthodes de travail en laboratoire de catégorie 3 et découverte de diverses techniques de laboratoire auprès de doctorants (13-23 novembre 2016) – Dans le cadre du programme PHC Ulysse ARBiC: Molecular characterization of Antibiotic Resistance and Biofilm formation in commensal and pathogenic Escherichia coli (validation de 2 modules) **- Optimiser le montage de projets et la réponse aux appels d'offres :** Connaître les clés pour monter un projet scientifique solide et expliquer la démarche de réponse aux différents types d'appels d'offres (20 et 21 mars 2017) **- Notions fondamentales en statistiques** : Revoir les bases en analyses statistiques et les appliquer à des cas scientifiques concrets en vue de les utiliser en routine (04-07 septembre 2018).

1.2. Modules Socio-professionnels

Lister les 4 modules effectués	- **Module OSP 9-1 : Enseignement dans le supérieur** : Fournir aux

(intitulé, descriptif succinct, date de validation)	doctorants les connaissances et outils pour améliorer leurs enseignements à l'université (21 et 24 mars 2017) - **Module SP41 : Maitriser l'information stratégique, un enjeu majeur pour la recherche :** Sensibiliser les doctorants à l'intérêt d'une démarche d'intelligence économique dans le domaine de la recherche afin de renforcer la compétitivité d'une entreprise ou le rayonnement d'un établissement de recherche (11 et 12 mai 2017) - **Module SP46 : Optimisation de documents et communications scientifiques :** Apprendre à oragniser ses connaissances avec le mindmapping et connaître les fonctions avancées et automatismes intégrés des outils d'édition de texte (04 et 05 mai 2017) - **Module SP21 : Création d'entreprise :** Module visant à exposer et définir les principales étapes pour réussir son projet de création d'entreprise (22 et 23 juin 2017)

2. Enseignement

Présentation de l'activité d'enseignement (thématiques enseignées, pratiques pédagogiques, présentation synthétique des enseignements, niveau (L, M, D), nature (cours, TD, TP), effectifs, volume horaire).	- 2 missions de doctorant contractuel en 2017 et 2018 (2x64h) : TP de microbiologie fondamentale pour les étudiants de Génie biologique en 3ème année de Polytech Clermont-Ferrand (une vingtaine d'élèves par TP) - Contenu du TP : isolements et cultures de bactéries, travail en conditions stériles, organisation de son emploi du temps, initiation au travail en équipe, impact de divers paramètres sur la culture des bactéries, réalisation de yaourts, observations microscopiques et colorations de GRAM, identification de bactéries par galeries API ou lactiques, initiation aux cultures en fermenteur ...

3. Diffusion de la culture scientifique ou transfert de technologie

Articles de vulgarisation, Conférences dans des établissements de l'enseignement secondaire, Participation à des émissions de radio, Fête de la Science, Journées Portes Ouvertes de l'Université, Participation à des associations de doctorants, Participation à des comités d'organisation de congrès, de journées scientifiques de l'école doctorale,…	- Présentation de la méthode d'analyse de formation de biofilms bactériens avec l'appareil Biofilm Ring Test à des élèves du BTS ANABIOTECH du Puy-en-Velay (2016-2017 et 2018) - Participation aux journées portes ouvertes 2019 de l'INRA de Theix (Mai 2019)

4. Compétences de recherche

4.1. Conception d'un projet de recherche

Conceptualiser et bâtir un projet de recherche. *Exemples : Suivi d'un module de gestion d'un projet de recherche, Action de candidature sur une proposition de thèse, Proposition d'un programme de travail,… (les citer).*	- Candidature sur mon sujet de thèse (Projet CoMBa) et passage devant le comité d'évaluation de l'école doctorale EDSVSAE en juin 2015 - Suivi d'une formation : **Optimiser le montage de projets et la réponse aux appels d'offres** (20 et 21 mars 2017)
Participation à la recherche de moyens financiers. *Exemple : Rédaction et obtention (ou pas) d'un projet de recherche (comme obtention d'un financement pour une 4ème de thèse)… Indiquer le titre et le/les organisme(s) sollicités.*	- Demande de la Bourse Municipale « Doctorants » 2018-2019 de la ville de Clermont-Ferrand pour financer ma fin de thèse.
Veille scientifique et méthodologique. *Exemples : Réalisation d'une étude bibliographique, Organisation d'un « Journal club » de l'UMR, Participation à un atelier INSERM, CNRS, INRA, Validation d'un module biologique,… Donner le/les titre(s).*	- Réalisation d'une étude bibliographie pour la rédaction de mon manuscrit de thèse. - Rédaction d'une revue de la littérature : **Molecular Determinants of Biofilm Formation in Diarrheagenic *Escherichia coli* (DEC) and Their Involvement in Surface Colonisation from Bacterial Adhesion to Sessile Development** – Soumission prévue en 2019 - Participation au Journal Club de l'UMR 454 MEDiS dans laquelle un doctorant présente un article scientifique en rapport avec son sujet de recherche
Formation à l'éthique (Validation du module en première année).	Dispensé car thèse débutée en 2015-2016
Analyse critique de la production scientifique. *Exemples : Participation à des réunions hebdomadaires d'unité/d'équipe, avec ou sans présentation de ses propres résultats, Rédaction d'articles, de reviews,… (les citer).*	- Participation aux réunions d'équipe mensuelle dans lesquelles j'ai présenté mes travaux à 4 reprises - Présentation de mes travaux lors de réunions projet avec les partenaires industriels - Présentation de mes travaux lors de mes 2 comités de thèse - Rédaction d'un article scientifique soumis à Scientific Reports en 2019 : **Differential homotypic and heterotypic interactions of antigen 43 variants in autotransporter-mediated bacterial autoaggregation** - Rédaction d'un article de revue : **Molecular Determinants of Biofilm Formation in Diarrheagenic *Escherichia coli* (DEC) and Their Involvement in Surface Colonisation from Bacterial Adhesion to Sessile Development** – Soumission prévue en 2019 - Rédaction en cours de deux autres articles à soumettre en 2019 don't les titres exacts ne sont pas encore établis - Co-auteur d'un article de revue publié en 2016 dans FEMS Microbiology Letters : **A secretome view of colonisation factors in Shiga toxin-encoding *Escherichia coli* (STEC) : from enterohaemorrhagic *E. coli* (EHEC) to related enteropathotypes**

Capacité à accepter la critique, à faire preuve d'humilité, de doute scientifique et d'éthique. *Exemple : Présentation et discussion de résultats en réunion ou à des congrès/séminaires scientifiques français et internationaux,… (les citer et expliciter).*	- Participation aux réunions d'équipe mensuelle dans lesquelles j'ai présenté mes travaux à 4 reprises - Réalisation de 2 comités de thèse (09/01/2017 et 01/12/2017) - Communication orales : **Caractérisation de l'Antigène 43 d'*E. coli* dans le processus de colonisation de surface.** Journées de l'école doctorale SVSAE 2018 **Caractérisation de l'Antigène 43 d'*E. coli* dans le processus de colonisation de surface.** Journée du réseau des microbiologistes Clermontois (2018) **Characterisation of EHEC abilities to colonise the extracellular matrix.** Teagasc d'Ashtown (2016) - Posters : **Characterisation of the Antigen 43 of *Escherichia coli* in the surface colonisation process.** Symposium BioPhysAdh à Toulouse (Septembre 2018) **Characterisation of the molecular determinants involved in the colonisation ability of *Escherichia coli* O157:H7 to the extracellular matrix.** Congrès VTEC à Florence (Mai 2018) **Characterisation of the molecular determinants involved in the colonisation ability of *Escherichia coli* O157:H7 to the extracellular matrix.** Congrès de la Microbiology Society à Birminhgam (Avril 2018). **Characterisation of the molecular determinants involved in the colonisation ability of *Escherichia coli* O157:H7 to the extracellular matrix.** Congrès du Réseau National Biofilms à Clermont-Ferrand (Décembre 2017) **Characterisation of the molecular determinants involved in the colonisation ability of *Escherichia coli* O157:H7 to the extracellular matrix.** Congrès FEMS à Valencia (Juillet 2017)

4.2. Réalisation d'un projet de recherche

Elaborer et mettre en œuvre un protocole de recherche dans des conditions maximales de sécurité et de maîtrise des aléas. *Exemple : Participation à la formation hygiène et sécurité du laboratoire (indiquer la date).*	- Formation Hygiène, sécurité et prévention du laboratoire (février 2016) - Formation au travail en laboratoire de confinement de catégorie P2* (Mars 2016) - Formation au travail en laboratoire de confinement de catégorie 3 dans le cadre du programme PHC Ulysse (Novembre 2016 – Teagasc d'Aschtown)
Garantir des méthodes reproductibles et des résultats fiables. *Exemple : Validation d'un module de statistiques (donner l'intitulé, le volume horaire…),…*	- Suivi de la formation Formation : **Notions fondamentales en statistiques** (04-07 septembre 2018 4x7h)
Rigueur, intégrité scientifique, traçabilité et validité des résultats. *Exemple : Tenue quotidienne d'un*	- Tenue quotidienne d'un cahier de laboratoire dans lequel je note clairement le protocole employé, les conditions expérimentales, les résultats obtenus, les discussions éventuelles et les pistes d'amélioration

cahier de laboratoire avec sommaire avec résultats clairement détaillés et plan d'expérimentation reproductible par les suivants,…	
Présenter, discuter des résultats et les intégrer dans un protocole de recherche. *Exemple : Rédaction de rapports d'étape d'un projet de recherche,…*	- Présentation de mes travaux lors de réunions projet avec les partenaires industriels et discuter des futures perspectives - Présentation de mes travaux lors de mes 2 comités de thèse (09/01/2017 et 01/12/2017) - Réunions informelles avec mon directeur de thèse pour discuter de mes travaux et définir des points d'améliorations et les futures perspectives du projet - Rédaction du rapport final du programme PHC Ulysse
Proposer des évolutions de son projet et rester à l'écoute des avis. *Exemple : Comités de suivi de thèse (donner les dates),…*	- Présentation de mes travaux lors de réunions projet avec les partenaires industriels et discuter des futures perspectives - Présentation de mes travaux lors de mes 2 comités de thèse (09/01/2017 et 01/12/2017)
Participation à d'autres projets de recherche/aide ponctuelle. *Exemple : Association à un article en étant co-auteur « au milieu de la liste »,...*	- Co-auteur d'un article de revue publié en 2016 dans FEMS Microbiology Letters : **A secretome view of colonisation factors in Shiga toxin-encoding *Escherichia coli* (STEC) : from enterohaemorrhagic *E. coli* (EHEC) to related enteropathotypes** - Aide ponctuelle aux travaux d'un doctorant lors de mon séjour au Teagasc d'Aschtown (novembre 2016)
Gérer ou prendre en compte les contraintes et moyens matériels, financiers, humains et juridiques relatifs à un projet, interagir et convaincre d'autres acteurs engagés dans le projet. *Exemples : Module Environnement de la recherche, Europe, Interaction avec partenaires industriels (les citer),…*	- Interaction régulières avec les deux partenaires industriels de mon projet : Lallemand et Biofilm Control - Suivi du **Module SP21 : Création d'entreprise** (22 et 23 juin 2017) - Suivi du **Module SP41 : Maitriser l'information stratégique, un enjeu majeur pour la recherche** (11 et 12 mai 2017)

4.3. Valorisation et transfert des résultats de la recherche

Prendre du recul sur l'état d'avancement et la portée de son projet, identifier la valeur ajoutée par rapport aux travaux antérieurs. *Exemples : Rédaction d'articles scientifiques en particulier en tant que premier auteur (les citer), Rédaction et dépôt de*	- Rédaction d'un article scientifique soumis à Scientific Reports en 2019 : **Differential homotypic and heterotypic interactions of antigen 43 variants in autotransporter-mediated bacterial autoaggregation** **-** Rédaction en cours de deux autres articles à soumettre en 2019 don't les titres exacts ne sont pas encore établis

brevets (les citer),…	
Effectuer une communication scientifique orale adaptée à différents publics, en utilisant les moyens audiovisuels appropriés. *Exemples : Présentation à des congrès nationaux ou internationaux (les citer), Participation à des Ecoles d'Eté (les citer),…*	**- Caractérisation de l'Antigène 43 d'*E. coli* dans le processus de colonisation de surface.** Journées de l'école doctorale SVSAE 2018 **- Caractérisation de l'Antigène 43 d'*E. coli* dans le processus de colonisation de surface.** Journée du réseau des microbiologistes Clermontois (2018) **- Characterisation of EHEC abilities to colonise the extracellular matrix.** Teagasc d'Ashtown (2016)
Transférer, vulgariser et valoriser ses résultats de recherche dans le respect de la propriété intellectuelle et d'une éventuelle confidentialité et en s'appuyant sur les services appropriés. *Exemples : Conférences grand public (les citer), Echanges lycéens/doctorants, ASTEP, Présentation aux journées Ecole doctorale, Revues générales (les citer), Fête de la Science, Modules OSP (les citer),...*	- **Caractérisation de l'Antigène 43 d'*E. coli* dans le processus de colonisation de surface.** Journées de l'école doctorale SVSAE 2018 **- Caractérisation de l'Antigène 43 d'*E. coli* dans le processus de colonisation de surface.** Journée du réseau des microbiologistes Clermontois (2018) - Présentation de la méthode d'analyse de formation de biofilms bactériens avec l'appareil Biofilm Ring Test à des élèves du BTS ANABIOTECH du Puy-en-Velay (2016-2017 et 2018) - Participation aux journées portes ouvertes 2019 de l'INRA de Theix (Mai 2019)
Contribuer au dialogue entre Sciences et Société. *Exemples : Module Sciences et Société, La main à la pâte, Modules OSP (les citer),…*	- Participation aux journées portes ouvertes 2019 de l'INRA de Theix (Mai 2019)
Comprendre le processus d'exploitation commerciale de résultats de recherche pour être capable d'identifier les occasions de contribuer à la coopération entre recherche universitaire, recherche industrielle et l'ensemble des secteurs de production. *Exemple : Modules OSP (les citer),...*	- Interaction régulières avec les deux partenaires industriels de mon projet : Lallemand et Biofilm Control - Suivi du **Module SP21 : Création d'entreprise** (22 et 23 juin 2017) - Suivi du **Module SP41 : Maitriser l'information stratégique, un enjeu majeur pour la recherche** (11 et 12 mai 2017)

5. Compétences techniques et méthodologiques

Enrichir et compléter ses connaissances et compétences	- **Formation au travail en laboratoire confiné de catégorie 3** (13-23 novembre 2016)

méthodologiques et techniques. *Exemples : Atelier EPST, Module de formation biologie, Suivi de cours de Master comme module de biologie (les citer),...*	**- Optimiser le montage de projets et la réponse aux appels d'offres** (20 et 21 mars 2017) - **Module SP46 : Optimisation de documents et communications scientifiques** (04 et 05 mai 2017) **- Notions fondamentales en statistiques** (04-07 septembre 2018). - Initiation à l'utilisation du microscope à force atomique (AFM) au LISBP de Toulouse (septembre 2018)
Maîtriser tous les aspects de la gestion de projet. *Exemple : Modules OSP (les citer),...*	- Suivi du **Module SP41 : Maitriser l'information stratégique, un enjeu majeur pour la recherche** - Suivi d'une formation : **Optimiser le montage de projets et la réponse aux appels d'offres** - Savoir gérer les problèmes rencontrés lors des 3 ans de thèse et adapter et orienter les travaux en fonction
Décomposer des problèmes complexes en questions plus simples, ordonner les diverses étapes du projet et planifier leur déroulement pas à pas. *Exemple : Participation aux Doctoriales (donner les dates de participation), Développement de méthodologies expérimentales et théoriques,…*	- Présentation de mes travaux lors de réunions projet avec les partenaires industriels et discuter des futures perspectives - Présentation de mes travaux lors de mes 2 comités de thèse (09/01/2017 et 01/12/2017) - Réunions informelles avec mon directeur de thèse pour discuter de mes travaux et définir des points d'améliorations et les futures perspectives du projet - Mise en place de protocoles expérimentaux au laboratoire - Savoir gérer les problèmes rencontrés lors des 3 ans de thèse et adapter et orienter les travaux en fonction
Développer des connaissances pointues dans son champ disciplinaire. *Exemple : Participation à des congrès, workshops, séminaires (les citer),…*	- Veille scientifique et méthodologique - Participations à de nombreux congrès, symposiums et séminaires : • Congrès Gut Microbiology INRA-Rowett 2016 à Clermont-Ferrand (juin 2016) • Congrès FEMS 2017 à Valence (juillet 2017) • 3rd International Symposium on Fermented Meat à Clermont-Ferrand (septembre 2017) • Congrès du Réseau National Biofilm à Clermont-Ferrand (décembre 2017) • Congrès annuel de la Microbiology Society à Birmingham (Avril 2018) • Congrès VTEC à Florence (mai 2018) • International Symposium on Biohysics of Bacterial Adhesion à Toulouse (Septembre 2018)
Faire le point sur les moyens disponibles et suggérer leur évolution, mettre en œuvre les solutions proposées en respectant et en gérant le budget alloué. *Exemples : Réunions de projet, d'équipe, d'unité, Comités de suivi de thèse (donner les dates), modules OSP (les citer),...*	- Présentation de mes travaux lors des réunions mensuelles d'équipe - Présentation de mes travaux lors de réunions projet avec les partenaires industriels et discuter des futures perspectives - Présentation de mes travaux lors de mes 2 comités de thèse (09/01/2017 et 01/12/2017)
Faire un bilan du projet.	- Présentation de mes travaux lors de réunions projet avec les

Exemples : Rédaction de rapports d'étape du projet, Comités de suivi de thèse,...	partenaires industriels et discuter des futures perspectives - Présentation de mes travaux lors de mes 2 comités de thèse (09/01/2017 et 01/12/2017) - Rédaction du rapport final du programme PHC Ulysse
Maîtriser sa communication écrite et orale *Exemple : Formation théâtre, Modules OSP (les citer), Rédaction de communications par affiches, Préparation de présentations orales,…*	- Communication orales : **Caractérisation de l'Antigène 43 d'*E. coli* dans le processus de colonisation de surface.** Journées de l'école doctorale SVSAE 2018 **Caractérisation de l'Antigène 43 d'*E. coli* dans le processus de colonisation de surface.** Journée du réseau des microbiologistes Clermontois (2018) **Characterisation of EHEC abilities to colonise the extracellular matrix.** Teagasc d'Ashtown (2016) - Posters : **Characterisation of the Antigen 43 of *Escherichia coli* in the surface colonisation process.** Symposium BioPhysAdh à Toulouse (Septembre 2018) **Characterisation of the molecular determinants involved in the colonisation ability of *Escherichia coli* O157:H7 to the extracellular matrix.** Congrès VTEC à Florence (Mai 2018) **Characterisation of the molecular determinants involved in the colonisation ability of *Escherichia coli* O157:H7 to the extracellular matrix.** Congrès de la Microbiology Society à Birminhgam (Avril 2018). **Characterisation of the molecular determinants involved in the colonisation ability of *Escherichia coli* O157:H7 to the extracellular matrix.** Congrès du Réseau National Biofilms à Clermont-Ferrand (Décembre 2017) **Characterisation of the molecular determinants involved in the colonisation ability of *Escherichia coli* O157:H7 to the extracellular matrix.** Congrès FEMS à Valencia (Juillet 2017) - **Module SP46 : Optimisation de documents et communications scientifiques** (04 et 05 mai 2017) - **Module OSP 9-1 : Enseignement dans le supérieur** (21 et 24 mars 2017) - 2 missions de doctorant contractuel en 2017 et 2018 (2x64h) : TP de microbiologie fondamentale pour les étudiants de Génie biologique en 3ème année de Polytech Clermont-Ferrand (une vingtaine d'élèves par TP)
Reformuler et structurer les résultats pour les rendre convaincants et accessible à tous. *Exemple : Présentation à des séminaires, congrès (les citer),...*	- Communication orales : **Caractérisation de l'Antigène 43 d'*E. coli* dans le processus de colonisation de surface.** Journées de l'école doctorale SVSAE 2018 **Caractérisation de l'Antigène 43 d'*E. coli* dans le processus de colonisation de surface.** Journée du réseau des microbiologistes Clermontois (2018) **Characterisation of EHEC abilities to colonise the extracellular matrix.** Teagasc d'Ashtown (2016) - Posters : **Characterisation of the Antigen 43 of *Escherichia coli* in the surface colonisation process.** Symposium BioPhysAdh à Toulouse (Septembre 2018) **Characterisation of the molecular determinants involved in the**

	colonisation ability of *Escherichia coli* O157:H7 to the extracellular matrix. Congrès VTEC à Florence (Mai 2018) **Characterisation of the molecular determinants involved in the colonisation ability of *Escherichia coli* O157:H7 to the extracellular matrix.** Congrès de la Microbiology Society à Birminhgam (Avril 2018). **Characterisation of the molecular determinants involved in the colonisation ability of *Escherichia coli* O157:H7 to the extracellular matrix.** Congrès du Réseau National Biofilms à Clermont-Ferrand (Décembre 2017) **Characterisation of the molecular determinants involved in the colonisation ability of *Escherichia coli* O157:H7 to the extracellular matrix.** Congrès FEMS à Valencia (Juillet 2017)
Utiliser des outils adaptés pour différents médias. *Exemples : Formation powerpoint, Module Communication scientifique, communiquer avec les médias,…*	- **Module SP46 : Optimisation de documents et communications scientifiques** (04 et 05 mai 2017)
Mobiliser sa force de conviction dans un contexte de négociation. *Exemple : Réunion de travail entre différents partenaires dans le cadre de dépôt de projet,...*	/
Intervenir en tant que formateur, en s'adaptant au public. *Exemple : Mission d'enseignement (nombre d'heures TD, niveau, intitulé, public), Formation DU Enseignement,…*	- 2 Missions de doctorant contractuel en 2017 et 2018 (2x64h) : TP de microbiologie fondamentale pour les étudiants de Génie biologique en 3ème année de Polytech Clermont- Ferrand (une vingtaine d'élèves par TP) - **Module OSP 9-1 : Enseignement dans le supérieur** (21 et 24 mars 2017) - Encadrement de Jennifer GRAY, étudiante Irlandaise en doctorat de Microbiologie qui est venue se former à la technique d'analyse de formations de biofilm bactériens au crystal violet, dans le cadre du programme PHC Ulysse (3 semaines, novembre/décembre 2017
Maîtriser les outils informatiques de gestion des données (bases de données, formats, …). *Citer les formations.*	- **Module SP46 : Optimisation de documents et communications scientifiques** (04 et 05 mai 2017) **- Formation Linux et Phylogénie** à l'INRA de Jouy en Josas (mai 2015) - Utilisation d'endnote pour la gestion des références bibliographiques - Maitrise du pack office - Bases de données protéiques et génétiques : Genbank, Uniprot, Interpro…
Identifier des outils informatiques et statistiques pertinents à l'analyse de données (tableurs, tests, outils de visualisation, ..) *Citer les formations.*	- Utilisation de logiciels d'analyse statistiques : GraphPAD, tableurs excel … - Utilisation de nombreux software de bioinformatiques pour traiter des données de séquences génétiques ou protéiques, les modéliser et les visualiser : clustal OMEGA, Phyre, I-Tasser, Pymol, …
Structurer des informations pour	- <u>Communication orales :</u>

leur donner de la valeur. *Exemples : Présentation à des congrès, Rédaction d'articles scientifiques, Présentation à des réunions scientifiques, Rédaction de rapports d'étape,...*	**Caractérisation de l'Antigène 43 d'*E. coli* dans le processus de colonisation de surface.** Journées de l'école doctorale SVSAE 2018 **Caractérisation de l'Antigène 43 d'*E. coli* dans le processus de colonisation de surface.** Journée du réseau des microbiologistes Clermontois (2018) **Characterisation of EHEC abilities to colonise the extracellular matrix.** Teagasc d'Ashtown (2016) - Posters : **Characterisation of the Antigen 43 of *Escherichia coli* in the surface colonisation process.** Symposium BioPhysAdh à Toulouse (Septembre 2018) **Characterisation of the molecular determinants involved in the colonisation ability of *Escherichia coli* O157:H7 to the extracellular matrix.** Congrès VTEC à Florence (Mai 2018) **Characterisation of the molecular determinants involved in the colonisation ability of *Escherichia coli* O157:H7 to the extracellular matrix.** Congrès de la Microbiology Society à Birminhgam (Avril 2018). **Characterisation of the molecular determinants involved in the colonisation ability of *Escherichia coli* O157:H7 to the extracellular matrix.** Congrès du Réseau National Biofilms à Clermont-Ferrand (Décembre 2017) **Characterisation of the molecular determinants involved in the colonisation ability of *Escherichia coli* O157:H7 to the extracellular matrix.** Congrès FEMS à Valencia (Juillet 2017) - Articles scientifiques : **Differential homotypic and heterotypic interactions of antigen 43 variants in autotransporter-mediated bacterial autoaggregation -** soumis à Scientific Reports en 2019 Rédaction en cours de deux autres articles à soumettre en 2019 don't les titres exacts ne sont pas encore établis - Rapports Rédaction du rapport final du programme PHC Ulysse
Tirer tous les bénéfices d'une compétence interdisciplinaire et confronter ses connaissances avec d'autres disciplines. *Exemples : Présenter ses travaux à des doctorants d'autres ED de ses travaux, Participer aux journées scientifiques de l'école doctorale,…*	- Participation aux JED - Présentation de mes travaux devant des non-spécialistes au Teagasc d'Aschtown - Journal Club de l'unité

6. Autres compétences - savoir être

6.1. Mobiliser un esprit d'innovation au quotidien

Faire preuve de curiosité *Exemples : Participer à des*	- Assisté à la conférence de Rolland Lehoucq dans le cadre du colloque « Science et Science-fiction » le 18 janvier à Clermont-Ferrand

séminaires, congrès d'autres disciplines, Être force de proposition dans des projets extérieurs à son projet doctoral, Découvrir une autre langue, Participer aux JED…	- Perfectionnement de l'anglais lors de mes séjours à l'étranger, congrès, contact avec étudiants étrangers… - Participation aux JED 2014 2015 2016 2017 et 2018
Gérer les risques liés à la confrontation avec l'inconnu. *Exemple : Se former à de nouvelles techniques (les citer),…*	- Séjour de formation à Aschtown (Irlande) dans le cadre du PHC Ulysse (Novembre 2016) - Initiation à l'utilisation du microscope à force atomique (AFM) au LISBP de Toulouse (septembre 2018) - Apprendre l'emploi de nombreux appareils : microscope à épifluorescence, cytomètre de flux, …
Avoir le réflexe de valoriser de manière pertinente ses découvertes et réalisations avec prise de risques. *Exemple : Participer aux modules Création d'entreprise, brevets, Créer après la soutenance sa propre start-up,…*	- **Module SP21 : Création d'entreprise** (22 et 23 juin 2017)
Dépasser les frontières de sa discipline. *Exemple : Participation à l'élaboration de projets (ANR, PEPS, FUI,…)…*	- Formation **Optimiser le montage de projets et la réponse aux appels d'offres** (20 et 21 mars 2017) - Assisté à la conférence de Rolland Lehoucq dans le cadre du colloque « Science et Science-fiction » le 18 janvier à Clermont-Ferrand

6.2. Travailler en équipe

Collaborer avec des collègues et partenaires aux statuts et aux missions diverses en cherchant à valoriser au plus juste la contribution de chacun. *Exemple à détailler : Travail régulier avec des techniciens, ingénieurs,…*	- Travail quotidien avec une technicienne, une assistante-ingénieur et d'autres doctorants au laboratoire - Interaction avec d'autres membres du laboratoire pour échanges sur les travaux ou sur de nouvelles méthodes - Interaction avec des chercheurs étrangers pour collaboration en vue d'articles scientifiques
Faire preuve de respect, d'écoute, de confiance, d'ouverture à la diversité, d'encouragement et d'humilité face aux propositions de ses collègues et partenaires. *Exemple : Participation à des conseils (unité de recherche, équipe,…),…*	- Réunions projet avec les partenaires industriels - Journal Club de l'unité

Se mettre au service d'un collectif pour l'aboutissement d'un projet. *Exemple : Faire partie d'un conseil scientifique (université, école doctorale, conseil de laboratoire),…* *Indiquer les propositions faites dans le cadre de cette fonction.*	- Participation aux décisions quant aux voies à suivre lors des réunions projet avec les partenaires industriels
Développer et entretenir un réseau scientifique et professionnel. *Exemple : Être Membre actif d'une association de doctorants, de réseaux sociaux professionnels (Linkedin,..),…*	- Membre actif sur LinkedIn - En contact avec d'anciens doctorants du laboratoire - Interaction avec des collaborateurs étrangers dans le cadre de la rédaction d'articles - Rencontres et contact avec d'autres chercheurs lors de congrès ou séminaires …

6.3. Coordonner

Encadrer des stagiaires. *Exemples : BTS, IUT, L3, M1, M2, autres...* *Citer les noms des stagiaires, l'intitulé de la formation, du sujet et de la durée des stages.*	- Encadrement de Jennifer GRAY, étudiante Irlandaise en doctorat de Microbiologie qui est venue se former à la technique d'analyse de formations de biofilm bactériens au crystal violet, dans le cadre du programme PHC Ulysse (3 semaines, novembre/décembre 2017)
Piloter un projet. *Exemples : Projet doctoral, Rédaction d'une publication, Animation des JED,…*	- Gestion du projet doctoral - Rédaction d'articles scientifiques, d'une revue, …

6.4. Travailler en autonomie

Mobiliser sa polyvalence. *Exemples : Ouverture d'esprit (citer des exemples), Obtention de diplômes autres (les citer),…*	- Ouverture d'esprit : Assisté à la conférence de Rolland Lehoucq dans le cadre du colloque « Science et Science-fiction » le 18 janvier à Clermont-Ferrand Séjour au Teagasc d'Aschtown (novembre 2016) - 2 missions de doctorant contractuel pour initiation à l'enseignement
Rendre compte régulièrement de ses avancées et progrès. *Exemples : Thèse CIFRE avec des rencontres régulières entreprise-universitaires, Réunions d'équipe, d'unité,…*	- Présentation de mes travaux lors des réunions mensuelles d'équipe - Présentation de mes travaux lors de réunions projet avec les partenaires industriels - Présentation de mes travaux lors de mes 2 comités de thèse (09/01/2017 et 01/12/2017) - Réunions informelles avec mon directeur de thèse pour discuter de mes travaux et définir des points d'améliorations et les futures perspectives du projet

Etablir un bilan complet de ses réalisations. Exemples : *Rapport final du projet de recherche, Rédaction du mémoire de doctorat,…* .	- Réunion de cloture du projet avec les partenaires industriels - Rédaction du manuscrit de thèse - Rédaction du rapport final du programme PHC Ulysse

6.5. S'adapter

Au contexte professionnel. *Exemples : Respect des horaires de travail selon la législation, Gestion d'une colonie d'animaux, de plantes transgéniques,…*	- Respect des horaires de travail : obligation de pointer à la badgeuse matin et soir - Anticiper ses demandes de congés - Participation aux diverses réunions du laboratoire
Aux personnes, à leurs méthodes de travail et à leurs besoins ou demandes. *Exemple : Respect de la vie du laboratoire (règlement intérieur, contraintes administratives, missions),...*	- Respect du règlement intérieur, des contraintes administratives, demandes de missions diverses, …
Aux rythmes imposés par l'environnement professionnel. *Exemple : Nécessité de fournir rapidement des données mises en forme pour une présentation orale, une visite HCERES, la rédaction d'une publication, Nécessité de venir un week-end avec un statutaire pour démarrer une expérience, rester plus tard le soir avec un statutaire,…*	- Devoir fournir des résultats, les mettre en forme et les présenter lors de réunions d'équipe ou de communications orale lors d'évènements - Devoir finir les expériences en temps et en heure avant la fin de la thèse en vue de la rédaction du manuscrit - Nombreuses heures supplémentaires - Venu plusieurs week end pour lancer des expériences et gagner du temps pour la semaine à venir

6.6. Utiliser une langue étrangère

A l'écrit. *Exemple : Rédaction d'articles, revues, rapports,..*	Rédactions d'articles scientifiques, d'une revue de bibliographie, du rapport final du projet PHC Ulysse, d'abstract en vue de présentation en congrès …
A l'oral. *Exemple : Présentation à des congrès internationaux, Rencontres avec d'autres chercheurs,...*	- Interaction avec de nombreux étudiants étrangers dans mon laboratoire - Interaction avec de nombreux chercheurs étrangers lors de congrès - Collaboration avec des étrangers en vue de publication d'article scientifique - Journal Club de l'unité - Séjour au Teagasc d'Aschtown (novembre 2016)

UNIVERSITÉ Clermont Auvergne

	- Journal Club de l'unité - Séjour au Teagasc d'Aschtown (novembre 2016)
Cours de langues. *Citer les langues, le type de cours, le nombre d'heures, le niveau acquis, TOEIC, autres,...*	/

Date et signature de la doctorante/du doctorant : 05/04/19	Date et signature de la directrice/du directeur de thèse : 5/4/2019	Date et signature de la directrice/du directeur de l'Ecole doctorale :

RÉSUMÉ

Les *Escherichia coli* entérohémorragiques (EHEC) O157:H7 sont des agents pathogènes alimentaires entrainant des colites hémorragiques et des syndromes hémolytiques et urémiques (SHU). Depuis le ruminant, le réservoir naturel, les EHEC peuvent contaminer certaines denrées alimentaires et *in fine* infecter l'Homme. La présence de composants de la matrice extracellulaire (ECM) le long de la chaine alimentaire pourrait participer à leur capacité de colonisation grâce à des protéines de surface bactérienne. Parmi ces protéines, l'antigène 43 (Ag43) est sécrété par le système de sécrétion de Type V, sous-type a (T5aSS) et appartient à la famille des SAAT (self-associating autotransporters). Ces protéines modulaires, composées d'un peptide signal N-terminal clivable, d'un domaine passager central exposé à la surface et d'un translocateur C-terminal membranaire, peuvent s'auto-associer selon un mécanisme de Velcro moléculaire. A l'origine, l'Ag43 était distribué en deux sous-familles SF-I et SF-II mais des analyses phylogénétiques ont révélé pour la première fois qu'il est en réalité distribué en 6 classes différentes nommées C1 à C6, basées sur le réarrangement de différents sous-domaines à l'intérieur du domaine passager de la protéine. Quant à la prévalence et la distribution chez les bactéries, l'*agn43* s'est avéré être présent essentiellement chez *E. coli* avec jusqu'à 5 copies du gène par génome, en différentes combinaisons de classes d'Ag43. Des analyses fonctionnelles des Ag43 ^{C1} à Ag43 ^{C4} ont démontré que des interactions homotypiques survenaient pour toutes les classes mais l'Ag43 ^{C3} présentait des différences significatives de cinétique de sédimentation et d'autoagrégation. L'étude des interactions hétérotypiques entre Ag43 de différentes classes a démontré qu'elles ne se produisaient que dans de rares cas. Chez *E. coli* O157:H7, l'Ag43 ^{C2} a été démontré comme contribuant à l'adhésion des bactéries aux collagènes I et III et à l'autoagrégation. De plus dans cette souche, ce phénotype s'est avéré être régulé par la méthylase Dam et le régulateur de transcription OxyR. Globalement, ces travaux apportent de nouvelles connaissances concernant la diversité de l'Ag43 et son rôle et sa régulation chez *E. coli* O157:H7. Ainsi, l'Ag43 pourrait être un important facteur de colonisation de la chaine alimentaire pour les EHEC et les autres *E. coli* diarrhéiques.

Mots clés : *Escherichia coli* entérohémorragiques; colonisation de la chaine alimentaire; Antigène 43; diversité; autoagrégation; régulation.

ABSTRACT

Enterohaemorrhagic *Escherichia coli* (EHEC) O157:H7 are anthropozoonotic agents leading to haemorrhagic colitis and haemolytic-uremic syndrome (HUS). From the ruminants, their natural reservoir, EHEC can contaminate some foodstuffs and consequently infect humans. The presence of extracellular matrix (ECM) components along the food chain could contribute to the colonisation process. Numerous proteins can be present at the bacterial cell surface in EHEC, among them, antigen 43 (Ag43) is secreted by the Type V, subtype a, secretion system (T5aSS) and belongs to the family of self-associating autotransporters (SAATs). These modular proteins, comprising a cleavable N-terminal signal peptide, a surface-exposed central passenger and an outer membrane C-terminal translocator, self-recognise in a Velcro-like handshake mechanism. Ag43 was originally considered as distributed into two subfamilies, namely SF-I and SF-II but phylogenetic network analyses revealed for the first time that it actually distribute into six distinct classes, namely C1 to C6 based on the shuffling of different subdomains within the Ag43 passengers. Regarding prevalence and distribution in Bacteria, *agn43* appeared to be essentially present in *E. coli* with up to 5 copies of the gene present in a single genome and in different combinations of Ag43 classes. Functional analyses of Ag43 ^{C1} to Ag43 ^{C4} revealed that homotypic interactions occurred for all Ag43 classes but Ag43 ^{C3} induced significant differences in the sedimentation kinetics and autoaggregation state. In contrast, heterotypic interactions occurred in a very limited number of cases. In *E. coli* O157:H7, Ag43 ^{C2} was shown to contribute to bacterial adhesion to collagens I and III and to the autoaggregation. Furthermore in this strain, the latter phenotype appeared to be regulated by the Dam methylase and the OxyR transcriptional regulator. Taken together, these results provide new insights of the diversity of Ag43 as well as its role and regulation in *E. coli* O157:H7. Thus, Ag43 may represent an important colonisation factor of the food chain by EHEC and other diarrheagenic *E. coli*.

Keywords: Enterohaemorrhagic *Escherichia coli;* food chain colonisation; Antigen 43; diversity; autoaggregation; regulation.