

Université de Montréal

**Les autotransporteurs auto-associatifs d'*Escherichia coli* :
de facteurs de virulence à déterminants sociaux**

Par

Jean-Philippe Côté

Département de pathologie et microbiologie
Faculté de médecine vétérinaire

Thèse présentée à la Faculté de médecine vétérinaire
en vue de l'obtention du grade de
philosophiae doctor (Ph.D.)
en sciences vétérinaires
option microbiologie

Juillet 2013

© Jean-Philippe Côté, 2013

Université de Montréal
Faculté de médecine vétérinaire

Cette thèse intitulée:
Les auto-transporteurs auto-associatifs d'*Escherichia coli*:
de facteurs de virulence à déterminants sociaux

Présentée par:
Jean-Philippe Côté

A été évaluée par un jury composé de:

France Daigle, présidente rapporteuse
Michael Mourez, directeur de recherche
Mario Jacques, membre du jury
Sébastien Faucher, examinateur externe
Daniel Dubreuil, représentant du doyen

Résumé

Les autotransporteurs monomériques représentent le système de sécrétion le plus simple et le plus utilisé chez les bactéries à Gram négatif. Les autotransporteurs monomériques sont des protéines modulaires qui contiennent toute l'information pour leur sécrétion dans leur séquence. Les phénotypes associés à l'expression d'un autotransporteur peuvent être très variés et, souvent, les autotransporteurs sont des protéines multifonctionnelles. C'est le cas notamment des autotransporteurs AIDA-I, TibA et Ag43 d'*Escherichia coli* qui promouvoient l'adhésion et l'invasion de cellules épithéliales, l'auto-agrégation des bactéries et la formation de biofilm. Ces trois autotransporteurs ont d'ailleurs été regroupés dans une même famille, appelée les autotransporteurs auto-associatifs (SAATs). À cause de leur fonctionnalité, les SAATs sont considérés comme étant d'importants facteurs de virulence d'*Escherichia coli*. Toutefois, il existe plusieurs différences entre les SAATs qui ne sont pas bien comprises, si bien que leur rôle pour les bactéries n'est toujours pas bien compris.

Nous avons donc d'abord caractérisé TibA, le membre des SAATs le moins bien étudié à l'aide d'une étude structure-fonction. Nous avons observé que TibA était une protéine modulaire et que son domaine fonctionnel était composé de deux modules : un module d'auto-agrégation en N-terminal et un module d'adhésion en C-terminal. En comparant nos résultats avec ceux obtenus pour les autres SAATs, nous avons réalisé que l'organisation des trois SAATs était très variée, c'est-à-dire que les trois SAATs sont composés de modules différents. Nous avons par ailleurs observé cet arrangement en modules lorsque nous avons analysé plusieurs séquences d'*aidA*, suggérant qu'un mécanisme d'échange et d'acquisition de modules était à la base de l'évolution des SAATs. Sans surprise, nous avons aussi observé que la famille des SAATs ne se limitait pas à AIDA-I, TibA et Ag43 et ne se limitait pas à *Escherichia coli*.

La comparaison a aussi révélé l'importance du phénotype d'auto-agrégation dans la fonctionnalité des SAATs. Nous avons donc entrepris une étude du mécanisme d'auto-agrégation. Nos résultats ont montré que l'auto-agrégation était le résultat d'une interaction

directe SAAT/SAAT et ont mis en évidence un mécanisme similaire à celui utilisé par les cadhérines eucaryotes. De plus, nous avons observé que, comme les cadhérines, les SAATs étaient impliqués dans des interactions homophiliques; un SAAT interagit donc spécifiquement avec lui-même et non avec un différent SAAT.

Finalement, les SAATs font parties des quelques protéines qui sont glycosylées chez *Escherichia coli*. Nous avons déterminé que le rôle de la glycosylation de TibA était de stabiliser la protéine et de lui donner la flexibilité nécessaire pour moduler sa conformation et, ainsi, être pleinement fonctionnelle.

Globalement, nos résultats suggèrent que les SAATs sont des molécules « cadhérines-like » qui permettent la reconnaissance de soi chez les bactéries. Une telle habilité à discriminer entre le soi et le non-soi pourrait donc être utilisée par les bactéries pour organiser les communautés bactériennes.

Mots clés

Escherichia coli; Sécrétion; Glycosylation; Autotransporteurs; SAAT; Adhésion; Auto-agrégation; Biofilm; Cadhérine;

Abstract

Autotransporters are versatile virulence factors of Gram-negative bacteria and use one of the simplest and most widespread secretion system in bacteria. The name autotransporter originate from the observation that all the information needed for the secretion of the protein is encoded in its own sequence, meaning that autotransporters do not need a specialized secretion apparatus. Many autotransporters are multifunctional proteins and can perform a large variety of functions. The self-associating autotransporters (SAATs), represented by AIDA-I, TibA and Ag43, are such multifunctional proteins and can mediate the adhesion and invasion of epithelial cells, the auto-aggregation of bacteria and the formation of biofilm. Because of these functionalities, SAATs are considered important virulence factors of *Escherichia coli*. However, there are many differences between the SAATs and we still do not know their exact role for the bacteria.

Therefore, we have realized a structure-function study of TibA, the least studied SAAT. Our study showed that TibA is a modular protein and that the functional domain of TibA is composed of two modules: an N-terminal module responsible for auto-aggregation and a C-terminal module responsible for adhesion. Our results showed that the organization of AIDA-I, TibA and Ag43 is different and that the SAATs represent different assemblies of modules. We also observed the modular organization when we analyzed various sequence of *aida*, suggesting that the SAATs have evolved by a mechanism of domain shuffling. Not surprisingly, we have found new SAATs in *Escherichia coli* and in other proteobacteria.

Our results also highlighted the importance of auto-aggregation in the functionality of the SAATs. We therefore assessed the mechanism of SAAT-mediated auto-aggregation of bacteria. Our results showed that SAATs mediate auto-aggregation of bacteria through direct SAAT/SAAT interactions and that these interactions were reminiscent of the interactions made by cadherin molecules in eukaryotes. We further observed that the SAATs were involved in homophilic interactions, as it is the case with cadherin molecules.

SAATs are part of the few proteins that are glycosylated in *Escherichia coli*. We therefore characterized the glycosylation of TibA and found that glycosylation of TibA stabilized the protein and allowed the protein to modulate its conformation, resulting in a fully functional protein.

Taken together, our results suggest that the SAATs may be cadherin-like molecules by bacteria in order to discriminate between self and non-self. Such an ability to discriminate self from non-self is rarely evoked in bacteria, but could play a role in the organization of multicellular communities.

Keywords

Escherichia coli; secretion; glycosylation; autotransporters; SAAT; Adhesion; Auto-aggregation; Biofilm; Cadherin;

Table des matières

Table des matières

Résumé	iii
Mots clés.....	iv
Abstract.....	v
Keywords.....	vi
Table des matières	vii
Liste des tableaux	x
Liste des figures.....	xi
Liste des sigles et des abréviations.....	xv
Remerciements.....	xvii
Recension des écrits.....	1
Introduction.....	2
1. Systèmes de sécrétion bactériens.....	4
1.1. Traversée de la membrane cytoplasmique.....	5
1.1.1. Système de sécrétion Sec	6
1.1.2. Système de sécrétion <i>TAT</i>	10
1.2. Traversée de la membrane externe	13
1.2.1. Système d'insertion dans la membrane externe (<i>BAM</i>)	13
1.2.2. Système de sécrétion de type II.....	16
1.2.3. Système de sécrétion de type V.....	17
1.3. Traversée des deux membranes	20
1.3.1. Système de sécrétion de type I	21
1.3.2. Système de sécrétion de type III.....	22
1.3.3. Système de sécrétion de type IV	23
1.3.4. Système de sécrétion de type VI.....	25
2. Autotransporteurs monomériques.....	26
2.1. Organisation	27
2.1.1. Séquence signale.....	28
2.1.2. Domaine passager	29
2.1.3. Région de jonction.....	30
2.1.4. Domaine β	31
2.2. Biogenèse	32
2.2.1. Sécrétion.....	32
2.2.2. Clivage	39
2.2.3. Glycosylation	40
2.2.4. Autres particularités	44
2.3. Fonctions	45
2.3.1. Adhésion.....	47
2.3.2. Auto-agrégation	50

2.3.3. Formation de biofilm	51
3. Autotransporteurs auto-associatifs d'<i>Escherichia coli</i>	52
3.1 Les pathotypes d'<i>E. coli</i>	53
3.2. AIDA-I.....	55
3.2.1. Implication dans la pathogenèse	55
3.1.2. Organisation et biogenèse d'AIDA-I	57
3.1.3. La famille AIDA.....	62
3.2. Ag43.....	63
3.3. TibA	65
3.4. Similarités et différences	66
4. Problématique et objectif.....	69
Méthodologie et Résultats	71
Article 1 – Structure-fonction de TibA.....	72
Avant-propos.....	72
Contributions	72
Abstract.....	74
Introduction	74
Material and Methods.....	76
Results	78
Discussion	82
Acknowledgements.....	83
References.....	84
Figure Legends.....	88
Table.....	91
Figures	92
Article 2 – Glycosylation de TibA.....	98
Avant-propos.....	98
Contributions	98
Abstract.....	100
Introduction	100
Materials and Methods.....	102
Results	106
Discussion	111
Acknowledgements.....	113
References.....	113
Figure legends.....	116
Table.....	119
Figures	120
Article 3 – Allèles d'AIDA-I.....	127
Avant-propos.....	127
Contributions	127
Abstract.....	129
Introduction	129
Materials and Methods.....	131
Results and Discussion	134

Acknowledgements.....	140
References.....	140
Figure Legends.....	144
Tables.....	147
Figures.....	149
Supplementary information	158
Article 4 – Oligomérisation d’AIDA-I	159
Avant-propos.....	159
Contributions	159
Abstract.....	161
Introduction	161
Experimental procedures	163
Results	168
Discussion	173
Acknowledgements.....	175
References.....	176
Figure legends.....	180
Table.....	184
Figures.....	185
Supplementary information	193
Article 5 – SAATs : molécules « cadhérine-like ».....	197
Avant-propos.....	197
Contributions	197
Abstract.....	199
Introduction	199
Results and discussion.....	201
Material and Methods.....	207
Acknowledgements.....	210
References.....	211
Figure legends.....	214
Table.....	216
Figures.....	218
Supplementary informations	223
Discussion.....	229
Conclusion	253
Références	255
Annexe	xviii
Résultats supplémentaires - Structure et oligomérisation d’AIDA-I.....	xix
1) Structure de AIDA-I.....	xix
2) Oligomérisation de AIDA-I.....	xxii
3) Interactions en <i>cis</i> vs en <i>trans</i>	xxviii
Référence.....	xxxii

Liste des tableaux

Recension des écrits

Tableau I: Exemple de fonctionnalité des autotransporteurs	46
Tableau II: Prévalence du gène codant pour AIDA-I chez des souches porcines isolées de cas de diarrhée	56

Article 1

Table I: Position and sequences of the insertion mutants	91
--	----

Article 2

Table I: Glycosylated and unglycosylated peptides in TibA	119
---	-----

Article 3

Table I: Primers used in this study	147
Table II: Strains used in this study	148

Article 4

Table I: Analytical ultracentrifugation of purified AIDA-I	184
--	-----

Article 5

Table I: Putative SAATs bearing the 19 amino-acid consensus sequence repeats	216
--	-----

Liste des figures

Recension des écrits

Figure 1 : Les différents systèmes de sécrétion chez les bactéries à Gram négatif	5
Figure 2 : Sécrétion à travers la membrane cytoplasmique par le système Sec	10
Figure 3 : Sécrétion à travers la membrane cytoplasmique par le système TAT	12
Figure 4 : Les différents modèles d'insertion de protéines de membrane externe par le complexe BAM	15
Figure 5 : Les variations du système de sécrétion de type V	18
Figure 6 : Organisation et structure des autotransporteurs monomériques	27
Figure 7 : Biogenèse des autotransporteurs monomériques	33
Figure 8 : Modèles pour la translocation du domaine passager dans le milieu extracellulaire	36
Figure 9 : Organisation et modélisation d'AIDA-I	58
Figure 10 : Biogenèse d'AIDA-I	60
Figure 11 : Similarités et différences entre les SAATs.....	67

Article 1

Figure 1 : Position of the insertion mutants in the passenger domain and the junction region of TibA	92
Figure 2 : Localization of the insertion mutants in the coil model of TibA	93
Figure 3 : Functionality of the insertion mutants	94
Figure 4 : Effect of the deletion of the N-terminal portion of the passenger domain of TibA	95
Figure 5 : Construction and expression of TibA lacking the proline-rich region	96
Figure 6 : Functional organization of SAATs	97

Article 2

Figure 1 : Effect of glycosylation on the expression level of TibA	120
Figure 2 : Localization of the peptides in TibA	121

Figure 3 : Effect of glycosylation on the functionality of TibA	122
Figure 4 : Effect of glycosylation on the generation of cytoplasmic and extracytoplasmic stress	123
Figure 5 : Effect of glycosylation on protease susceptibility	124
Figure 6 : Effect of glycosylation on purified TibA	125
Figure 7 : Effect of glycosylation on the modulation of conformation by salt	126

Article 3

Figure 1 : Expression of AIDA-I	149
Figure 2 : Polymorphism in the <i>aah/aidA</i> operon	150
Figure 3 : Phylogenetic analysis of <i>aah</i> and <i>aidA</i>	151
Figure 4 : Promoter region upstream of <i>aah</i>	152
Figure 5 : Regions of variability and conservation in <i>aah/aidA</i>	153
Figure 6 : Localization of <i>aidA</i> on large plasmids	154
Figure 7 : Colocalization of <i>aidA</i> and other virulence factor	155
Figure 8 : Recombination of the sequence from strain 14466	156
Figure 9 : Functionality of the three alleles	157
Figure S1 : Polymorphism of the <i>agn43</i> gene	158

Article 4

Figure 1 : Auto-aggregation of bacterial cells expressing AIDA-I	185
Figure 2 : Inhibition of auto-aggregation by purified AIDA-I and not by cellular extracts	186
Figure 3 : Auto-aggregation of fluorescent beads mediated by purified AIDA-I	187
Figure 4 : In vitro evidence of AIDA-AIDA interactions	188
Figure 5 : Chemical cross-linking of AIDA-I	189
Figure 6 : Modulation of AIDA-AIDA interaction and bacterial auto-aggregation	190
Figure 7 : Conformation change in AIDA-I	191
Figure 8 : Effect of bile salt on AIDA-AIDA interaction and auto-aggregation	192
Figure S1: Auto-aggregation of bacteria expressing AIDA-I	193

Figure S2: AIDA-AIDA interactions monitored by Surface Plasmon Resonance (SPR) on a very low-density AIDA-I surface	194
Figure S3: Sedimentation profile of purified wild-type AIDA-I in TBS observed by AUC	195
Figure S4: Sedimentation profiles observed by AUC	196

Article 5

Figure 1 : SAAT-SAAT interactions are specific	218
Figure 2 : Effect of length on the specificity	219
Figure 3 : SAAT specificity of interaction allows the segregation of two similar <i>E. coli</i>	220
Figure 4 : Adhesion of bacteria in aggregates on Hep-2 cells	221
Figure 5 : Autoaggregation by the SAATs increases adhesion of bacteria	222
Figure S1 : Interaction between AIDA-I, TibA and AIDA-I Δ N by ELISA assay	223
Figure S2 : Rupture force of the AIDA-I/AIDA-I and TibA/TibA interaction probed by single-molecule force spectroscopy	224
Figure S3 : Cah is a deletion variant of Ag43	225
Figure S4 : Aggregates are not sensitive to shear stress	226
Figure S5 : i3 mutant does not adhere under shear stress conditions	227
Figure S6 : Specificity of interaction during adhesion under shear stress conditions ...	228

Discussion

Figure 1 : Duplications dans le domaine passager du SAAT de <i>C. rodentium</i>	239
Figure 2 : Modèle d'interaction de l'autotransporteur Hap	243
Figure 3 : Structure et mécanisme d'interactions des cadhérines	246
Figure 4 : Les SAATs comme gènes de barbe verte	251

Annexe - Résultats supplémentaires

Figure 1 : AIDA-I est monomérique en présence de détergent	284
Figure 2 : Structure d'AIDA-I monomérique	285
Figure 3 : Oligomérisation d'AIDA-I en absence de détergent	287

Figure 4 : Interaction du mutant i10 observé par FRET	289
Figure 5 : Caractérisation des interactions AIDA-I/AIDA-I par FRET	290
Figure 6 : AIDA-I interagit en <i>cis</i> et en <i>trans</i>	291
Figure 7 : L'interaction préalable en <i>cis</i> est nécessaire pour l'interaction en <i>trans</i>	293
Figure 8 : Modèle pour l'interaction d'AIDA-I à la surface des bactéries	293

Liste des sigles et des abréviations

<i>ABC</i>	<i>ATP-binding cassette</i>
ADN	Acide déoxyribonucléique
ADP	Adénosine di-phosphate
Ala, A	Alanine
APEC	Avian pathogenic <i>Escherichia coli</i>
ARN	Acide ribonucléique
Arg, R	Arginine
Asn, N	Asparagine
Asp, D	Aspartate
ATP	Adénosine tri-phosphate
<i>BAM</i>	<i>β-barrel assembly machinery</i>
b-OG	n-octyl- β -D-glucoside
CCM/CMC	Concentration critique micellaire
CD	Dichroïsme circulaire
<i>CFU</i>	<i>Colony forming unit</i>
Cys, C	Cystéine
DAEC	Diffusely adhering <i>Escherichia coli</i>
DOC	Déoxycholate
EHEC	Enterohemorrhagic <i>Escherichia coli</i>
EPEC	Enteropathogenic <i>Escherichia coli</i>
ETEC	Enterotoxigenic <i>Escherichia coli</i>
<i>FPM</i>	Force proton motrice
<i>FRET</i>	<i>Fluorescence</i> ou <i>Forster resonance energy transfert</i>
Gln, Q	Glutamine
Glu, E	Glutamate
Gly, G	Glycine
His	Histidine
<i>IBC</i>	<i>Intracellular bacterial community</i>
IM	Membrane interne

IPTG	Isopropyl- β -D-thiogalactopyranoside
LPS	lipopolysaccharide
<i>MED</i>	<i>Membrane embedded domain</i>
<i>MFP</i>	<i>Membrane fusion protein</i>
<i>MRE</i>	<i>Mean residual ellipticity</i>
NTP	Nucléotide tri-phosphate
OD	Densité optique
OM	Membrane externe
OMP	Protéine de membrane externe
<i>POTRA</i>	<i>Polypeptide transport associated domain</i>
PP	Domaine riche en proline
Pro, P	Proline
SAAT	<i>Self-associating autotransporter</i>
Ser	Sérine
<i>SPATE</i>	Serine protease autotransporter of <i>Enterobacteriaceae</i>
SPR	Résonance plasmonique de surface
<i>SRP</i>	<i>Signal recognition particle</i>
SST1	Système de sécrétion de type 1
SST2	Système de sécrétion de type 2
SST3	Système de sécrétion de type 3
SST4	Système de sécrétion de type 4
SST5	Système de sécrétion de type 5
SST6	Système de sécrétion de type 6
STEC	Shiga toxin producing <i>Escherichia coli</i>
<i>TAM</i>	<i>Translocation and assembly module</i>
<i>T4CP</i>	<i>Type-4 coupling protein</i>
Trp	Tryptophane
TPS	Système à deux partenaires
<i>TU</i>	Unité de translocation
UDP	Uridine di-phosphate
UPEC	Uropathogenic <i>Escherichia coli</i>

Remerciements

Je tiens d'abord à remercier Dr Michael Mourez qui a été le meilleur directeur de recherche que j'aurais pu espérer. J'ai énormément grandi pendant toutes ces années passées dans le laboratoire grâce à Michael.

J'aimerais aussi remercier les membres de mon comité conseil, les Drs France Daigle et Mario Jacques, pour leur conseil et leur disponibilité, ainsi que toutes les personnes avec qui j'ai collaborées.

J'aimerais aussi remercier tous les gens que j'ai rencontrés dans le laboratoire et avec qui ça a été un plaisir de travailler. J'aimerais particulièrement remercier Frédéric Berthiaume qui a été d'une très grande aide tout au long de mon doctorat et, grâce à qui, il y avait toujours de la bonne musique dans le labo. J'aimerais ensuite remercier Victoria Girard avec qui j'ai travaillé au début de mon doctorat et qui m'a enseigné beaucoup. Finalement, j'aimerais remercier Marie-Ève Charbonneau pour tous les conseils et sans qui j'aurais eu beaucoup de difficultés à me rendre jusqu'à la fin.

J'aimerais aussi remercier les membres du GREMIP et du CRIP que j'ai croisés pendant toutes ces années qui m'ont épaulé et conseillé et qui ont rendu mon séjour au GREMIP aussi enrichissant.

J'aimerais aussi remercier les organismes subventionnaires pour le soutien financier : les Fonds québécois de la recherche sur la nature et la technologie (FQRNT), le Conseil de recherche en sciences naturelles et en génie du Canada (CRSNG), la Faculté de médecine vétérinaire, le CRIP et la faculté des études supérieures et postdoctorales.

Finalement, j'aimerais remercier ma famille et mes amis pour leur support, leur compréhension et leurs encouragements. J'aimerais aussi particulièrement remercier Daphnée Lamarche pour l'ensemble de son oeuvre et pour son incroyable support jusqu'à la toute fin.

Recension des écrits

Introduction

Au cours de leur cycle de vie, les bactéries vivent majoritairement en communauté, comme dans un biofilm, plutôt que seules sous forme planctonique (1). Il n'est donc pas étonnant qu'elles consacrent autant d'énergie à synthétiser des systèmes leur permettant de communiquer et d'interagir avec leur pairs et leur environnement. Ces interactions se font par des molécules qui se retrouvent à la surface des bactéries ou dans le milieu extracellulaire et les bactéries ont élaboré plusieurs systèmes permettant leur sécrétion (2). Le système de sécrétion de type V, représenté par les autotransporteurs monomériques, est le système le plus simple et le plus utilisé chez les bactéries à Gram négatif.

Les autotransporteurs monomériques sont des protéines modulaires qui sont extrêmement simples. Ils sont appelés autotransporteurs puisque toute l'information requise pour leur sécrétion est retrouvée dans un seul polypeptide (3). Les autotransporteurs n'ont donc pas besoin d'un système spécialisé pour leur sécrétion et utilisent uniquement les systèmes généraux de la bactérie (4). Ils sont composés d'une séquence signal permettant la translocation à travers la membrane interne par le système Sec et d'un tonneau β qui utilise le système *BAM* pour son insertion dans la membrane externe. Le domaine fonctionnel d'un autotransporteur, appelé domaine passager, se retrouve alors à l'extérieur de la bactérie et est généralement associé à la pathogenèse de diverses façons et à différentes étapes du développement d'une infection (5). Certains domaines passagers permettent l'adhésion et l'invasion de cellules épithéliales, tandis que d'autres sont impliqués dans la motilité intracellulaire, la résistance au sérum ou possèdent une activité cytotoxique. Conséquemment, il existe plusieurs sous-familles d'autotransporteurs monomériques, notamment les autotransporteurs auto-associatifs (SAATs; pour *self-associating autotransporters*) (6).

La famille des SAATs est composée de trois autotransporteurs provenant d'*Escherichia coli* : l'adhésine impliquée dans l'adhésion diffuse, AIDA-I; le locus

d'invasion B des *E. coli* entérotoxigéniques (ETEC), TibA; et le facteur d'agrégation, antigène 43, Ag43 (6). Ces autotransporteurs sont retrouvés principalement chez des souches d'*E. coli* intestinales qui causent des diarrhées. *E. coli* est d'ailleurs un important agent pathogène responsable de nombreux cas de diarrhée à travers le monde (7). Ces trois autotransporteurs ont été regroupés pour leur capacité à promouvoir l'auto-agrégation des bactéries via un mécanisme d'auto-association, bien que ce mécanisme n'ait jamais été démontré. Ces autotransporteurs partagent en plus d'importantes caractéristiques structurales, au-delà de l'organisation normale d'un autotransporteur. Notamment, leur domaine passager est composé uniquement de répétitions imparfaites d'une séquence de 19 acides aminés. De plus, les SAATs font partie des rares protéines qui sont glycosylées chez *E. coli*.

Toutefois, malgré les nombreuses similitudes entre les SAATs, il existe aussi plusieurs différences. Principalement, bien qu'ils aient la même organisation, certains SAATs possèdent des domaines additionnels et il existe des différences dans la biogenèse au niveau de la glycosylation et du clivage. De plus, les trois SAATs possèdent un nombre différent de répétitions de la séquence consensus de 19 acides aminés. Finalement, bien que AIDA-I, TibA et Ag43 aient tous été associés à la pathogenèse, la façon à laquelle ils contribuent à la virulence de *E. coli* semble être différentes.

Toutes ces différences mettent en évidence le manque de compréhension à propos des SAATs. Le but de ce travail de doctorat était donc de caractériser les SAATs afin de déterminer leur implication dans la pathogenèse. Une étude approfondie de structure-fonction de TibA nous a d'abord permis de déterminer l'importance de l'auto-agrégation dans la fonctionnalité des SAATs et a mis en évidence un mécanisme permettant l'évolution des SAATs. En étudiant le mécanisme d'auto-agrégation et l'évolution des SAATs, nous avons déterminé que les SAATs constituaient une grande famille d'autotransporteurs retrouvés chez les protéobactéries et que leur rôle ne se limitait pas à la pathogenèse. Nos résultats suggèrent que les SAATs ne sont pas seulement des facteurs de virulence, mais également des déterminants sociaux permettant d'organiser les communautés bactériennes.

1. Systèmes de sécrétion bactériens

L'interaction entre les bactéries ou entre les bactéries et leur environnement, incluant les cellules d'un organisme hôte, se fait principalement par des molécules sécrétées ou qui se retrouvent à la surface des bactéries (2). Pour ce faire, ces molécules doivent traverser complètement l'enveloppe bactérienne composée d'une (bactéries à Gram positif) ou deux membranes (bactéries à Gram négatif) (8). Les bactéries ont donc développé plusieurs systèmes de sécrétion pour permettre à des protéines synthétisées dans le cytoplasme de se retrouver dans le milieu extracellulaire. Ces protéines peuvent alors demeurer associées à la surface de la bactérie ou bien être sécrétées dans le milieu extracellulaire ou directement dans la cellule cible, qu'elle soit procaryote ou eucaryote. Ces protéines sécrétées peuvent ensuite agir de différentes façons : par exemple, en favorisant l'adhésion des bactéries aux cellules hôtes, en modulant la signalisation des ces cellules hôtes ou bien en permettant la communication entre les bactéries. Les fonctions associées aux protéines sécrétées sont généralement essentielles pour permettre la survie et la prolifération des bactéries dans l'environnement ainsi que dans l'organisme hôte (2, 9). Étant donné l'importance de ces systèmes de sécrétion pour le pouvoir pathogène des bactéries, leur inhibition est une voie prometteuse pour le développement de nouveaux antibiotiques. Par conséquent, il est essentiel de bien comprendre le fonctionnement de ces systèmes de sécrétions.

Cette section traite des différents systèmes de sécrétion retrouvée chez les bactéries, en particulier chez les bactéries à Gram négatif (Fig. 1). Étant donné qu'elles possèdent une seule membrane, la sécrétion chez les bactéries à Gram positif est relativement plus simple. En fait, les bactéries à Gram positif partagent les mêmes systèmes généraux de sécrétion à travers la membrane cytoplasmique que les bactéries à Gram négatif. Dans cette section, les systèmes sont divisés selon qu'ils permettent la translocation à travers la membrane cytoplasmique, la membrane externe ou bien les deux membranes en une seule étape. La nomenclature est celle utilisée par les principaux groupes travaillant sur la sécrétion de protéines chez les bactéries (10, 11).

Afin de simplifier, cette section traite des systèmes de sécrétion généraux et des systèmes permettant la sécrétion de protéines dans le milieu extracellulaire chez les bactéries à Gram négatif. Conséquemment, certains systèmes, comme le système chaperon-placier (*chaperone-usher*) spécifique à la formation de certains types de pili (12) et le système *Esx* chez les mycobactéries (13) par exemple, ne seront pas couverts.

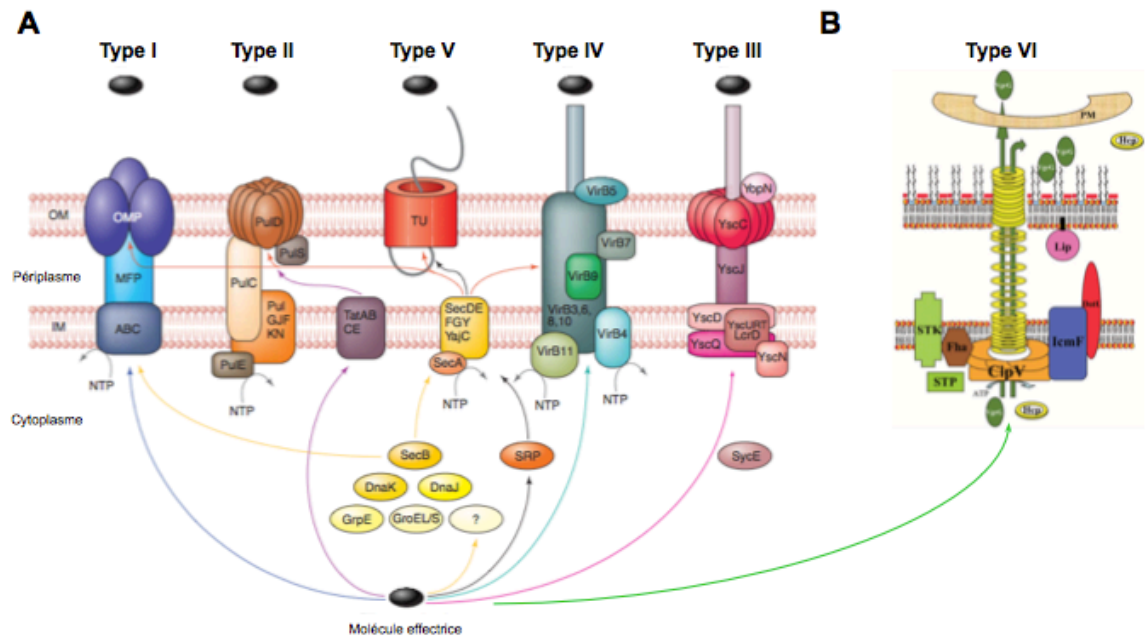


Figure 1 : Les différents systèmes de sécrétion chez les bactéries à Gram négatif [figure adaptée des références (14) et (15)] Représentation schématique des systèmes Sec et TAT ainsi que des systèmes sécrétion de type I à V (A) et de type VI (B). Les molécules effectrices synthétisées dans le cytoplasme sont prises en charge par les différentes chaperonnes cytoplasmiques et exportées dans le périplasmé par les systèmes généraux Sec et TAT ou directement dans le milieu extracellulaire par le *SST1*, *SST3*, *SST4* ou *SST6*. Les molécules qui se retrouvent dans le périplasmé sont ensuite exportées dans le milieu extracellulaire par le *SST2* et le *SST5* (et occasionnellement par le *SST1* et *SST4*).

1.1. Traversée de la membrane cytoplasmique

La première barrière à laquelle les protéines sécrétées font face est la membrane cytoplasmique. Deux voies de sécrétion principales existent pour traverser cette membrane, soit le système général de sécrétion Sec et le système de sécrétion TAT (twin-arginine translocation), le plus important étant le système Sec (16). Ces deux

systèmes sont retrouvés dans tous les domaines de la vie. La différence principale entre les deux est que le système Sec permet la translocation de polypeptides sous une forme dépliée alors que le système *TAT* permet la translocation de protéines complètement repliées.

1.1.1. Système de sécrétion Sec

Le système Sec est un composant essentiel à la viabilité des bactéries et est retrouvé dans tous les domaines de la vie (17). Il s'agit du principal système utilisé pour la translocation de protéines à travers la membrane cytoplasmique (18). Il sert aussi à l'intégration de protéines dans la membrane cytoplasmique et, chez les bactéries à Gram négatif, il constitue la première étape de plusieurs autres systèmes de sécrétion. Comme tous les systèmes de sécrétion, le système Sec est composé d'éléments permettant la reconnaissance des substrats, d'un moteur fournissant l'énergie et d'un conduit permettant la translocation.

1.1.1.1. Le translocon et autres composants

La composante centrale du système Sec est le translocon, un complexe hétérotrimérique composé des protéines SecY, SecE et SecG qui forment un pore dans la membrane plasmique et permettent le passage des polypeptides (17). Le pore est formé par SecY tandis que SecE est principalement requis pour la stabilité de SecY et ce complexe SecYE constitue l'unité minimale essentielle pour la translocation (19-21). La structure du translocon provenant de l'archéobactérie *Methanococcus jannaschii* montre que la protéine SecY forme le pore à elle-seule (22). SecY adopte une structure ressemblant à un sablier avec le centre de la structure formant un anneau de constriction. De plus, du côté périplasmique, une hélice α vient fermer le pore qui a une dimension de 5-8 Å. Cette structure représente le translocon dans un état fermé et celui-ci doit donc subir d'importants changements de conformation, le déplacement de l'hélice et l'ouverture du pore entre autres, pour devenir fonctionnel (23, 24). D'ailleurs, une étude récente visant à déterminer la taille maximale du canal a démontré que des molécules

avec un diamètre aussi large que 22-24 Å pouvaient être sécrétées efficacement (25). Ceci veut donc dire que le pore est assez large pour permettre le passage de polypeptides ayant adoptés des structures secondaires minimales ou arborant des modifications post-traductionnelles simples, comme des sucres simples. Des structures provenant d'autres organismes et utilisant des conditions différentes ont ensuite montré le translocon dans un état ouvert. Ses structures montrent, entre autres, un déplacement de l'hélice α et un relâchement de l'anneau de constriction pour ouvrir le pore (26) ainsi qu'une ouverture latérale permettant l'entrée de la séquence signale dans le système et servant pour l'insertion des protéines membranaires (27). Finalement, quelques études suggèrent que le translocon forme des dimères dans des conditions où il serait dans un état actif (28, 29). Un premier modèle veut que les deux translocons soient face-à-face, tandis qu'un deuxième modèle prévoit une interaction dos-à-dos. Aucun des deux modèles n'expliquent en totalité le fonctionnement du translocon, ce qui fait que l'état d'oligomérisation de SecYEG est toujours controversé.

De son côté, la protéine SecG complète le translocon en interagissant avec SecY, mais n'est pas essentielle pour le fonctionnement du translocon (21). Toutefois, SecG semble stimuler la translocation en facilitant l'interaction de SecA avec SecY (30). SecA est le moteur du translocon et fournit l'énergie nécessaire à la translocation (31). SecA fait partie de la famille de protéines contenant un motif DExD, où x peut être n'importe quel acide aminé (32). Cette famille est composée, entre autres, d'hélicases permettant de dérouler l'ADN ou l'ARN de façon ATP-dépendante. SecA oscille entre le cytoplasme et la périphérie de la membrane et interagit avec les séquences signales des polypeptides à être exportés ainsi qu'avec la chaperonne cytoplasmique SecB (33). SecA entrainerait donc les polypeptides à être exportés à la membrane grâce à son interaction avec les séquences signales et avec le translocon. De plus, il semble que SecA ait une activité de chaperonne vis à vis les polypeptides n'ayant pas de séquences signales. En effet, SecA favorise le repliement des protéines cytoplasmiques, ce qui aide probablement à discriminer les substrats à être exportés (34).

D'autres composants interagissent avec le translocon SecYEG, comme les protéines SecD, SecE et YidC (35). SecD et SecE sont non essentielles pour le fonctionnement du translocon, mais améliorent grandement l'efficacité de la sécrétion. Ce sont deux protéines de membrane interne qui forment un complexe ensemble. SecD contient un large domaine périplasmique, ce qui suggère que le complexe SecDE joue un rôle post-translocation (36). Le modèle d'action est donc que le domaine périplasmique de SecD interagit avec les polypeptides sortant du translocon et les maintient dans le périplasma afin d'éviter un retour en arrière (37, 38). YidC, de son côté, interagit avec SecYEG via SecDE et est impliqué dans l'intégration des protéines dans la membrane (39, 40).

1.1.1.2. La séquence signale

La majorité des protéines destinées aux espaces extra-cytoplasmiques sont synthétisées sous la forme d'une pré-protéine. Elles contiennent une séquence signale N-terminale qui les dirige vers le système de sécrétion Sec et qui est clivée lors de la translocation par une « signal peptidase », à l'exception des protéines membranaires pour lesquelles la séquence signale n'est pas clivée (41, 42). La séquence signale typique pour le système de sécrétion Sec est composée de 20 à 30 acides aminés et est divisée en trois domaines (43) : (i) un domaine n, en N-terminal, chargé positivement, (ii) un domaine h, au centre, riche en résidus hydrophobes et (iii) un domaine c, en C-terminal, plutôt hydrophile. À la fin du domaine c, on retrouve le motif de clivage reconnu par la « signal peptidase » qui est formé de petits acides aminés non chargés, Ala le plus souvent, en position -1 et -3 avant le site de clivage.

La séquence signale permet aux pré-protéines d'être reconnues directement au cours du processus de traduction. En effet, dès la sortie du ribosome, les polypeptides peuvent être reconnus par la ribonucléoprotéine *SRP* (*Signal Recognition Particule*) (44). *SRP* est une chaperonne cytoplasmique, formée de la protéine *Ffh* (*fifty-four homologue*) et de l'ARN 4.5S (45). Alternativement, le polypeptide nouvellement synthétisé peut aussi être reconnu par la chaperonne cytoplasmique SecB et le *TF*

(*Trigger factor*) (44, 46). En fait, *SRP* et le *TF* sont en compétition pour la liaison avec le polypeptide naissant directement à la sortie du ribosome et il semble que la liaison d'un facteur inhibe la liaison avec le deuxième (47, 48). *SRP* reconnaît principalement les séquences signales plus hydrophobes qui sont retrouvées majoritairement chez les protéines membranaires (49, 50). D'ailleurs, la majorité des substrats de *SRP* sont des protéines de membrane interne. La liaison de la séquence signal avec *SRP* provoque l'arrêt de la traduction et entraîne le complexe polypeptide naissant-ribosome vers le translocon. Le *TF*, quant à lui, reconnaît une séquence qui revient fréquemment dans un polypeptide déplié et prévient son repliement prématuré (51). C'est cependant la protéine SecB qui va emmener les polypeptides destinés au périplasma à la machinerie de sécrétion (52). Il existe donc deux voies de sécrétion : une voie co-traductionnelle *SRP*-dépendante et une voie post-traductionnelle SecB-dépendante. Le choix entre les deux voies s'effectue directement à la sortie du ribosome.

1.1.1.3. Voie co-traductionnelle

La voie co-traductionnelle débute donc avec la liaison du polypeptide naissant à *SRP* directement à la sortie du ribosome (53) (Fig. 2). Cette liaison provoque un arrêt ou un ralentissement de la traduction et *SRP* entraîne alors le complexe ribosome-polypeptide naissant vers la membrane, où *SRP* reconnaît la protéine de membrane interne FtsY. FtsY se lie ensuite à SecY, ce qui provoque la dissociation de *SRP* du complexe et permet au ribosome d'interagir avec le translocon afin de coupler la traduction avec la sécrétion. Dans certains cas, l'insertion dans la membrane se fait avec l'aide de la protéine YidC (39).

1.1.1.4. Voie post-traductionnelle

Dans la voie post-traductionnelle, le polypeptide complètement synthétisé est maintenu dans un état déplié dans le cytoplasme par différentes chaperonnes cytoplasmiques (54) (Fig. 2). C'est généralement la protéine SecB qui prend en charge

les polypeptides destinés au système Sec, mais d'autres chaperonnes périplasmiques, comme DnaK et DnaJ, peuvent aussi intervenir (55, 56). Le complexe SecB-polypeptide est ensuite recruté à la membrane grâce à une interaction avec SecA, qui interagît préalablement avec le translocon. Par la suite, SecA fixe une molécule d'ATP, provoquant la dissociation de SecB et l'initiation de la sécrétion (57). Plusieurs cycles d'hydrolyse d'ATP par SecA s'ensuivent et, à chaque hydrolyse d'ATP, SecA transfère environ 30 acides aminés au translocon (58).

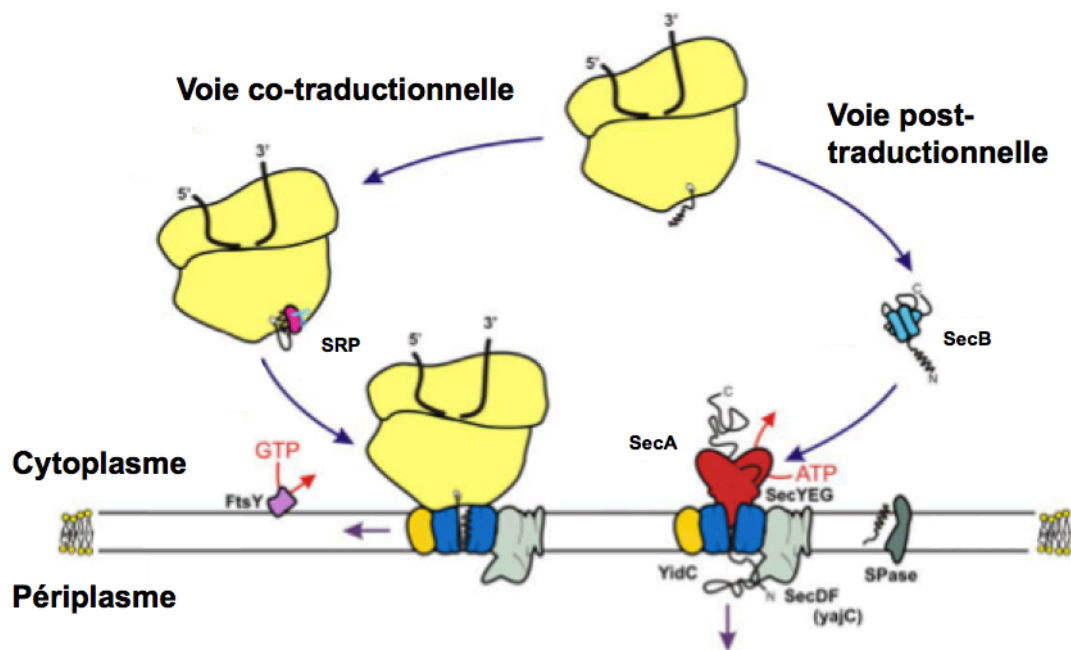


Figure 2 : Sécrétion à travers la membrane cytoplasmique par le système Sec [figure adaptée de la référence (59)] Le système Sec (SecYEG) peut transloquer des protéines dépliées dans le périplasma par une voie co-traductionnelle *SRP*-dépendante ou bien par une voie post-traductionnelle utilisant SecA et SecB.

1.1.2. Système de sécrétion *TAT*

Un deuxième système de sécrétion localisé dans la membrane plasmique, le système *TAT*, complète le système Sec. En effet, contrairement à Sec qui permet la sécrétion de polypeptides dépliés, le système *TAT* permet la translocation de protéines

complètement repliées à travers la membrane cytoplasmique (60). Le système *TAT* est moins important que *Sec* et n'est pas retrouvé chez toutes les bactéries. Cependant, lorsqu'il est retrouvé chez des bactéries pathogènes, celui-ci joue généralement un rôle dans la virulence (61). Les substrats sécrétés par le système *TAT* sont des protéines qui ont besoin de co-facteurs, qui font partie d'un complexe hétéro-oligomérique ou dont les conditions requises pour le repliement sont retrouvées dans le cytoplasme. Dans la plupart des cas, le système *TAT* est utilisé pour un plus petit nombre de protéines que *Sec*. Cependant, il peut aussi être très important chez certaines bactéries et archéobactéries, notamment chez les streptomycètes et chez certaines bactéries halophiles, où près de la moitié des protéines sécrétées utilisent le système *TAT* (60, 62, 63). Par exemple, *Streptomyces coelicolor* sécrète plus de 150 protéines par le système *TAT*, comparativement à *E. coli* qui utilise *TAT* pour environ 25 protéines.

La séquence signal des substrats destinés au système *TAT* possède une organisation similaire à celle reconnue par le système *Sec*, avec une région n, une région h et une région c. Toutefois, il existe des différences importantes qui distinguent les deux voies de sécrétion (60). Notamment, la région n est généralement plus longue pour les séquences signales du système de sécrétion *TAT*, la région h est moins hydrophobe et la région c, riche en acides aminés chargés positivement, sert de motif d'exclusion pour le système *Sec* (64). Plus important encore, ces séquences signales possèdent un motif « doubles arginines » caractéristique qui se trouve entre la région n et h et qui correspond à la séquence Z-R-R-x-F-L où Z est un résidu polaire et x représente n'importe quel résidu (65).

La sécrétion de protéines repliées par le système *TAT* représente un défi important pour la bactérie comparativement à un système sécrétant des polypeptides dépliés qui, peu importe leur taille finale, utilise un complexe de taille unique. En effet, le système *TAT* permet autant la sécrétion de très petites protéines que de gros complexes protéiques pouvant atteindre des tailles allant jusqu'à 70 Å chez *E. coli* (66). Le processus doit donc être dynamique et bien régulé afin de conserver la perméabilité membranaire.

La sécrétion par le système *TAT* se fait de façon post-traductionnelle puisque les substrats doivent être repliés avant leur transport. Les substrats interagissent donc d'abord avec des chaperonnes cytoplasmiques qui permettent un repliement adéquat avant d'être dirigés vers le système *TAT*. Dans certains cas, des chaperonnes cytoplasmiques générales comme DnaK et SlyD interagissent avec la séquence signalale des ces substrats alors que certains substrats possèdent leur propre chaperonne (67, 68). Par exemple, c'est le cas pour la *TMAO* réductase (TorA) de *E. coli* qui possède sa chaperonne spécifique, TorD. TorD permet un repliement efficace de TorA et prévient une dégradation prématurée de la séquence signalale (68).

Lorsqu'un substrat est correctement replié et prêt pour sa sécrétion, celui-ci interagit avec un complexe formé par les protéines TatB et TatC (Fig. 2B). TatBC forme un complexe membranaire qui reconnaît la séquence signalale des substrats. C'est TatC qui reconnaît et se lie au motif « doubles arginines » de la séquence signalale (69). La protéine TatA est alors recruté au complexe TatBC (70) et, au contact de TatBC, Tata va polymériser pour former un oligomère de taille variable qui constitue le pore membranaire permettant la translocation du substrat (71). L'oligomérisation de TatA nécessite la force proton motrice (*FPM*) et varie en fonction du substrat présent afin de conserver la perméabilité membranaire. Finalement, lorsque le substrat est dans le périplasma, la séquence signalale est clivée et TatA se dépolymérise (72, 73). Chez certaines bactéries, TatB n'est pas présent et le système est composé seulement de TatA et TatC (74).

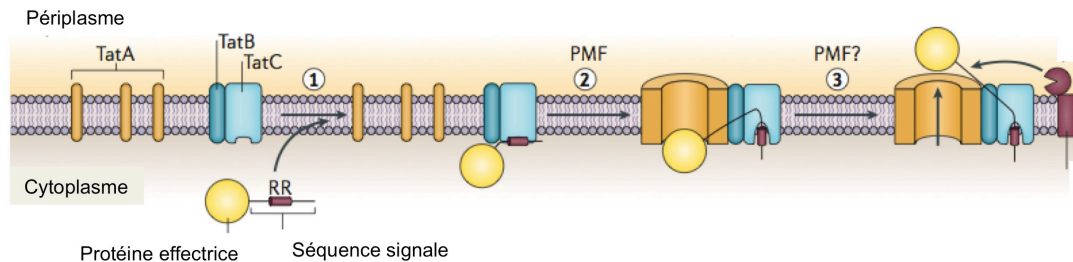


Figure 3 : Sécrétion à travers la membrane cytoplasmique par le système *TAT* [Figure adaptée de la référence (60)] Le système *TAT* permet la translocation de protéines repliées par une pore formé par l'oligomérisation de TatA dont la grosseur varie selon le substrat à être sécrété.

1.2. Traversée de la membrane externe

Dans le cas des protéines sécrétées par les bactéries à Gram négatif, celles-ci doivent traverser le périplasma pour se rendre à la membrane externe et être sécrétées dans le milieu extracellulaire. Suite à la translocation à travers la membrane interne, les protéines qui sont destinées à la membrane externe ou au milieu extracellulaire sont donc prises en charge par différentes chaperonnes périplasmiques (75). Les principales chaperonnes périplasmiques sont SurA, Skp et DegP, bien qu'il existe d'autres chaperonnes, comme FkpA, PpiA et PpiD, qui pourraient avoir un rôle secondaire dans le transport périplasmique des protéines de membrane externe. SurA semble être la chaperonne principale, tandis que Skp et DegP représenteraient une voie alternative. Il semble que cette seconde voie alternative soit moins importante et servirait surtout en cas de stress important. Les protéines membranaires sont ensuite insérées dans la membrane externe par le complexe *BAM* (*β-barrel assembly machinery*) alors que les protéines sécrétées vont l'être par les systèmes de sécrétion de type 2 ou de type 5.

1.2.1. Système d'insertion dans la membrane externe (*BAM*)

La majorité des protéines de membrane externe forment des tonneaux β et nécessitent le complexe *BAM* (*β-barrel assembly machinery*) pour leur insertion dans la membrane externe. La composante centrale du complexe *BAM* est la protéine BamA. BamA est essentiel et est présent chez toutes les bactéries à Gram négatif, de même que dans les mitochondries et les chloroplastes des cellules eucaryotes (76). L'importance de BamA dans la biogenèse des protéines de membrane externe a été découverte chez *Neisseria meningitidis* où il a été démontré que la délétion de BamA entraînait une accumulation de protéines, normalement destinées à la membrane externe, dépliées dans le périplasma (77). La structure de BamA consiste en un tonneau β ancré dans la membrane externe sur lequel est greffé, dans le périplasma, cinq domaines *POTRA* (*polypeptide transport-associated*) (78). Chacun des domaines *POTRA* adopte une structure similaire qui consiste en un feuillet composé de trois brins β sur lequel repose deux hélices α antiparallèles (78, 79). Ces domaines *POTRA* interagissent avec les

substrats du complexe *BAM* ainsi qu'avec les autres composantes du complexe *BAM* (78, 80, 81).

Les autres composantes du complexe *BAM* (BamB, C, D et E) sont des lipoprotéines de membrane externe contenant un large domaine périplasmique (82). Du nombre, seul BamD est essentiel pour la viabilité des bactéries, tandis que la délétion des autres affecte de façon plus ou moins importante l'intégrité membranaire. Le rôle des lipoprotéines accessoires n'est toujours pas bien défini, mais il semble que certaines pourraient recruter les substrats au complexe *BAM*. Par exemple, BamB, par sa structure en hélice β , pourrait stabiliser les brins β naissants des substrats (83). De plus, un orthologue de BamD dans la mitochondrie interagit avec les substrats du système d'insertion des protéines membranaires dans la mitochondrie (84).

Le signal de recrutement pour le système *BAM* se retrouve dans l'extrémité C-terminale de la protéine et correspond au dernier brin β (80). Il s'agit d'un motif où les résidus 1 (généralement une phénylalanine), 3 (généralement une tyrosine), 5, 7 et 9 à partir de l'extrémité C-terminale sont des résidus hydrophobes. Dans le périplasma, les protéines de membrane externe sont prises en charge par les chaperonnes SurA, Skp et DegP, qui les dirigent vers le complexe *BAM* (85). Toutefois, une interaction avec le complexe *BAM* a été montrée seulement pour SurA, ce qui suggère que SurA est la principale chaperonne pour délivrer les protéines de membrane externe au complexe *BAM* (75). Le rôle des chaperonnes périplasmiques est de conserver les protéines de membrane externe dans un état déplié. Cela veut donc dire que le repliement se fait, du moins en partie, par le complexe *BAM*. Les domaines *POTRA* sont de bons candidats pour le repliement des substrats, puisqu'ils interagissent avec ceux-ci par un processus de β -augmentation, où le brin β du substrat est stabilisé par le feuillet du domaine *POTRA* (78). De plus, il a été montré que les domaines *POTRA* interagissaient avec des brins β autres que le brin C-terminal (79).

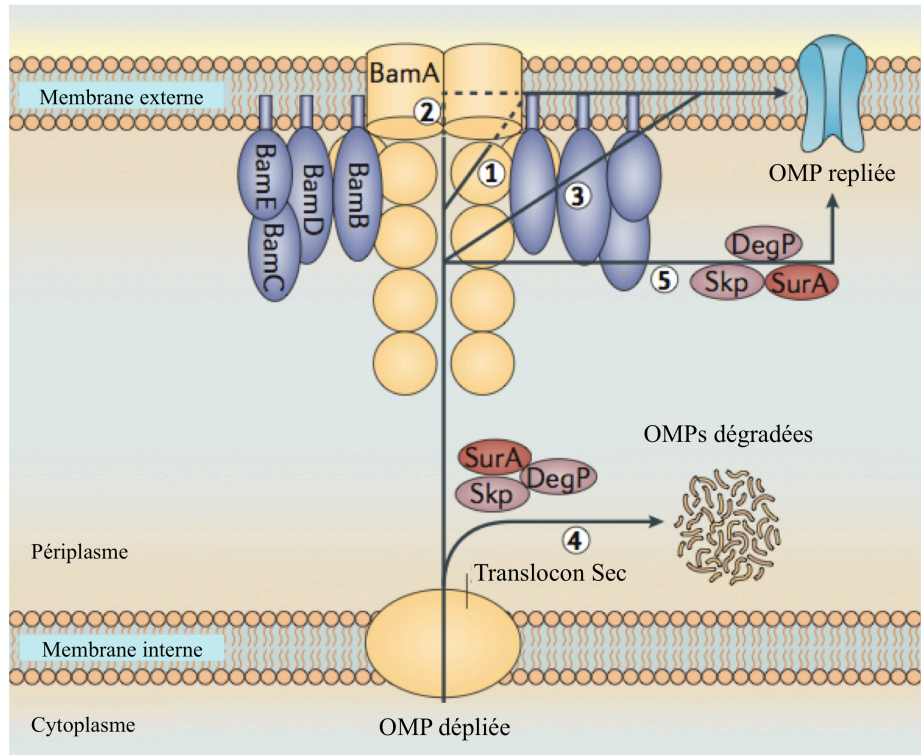


Figure 4 : Les différents modèles d'insertion de protéines de membrane externe par le complexe *BAM* [figure adaptée de la référence (4)]. (1) Dans le premier modèle, les protéines de membrane externe (*OMPs*) sont correctement repliées par les domaines *POTRA* et les protéines accessoires de *BAM* et sont incorporées dans la membrane à partir du tonneau de BamA. (2) Dans le deuxième modèle, l'oligomérisation de BamA permet la formation d'un pore central qui sert à l'insertion des *OMPs* dans la membrane. (3) Un modèle alternatif suggère que le tonneau de BamA serve de modèle pour le repliement et l'insertion des brins β de la nouvelle *OMP*. (4) Les protéines qui ne sont pas correctement repliées dans le périplasme sont alors dégradées par DegP. (5) Un dernier modèle suggère que les *OMPs* sont repliées par le complexe Bam, mais sont insérées dans la membrane par les chaperonnes périplasmiques, notamment DegP.

Le mécanisme d'insertion des protéines dans la membrane par le complexe *BAM* est plus nébuleux que la reconnaissance des substrats et plusieurs modèles ont été développés (86) (Fig. 4). Dans un premier modèle, le tonneau de BamA sert de pore pour l'incorporation des nouvelles protéines dans la membrane externe. Dans ce modèle, les domaines *POTRA* et les autres composantes du complexe serviraient au recrutement et à l'intégration des nouvelles protéines dans le tonneau de BamA. Ce modèle implique toutefois l'ouverture du tonneau de BamA afin de permettre la diffusion de la nouvelle

protéine dans la membrane, ce qui est peu probable. Dans un deuxième modèle, BamA oligomérisé afin de former une cavité centrale qui permet l'insertion des nouvelles protéines dans la membrane. Dans ce cas-ci, les domaines *POTRA* et les autres composantes participeraient au repliement de la nouvelle protéine et à son insertion dans la membrane. Ce modèle provient du fait que BamA peut oligomériser *in vitro* (80). Alternativement, le tonneau de BamA pourrait servir de modèle pour les nouvelles protéines et celles-ci pourraient se replier à proximité de BamA. Finalement, dans un dernier modèle, le complexe *BAM* sert seulement à replier les nouvelles protéines. Lorsque bien repliées, les nouvelles protéines sont insérées dans la membrane par un facteur accessoire qui pourrait être DegP. Il a été observé que DegP forme des cages composées de 12 ou 24 sous-unités et que ces cages pouvaient contenir une protéine de membrane externe correctement repliée (87, 88). De plus, les cages contenant une protéine repliée peuvent interagir avec la membrane. Donc, dans ce modèle, DegP reçoit la protéine correctement repliée du complexe *BAM* et entraîne celle-ci à la membrane pour faciliter son insertion. L'intérêt de ce modèle est que DegP est à la fois une chaperonne et une protéase. Il pourrait donc s'agir d'un mécanisme de contrôle de qualité où la nouvelle protéine est insérée dans la membrane si elle est correctement repliée ou bien elle est dégradée si elle n'est pas adéquatement repliée.

1.2.2. Système de sécrétion de type II

Le système de sécrétion de type II (*SST2*) permet la translocation de protéines repliées du périplasme au milieu extracellulaire (89). Le *SST2* a été le premier système de sécrétion spécifique décrit chez les bactéries à Gram négatif et a été considéré à l'origine comme étant la branche terminale principale du système général de sécrétion (90). Le *SST2* forme un complexe de 12 à 15 protéines qui traverse les deux membranes et permet, entre autres, la sécrétion de plusieurs toxines, protéases ou lipases. Par exemple, *Vibrio cholerae* utilise un *SST2* pour la sécrétion de la toxine cholérique (91). Le système modèle pour l'étude du *SST2* est celui permettant la sécrétion de la pullulanase, PulA, chez *Klebsiella oxytoca* (92).

Le *SST2* est composé de quatre parties : (I) un complexe dans la membrane externe, (II) un pseudopilus dans le périplasma, (III) un complexe dans la membrane interne et (IV) une ATPase dans le cytoplasme. Le complexe dans la membrane externe est formé majoritairement par la sécrétine GspD, qui s'oligomérisent et forme un pore d'environ 55 Å, assez grand pour laisser passer des substrats repliés (93). L'assemblage et l'oligomérisation correcte de la sécrétine se fait avec l'aide d'une pilotine, GspS (94). Le complexe dans la membrane interne est composé de quatre protéines, GspC-F-L-M et fait le lien entre l'ATPase cytoplasmique, GspE, qui fournit l'énergie au système, et la sécrétine de la membrane externe. Finalement, le pseudopilus qui est composé de la pseudopiline majeure GspG et des pseudopilines mineures GspH-I-J-K est ancré dans le complexe de la membrane interne. Le pseudopilus présente une homologie structurale avec les pili de type IV et fait en sorte que le *SST2* et les pili de type IV sont souvent comparés l'un à l'autre (95, 96).

La séquence signale des substrats du *SST2* n'a toujours pas été bien identifiée (97). Cependant, suite à la translocation des substrats dans le périplasma par les systèmes Sec ou TAT, le substrat replié interagit avec la partie périplasmique de la sécrétine et avec l'extrémité du pseudopilus (98, 99). Les substrats vont ensuite être poussés vers l'extérieur grâce au pseudopilus qui agirait de façon similaire à un piston.

1.2.3. Système de sécrétion de type V

Le système de sécrétion de type V (*SST5*) est particulier (100). Il est représenté majoritairement par les autotransporteurs monomériques (qui seront décrits plus en détail dans une section subséquente). Les autotransporteurs monomériques (type Va) constituent le système de sécrétion le plus simple décrit dans la littérature. Toute l'information nécessaire à la sécrétion est contenue dans un seul polypeptide, d'où le terme autotransporteur puisque la protéine promouvoit sa propre sécrétion. Ils n'ont donc pas besoin d'un système de sécrétion spécifique et utilisent seulement les systèmes généraux de la bactérie. Un autotransporteur est composé d'une séquence signale qui permet la translocation à travers la membrane interne, d'un tonneau β qui s'insère dans la membrane externe et qui promouvoit la sécrétion d'un domaine fonctionnel, appelé

domaine passager, dans le milieu extracellulaire (3). Il existe plusieurs variantes aux autotransporteurs monomériques qui font aussi partie du *SST5* (Fig. 5) (101, 102).

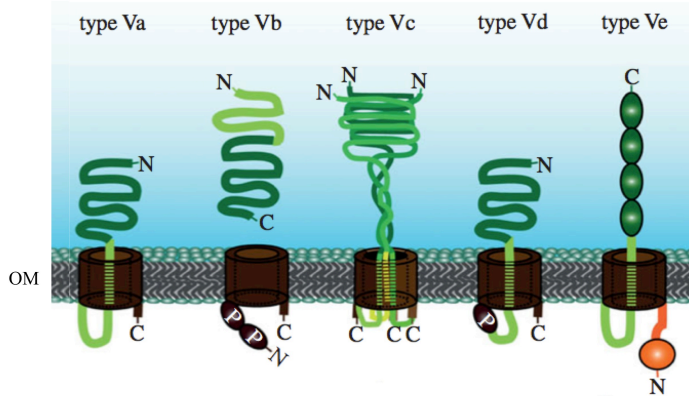


Figure 5 : Les variations du système de sécrétion de type V [figure adaptée de la référence (101)]. Représentation schématique des différentes variations du *SST5* : type Va, les autotransporteurs monomériques; type Vb, les systèmes à deux partenaires; type Vc, les autotransporteurs trimériques, type Vd, le système hybride entre les types Va et Vb; type Ve; les autotransporteurs inversés.

Dans le système à deux partenaires (*TPS*; type Vb), le tonneau β (TpsB) et le domaine passager (TpsA) sont codés par deux polypeptides différents (103, 104). Les exemples les plus étudiés de systèmes à deux partenaires sont l'hémagglutinine filamenteuse FHA de *Bordetella pertussis* (105) et les adhésines de haut poids moléculaire HMW1 et HMW2 de *Haemophilus influenzae* (106). Les systèmes *TPS* permettent généralement aux bactéries d'interagir avec leur environnement. Les protéines TpsA ont été associées à une multitude de fonctions différentes, allant de l'adhésion et l'invasion jusqu'à la survie intracellulaire en passant par l'inhibition de la phagocytose (107). Les protéines TpsA varient donc beaucoup autant en séquence qu'en longueur et peuvent atteindre jusqu'à 5000 acides aminés. Toutefois, malgré les différences de séquence, les protéines TpsA adoptent généralement une structure en forme d'hélice β (108). Les protéines TpsA possèdent également une séquence signal permettant la translocation à travers la membrane interne par le système Sec et possèdent une région *TPS* conservée. Cette région TPS d'environ 300 acides aminés constitue le

signal pour la reconnaissance et l'interaction avec son partenaire, la protéine TpsB (109, 110).

Les protéines TpsB font partie de la superfamille BamA des protéines de membrane externe (111). La structure de FhaC, qui permet la sécrétion de FHA, montre un tonneau β monomérique composé de 16 brins β et d'une partie périplasmique composée de deux domaines *POTRA* qui servent à la reconnaissance de la région TPS de la protéine TpsA (112). La diversité entre les TPS se voit aussi au niveau de TpsB. La structure de HMW1B, impliquée dans la sécrétion de l'adhésine HMW1, montre un tonneau composé de 10 brins β ayant un seul domaine *POTRA* (113). De plus, il semble que HMW1B se retrouve sous forme de dimère dans la membrane (114). Les différents TPS se rejoignent toutefois au niveau du mécanisme utilisé pour la sécrétion qui semble être plus simple et plus indépendant comparativement aux autotransporteurs. En effet, il semble que le composant TpsB soit suffisant pour permettre la sécrétion de son partenaire et n'a donc pas besoin de l'aide de systèmes accessoires présents dans la membrane, comme le système *BAM* (115, 116). Par contre, la sécrétion requiert toujours la présence des chaperonnes périplasmiques.

Le type Vc représente les autotransporteurs trimériques (117). Ceux-ci sont retrouvés autant chez des bactéries pathogènes que chez des bactéries environnementales (118). Dans tous les cas, il semble que la fonction principale d'un autotransporteur trimérique soit l'adhésion, bien que certains puissent en plus promouvoir la formation de biofilm et l'auto-agrégation par exemple (119). Les plus connus sont les adhésines YadA de *Yersinia enterocolitica* (120), BadA de *Bartonella henselae* (121) et Hia de *Haemophilus influenzae* (122). L'organisation des autotransporteurs trimériques est similaire à celle des autotransporteurs monomériques. On retrouve une séquence signale Sec en N-terminal, un domaine β en C-terminal et un domaine passager entre les deux. La différence est que les protéines doivent oligomériser pour former un homotrimère afin d'être stables et fonctionnelles (123). Le domaine β forme un tonneau β de 12 brins, où chaque monomère contribue pour quatre brins β (124). De son côté, le domaine passager est séparé en deux parties, une tête et une tige. La tête est responsable pour la

fonctionnalité de la protéine tandis que la tige forme une structure « coiled-coil » étendue qui éloigne le domaine d'adhésion de la membrane (125, 126).

Finalement, il a été proposé que deux nouveaux sous-groupes, les types Vd et Ve, fassent partie du *SST5* (101, 102). Le type Vd représente un hybride entre un autotransporteur monomérique et un système à deux partenaires (127). Le domaine passager est attaché au domaine β , mais ce dernier ressemble à un domaine β d'un système à deux partenaires. Le domaine β est composé de 16 brins β et contient un domaine *POTRA* périplasmique. La raison de la présence de ce domaine *POTRA* est plutôt incomprise, puisque les domaines *POTRA* des protéines TpsB servent à reconnaître la protéine TpsA via la région *TPS*. Dans le cas des protéines du type Vd, comme le domaine passager est fusionné au domaine β , cette étape de reconnaissance n'est pas requise. De plus, il n'y a pas de région *TPS* dans le domaine passager des protéines du type Vd. Il a été suggéré que le domaine *POTRA* agissait comme chaperonne en limitant le repliement du domaine passager dans le périplasma et en promouvant sa sécrétion dans le milieu extracellulaire. Le modèle pour le type Vd est la protéine PlpD de *Pseudomonas aeruginosa* et plus de 200 orthologues ont été identifiés (127). Le type Ve quant à lui ressemble à un autotransporteur monomérique, mais est plutôt inversé (128). Ce type d'autotransporteur est représenté par l'intimine de *E. coli* et l'invasine de *Yersinia pestis*. Bien que ces protéines soient connues depuis longtemps, des études récentes suggèrent de les regrouper dans le *SST5*. Le tonneau β est donc retrouvé dans la partie N-terminale du polypeptide alors que le domaine passager est en C-terminal. Les protéines faisant partie du type Ve utilisent donc probablement un mécanisme de sécrétion similaire aux autotransporteurs monomériques à l'exception que le repliement se fait à partir de la partie C-terminale.

1.3. Traversée des deux membranes

Les autres systèmes de sécrétion (Systèmes de type 1-3-4-6) traversent les deux membranes et permettent la translocation à travers les deux membranes en une seule

étape. Dans certains cas, ces systèmes permettent la sécrétion de protéines effectrices directement dans la cellule cible.

1.3.1. Système de sécrétion de type I

Le système de sécrétion de type I (*SSTI*) est formé d'un complexe qui traverse les deux membranes et permet la sécrétion d'un substrat directement du cytoplasme au milieu extracellulaire (129). Il existe plusieurs variations du *SSTI* qui permettent la sécrétion de polysaccharides et d'une variété d'autres petites molécules (130). Les substrats sont très variés, incluant des adhésines, des protéases et des toxines, et peuvent atteindre un poids moléculaire allant jusqu'à 800 kDa (131). En fait, n'importe quelle protéine pourrait potentiellement être sécrétée par le *SSTI* si elle possédait une séquence signal pour ce système (132). Les substrats les plus connus utilisant le *SSTI* sont l'hémolysine (HlyA) des *E. coli* uropathogènes (UPEC) (133), l'adénylate-cyclase de *Bordetella pertussis* (134) et la Colicine V de *E. coli* (135).

Le système est formé de trois composants : un transporteur *ABC* (ATP-binding cassette) dans la membrane interne, une protéine membranaire de fusion (*MFP*) dans le périplasme ainsi qu'une protéine de membrane externe, homologue à TolC. Par ailleurs, TolC est aussi utilisé dans d'autres systèmes, tels que les systèmes de pompes à efflux (136). La séquence signal pour le *SSTI* est retrouvée dans la région C-terminale des substrats et n'est pas clivée lors de la sécrétion (137). Elle peut faire jusqu'à 50 acides aminés, mais aucun motif particulier n'a été identifié.

Dans le cytoplasme, les substrats sont maintenus dans un état déplié par des chaperonnes cytoplasmiques telles que GroEL et SecB (138). Il semble que la sécrétion des substrats se fasse lorsqu'ils sont dépliés et que le repliement de la protéine se fasse une fois dans le milieu extracellulaire. De toute façon, le pore formé par TolC dans la membrane externe n'est pas assez gros pour permettre le passage de protéines repliées (139). La séquence signal du substrat interagit avec le transporteur ABC, ce qui déclenche la formation du complexe *ABC-MFP-TolC* et la création du canal permettant

la sécrétion. L'énergie nécessaire à la sécrétion vient de l'hydrolyse d'ATP par le transporteur *ABC* ainsi que de la *FPM* (140).

1.3.2. Système de sécrétion de type III

Le système de sécrétion de type III (*SST3*) permet la sécrétion de protéines du cytoplasme directement dans le cytosol de la cellule cible (141). Le *SST3* est un des systèmes de sécrétion les plus complexes et est composé d'au moins 20 composants structuraux. Plusieurs importants pathogènes, comme *Pseudomonas*, *Yersinia*, *Shigella*, *Salmonella* et *E. coli*, possèdent un ou plusieurs *SST3*. Par ailleurs, le flagelle des bactéries à Gram négatif possède un système de sécrétion très similaire au *SST3* qui permet son assemblage (142). Le nombre et la fonction des effecteurs du *SST3* varient grandement d'une bactérie à l'autre et même d'un système à l'autre lorsqu'une bactérie possède plus d'un *SST3* (143). Le profil d'effecteurs reflète le rôle du *SST3* dans le mode de vie de la bactérie. Malgré les différences, il y a certains points communs et plusieurs effecteurs vont provoquer un réarrangement du cytosquelette de la cellule hôte, un débalancement de l'intégrité membranaire ou induire l'apoptose (144).

Le *SST3* est composé de trois parties : un corps basal, une seringue et un translocon (145). Le corps basal traverse complètement l'enveloppe bactérienne et forme deux anneaux, un dans la membrane interne et le deuxième dans la membrane externe. Dans le cytoplasme, le corps basal est associé à une ATPase qui fournit l'énergie au système. À l'intérieur du corps basal, on retrouve la seringue, qui s'étend de la membrane interne jusque dans le milieu extracellulaire. La seringue agit comme canal pour la sécrétion des effecteurs, mais comme le diamètre n'est que de 2-3 nm, les effecteurs doivent être dépliés lors de la sécrétion (146, 147). Finalement, au bout de la seringue, on retrouve le translocon, qui ne s'assemble qu'au moment du contact avec la cellule cible (148, 149).

Beaucoup de composants structuraux, comme les sous-unités formant la seringue, sont aussi sécrétés par le système au fur et à mesure de son assemblage. Le

SST3 doit donc être très bien régulé (150). D'abord, la seringue est assemblée jusqu'à l'obtention d'une longueur bien définie (151, 152). Cette longueur varie d'un *SST3* à l'autre, mais reflète la distance à parcourir jusqu'à la membrane de l'hôte selon les autres structures présentes à la surface de la bactérie (153). Une fois que la bonne longueur est atteinte et au contact de la cellule hôte, les protéines formant le translocon sont assemblées et les effecteurs sont ensuite sécrétés. À cause de cette hiérarchie, plusieurs effecteurs sont entreposés dans le cytoplasme de la bactérie en attente de leur sécrétion. Ces effecteurs sont maintenus dans un état compétant avec la sécrétion par des chaperonnes spécifiques à ces effecteurs du *SST3* (154). La séquence signal des effecteurs du *SST3* se retrouve en position N-terminal et n'est pas clivée (155). Comme aucun motif consensus au niveau de la séquence primaire n'a été identifié, il a été suggéré que le signal se trouverait au niveau de l'ARN (156, 157).

1.3.3. Système de sécrétion de type IV

Le système de sécrétion de type IV (*SST4*) est un autre système permettant la sécrétion d'effecteurs directement dans la cellule cible. Par contre, le *SST4* est très versatile et peut aussi bien permettre la translocation d'ADN que de protéines. Les *SST4* sont divisés en trois groupes (158). Le premier groupe, et le plus répandu, est le système de conjugaison bactérien qui permet le transfert d'ADN d'une bactérie à une autre. Celui-ci est aussi retrouvé chez les bactéries gram-positives et les archéobactéries. Un deuxième groupe permet l'acquisition et la relâche d'ADN à partir du milieu extracellulaire. Ce groupe correspond au système d'import d'ADN (ComB) chez *Helicobacter pylori* et au système de relâche d'ADN (GGI) chez *Neisseria gonorrhoeae* (159, 160). Finalement, un dernier groupe permet la sécrétion de protéines. Les *SST4* correspondants au dernier groupe sont généralement retrouvés chez des bactéries pathogènes et jouent un rôle important dans la virulence. *Bordetella pertussis*, *Legionella pneumophila* et *Bartonella hensalae* sont des exemples de bactéries qui utilisent ce type de *SST4*. En général, les substrats sont transloqués à partir du cytoplasme de la bactérie directement dans le cytosol de la cellule cible. La sécrétion de la toxine pertussis est une exception puisqu'elle est sécrétée dans le milieu

extracellulaire à partir du périplasme par le *SST4* Ptl de *Bordetella pertussis* (161).

Le *SST4* forme un complexe multiprotéique, composé de 12 protéines (nommées VirB1-11 et VirD4 selon le système modèle d'*Agrobacterium tumefaciens*) et est composé de trois structures : un canal pour la translocation, un pilus extracellulaire et une *type IV coupling protein (T4CP)* (162). Les protéines VirB7-B9 et B10 forment le corps du complexe qui traverse les deux membranes et chacune de ces protéines est présente en 14 copies (163, 164). Ces protéines forment donc un imposant complexe composé de 2 anneaux, dans la membrane interne et dans la membrane externe, créant un canal allant du cytoplasme au milieu extracellulaire. Le pore dans la membrane externe est formé par VirB10 et est relativement petit, 10 Å, ce qui suggère qu'il doit y avoir un changement de conformation lors de la sécrétion des substrats et l'assemblage du pilus (163). Le pilus du *SST4* est formé par les pilines majeures et mineures, VirB2 et VirB5 (165, 166). VirB5 semble se retrouver au bout du pilus et servirait à l'adhésion à la cellule cible (167). Finalement, la *T4CP*, représentée par l'ATPase VirD4, est le récepteur pour les substrats du *SST4*.

Le signal reconnu par la *T4CP* semble être un groupe de résidus hydrophobes retrouvé en C-terminal des substrats (168, 169). Toutefois, dans plusieurs cas, des domaines internes semblent aussi être importants pour la sécrétion des substrats (170, 171). Comme pour le *SST3*, plusieurs effecteurs du *SST4* requièrent la présence de chaperonnes spécifiques dans le cytoplasme (172, 173). Une fois reconnus par VirD4, les substrats sont transférés au canal avec l'aide de quelques autres ATPases et il semble que les substrats traversent alors l'enveloppe bactérienne par le canal formé par le pilus. Dans certains cas, aucun pilus n'est assemblé, mais ceci corrèle généralement avec la fonction du *SST4*. C'est le cas par exemple du système Dot/Icm de *L. pneumophila* qui est fonctionnel à l'intérieur des cellules hôtes et n'a donc pas besoin d'une longue structure pour atteindre la cellule cible (174).

1.3.4. Système de sécrétion de type VI

Le système de sécrétion de type VI (*SST6*) est le plus récent système de sécrétion décrit (175, 176). Le *SST6* est retrouvé chez plusieurs bactéries à Gram négatif et autant chez des bactéries pathogènes comme *Vibrio cholerae* et *Pseudomonas aeruginosa* que chez des bactéries environnementales comme *Rhizobium leguminarosium* (177, 178). Le *SST6* permet lui aussi la translocation d'effecteurs directement dans la cellule cible. La particularité du *SST6* repose sur le fait que sa cellule cible peut aussi bien être une cellule eucaryote qu'une autre bactérie. En fait, la structure du *SST6* ressemble à la structure d'un bactériophage inversé (179). Cette ressemblance provient de deux composantes majeures du *SST6* qui se retrouve dans le milieu extracellulaire : VgrG et Hcp (179). Hcp oligomérisé et forme un cylindre à l'extérieur de la bactérie (180, 181). Pour sa part, VgrG forme un complexe trimérique qui ressemble à la queue d'un bactériophage T4 qui perce la membrane bactérienne (182). La sécrétion de Hcp et de VgrG par le *SST6* est interdépendante, ce qui suggère que ces deux protéines font partie de la machinerie du *SST6*. Le modèle propose que VgrG se situerait au sommet du cylindre formé par Hcp, constituant ainsi un appareil servant à percer la membrane de la cellule cible et, par la même occasion, le canal de sécrétion. Toutefois, l'interaction entre Hcp et VgrG n'a toujours pas été démontrée.

Outre Hcp et VgrG, le *SST6* est composé d'environ 15 autres protéines et forme un complexe qui traverse les deux membranes de la bactérie (177). Il est intéressant de noter que deux composants de la membrane interne du *SST6*, IcmF et DotU, sont aussi retrouvés dans le *SST4*, ce qui suggère que le *SST6* a évolué en acquérant des parties de différents systèmes (177, 183). Deux autres protéines, VipA et VipB forment des tubes contractiles dans le cytoplasme de la bactérie qui ressemblent à la gaine d'un bactériophage (184, 185). Le modèle suggère donc que la contraction de VipA/B pousserait le complexe Hcp/VgrG à travers la membrane de la cellule cible.

Les premiers effecteurs identifiés du *SST6* ont été les protéines VgrG. Dans plusieurs systèmes, une protéine VgrG possède un domaine effecteur C-terminal

additionnel qui, par exemple, permet la réorganisation de l'actine (182, 186). VgrG servirait donc à la fois comme composant structural et comme protéine effectrice. De « vraies » protéines effectrices ont ensuite été identifiées chez *Pseudomonas aeruginosa* (187). Les protéines Tse1-3 ont la particularité toutefois de cibler d'autres bactéries et sont retrouvées en association avec une protéine d'immunité qui protège la bactérie contre une auto-intoxication. Ces couples effecteurs/protéines d'immunité ont été utilisés pour découvrir des effecteurs similaires dans plusieurs autres types de bactéries (188). Le conduit formé par le trimère de VgrG serait toutefois trop étroit pour permettre le passage de protéines effectrices et il a donc été proposé que VgrG pourrait se dissocier de Hcp suite à la perforation de la membrane (189).

2. Autotransporteurs monomériques

Comme il a été mentionné précédemment, les autotransporteurs monomériques font partie du système de sécrétion de type V (type Va) (100). Le terme autotransporteur a été utilisé originalement puisque toute l'information pour la sécrétion de la protéine est contenue dans un même polypeptide (190). Un autotransporteur contient une séquence signale Sec-dépendante pour la translocation à travers la membrane interne et possède un tonneau β qui permet la translocation d'un domaine fonctionnel à travers la membrane externe. Ce système de sécrétion correspond au système le plus simple et le plus utilisé chez les bactéries à Gram négatif. Toutefois, ce mécanisme de sécrétion est un peu plus complexe que ce que son apparente simplicité le suggère. Le terme « autotransporteur » a même été contesté lorsqu'il a été déterminé que le tonneau β seul n'était pas suffisant pour la sécrétion du domaine passager (191). Toutefois, étant donné que seuls les systèmes généraux de la bactéries sont utilisés, le terme autotransporteurs est toujours approprié. La présente section décrit l'organisation, la biogenèse et la fonctionnalité des autotransporteurs monomériques.

2.1. Organisation

Les autotransporteurs monomériques sont des protéines modulaires qui ont tous une organisation similaire. Ils sont composés, dans l'ordre à partir du N-terminal : d'une séquence signal; d'un domaine fonctionnel, appelé domaine passager; d'une région de jonction; suivi d'un domaine β (Figure. 6A). Le domaine passager et la région de jonction constituent la partie extracellulaire de la protéine, tandis que le domaine β constitue le domaine transmembranaire.

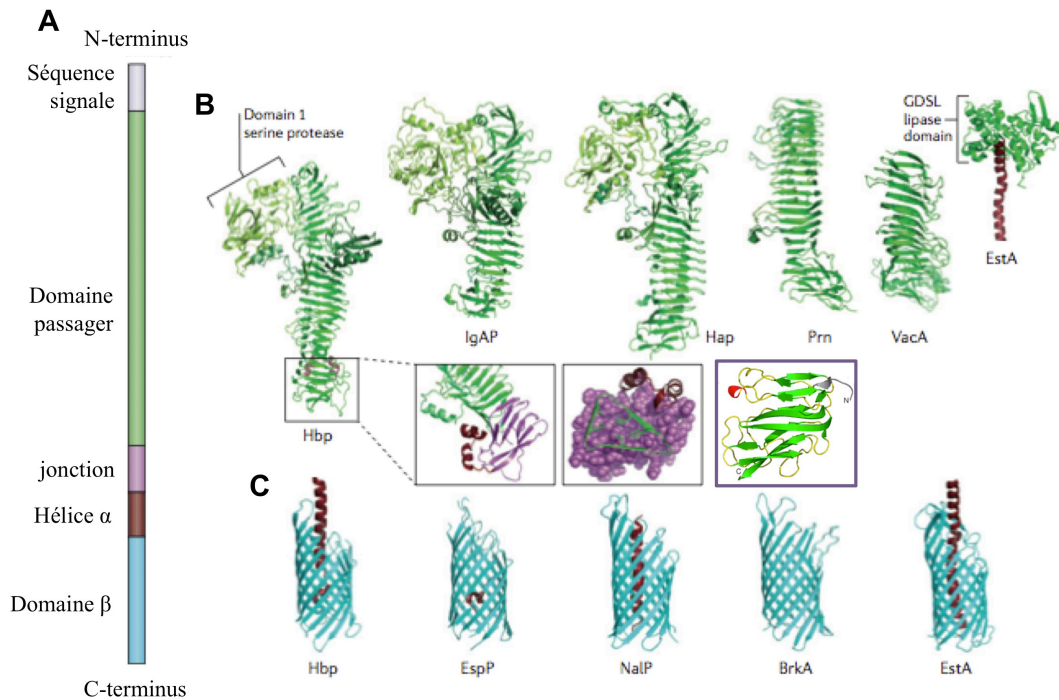


Figure 6 : Organisation et structure des autotransporteurs monomériques [figure adaptée des références (4) et (192)]. (A) Schéma représentant l'organisation des autotransporteurs monomériques comprenant, à partir du N-terminus, la séquence signal Sec-dépendante en gris, le domaine passager en vert, la région de jonction en mauve et le domaine β en rouge et bleu. (B) Structures des domaines passagers de Hbp, IgA protéase, Hap, pertactine (Prn), VacA et EstA. Les encadrés représentent les structures de la région de jonction de Hbp et de IcsA (encadré mauve). (C) Structures des domaines β (incluant l'hélice α) des autotransporteurs Hbp, EspP, NalP, BrkA et EstA.

2.1.1. Séquence signale

La plupart des autotransporteurs possèdent une séquence signale typique pour le système Sec, bien qu'elle soit plutôt variable, et utilise la voie de sécrétion post-traductionnelle. Environ 10% des autotransporteurs, ainsi que certaines protéines TpsA, possèdent une séquence signale allongée de 50-60 acides aminés (3). Cette longue séquence signale est composée de cinq régions, N1, H1, N2, H2 et C. La partie N1-H1 constitue une extension à la séquence signale normale (N2-H2-C). Cette extension est très bien conservée entre les autotransporteurs, ce qui suggère qu'elle pourrait avoir une fonction additionnelle (193). Toutefois, le rôle de cette extension est toujours controversé. Il a d'abord été observé que des autotransporteurs ayant une séquence signale longue, Hbp et EspP, étaient recrutés au système Sec via la voie *SRP* (194, 195). Cette idée est plutôt unique étant donné que les substrats de *SRP* sont pratiquement tous des protéines de membrane interne. Par la suite, des études faites avec les autotransporteurs IcsA, Pet et EspP, ainsi qu'avec la protéine FHA (TpsA), ont plutôt montré que, malgré la présence de cette extension de la séquence signale, ces autotransporteurs utilisaient bel et bien la voie post-traductionnelle (193, 196-198).

Par ailleurs, la délétion de cette extension n'a pas affecté la sécrétion ni le repliement du domaine fonctionnel de l'autotransporteur Pet à l'extérieur de la bactérie, ce qui suggère que l'extension de la séquence signale n'est pas requise pour la sécrétion, le repliement et la fonctionnalité des autotransporteurs (199). Une autre hypothèse suggère que cette extension de la séquence signale ralentisse la translocation à travers la membrane interne et le relargage dans le périplasme (193, 200, 201). Dans le cas de l'autotransporteur Hbp, la séquence signale, possiblement via son extension, interagit avec la protéine YidC ce qui retarde le relâchement de la protéine dans le périplasme et prévient ainsi le mauvais repliement de la protéine dans le périplasme (202). Il a aussi été montré que EspC, un autre autotransporteur faisant partie de la famille des *SPATEs* (voir plus loin) comme Hbp, avait aussi besoin de YidC (202). Toutefois, la seule étude faite avec un autotransporteur non-*SPATEs*, IcsA de *Shigella flexneri*, montre que YidC n'a pas d'influence sur sa biogenèse (198). Bref, il semble clair maintenant que cette

extension de la séquence signal soit associée à un ralentissement de la translocation à travers la membrane interne, toutefois le mécanisme et la raison derrière ce ralentissement restent toujours à déterminer, puisque seulement 10% des autotransporteurs possèdent cette extension de la séquence signal.

Finalement, certains autotransporteurs possèdent un motif typique de lipoprotéines. C'est le cas notamment des autotransporteurs NalP de *Neisseria meningitidis*, SphB1 de *Bordetella pertussis* et de CapA de *Campylobacter jejuni* (203-205). Après le passage dans le périplasme, ces protéines sont dirigées vers une acyltransférase qui attache une chaîne lipidique à la première cystéine de la protéine. Ces autotransporteurs sont ensuite envoyés vers la membrane externe par le système Lol (localisation of lipoproteins), le système qui dirige les lipoprotéines vers la membrane externe (206). Le rôle de la lipodation n'est pas clairement défini, mais elle est essentielle pour le fonctionnement de SphB1 (206).

2.1.2. Domaine passager

Le domaine passager est le domaine fonctionnel d'un autotransporteur et se retrouve dans le milieu extracellulaire. Il est généralement relié à la virulence (207). Il n'existe pas de consensus quant à la séquence, à la taille et à la fonction associée au domaine passager. Dans certains cas, le domaine passager peut être très petit (quelques acides aminés pour TynE de *Bordetella parapertussis*) ou bien il peut être énorme dans d'autres cas (~10 000 acides aminés pour BigE de *Polynucleobacter necessarius*) (208). Malgré cette diversité, la majorité des domaines passagers adoptent une structure similaire (5, 108). En effet, plus de 97% des domaines passagers adopteraient une structure formant une hélice β droite (209).

La première structure d'un domaine passager a été celle de la pertactine, une adhésine de *Bordetella pertussis* (210). Le domaine passager en entier adopte la structure d'une hélice β droite composé de 16 tours; chaque tour étant composé par 25 acides aminés qui forment trois brins β attachés ensemble par des boucles de longueurs

variables. Depuis, la structure de plusieurs autres domaines passagers a été déterminée. C'est le cas pour les autotransporteurs Hbp (211) et EspP (212) de *E. coli*, IgA1P (213) et Hap (214) de *Haemophilus influenzae*, et VacA de *Helicobacter pylori* (215) (Fig. 6B). Dans tous les cas, la partie centrale du domaine passager est constituée d'une hélice β droite. Toutefois, plusieurs de ces autotransporteurs sont des protéases et l'hélice β à elle seule ne peut pas expliquer l'activité protéolytique. Dans ces cas, un domaine protéase est rajouté à l'extrémité N-terminale du domaine passager et ressort de l'hélice β . Par exemple, le domaine passager de l'autotransporteur Hap adopte une structure ressemblant à une hache, avec l'hélice β formant le manche et le domaine protéase formant la tête (214). Les fonctions associées aux autotransporteurs sont très diversifiées et, souvent, un même autotransporteur peut être responsable de plusieurs fonctions différentes. Il semble donc que l'hélice β soit la base d'une structure qui peut être très versatile. Cette architecture, avec l'hélice β comme partie centrale, explique bien comment, malgré d'importantes contraintes mécanistiques de sécrétion (voir plus loin), les autotransporteurs peuvent être associés à autant de fonctions différentes. À cette base peut ainsi se greffer différentes boucles et différents domaines qui vont ressortir de l'hélice et contenir les résidus importants pour la fonctionnalité de la protéine. Ce ne sont cependant pas tous les autotransporteurs qui possèdent cette structure en hélice β . La récente structure de l'autotransporteur EstA de *Pseudomonas aeruginosa* montre un domaine passager globulaire formé majoritairement d'hélice α et représente l'exception (216).

2.1.3. Région de jonction

En position C-terminale du domaine passager, et toujours dans le milieu extracellulaire, on retrouve une région de jonction, aussi appelée domaine auto-chaperonne (217). Cette région de jonction est le prolongement de l'hélice β et sa structure est aussi très bien conservée malgré une séquence plutôt variable (5). La structure de la région de jonction seule de l'autotransporteur IcsA de *Shigella flexneri* a été déterminée et s'ajoute aux nombreuses structures des régions de jonction qui ont été cristallisées en même temps que les domaines passagers (192) (Fig. 6B). Comme le

domaine passager, la région de jonction est composée majoritairement de brins β . Toutefois, au lieu d'adopter une structure en hélice β , elle adopte plutôt un repliement de type β -sandwich, qui consiste en deux feuilletts β arrangés de façon anti-parallèle.

Selon les modèles de sécrétion, la région de jonction est la première région à être exposée dans le milieu extracellulaire (218). De plus, il semble que cette région soit plus résistante à la dégradation et à la dénaturation et donc plus stable que le domaine passager (209). Finalement, chez plusieurs autotransporteurs, AIDA-I, Pet, IcsA, BrkA entre autres, la région de jonction est importante pour le repliement du domaine passager dans le milieu extracellulaire (217, 219-221). Le modèle qui en découle est donc que la région de jonction, en étant la première région qui se retrouve dans le milieu extracellulaire et en étant plus stable, servirait de base pour le repliement du domaine passager. Il s'agit donc d'un modèle de repliement vectoriel, où le domaine passager se replie au fur et à mesure qu'il est sécrété (voir plus loin).

2.1.4. Domaine β

Finalement, le domaine β , retrouvé dans la région C-terminale d'un autotransporteur, adopte la structure d'un tonneau β dans la membrane externe. Plusieurs structures de domaine β sont disponibles et sont très similaires, voir même superposables, malgré le faible consensus dans la séquence primaire (216, 222-225) (Fig. 6C). La structure ressemble aux autres tonneaux β retrouvés dans la membrane externe et est formée par 12 brins β reliés entre eux par des boucles plus ou moins longues. L'extérieur du tonneau est formé par des résidus hydrophobes alors que l'intérieur est formé par des résidus plutôt hydrophiles. Le pore ainsi formé est relativement petit, d'une grosseur de 10 X 12,5 Å (222). À l'intérieur du tonneau β , on retrouve une hélice α qui est connectée au tonneau par une boucle périplasmique. Son rôle à l'intérieur du tonneau est de bloquer le pore afin de maintenir l'intégrité membranaire, mais elle sert aussi à stabiliser le tonneau β (223, 226). Sans hélice α , le domaine β n'est pas stable et le tonneau n'arrive pas à adopter sa bonne conformation.

2.2. Biogenèse

2.2.1. Sécrétion

Les autotransporteurs monomériques, comme toutes les autres protéines, sont synthétisés dans le cytoplasme. La première étape de la sécrétion d'un autotransporteur monomérique constitue donc le passage de la membrane interne. Comme mentionné précédemment, les autotransporteurs possèdent une séquence signal qui entraîne la protéine vers le système Sec de façon fort probablement post-traductionnelle. Les autotransporteurs sont ensuite transportés à la membrane externe de la même façon que les autres protéines de membrane externe, c'est-à-dire grâce aux chaperonnes périplasmiques et au complexe Bam. Finalement, le domaine β est inséré dans la membrane externe et le domaine passager est sécrété dans le milieu extracellulaire (Fig. 7).

2.2.1.1. Passage du périplasma

Tout d'abord, les autotransporteurs monomériques nécessitent la présence des chaperonnes périplasmiques et suivent donc les mêmes règles que les autres protéines de membrane externe. L'interaction avec les chaperonnes périplasmiques est nécessaire afin de conserver une conformation compatible avec l'insertion et le passage de la membrane externe. Ces interactions avec les chaperonnes périplasmiques se font directement à la sortie du translocon et, possiblement, même avant le clivage de la séquence signal (86).

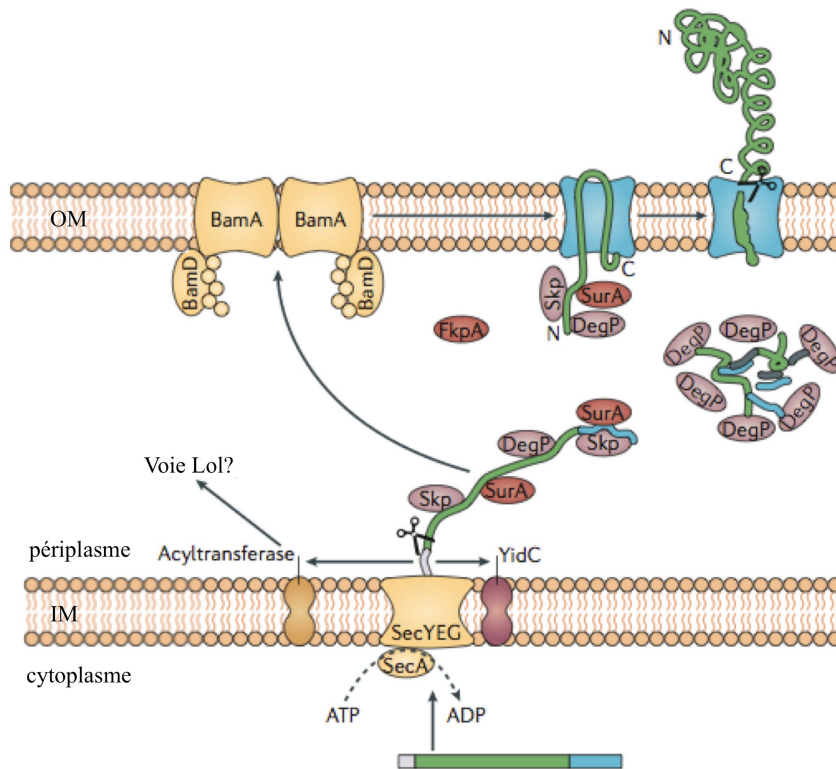


Figure 7 : Biogenèse des autotransporteurs monomériques [figure adaptée de la référence (4)]. Les autotransporteurs monomériques sont synthétisés dans le cytoplasme. Ils sont composés d'une séquence signal (grise), d'un domaine extracellulaire (vert; domaine passager + région de jonction) et d'un domaine transmembranaire (bleu; domaine β). Suite à la translocation à travers la membrane interne par le système Sec, les autotransporteurs sont pris en charge par les chaperonnes périplasmiques. Certains autotransporteurs possèdent un motif de lipoprotéines et sont probablement dirigés vers la voie Lol tandis que certains autotransporteurs semblent interagir avec YidC. Ils sont ensuite insérés dans la membrane externe avec l'aide le complexe Bam et le repliement du domaine passager se fait dans le milieu extracellulaire. Finalement, dans plusieurs cas, le domaine extracellulaire est clivé et relâché dans le milieu.

Il a d'abord été démontré que l'inactivation de ces chaperonnes chez *Shigella flexneri* avait un impact important sur la biogenèse de l'autotransporteur IcsA (227). Une interaction directe entre les autotransporteurs et les trois chaperonnes périplasmiques principales, soit SurA, Skp et DegP, a ensuite été démontrée pour les autotransporteurs EspP et Hbp (228-230). Il est intéressant de noter qu'autant le domaine passager que le domaine β bénéficie de l'aide des chaperonnes périplasmiques. Selon les études, le

tonneau β interagît avec Skp seul ou bien avec Skp et SurA, transitoirement et séquentiellement, tandis que le domaine passager peut interagir avec SurA, Skp et DegP (231). Les résultats sont variables selon les approches utilisées, mais reflète la redondance entre les différentes chaperonnes et la nature transitoire des interactions. D'ailleurs, l'interaction du domaine passager d'EspP avec les chaperonnes périplasmiques est séquentielle, c'est-à-dire qu'il interagit d'abord avec Skp et ensuite avec SurA. Le modèle suggéré, pour EspP du moins, est donc que Skp est la chaperonne principale et celle-ci interagirait avec le domaine β et le domaine passager directement à la sortie du translocon Sec. La protéine est ensuite transférée à SurA qui entraîne la protéine vers le complexe *BAM*. Ces interactions avec les chaperonnes périplasmiques font en sorte que le domaine β et le domaine passager demeurent dans un état déplié compatible avec le passage dans le périplasme et la sécrétion dans le milieu extracellulaire.

Il est à noter toutefois que les chaperonnes cytoplasmiques n'aident pas au repliement et à la localisation du domaine β (227, 232). Dans le cas de IcsA, l'inactivation des chaperonnes périplasmique n'affecte pas l'insertion du domaine β dans la membrane externe, mais affecte plutôt la sécrétion du domaine passager à travers la membrane externe (227). Une observation similaire a été faite avec EspP : bien que l'inactivation des chaperonnes périplasmiques entraînent une diminution de la sécrétion du domaine passager dans le milieu extracellulaire, l'insertion du domaine β dans la membrane externe n'est pas affectée (228). Il semble donc que le rôle des chaperonnes cytoplasmiques soit de conserver le domaine passager dans un état déplié compatible avec la translocation à travers la membrane externe en plus de protéger la protéine contre la dégradation par des protéases périplasmiques. Il est à noter que ces résultats ont été obtenus chez *E. coli* et il est possible que ce soit différent chez d'autres bactéries. Par exemple, pour l'autotransporteur IgAP chez *Neisseria meningitidis*, les chaperonnes périplasmiques n'ont pas la même importance et leur inactivation n'affecte pas l'insertion du tonneau β dans la membrane externe ni la sécrétion du domaine passager dans le milieu extracellulaire (233).

2.2.1.2. Insertion dans la membrane externe

Suite au passage dans le périplasma, le domaine β est inséré dans la membrane externe. Jusqu'à la découverte relativement récente du complexe *BAM*, il avait été proposé que le domaine β d'un autotransporteur, comme les autres protéines de membrane externe, se repliait et s'insérait spontanément dans la membrane. Suite à la découverte de *BAM*, il a été démontré que celui-ci était requis pour la biogenèse de plusieurs autotransporteurs, notamment IcsA, AIDA-I, Pet, Hbp, BrkA et IgAP (77, 234, 235). Des interactions directes avec le complexe *BAM* ont d'ailleurs été démontrées pour les autotransporteurs EspP et Hbp (228-230). Comme SurA est la seule chaperonne pour laquelle l'interaction avec BamA a été démontrée, il a été proposé que SurA transfère la protéine dans un état déplié à BamA qui va alors se charger de son repliement et son insertion dans la membrane. D'ailleurs, comme il a été mentionné précédemment, Skp semble interagir initialement avec les autotransporteurs à la sortie du translocon Sec et les transfère ensuite à SurA qui va ensuite les transporter jusqu'à BamA.

Encore une fois, comme pour les chaperonnes périplasmiques, l'interaction de EspP avec les composantes du complexe *BAM* est transitoire et séquentielle. Il semble que le domaine β soit transféré de SurA à BamA, via les domaines *POTRA*, avant d'être ensuite transféré aux autres composantes de *BAM* qui pourraient jouer un rôle dans les étapes tardives d'insertion dans la membrane (231). BamD semble être le dernier à interagir avec EspP et pourrait donc être le facteur responsable de l'insertion dans la membrane (231). Finalement, seul BamA et BamD sont requis pour l'insertion du domaine β d'un autotransporteur dans la membrane externe (235). Encore une fois, les autotransporteurs suivent donc les mêmes règles que les autres protéines de membrane externe et utilisent le complexe *BAM* pour leur insertion dans la membrane externe. D'ailleurs, tous les autotransporteurs contiennent le motif C-terminal de reconnaissance pour le système *BAM*.

2.2.1.3. Sécrétion du domaine passager

La dernière étape constitue la translocation du domaine passager dans le milieu extracellulaire. Cette étape suscite beaucoup d'intérêt dans la communauté scientifique et, depuis la découverte de ce mode de sécrétion, plusieurs modèles se sont succédés pour expliquer la sécrétion du domaine passager. Bien que l'arrangement « tonneau β -domaine passager » est à priori très simple, le mécanisme par lequel le domaine passager traverse la membrane externe par le tonneau β est plus complexe que ce qui était imaginé au début. Une première complication est qu'il n'y a pas d'énergie de disponible à la membrane externe. De plus, le tonneau β est très petit et est bloqué par une hélice α , mais semble néanmoins être responsable de la translocation du domaine passager. La majorité des modèles ont donc été contredits par les études subséquentes et deux modèles principaux subsistent aujourd'hui : le modèle « hairpin » et le modèle BamA (4) (Fig. 8).

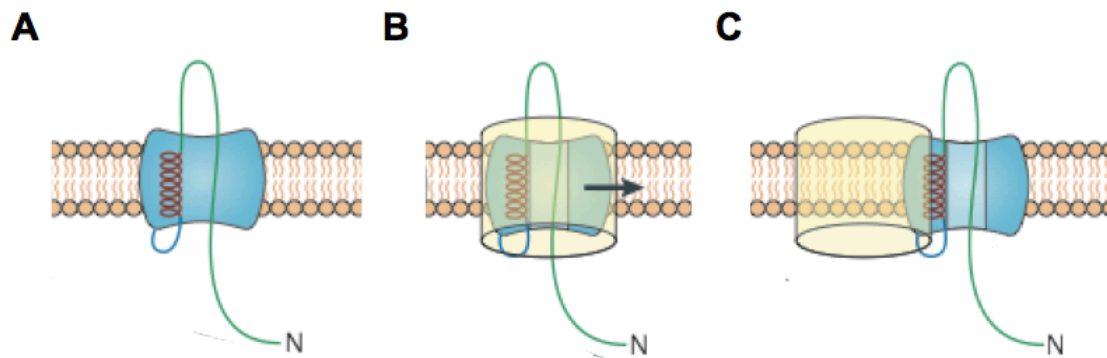


Figure 8 : Modèles pour la translocation du domaine passager dans le milieu extracellulaire [Figure adaptée de la référence (4)] (A) Dans le modèle « hairpin », la partie C-terminale du domaine passager est sécrétée en premier à travers le pore du tonneau β . (B) Dans le modèle BamA, BamA (en jaune) participe aussi à la sécrétion du domaine passager. (C) Un modèle alternatif rejoint le modèle « hairpin » et le modèle BamA où le tonneau β est maintenu dans une conformation étendue pour permettre la sécrétion de domaine passager.

Dans le modèle « hairpin », il y a formation d'une boucle dans la partie C-terminale du domaine passager à la sortie du tonneau β . Ce modèle implique d'abord que la partie C-terminale du domaine passager, qui correspond à la région de jonction, se retrouve en premier à l'extérieur. Plusieurs études ont montré, en bloquant la sécrétion complète du domaine passager, que la région de jonction se retrouve effectivement à l'extérieur en premier (209, 218, 236, 237). Ce modèle implique aussi que deux brins du polypeptide se retrouvent dans le tonneau β en même temps et ce malgré la présence de l'hélice α . Le problème avec ceci est que le tonneau β est trop petit pour accommoder deux brins. Ensuite, il est très peu probable que l'hélice α se retire du tonneau β dans le périplasma pendant la sécrétion du domaine passager. En effet, l'hélice α est très importante pour la stabilité du tonneau β et il semble qu'elle soit déjà à l'intérieur du tonneau avant même son repliement complet (238-240). De plus, l'hélice α occupe toute l'espace disponible à l'intérieur du pore. Alternativement, il est possible que les deux brins, incluant l'hélice α , soient dans un état déplié pendant la sécrétion. Toutefois, lorsque des boucles de petites tailles sont introduites dans le domaine passager par l'ajout de cystéines, de petites structures sont compatibles avec la sécrétion, ce qui suggère que de petites structures secondaires sont tolérées et contraste avec ce modèle (241-243). Il a été proposé que le tonneau β pouvait légèrement s'étirer afin d'agrandir le pore pour permettre la sécrétion (244). Un même type d'étirement a d'ailleurs été observé pour la protéine de membrane externe FimD, le canal permettant la sécrétion et l'assemblage des fimbriae de type 1 (245). La structure de FimD avec un substrat est différente de celle sans substrat et montre un tonneau légèrement étiré pour accommoder son substrat. Cela suggère donc qu'il est possible que le domaine β , grâce à une déformation, puisse permettre la sécrétion du domaine passager malgré la présence de l'hélice α à l'intérieur du tonneau.

Dans le modèle BamA, le complexe *BAM* ne sert pas seulement à l'insertion du domaine β dans la membrane externe, mais participe également à la sécrétion du domaine passager. La façon à laquelle BamA participe à la sécrétion du domaine passager n'est toutefois pas clairement établie. Il est possible que le tonneau de BamA serve de pore pour la translocation du domaine passager. Ceci impliquerait toutefois que

le tonneau de BamA et celui de l'autotransporteur doivent s'ouvrir latéralement pour effectuer le transfert du domaine passager, ce qui serait physiquement très défavorable. Ceci impliquerait aussi que le domaine β de l'autotransporteur ne soit pas impliqué dans la sécrétion du domaine passager et serve uniquement de signal de localisation dans la membrane externe. Or, il a été montré récemment que la substitution d'un domaine β par un autre affecte la sécrétion du domaine passager, suggérant ainsi que le domaine β joue un rôle actif dans la sécrétion du domaine passager (241). Alternativement, il a aussi été suggéré que le domaine β s'insère dans le tonneau de BamA. Toutefois, il est peu probable que le tonneau de BamA soit suffisamment grand pour accommoder le domaine β d'un autotransporteur. Finalement, la dernière hypothèse, qui rejoint le modèle « hairpin » avec le modèle BamA, est que l'interaction côte-à-côte entre le domaine β et BamA permet au domaine β d'adopter une conformation partielle et étendue compatible avec la sécrétion du domaine passager.

Une autre question qui suscite beaucoup d'intérêt est de déterminer d'où vient l'énergie nécessaire à la sécrétion du domaine passager dans le milieu extracellulaire. Comme il a été mentionné plus tôt, au niveau de la membrane externe les sources d'énergie normalement utilisées, c'est-à-dire l'ATP et la *FPM*, ne sont pas disponibles. L'hypothèse principale est que c'est le repliement du domaine passager dans le milieu extracellulaire qui fournit l'énergie nécessaire à la sécrétion (209, 218, 236). Comme il a déjà été mentionné, le repliement du domaine passager se fait de façon vectorielle à partir de la région de jonction qui est plus stable. Donc, à mesure que la région de jonction est sécrétée, elle se replie, entraîne ainsi la sécrétion du reste du domaine extracellulaire (246). Plusieurs études ont d'ailleurs montré qu'en bloquant la sécrétion de la partie N-terminale du domaine passager, la région de jonction ainsi que la partie C-terminale du domaine passager se retrouvaient dans le milieu extracellulaire en premier et dans une conformation bien repliée (218, 237). La plus grande stabilité de la région de jonction est très importante pour ce processus. En fait, c'est la grande stabilité de la région de jonction couplée à la faible stabilité du domaine passager qui permet la sécrétion. Remarquablement, la déstabilisation de la région C-terminale ou bien la stabilisation de la région N-terminale abolissent la sécrétion du domaine passager dans

le milieu extracellulaire (247). La protéine contient donc en elle-même l'énergie nécessaire à sa propre sécrétion, sous la forme d'une distribution anisotrope de la stabilité (248).

Finalement, il a tout récemment été proposé qu'un nouveau complexe serait essentiel à la sécrétion des autotransporteurs (249). Le complexe a été nommé *TAM* (translocation and assembly module) et est retrouvé dans toutes les classes de protéobactéries. Il est formé par les protéines TamA, une protéine de membrane externe de la superfamille BamA, et TamB, une protéine de membrane interne qui s'associent ensemble pour former un complexe traversant les deux membranes. Selon cette étude, ce complexe serait essentiel à la sécrétion des autotransporteurs puisque la délétion de TamA ou de TamB entraîne une perte de la sécrétion de plusieurs autotransporteurs chez *E. coli* et chez *Citrobacter rodentium*. Les auteurs de l'étude suggèrent que TamA aurait un rôle similaire aux protéines TpsB des systèmes à deux partenaires et servirait donc à la sécrétion du domaine passager. Ce complexe *TAM* ne semble toutefois pas faire l'unanimité dans la communauté scientifique, et d'autres études sont requises pour savoir si celui-ci participe vraiment à la sécrétion des autotransporteurs.

2.2.2. Clivage

Beaucoup d'autotransporteurs monomériques sont clivés lorsque le processus de sécrétion est complètement terminé. Plusieurs domaines passagers sont des toxines, comme la toxine VacA de *Helicobacter pylori* (250) ou bien l'entérotoxine Pet de *E. coli* (251), et doivent être relâchés dans le milieu extracellulaire pour exercer leur fonction. Par contre, dans certains cas, AIDA-I par exemple (252), le domaine passager reste associé de façon non covalente à la surface bactérienne. Le rôle du clivage dans ces cas-là reste inconnu. Les mécanismes de clivage sont très nombreux et très variés. Dans certains cas, le clivage est effectué par d'autres protéines présentes dans la membrane externe. C'est le cas par exemple pour l'autotransporteur IcsA chez *Shigella flexneri* qui est clivé par la protéase IcsP (253) et pour les autotransporteurs App, AusI, MspA et IgAP qui sont clivés par la protéine NalP, un autre autotransporteur, chez *Neisseria*

meningitidis (203, 254, 255). Dans la majorité des cas, par contre, les autotransporteurs sont clivés par des mécanismes autocatalytiques. Pour Hap de *Haemophilus influenzae*, le clivage est autocatalytique et intermoléculaire, ce qui veut dire qu'il est effectué par une deuxième molécule de Hap (256). Pour les *SPATEs*, comme EspP ou bien EspC, le clivage est intramoléculaire et implique une paire de résidus Asp et Asn conservés dans le tonneau qui clive l'hélice α (223, 257). AIDA-I de son côté est clivé par une paire de résidus Asp/Glu situés dans une boucle juste en dessous du site de clivage situé dans la région de jonction (258). Finalement, dans plusieurs cas, comme Ag43 (259) ou bien VacA (260), le mécanisme de clivage est encore inconnu.

2.2.3. Glycosylation

Quelques autotransporteurs monomériques sont modifiés par glycosylation. C'est le cas des autotransporteurs AIDA-I (261), TibA (262), Ag43 (263) et EhaJ (264), qui sont tous retrouvés chez *E. coli*. La glycosylation de protéines est une des modifications les plus courantes chez les eucaryotes. Chez les bactéries, la glycosylation de protéines a été découverte relativement récemment, compte tenu du fait que l'importance des polysaccharides pour les bactéries est connue depuis longtemps, surtout en ce qui concerne les polysaccharides de surface comme les lipopolysaccharides (LPS) ou la capsule (265, 266).

L'ajout d'un sucre peut avoir différents effets sur une protéine (267, 268). D'abord, l'ajout de sucres se fait en général sur des résidus aspartate (N-glycosylation) ou bien sérine et thréonine (O-glycosylation) et modifie donc les propriétés physico-chimiques de la protéine, notamment la solubilité et la charge nette. La glycosylation peut contribuer, entre autres, au bon repliement d'une protéine, à une augmentation de sa stabilité ou à protéger contre la dégradation. D'ailleurs, chez les archéobactéries, la glycosylation de protéines est très fréquente et il a été suggéré que la glycosylation intensive de protéines serait primordiale à la survie en conditions extrêmes (269). De plus, la glycosylation de protéines est souvent associée aux interactions protéines-

protéines. D'ailleurs, la majorité des protéines glycosylées chez les bactéries sont retrouvées à la surface cellulaire (270).

Il existe plusieurs voies de glycosylation chez les bactéries (270, 271). Les bactéries possèdent aussi bien des systèmes généraux de glycosylation (N- ou O-glycosylation) que des systèmes spécifiques. *Campylobacter jejuni* est l'organisme modèle pour l'étude de la glycosylation de protéines chez les bactéries (272). *C. jejuni* possède un système général de N-glycosylation et un système de O-glycosylation de la flagelline. Dans les deux cas, un locus complet de plusieurs gènes est responsable de la glycosylation. Par contre, le système de O-glycosylation sert à la glycosylation d'un seul substrat, la flagelline, alors que le système général de N-glycosylation permet la glycosylation de plus de 60 protéines différentes (273). Plusieurs autres bactéries possèdent des systèmes complexes de glycosylation comme ceux de *C. jejuni*. Le système général de N-glycosylation est retrouvé chez plusieurs autres protéobactéries (271), tandis qu'il existe des systèmes de O-glycosylation de la flagelline ou des pilines dans un grand nombre de bactéries comme *Helicobacter pylori*, *Pseudomonas aeruginosa*, *Neisseria meningitidis* ou bien *Listeria monocytogenes* (271, 274). Alternativement, il existe des systèmes de glycosylation beaucoup plus simple, comme la glycosylation du TpsA HMW1 par la glycosyltransférase HMW1c chez *H. influenzae* (275). Dans ce cas précis, la glycosylation d'un seul substrat repose sur une seule glycosyltransférase spécifique qui utilise un sucre simple récupéré de la voie de biosynthèse d'autres oligosaccharides. HMW1c ajoute des résidus glucoses ou galactoses à HMW1 qui proviennent de la voie de biosynthèse des lipooligosaccharides. Surprenamment, HMW1c est une enzyme cytoplasmique, ce qui suggère que la glycosylation de HMW1 se fait dans le cytoplasme avant sa sécrétion, ce qui est plutôt rare.

La glycosylation des autotransporteurs AIDA-I (261), TibA (262), Ag43 (263) et EhaJ (264) se fait aussi via un système simple. Dans les cas de AIDA-I, TibA et EhaJ, ces trois autotransporteurs sont retrouvés sous la forme d'un opéron de deux gènes : *aah-aidA*, *tibC-tibA* et *ehaJ-egtA*. Aah, TibC et EgtA représente les glycosyltransférases

spécifiques de AIDA-I, TibA et Ag43 et leur inactivation supprime la glycosylation des trois autotransporteurs. Ces trois glycosyltransférases présentent de fortes homologies suggérant qu'elles fonctionnent de façon similaire (276). D'ailleurs, Aah et TibC sont fonctionnellement interchangeables et Aah peut aussi glycosylé l'Ag43 (277). Dans le cas d'Ag43, aucune glycosyltransférase spécifique n'a été identifiée. Toutefois, la forme glycosylée de Ag43 a été identifiée dans quelques souches sauvages d'*E. coli* (263, 278).

Il est connu depuis un certain temps que AIDA-I est modifié par des résidus heptoses provenant de la voie de biosynthèse des LPS, via le précurseur UDP-glycéro-manno-heptopyranose (261). Il s'agit fort probablement de O-glycosylation étant donné le grand nombre de résidus sérines présents dans la protéine. L'interchangeabilité entre Aah et TibC suggère que TibA est aussi modifié par des résidus heptoses. Étant donné que le précurseur de la glycosylation est retrouvé dans le cytoplasme, il a été suggéré que la glycosylation s'effectuait dans le cytoplasme, ce qui a d'ailleurs été confirmé plus tard (279). La glycosylation dans le cytoplasme est plutôt rare chez les bactéries, étant donné que l'ajout de sucres complexes est incompatible avec la sécrétion à travers la membrane interne par le système Sec. Toutefois, l'ajout d'un sucre simple, comme un heptose, peut être toléré par le système Sec.

Des études ont démontré qu'une molécule AIDA-I pouvaient contenir jusqu'à 19 heptoses et 16 sites de glycosylation ont été identifiés (261, 279). Les sites de glycosylation sont majoritairement des sérines, bien qu'une thréonine ait aussi été identifiée, confirmant qu'il s'agit bien de O-glycosylation. Plusieurs sites de glycosylation ont également été identifiés dans Ag43, lorsque ce dernier est co-exprimé avec Aah (280). Encore une fois, il s'agit surtout de sérines et de quelques thréonines. Dans tous les cas, les sites de glycosylation se retrouvent dans le domaine passager, dans une région constituée uniquement de répétitions d'une séquence consensus de 19 acides aminés. Il est à noter que la glycosylation par Aah est hétérogène, ce qui veut dire que les molécules d'AIDA-I ne sont pas toutes glycosylées de la même façon (279).

Les auteurs de l'étude de la glycosylation d'Ag43 par Aah ont essayé de trouver un motif de glycosylation, mais ni sont pas parvenus (280). En fait, les auteurs ont

retrouvé un motif, mais celui-ci correspondait à la séquence consensus de 19 acides aminés. Il est à noter que, pour tous les systèmes de O-glycosylation identifiés dans la littérature, aucun motif n'a encore été identifié (281). Cette séquence consensus répétée aurait donc pu constituer un motif permettant la glycosylation de AIDA-I par Aah. Toutefois, récemment, il a été démontré que Aah reconnaissait un élément structural plutôt qu'un motif dans la séquence (276). En fait, chaque répétition de la séquence consensus de 19 acides aminés forme un tour de l'hélice β du domaine passager composé de trois brins β interconnectés par des boucles (108). Il a alors été montré que Aah pouvait glycosyler une sérine retrouvée dans une boucle flanquée de deux brins β et ce indépendamment de la séquence des deux brins β . La question d'actualité est désormais de déterminer comment Aah réussit à reconnaître spécifiquement AIDA-I, au détriment d'autres protéines présentes dans le cytoplasme ayant potentiellement ce même motif structural. Or, d'après des résultats récents du laboratoire, il semble que la glycosylation de AIDA-I se fasse de façon co-translationnelle (Charbonneau ME et al., en préparation).

Le rôle de la glycosylation pour ces autotransporteurs demeure toujours incertain. Dans le cas d'AIDA-I et de TibA, la glycosylation est nécessaire pour le phénotype d'adhésion, mais est non essentielle pour l'auto-agrégation et la formation de biofilm (262, 279, 282). Pour EhaJ, la glycosylation augmente la capacité de formation de biofilm sans toutefois affecter la capacité d'adhérer à la matrice extracellulaire (264). Finalement, pour Ag43, les conclusions sont plus compliquées étant donné qu'il existe différents allèles d'Ag43 qui sont associés à différentes fonctionnalités et que l'effet de la glycosylation est différent selon les allèles (263). De plus, les études portant sur la glycosylation de Ag43 ont été effectuées en co-exprimant Ag43 avec Aah, donc dans un contexte artificiel. Il semble toutefois que l'effet de la glycosylation sur Ag43 ne modifie pas ses capacités d'auto-agrégation et de formation de biofilm. En revanche, pour les allèles permettant l'adhésion, la glycosylation modifie le comportement de Ag43, en augmentant sa capacité d'adhésion à la matrice extracellulaire ou bien aux cellules selon les allèles (263) (278).

L'effet de la glycosylation sur la fonctionnalité de ces autotransporteurs pourrait toutefois être indirect. La perte de la glycosylation d'AIDA-I entraîne une diminution importante de la quantité de protéines retrouvées à la surface de la bactérie (279). En fait, la forme non glycosylée d'AIDA-I est plus sensible à la dégradation par des protéases et induit un stress extracytoplasmique important. Cette étude suggère donc que la glycosylation influence la stabilité d'AIDA-I et que l'effet observé sur l'adhésion résulte de la mauvaise expression d'AIDA-I non glycosylée à la surface de la bactérie. Un effet de la glycosylation sur la stabilité d'Ag43 a aussi été observé (280). Dans ce cas-ci, la forme non glycosylée de Ag43 est bien exprimée, mais est plus sensible à la dénaturation chimique ou thermique que la forme glycosylée. De plus, le repliement de la forme glycosylée se fait plus rapidement que pour la forme non glycosylée. Par contre, il est à noter que cette étude utilise l'Ag43 glycosylé par Aah, la glycosyltransférase de AIDA-I, et utilise seulement le domaine passager de Ag43. Donc, bien qu'il soit maintenant clair que la glycosylation a un effet sur la conformation d'AIDA-I et Ag43, le rôle de celle-ci sur la fonctionnalité de la protéine est toujours incertain. Il est possible que la glycosylation affecte autant la conformation que la fonctionnalité de la protéine. C'est le cas par exemple pour l'autotransporteur trimérique EmaA de *Aggregatibacter actinomycetemcomitans* (283).

2.2.4. Autres particularités

Quelques autotransporteurs présentent des particularités lors de leur biogénèse qui leur sont propres et qui sont essentielles pour leur fonctionnement. C'est le cas par exemple pour IcsA de *Shigella flexneri*. IcsA est la protéine qui permet la polymérisation de l'actine à un pôle de la bactérie et donc qui permet la mobilité intracellulaire de la bactérie (284). Pour exercer sa fonction, IcsA doit donc se localiser à un seul pôle de la bactérie (285). Le mécanisme de la localisation polaire n'est pas encore clairement établi, mais il semble qu'elle se produise dans le cytoplasme de la bactérie indépendamment de la sécrétion de la protéine (198). De plus, IcsA conserve sa localisation polaire lorsqu'exprimé chez *E. coli*, *Salmonella Typhimurium*, *Yersinia pseudotuberculosis* et *Vibrio cholerae*, ce qui suggère que le facteur responsable de la

localisation est conservé chez les bactéries à Gram négatif (286). Il a été suggéré que la chaperonne cytoplasmique DnaK et le composant de la machinerie de division cellulaire FtsQ pourraient être impliqués dans la localisation de IcsA (287, 288). Toutefois, la façon par laquelle ces protéines participent à la localisation de IcsA demeure inconnue.

Une localisation polaire a aussi été montrée pour d'autres autotransporteurs, comme SepA de *S. flexneri*, AIDA-I de *E. coli* et BrkA de *B. pertussis* (289). Par contre, contrairement à IcsA, la raison de cette localisation polaire est beaucoup moins évidente pour ces autotransporteurs. Par exemple, AIDA-I permet l'adhésion des bactéries aux cellules épithéliales ainsi que l'auto-agrégation des bactéries (290, 291). Dans ce cas-là, l'avantage d'une localisation polaire est plutôt difficile à prédire et une confirmation est nécessaire avant d'affirmer qu'AIDA-I possède réellement une localisation particulière.

2.3. Fonctions

Une très grande diversité de fonctions a été attribuée aux autotransporteurs monomériques (100). Le tableau 1 dresse une liste non exhaustive des différentes familles d'autotransporteurs et des différentes fonctions associées aux autotransporteurs. Par exemple, IcsA permet la polymérisation de l'actine et la motilité intracellulaire de *Shigella flexneri* (284); BrkA rend *Bordetella pertussis* résistant au sérum (292); VacA de *Helicobacter pylori* entraîne la formation de vacuoles à l'intérieur des cellules cibles (250); IgA1P de *Neisseria meningitidis* aide la colonisation en clivant les IgA mucoales (293). Il existe toutefois quelques sous-familles qui regroupent plusieurs autotransporteurs ayant des fonctions similaires. Une première famille se nomme les SPATEs (Serine Protease AutoTransporters of *Enterobacteriaceae*) et regroupe plusieurs autotransporteurs possédant une activité sérine protéase comme fonctionnalité principale. Cette sous-famille est composée entre autres des autotransporteurs Hbp, EspC, EspP, Vat, Sat, Pic, Pet, Tsh, EatA et EpeA retrouvés chez différentes souches de *E. coli* intestinales (EPEC, ETEC, EHEC) ou extra-intestinales (APEC, UPEC) (294). Bien que chaque membre de cette famille possède un domaine sérine protéase similaire, celui-ci est utilisé de diverses façons par les différents membres de la famille (295, 296).

Tableau I : Exemples de fonctionnalité des autotransporteurs

Famille	Autotransporteur	Organisme	Fonctions	F
SPATEs	EspP	<i>Escherichia coli</i>	Dégradation du complément, biofilm	(
	EpeA	<i>E. coli</i>	Activité mucinase	(
	Pet	<i>E. coli</i>	Inflammation, activité cytotoxique	(
	Pic	<i>E. coli</i>	Activité mucinase, résistance au sérum	(
	EspC	<i>E. coli</i>	Activité cytotoxique	(
	Hbp	<i>E. coli</i>	Hémagglutinine, adhésion	(
	Sat	<i>E. coli</i>	Activité cytotoxique, autophagie	(
	Vat	<i>E. coli</i>	Formation de vacuoles	(
	EatA	<i>E. coli</i>	Activité mucinase	(
	SPATEs-like	SepA	<i>Shigella flexneri</i>	Inflammation
IgA1P		<i>Neisseria meningitidis</i>	Dégradation des IgA1	(
Hap		<i>Haemophilus influenzae</i>	Adhésion, AAG ¹ , biofilm	(
Protéases non-SPATEs	NalP	<i>Neisseria meningitidis</i>	Maturation d'autotransporteurs	(
	MapA	<i>Moraxella catarrhalis</i>	Phosphatase acide	(
	NanB	<i>Pseudomonas haemolytica</i>	Activité sialidase	(
	SphB1	<i>Bordetella pertussis</i>	Maturation de l'adhésine FHA	(
SAATs	AIDA-I	<i>E. coli</i>	Adhésion, invasion, AAG, biofilm	(
	TibA	<i>E. coli</i>	Adhésion, invasion, AAG, biofilm	(
	Ag43	<i>E. coli</i>	Adhésion, invasion, AAG, biofilm	(
	Cah	<i>E. coli</i>	AAG	(
	YejO	<i>E. coli</i>	Inconnu	(
	YpjA (EhaD)	<i>E. coli</i>	Biofilm	(
Adhésines	Pertactine	<i>Bordetella pertussis</i>	Adhésion / Résistance aux neutrophiles	(
	BapC	<i>Bordetella pertusis</i>	Adhésion	(
	EhaA	<i>E. coli</i>	Adhésion, AAG, biofilm	(
	AlpA	<i>Helicobacter pylori</i>	Adhésion	(
	AutA	<i>Neisseria meningitidis</i>	Adhésion	(
	MisL	<i>Salmonella enterica</i>	Adhésion / invasion	(
	ShdA	<i>Salmonella enterica</i>	Adhésion / invasion	(
	PmpD	<i>Chlamidia trachomatis</i>	Adhésion / invasion	(
	App	<i>Neisseria meningitidis</i>	Adhésion	(
	CapA	<i>Campylobacter jejuni</i>	Adhésion	(
	BabA	<i>Helicobacter pylori</i>	Adhésion	(
	YapH	<i>Yersinia enterocolitica</i>	Adhésion, AAG, biofilm	(
	rOmpA/B	<i>Rickettsia rickettsii</i>	Adhésion / invasion	(
	Autres	BrkA	<i>Bordetella pertussis</i>	Résistance au sérum, adhésion
IcsA		<i>Shigella flexneri</i>	Polymérisation de l'actine/ motilité intracellulaire	(
VacA		<i>Helicobacter pylori</i>	Formation de vacuoles	(
McaP		<i>Moraxella catarrhalis</i>	Hydrolyse de lipides, adhésion	(
YfaL (EhaC)		<i>E. coli</i>	Biofilm	(
EstA		<i>Pseudomonas aeruginosa</i>	Production de rhamnolipides	(

¹ AAG = auto-agrégation

Par exemple, l'autotransporteur Pet a une activité cytotoxique et doit être internalisé afin de rejoindre son substrat, la fodrine (une protéine qui est responsable de la polymérisation de l'actine), ce qui provoque un réarrangement de l'actine. D'un autre côté, Hbp (Hemoglobin binding protein) a une très bonne affinité pour l'hémoglobine et semble donc participer à l'acquisition du fer.

Un deuxième groupe correspond aux SAATs (Self-Associating AutoTransporters) de *E. coli* (6). Selon la classification initiale, cette famille est composée des autotransporteurs AIDA-I, TibA et Ag43. Ces trois autotransporteurs sont des protéines multifonctionnelles capables de médier l'adhésion et l'invasion de cellules épithéliales, l'auto-agrégation des bactéries ainsi que la formation de biofilm. La multifonctionnalité observée pour les SAATs semble être un trait commun aux autotransporteurs puisque plusieurs sont associés à plus d'une fonction (304). De plus, la séparation entre les différentes sous-familles n'est pas toujours aussi bien définie. Dans le cas de Hap de *Haemophilus influenzae*, la structure montre que le domaine passager arbore un domaine protéase similaire à celui retrouvé chez les SPATEs en plus de posséder un domaine « SAAT » qui permet l'auto-agrégation des bactéries (214). Les auteurs de cette dernière étude suggèrent que plusieurs autres autotransporteurs qui médient l'auto-agrégation possèdent un domaine « SAAT ».

Toutefois, AIDA-I, TibA et Ag43 ont des caractéristiques communes qui leur sont propres et qui les différencient des autres autotransporteurs qui médient l'auto-agrégation comme Hap. La famille des SAATs sera décrite en détail dans la prochaine section. De plus, comme plusieurs autotransporteurs partagent ces mêmes fonctions, les concepts d'adhésion, d'auto-agrégation et de formation de biofilm seront brièvement introduits dans la présente section.

2.3.1. Adhésion

L'adhésion bactérienne, que ce soit sur des cellules d'un organisme hôte ou bien sur une surface abiotique, joue un rôle très important pour les bactéries (305). Par exemple, l'adhésion des bactéries sur une surface abiotique constitue la première étape de la formation d'un biofilm. D'un autre côté, dans le cas des bactéries pathogènes,

l'adhésion constitue généralement une des premières étapes cruciales pour le développement d'une infection. L'adhésion empêche l'élimination des bactéries par les mécanismes physiques de l'hôte, comme le péristaltisme du système digestif, et contribue donc à l'établissement de l'infection. De plus, l'adhésion locale des bactéries à un site précis permet aux bactéries de sécréter leurs toxines et autres facteurs de virulence à un endroit précis de l'organisme hôte, ce qui maximise le développement de l'infection. L'adhésion des bactéries se fait donc à un endroit précis dans l'organisme hôte par l'entremise d'interaction spécifique adhésine-récepteur.

Il existe principalement deux classes d'adhésines : les adhésines fimbriaires, qui sont de longues structures polymériques et les adhésines afimbriaires, qui sont plutôt des protéines monomériques ou bien oligomériques (304). Il existe une quantité impressionnante d'adhésines, ce qui reflète l'importance de l'adhésion pour les bactéries. Sans faire une liste exhaustive, les adhésines fimbriaires les plus connues sont les pili et fimbriae (12) ainsi que les curlis (306), tandis que le *SST5* est particulièrement utilisé pour la sécrétion d'adhésine afimbriaires (304, 307). Notamment, les autotransporteurs trimériques ont tous une fonction d'adhésion ainsi que plusieurs autotransporteurs monomériques (Tableau 1), systèmes à deux partenaires ainsi que l'intimine, par exemple, sont des adhésines. Un des meilleurs exemples d'un autotransporteur étant une adhésine est la pertactine de *Bordetella pertussis* (301). La pertactine est d'ailleurs un composant du vaccin contre le pertussis (308).

Dans bien des cas, les mécanismes d'adhésion sont très bien compris et sont très variés. Par exemple, l'adhésion des UPEC dans la vessie par les fimbriae de type 1 se fait via la sous-unité terminale FimH (12, 309). FimH reconnaît l'uroplakine 1a, une glycoprotéine présente sur les cellules urothéliales. En fait, FimH reconnaît spécifiquement les résidus mannose de l'uroplakine 1a et la liaison de FimH à l'uroplakine 1 est maximisée par la présence de forces de cisaillement (shear stress), c'est-à-dire dans des conditions retrouvées dans le système urinaire. Un autre exemple est celui de l'intimine (310). Les EHEC et les EPEC adhèrent intimement avec les cellules de l'intestin, ce qui provoque la formation d'un piédestal. C'est l'intimine qui

procure cette adhésion intime entre les bactéries et les cellules intestinales. L'intimine reconnaît la protéine Tir à la surface des cellules de l'hôte. La particularité de Tir est qu'il s'agit d'une protéine sécrétée par le *SST3* de la bactérie dans la cellule de l'hôte. Donc, les EHEC et EPEC ont développé une stratégie où ils produisent l'adhésine responsable de l'adhésion intime ainsi que le récepteur de cette adhésine, qu'elles injectent dans la cellule hôte. Ces deux exemples illustrent bien le fait que les bactéries ont adapté différentes stratégies pour adhérer précisément au bon endroit. Par ailleurs, de légères modifications dans la séquence primaire de FimH modifient sa spécificité envers d'autres types d'oligomannose, ce qui modifie par le fait même le tropisme cellulaire (311). De plus, les bactéries peuvent exprimer plusieurs adhésines différentes qui possèdent un tropisme différent (12). Par exemple, les UPEC utilisent le fimbriae de type 1 pour coloniser la vessie et un pilus de type P pour coloniser les reins. Dans ce cas, les deux adhésines sont exprimées différemment et c'est l'action concertée des deux adhésines à deux endroits différents qui permet une colonisation maximale du système urinaire. Finalement, plusieurs adhésines se lient à des molécules de la matrice extracellulaire plutôt qu'à des récepteurs, comme la fibronectine ou le collagène. Cette liaison à la matrice extracellulaire peut ainsi faire un pont entre les bactéries et les cellules de l'hôte.

L'adhésion des bactéries sur des cellules de l'organisme hôte peut aussi entraîner l'internalisation des bactéries. Plusieurs bactéries sont intracellulaires et possèdent donc des mécanismes spécifiques pour leur internalisation dans la cellule. Encore là, les mécanismes sont nombreux et bien compris. Dans certains cas, la liaison d'une adhésine à son récepteur entraîne une cascade de signalisation à l'intérieur de la cellule hôte, ce qui provoque l'internalisation des bactéries. Par exemple, les UPEC sont capables d'envahir les cellules urothéliales suite à la liaison de FimH aux cellules (309). La liaison provoque l'activation de certaines GTPases de la famille Rho et l'invasion des bactéries. Suite à l'invasion, les bactéries s'évadent des vésicules d'endocytose et se retrouvent dans le cytosol de la cellule où elles se répliquent et forment un biofilm intracellulaire, appelé *IBC* (Intracellular bacterial communities) (309, 312). Cette invasion des UPEC dans les cellules urothéliales contribue à la récurrence des infections

urinaires. D'un autre côté, *Salmonella* et *Shigella*, suite à l'adhésion aux cellules épithéliales de l'intestin, utilisent un *SST3* pour envahir les cellules (305). La stratégie consiste donc à adhérer aux cellules et, ensuite, à injecter directement dans ces cellules des effecteurs qui vont provoquer l'invasion des bactéries. Plusieurs autres bactéries utilisent des mécanismes différents pour leur adhésion et leur invasion. Pour l'autotransporteur AIDA-I, certains mutants ayant perdu la capacité de promouvoir l'adhésion des bactéries sont toujours capables de promouvoir l'invasion, ce qui suggère que AIDA-I utilise deux mécanismes différents pour promouvoir l'adhésion et l'invasion (313).

2.3.2. Auto-agrégation

Plusieurs molécules ayant une fonction d'adhésion permettent aussi l'auto-agrégation des bactéries. L'auto-agrégation est définie comme étant la formation de larges agrégats cellulaires et est facilement observable par la sédimentation des cellules dans une culture statique. Pratiquement, l'auto-agrégation se voit macroscopiquement en mesurant la densité optique dans le haut d'une culture de bactéries dans des conditions statiques. Il s'agit d'un phénomène autant présent chez les bactéries à Gram positif que les bactéries à Gram négatif et le phénomène est aussi observable chez plusieurs types de levures et autres organismes eucaryotes unicellulaires. Comme mentionné précédemment, plusieurs adhésines sont aussi responsables de l'agrégation des bactéries. C'est le cas notamment des curlis (306), de plusieurs types de pili et fimbriae (314, 315) ainsi que de plusieurs autotransporteurs monomériques, systèmes à deux partenaires et autotransporteurs trimériques (6, 307). En général, il semble que les mécanismes d'auto-agrégation et d'adhésion soient différents.

Les mécanismes par lesquels ces adhésines permettent l'auto-agrégation des bactéries sont rarement aussi bien définis que les mécanismes d'adhésion. En général, il est suggéré que l'adhésine interagit directement avec elle-même et que c'est cette interaction qui est responsable de l'auto-agrégation. Par exemple, dans les cas SAATs, il a été suggéré par quelques groupes que l'auto-agrégation était le résultat d'une auto-

association. Cette suggestion provient de la visualisation de l'agrégation de bactéries fluorescentes exprimant un SAAT (6). Ainsi, seules les bactéries exprimant un SAAT se retrouvent dans l'agrégat, suggérant un mécanisme auto-associatif. Toutefois, il n'existe aucune preuve directe de l'auto-association. Or, il existe d'autres mécanismes qui sont non auto-associatifs et qui expliquent cette spécificité d'agrégation. Par exemple, *Saccharomyces cerevisiae* agrège spécifiquement grâce à l'adhésine Flo1 (316). Toutefois, l'adhésine Flo1 n'interagit pas directement avec elle-même, mais plutôt avec un composant présent à la surface de l'autre cellule et c'est la participation des molécules Flo1 des deux cellules qui entraîne l'agrégation spécifique. Il existe seulement quelques exemples pour lesquels il a été démontré que l'auto-agrégation était le résultat d'une interaction directe. C'est le cas par exemple pour l'hémagglutinine filamenteuse FHA de *B. pertussis* (317).

2.3.3. Formation de biofilm

Il est maintenant clair que les bactéries vivent majoritairement en communauté, comme dans un biofilm, plutôt que seules sous forme planctonique. Un biofilm est formé par un agrégat de bactéries qui est entouré d'une matrice composée de substances polymérique extracellulaire (318). La matrice extracellulaire est produite par les bactéries présentes dans le biofilm et, en général, les biofilms sont retrouvés aux interfaces solide-liquide et solide-air. La formation de biofilm est relativement complexe et se résume en quatre étapes principales. La première étape constitue l'attachement des bactéries à une surface abiotique. Suite à l'adhésion, les bactéries vont interagir ensemble pour former des microcolonies et les bactéries vont commencer à produire les composants de la matrice du biofilm. Il y a ensuite une étape de maturation du biofilm où le biofilm adopte sa structure finale. Finalement, la dernière étape consiste au détachement et à la dispersion des bactéries du biofilm pour permettre la colonisation de nouvelles surfaces. La formation de biofilm est généralement associée à un changement du profil d'expression des gènes des bactéries, pour favoriser la production de la matrice et la production d'adhésines.

Les adhésines peuvent être impliquées à différentes étapes de la formation d'un biofilm. Évidemment, plusieurs adhésines ont un rôle dans l'étape initiale d'adhésion aux surfaces abiotiques. Il a été montré par exemple que la délétion des fimbriae de type 1 réduisait l'adhésion des bactéries à différentes surfaces abiotiques (319). Les curlis, en plus de promouvoir l'adhésion initiale des bactéries, interagissent avec la cellulose et jouent donc aussi un rôle dans l'interaction des bactéries avec la matrice du biofilm (306). Plusieurs adhésines, par leurs propriétés d'auto-agrégation, ont aussi un rôle lors de la formation de la microcolonie et lors de la maturation du biofilm. Les adhésines afimbriaires, comme les autotransporteurs, pourraient d'ailleurs jouer un rôle important au niveau des interactions entre les bactéries dans le biofilm par leur structure plus courte. Il est donc probable que les étapes initiales de formation d'un biofilm soient réalisées par les adhésines fimbriaires plus longues qui permettent une bonne adhésion aux surfaces abiotiques, tandis que les autotransporteurs permettent les contacts dans le biofilm mature. D'ailleurs, il a été observé que l'expression d'Ag43 est augmentée dans les biofilms intracellulaires formés par les UPEC (312).

3. Autotransporteurs auto-associatifs d'*Escherichia coli*

Les autotransporteurs auto-associatifs (SAATs) sont une famille d'autotransporteurs monomériques regroupés pour leurs similarités fonctionnelles (6); ils sont tous capables de promouvoir l'auto-agrégation des bactéries. Le groupe est composé des autotransporteurs AIDA-I (adhesin involved in diffuse adherence), TibA (enterotoxigenic invasion locus B) et Ag43, qui sont retrouvés dans différentes souches d'*E. coli*. La prochaine section présente donc ces trois autotransporteurs en insistant sur leurs similitudes et leurs différences.

3.1 Les pathotypes d'*E. coli*

E. coli est avant tout une bactérie commensale qui réside dans le système gastro-intestinal de la majorité des animaux (7, 320). Toutefois, il existe aussi plusieurs souches d'*E. coli* qui sont pathogènes qui se divisent en deux grandes catégories qui elles-mêmes sont sous-divisées selon la présence de certains facteurs de virulence. D'ailleurs, bien qu'Ag43 soit retrouvé dans la majorité des souches, même les souches commensales et de laboratoire, les SAATs semblent être particulièrement associées aux souches causant des maladies intestinales.

Les souches pathogènes d'*E. coli* sont divisées en deux grandes catégories : les souches causant des maladies intestinales et les souches causant des maladies extra-intestinales (extraintestinal pathogenic *E. coli* – ExPEC) (321). Tout d'abord, les ExPECs sont composés notamment des souches uropathogènes (uropathogenic *E. coli* – UPEC), des souches causant des méningites et des septicémies (neonatal meningitidis *E. coli* – NMEC) ainsi que des souches aviaires (avian pathogenic *E. coli* – APEC). Les UPECs colonisent le système urinaire, d'abord la vessie et ensuite les reins si l'infection n'est pas traitée, et sont responsables de la majorité des infections urinaires. Ses souches sont caractérisées par la présence de plusieurs adhésines, particulièrement les fimbriaes de type I et de type P, ainsi que la présence de plusieurs toxines et systèmes d'acquisition du fer. Les APECs sont des souches similaires aux UPECs, mais qui causent des infections du système respiratoire des oiseaux. Finalement, les NMECs sont des souches qui se retrouvent dans le sang et dans le système nerveux central et causent ainsi des méningites chez les nouveaux-nés.

Les souches causant des maladies intestinales quant à elles sont divisées en plusieurs pathotypes (7, 320, 321). Tout d'abord, il y a les souches entéropathogènes (enteropathogenic *E. coli* – EPEC). Les EPECs sont caractérisées par des lésions attachantes-effaçantes où les bactéries adhèrent de façon intime à la paroi cellulaire ce qui provoque la formation d'un piédestal et l'effacement des microvillosités. Les EPEC s'approchent d'abord des entérocytes à l'aide de fimbriaes, comme le *Bundle-forming*

pili (*BFP*). Ensuite, il y a l'adhésion intime, grâce à l'intimine et à un *SST3* très bien conservé qui injecte dans les entérocytes le récepteur de l'intimine, et l'effacement des microvillosités, grâce aux effecteurs du *SST3* qui ciblent le cytosquelette. Les facteurs de virulence des EPEC se trouvent majoritairement sur le *LEE* (*Locus of enterocytes effacement*), bien que certains effecteurs du *SST3* ne se retrouvent pas sur le *LEE*.

Les souches entérohémorragiques (enterohemorrhagic *E. coli* – EHEC) sont aussi caractérisées par la présence des lésions attachantes-effaçantes. Les EHEC possèdent donc, comme les EPEC, un *SST3* et l'intimine, à la différence que les EHEC injectent environ deux fois plus d'effecteurs que les EPEC. Toutefois, les EHEC se distinguent des EPEC par la présence de plusieurs toxines, dont la toxine shiga, d'où l'appellation STEC (Shiga toxin producing *E. coli*) et l'hémolysine HlyA. Dans environ 5-10% des cas, une infection par des EHEC, à cause de la toxine shiga, peut entraîner le développement d'un syndrome hémolytique et urémique. La souche connue O157 :H7 fait partie des EHEC.

Ensuite, les souches entérotoxigéniques (enterotoxigenic *E. coli* – ETEC) sont caractérisées par la présence des toxines *HT* (*Heat-labile toxin*) et *ST* (*Heat-stable toxin*). L'adhésion des ETEC aux entérocytes se fait grâce aux nombreux facteurs d'adhérence présents, appelés facteurs de colonisation (*CF*), ainsi qu'à certains facteurs d'invasion, comme TibA. Ensuite, ce sont les toxines qui provoquent le développement de la diarrhée.

Finalement, il existe quelques autres pathotypes; les souches entéroinvasives (enteroinvasive *E. coli* – EIEC), les souches entéroagréatives (enteroaggregative *E. coli* – EAEC) et les souches à adhérence diffuse (Diffusely adhering *E. coli* – DAEC). Ces trois pathotypes sont décrits selon leur profile d'adhésion/invasion. Les EIEC sont des pathogènes intracellulaires facultatifs et sont caractérisées par leur capacité d'invasion. *Shigella flexneri* est généralement classé avec les souches EIEC. Les EAEC sont caractérisées par une adhésion agrégative qui est causé par la présence de fimbriaes d'adhérence agrégative (*AAF*). Les DAEC, de leur côté, sont caractérisées par une adhésion diffuse qui est causée par la présence de plusieurs adhésines afimbriales (*AFA*) ainsi que l'autotransporteur AIDA-I.

3.2. AIDA-I

L'autotransporteur AIDA-I est un des autotransporteurs monomériques les plus étudiés. Il s'agit d'un important modèle d'étude pour la biogenèse et la fonctionnalité des autotransporteurs. D'ailleurs, dans la classification des autotransporteurs, il y a une famille nommée AIDA.

3.2.1. Implication dans la pathogenèse

AIDA-I a été initialement découvert dans un cas de diarrhée infantile causée par la souche de *E. coli* 2787 (O126 :H27) (290). La souche 2787 est une souche de *E. coli* avec un phénotype d'adhérence diffuse (DAEC) et AIDA-I a été identifié comme étant le principal facteur permettant l'adhérence diffuse (290). Par contre, AIDA-I est rarement retrouvé dans des souches *E. coli* infectant l'humain, mais se retrouve surtout dans des souches infectant le porc. En particulier, AIDA-I est associé à des souches causant des diarrhées post-sevrage et la maladie de l'œdème. Le gène codant pour AIDA-I est localisé sur un plasmide (290) et il est donc possible qu'il y ait eu un transfert entre souches de *E. coli* porcines et humaines. D'ailleurs, les deux seules séquences existantes d'AIDA-I, soit la séquence provenant de 2787 et une séquence provenant d'une souche porcine isolée en Chine (322), sont identiques.

Des études réalisées en France ont noté la présence du gène codant pour AIDA-I dans seulement 2 à 4 % des souches isolées de patients atteints de diarrhée (323, 324). En revanche, lors d'études réalisées dans différents pays, notamment au Canada, aux États-Unis, en Chine, en Allemagne et en Belgique, la prévalence du gène AIDA-I dans des isolats de porcs malades atteint jusqu'à 63% selon les études (Tableau I). De plus, dans toutes ces études, aucun isolat provenant de porcs sains ne possédait le gène codant pour AIDA-I. Ceci suggère donc que AIDA-I pourrait être un important facteur de virulence chez le porc.

Tableau II : Prévalence du gène codant pour AIDA-I chez des souches porcines isolées de cas de diarrhée

Pays	Nombre d'isolats de diarrhée	Nombre de souches AIDA+	% de souches AIDA+	Ref
Allemagne	44	18	41	(325)
Belgique	174	110	63	(326)
Canada	170	15	9	(327)
Chine	320	21	6,5	(322)
Corée	604	45	7	(328)
USA	304	82	27	(329)

Des études ont aussi mis en évidence la présence du gène AIDA-I dans des souches ETEC (330) et STEC (325), montrant que AIDA-I n'est pas limité aux souches DAEC. Finalement, une infection expérimentale chez des porcelets nouveaux-nés utilisant la souche ETEC PD20, qui possède AIDA-I et la toxine STb comme seuls facteurs de virulence, a mis en évidence l'importance d'AIDA-I comme facteur de virulence (331). Dans cette étude, le développement de la diarrhée chez les porcelets nécessitait la présence d'AIDA-I ; la toxine STb seule n'étant pas capable de causer la diarrhée. Il est peu probable que AIDA-I soit la cause directe des symptômes de diarrhée, mais l'adhésion des bactéries aux cellules intestinales, via AIDA-I, améliore possiblement l'efficacité de la toxine en permettant une sécrétion et une concentration de la toxine à proximité des cellules (330).

Comme mentionné précédemment, AIDA-I a été initialement découvert pour son implication dans l'adhésion aux cellules (290). Dans cette optique, différents groupes ont essayé de trouver un récepteur pour AIDA-I. Ainsi, une première étude a réussi à identifier une glycoprotéine membranaire de 119 kDa (gp119) à partir des cellules HeLa (332). Ce récepteur potentiel n'a cependant pu être identifié avec plus de précision, mais il semble être plutôt ubiquitaire puisqu'AIDA-I permet l'adhésion à différents types de cellules ayant des origines différentes, comme les cellules CV-1 et CHO (332). Une deuxième étude a identifiée un récepteur dans le mucus présent à la surface des cellules (333). Toutefois, les auteurs de cette étude ont identifié comme récepteur une protéine

cytoplasmique et l'implication de ce dernier dans l'adhésion n'a pas été confirmée. Finalement, AIDA-I ne semble pas être impliqué dans l'adhésion aux protéines de la matrice extracellulaire (263).

Une autre étude faite dans notre laboratoire a montré qu'AIDA-I interagissait avec une molécule présente dans le sérum porcin et a identifié cette molécule comme étant l'apolipoprotéine A1 (ApoA1) (Mélissa René, Jason Létourneau, résultats non publiés). Il est intéressant de noter que, chez le porc, l'ApoA1 est aussi sécrétée dans la lumière intestinale. L'ApoA1 pourrait donc être un récepteur présent dans le mucus intestinal chez le porc, ce qui expliquerait la présence d'AIDA-I principalement dans des souches porcines. Finalement, comme ses homologues TibA et Ag43, AIDA-I permet aussi l'auto-agrégation des bactéries et la formation de biofilm (291) ainsi que l'invasion de cellules épithéliales (313).

3.1.2. Organisation et biogenèse d'AIDA-I

AIDA-I est synthétisé sous la forme d'une pré-pro-protéine de 1287 acides aminés dans le cytoplasme de la bactérie (334). L'organisation de AIDA-I suit l'organisation normale d'un autotransporteur ; AIDA-I est composé d'une séquence signale, d'un domaine passager, d'une région de jonction et d'un domaine β (Fig. 9A). AIDA-I possède d'abord une séquence signale longue de 49 acides aminés, contenant l'extension de la séquence signale conservée chez plusieurs autotransporteurs.

Le domaine passager s'étend des résidus 50 à 707 et, selon les prédictions, forme une hélice β droite semblable aux autres domaines passagers d'autotransporteurs monomériques, mais ne contiendrait aucune structure ou boucle émergeant de cette hélice (Fig. 9B-E). D'ailleurs, une évaluation de la structure secondaire par dichroïsme circulaire a montré que le domaine passager était composé majoritairement de brins β (332). La particularité du domaine passager d'AIDA-I est qu'il est composé uniquement de répétitions imparfaites d'une séquence de 19 acides aminés. Il est formé de 35 répétitions de la séquence consensus de 19 acides aminés : SGGxAxxTx•NSGGxQx•x,

étude structure-fonction sur AIDA-I a été effectuée et montre que les trois surfaces ne sont pas fonctionnellement équivalentes, une face étant possiblement associée à un phénotype particulier (313). Par ailleurs, la conclusion principale de cette étude structure-fonction est que AIDA-I est une protéine modulaire composée de différents modules étant responsable des différents phénotypes. Par exemple, la partie N-terminale du domaine passager est formée par un domaine d'adhésion et la délétion de ce domaine d'adhésion n'a pas d'effet sur les autres phénotypes.

Ensuite, la région de jonction se retrouve entre les résidus 708 et 952. Comme pour les autres autotransporteurs monomériques, la région de jonction est essentielle pour la sécrétion d'AIDA-I (219). C'est aussi au niveau de la région de jonction, plus précisément entre les résidus Ser 847 et Ala 848, que le clivage a lieu (335). Suite au clivage, la partie extracellulaire est appelée AIDA-I mature tandis que la partie membranaire est appelée AIDAc (Fig. 9A). Malgré l'événement de clivage, AIDA-I mature et AIDAc restent fortement associés et AIDA-I mature n'est pas relâché dans le milieu extracellulaire. Les deux parties restent associées de façon non-covalente et la seule façon répertoriée de les séparer est en dénaturant la partie AIDA-I mature, via un traitement à la chaleur par exemple.

Finalement, le domaine β est formé des résidus 953 à 1287 et est prédit comme étant un tonneau β classique d'un autotransporteur contenant une hélice α au centre. Il avait originalement été prédit que le domaine β de AIDA-I formerait un tonneau composé de 14 brins β (219), mais depuis la publication de structures d'autres domaines β d'autotransporteur, une nouvelle modélisation suggère que le domaine β forme bel et bien un tonneau composé de 12 brins β centré d'une hélice α (336).

La biogenèse d'AIDA-I suit les mêmes règles que celle des autres autotransporteurs, à quelques particularités près. D'abord, très peu d'informations sont disponibles quant au moment où AIDA-I est exprimé par la bactérie. Récemment, deux études se sont penchées sur les mécanismes de régulation d'AIDA-I, avec des résultats légèrement différents (337, 338). Une des études a été effectuée avec la souche 2787

(337) alors que l'autre a examiné l'expression de AIDA-I à partir d'un plasmide dans une souche de *E. coli* de laboratoire (338), ce qui explique possiblement les différences. Néanmoins, malgré les différences mineures, les conclusions principales des deux études se rejoignent. Il semble que l'opéron *aah-aidA* soit transcrit en un ARN bicistronique et soit induit au début de la phase stationnaire par le régulateur RpoS dans des conditions pauvres en glucose.

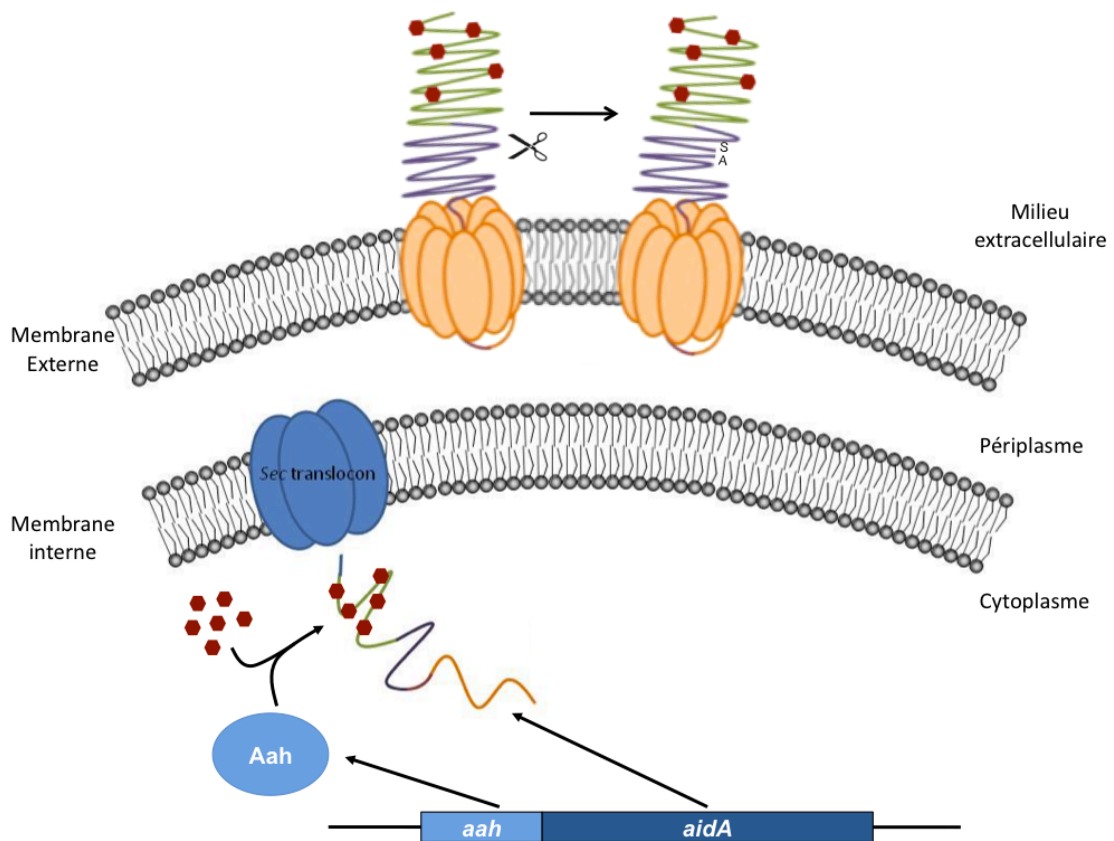


Figure 10 : Biogenèse d'AIDA-I Aah et AIDA-I (gris, SS ; vert, domaine passager ; mauve, région de jonction ; orange, domaine β) sont transcrits sous la forme d'un ARN bicistronique et sont donc traduits en même temps dans le cytoplasme. Dans le cytoplasme, AIDA-I est glycosylé par Aah par l'ajout de résidus heptoses (en rouge) avant même sa sécrétion. AIDA-I est alors exporté dans le périplasm par le système Sec et est ensuite entraîné à la membrane externe où le domaine β s'insère dans la membrane. Le domaine passager est alors sécrété dans le milieu extracellulaire et, suite au repliement, un clivage autocatalytique se produit séparant AIDA-I mature de AIDA-C. Toutefois, les deux parties restent fortement associées de façon non-covalente.

La première étape dans la biogenèse d'AIDA-I constitue sa glycosylation par Aah dans le cytoplasme (Fig. 10). Aah modifie AIDA-I par l'ajout de résidus heptoses provenant de la voie des LPS (276, 279, 339). C'est le domaine passager qui est glycosylé et les sucres sont ajoutés surtout sur des résidus sérines, donc il s'agit de O-glycosylation. AIDA-I est ensuite exporté dans le périplasma par le système Sec, possiblement par la voie post-traductionnelle comme les autres autotransporteurs, quoique ceci n'ait pas été confirmé. Une étude a évalué le rôle de l'extension de la séquence signale pour AIDA-I (339). Il a été observé que la délétion de l'extension de la séquence signale entraînait une diminution de la quantité d'AIDA-I présent à la surface de la bactérie. Les auteurs ont suggéré que l'extension permettait le ralentissement du passage dans le périplasma, comme pour les autres autotransporteurs, ce qui améliore l'efficacité de la glycosylation et ainsi l'expression d'AIDA-I à la surface des bactéries.

Ensuite, dans le périplasma, il est fort probable que AIDA-I nécessite la présence des chaperonnes périplasmiques comme les autres autotransporteurs de *E. coli*. Encore une fois, le rôle des chaperonnes périplasmiques sur la biogenèse d'AIDA-I n'a pas été démontré. Il a été montré par contre que le repliement *in vitro* du domaine β avec la région de jonction ne nécessitait pas la présence des chaperonnes SurA, Skp et DegP (232). De plus, toujours *in vitro*, aucune interaction entre le domaine β et Skp n'a été observée.

Finalement, le domaine β est inséré dans la membrane externe avec l'aide du complexe *BAM*. Le brin C-terminal du tonneau contient le motif de reconnaissance du complexe *BAM* et la déplétion de BamA affecte la sécrétion de AIDA-I dans le milieu extracellulaire (234). Par la suite, le domaine passager est sécrété dans le milieu extracellulaire où il se replie et AIDA-I est clivé entre les résidus Ser 846 et Ala 847 pour former AIDA-I mature et AIDAc. Il s'agit d'un clivage auto-catalytique intramoléculaire qui est effectué par une paire de résidus Asp 878 et Glu 897 qui, selon une modélisation, sont situés dans une boucle juste en dessous du site de clivage (258). Il est à noter que le clivage ne sert pas au relâchement de AIDA-I mature dans le milieu extracellulaire, puisque AIDA-I mature reste fortement associé à AIDAc (252). Finalement, le rôle du clivage n'a toujours pas été déterminé puisqu'un mutant non clivé est aussi bien exprimé et aussi fonctionnel *in vitro* que AIDA-I sauvage (252). Il est

possible toutefois que le clivage ait un rôle dans un contexte *in vivo* qui ne puisse pas être identifié par nos essais fonctionnels *in vitro*. Il est donc évident que le rôle de AIDA-I dans la pathogenèse n'est toujours pas bien compris.

3.1.3. La famille AIDA

Selon la classification initiale des autotransporteurs monomériques, il existe une grande famille d'autotransporteurs appelée AIDA (100). Cette famille est la plus représentée et inclut les autotransporteurs AIDA-I, TibA et Ag43 de *E. coli*, IcsA de *S. flexneri*, ShdA et MisL de *Salmonella enterica* ainsi que plusieurs autres. Cette classification est basée sur une homologie avec la partie C-terminal d'AIDA-I, correspondant principalement au domaine β et à la région de jonction (domaine COG3468). Ceci fait en sorte que la famille AIDA est très variée et il y a très peu de similarité au niveau de la séquence du domaine passager. De plus, les fonctions associées avec les membres de la famille AIDA sont donc très variées et il n'y a pas de point commun quant à la fonctionnalité des membres de la famille AIDA. Une étude plus récente a identifié tous les autotransporteurs présents dans les génomes d'*E. coli*, disponibles au moment de l'étude en 2010, et les a classés dans trois familles distinctes : les SPATES, les autotransporteurs trimériques et la famille AIDA (303). Dans cette classification, les autotransporteurs qui ne figurent pas dans les deux autres familles, c'est-à-dire les SPATES et les autotransporteurs trimériques qui sont plus facilement identifiables (par la présence d'un domaine protéase ou bien d'un domaine β trimérique), tombent dans la famille AIDA. Encore là, la seule homologie entre les membres de la famille AIDA se retrouve dans le domaine β et la région de jonction. Par contre, en ce limitant à *E. coli*, les membres de la famille AIDA ont maintenant quelques points en commun. D'abord, ils possèdent tous des répétitions internes dans leur domaine passager, bien que la séquence et la longueur de ces répétitions soient très variables. Ensuite, les auteurs ont remarqué que tous sont impliqués dans l'adhésion et l'invasion de cellules de l'hôte, l'adhésion aux composants de la matrice extracellulaire, la formation d'agrégats bactériens ou bien la formation de biofilm. Toutefois, cette étude est limitée à *E. coli* et ces observations ne semblent pas se prêter aux autotransporteurs

de type AIDA retrouvés chez d'autres espèces bactériennes. Ceci fait en sorte que plusieurs autotransporteurs sont classés dans la famille AIDA uniquement à cause du domaine β et de la région de jonction, ce qui ne reflète pas leur fonctionnalité. Par exemple, IcsA qui fait partie de la famille AIDA-I permet la polymérisation de l'actine à un pôle de *S. flexneri* et ainsi diffère grandement d'AIDA-I (284). Les membres de la famille AIDA peuvent donc être assez éloigné de l'autotransporteur AIDA-I. Par conséquent, dans les banques de données, beaucoup d'autotransporteurs sont appelés « AIDA-like » même si le domaine passager est complètement différent et possède une fonctionnalité différente.

3.2. Ag43

Ag43 a été découvert bien avant de savoir qu'il s'agissait d'un autotransporteur. L'intérêt initial pour cette protéine vient du fait qu'il s'agissait de la protéine de membrane externe soumise à la variation de phase qui est la plus abondante chez *E. coli* (259, 340). Au moment de sa découverte, le gène codant pour Ag43 s'appelait *flu* puisque celui-ci était responsable de la floculation des bactéries (340). L'intérêt pour Ag43 a donc d'abord été à propos de sa régulation et du mécanisme responsable de la variation de phase.

La variation de phase touche surtout des molécules retrouvées à la surface des bactéries et est caractérisée par la présence, dans une population clonale, de cellules exprimant le gène (ON) et d'autres qui ne l'expriment pas (OFF) (341). Il s'agit d'un processus qui est héréditaire et réversible d'une génération à l'autre. Par exemple, une bactérie en phase ON va produire majoritairement des descendants qui vont être aussi en phase ON, à l'exception d'une petite fraction qui va se retrouver en phase OFF. La fréquence du passage d'un phénotype à l'autre est variable et peut atteindre jusqu'à 10% de la population. Il s'agit donc d'un système permettant aux bactéries de s'adapter rapidement à un environnement variable, puisque la fréquence de variation de phase est beaucoup plus élevée que celle d'une mutation spontanée qui entraînerait un changement

phénotypique. La variation de phase est entre autres utilisée par plusieurs bactéries pour échapper au système immunitaire de l'hôte en faisant varier ses protéines de surface.

Il existe différents mécanismes de variation de phase. Dans le cas d'Ag43, la variation de phase se fait via les protéines régulatrices OxyR et Dam (259, 342, 343). OxyR est un régulateur transcriptionnel qui contrôle l'expression de protéines lors d'un stress oxydatif (344). Il devient actif suite à une exposition au peroxyde et entraîne alors la transcription de plusieurs gènes afin de protéger la cellule contre l'oxydation. OxyR est un répresseur de l'expression d'Ag43 (OFF). Toutefois, dans le site de liaison d'OxyR à l'ADN, il y a trois sites GATC qui sont sensibles à la méthylation par Dam. Lorsqu'un de ces sites est méthylé, OxyR ne peut plus se lier à l'ADN et le gène codant pour Ag43 est alors transcrit (ON). La liaison d'OxyR à l'ADN empêche toutefois la méthylation par Dam. La variation de phase se produit donc au moment où il y a réplication de l'ADN, où il y a génération d'ADN non méthylé, et c'est la compétition entre OxyR et Dam qui détermine la quantité de cellules en phase ON et OFF. L'implication d'OxyR dans la régulation de Ag43 suggère qu'Ag43 est régulé par le stress oxydatif. Toutefois, autant les formes oxydées (actives) que réduites (inactives) d'OxyR peuvent se lier à la région promotrice d'Ag43, ce qui l'hypothèse de la régulation par le stress oxydatif (344). Le rôle de cette régulation par variation de phase pour Ag43 n'est toutefois toujours pas établi.

Ag43 est un autotransporteur monomérique faisant partie de la famille AIDA (100). Toutefois, contrairement à la majorité des membres de cette famille, la ressemblance ne se limite pas seulement au domaine β et à la région de jonction. En effet, comme pour AIDA-I, le domaine passager d'Ag43 est composé de répétitions de la même séquence consensus de 19 acides aminés (6). Comme pour AIDA-I, l'expression d'Ag43 permet aux bactéries d'auto-agrégier, de former du biofilm et d'adhérer à des cellules en culture (263, 345, 346). Finalement, comme AIDA-I, Ag43 est clivé et les deux parties restent associées fortement ensemble et la seule façon d'extraire le domaine passager est de le dénaturer en effectuant, par exemple une extraction à la chaleur (347). Le clivage se fait à l'extrémité C-terminale du domaine passager par un mécanisme encore inconnu.

Par contre, contrairement à AIDA-I qui est retrouvé dans un type de souche précis, Ag43 est retrouvé dans la majorité des souches d'*E. coli* commensales ou pathogènes (348). Il existe plusieurs allèles d'Ag43 qui se divisent en deux grandes catégories (349). En fait, une des catégories correspond à l'autotransporteur Cah (299). Cah a été décrit comme étant un homologue d'Ag43 liant le calcium et est donc fonctionnellement et structurellement similaire à Ag43. Il est intéressant de noter que les allèles correspondant à Cah sont fortement associés à des souches pathogènes, suggérant un rôle pour Cah/Ag43 dans la pathogénèse. Par ailleurs, Ag43 est particulièrement présent chez les UPEC et il semble qu'Ag43 contribue à la pathogénèse des UPEC. Les maladies urinaires sont reconnues pour leur persistance et il a été montré que les UPEC pouvaient se retrouver à l'intérieur des cellules urothéliales et former des communautés intracellulaires appelées *IBC* (312). Or, Ag43 est bien exprimée dans ces *IBC* et il est donc possible qu'Ag43 contribuent aux *IBC* par sa capacité d'auto-agrégation/formation de biofilm. Finalement, une autre étude a montré qu'Ag43 promouvait la colonisation à long terme dans un modèle d'infection urinaire (350).

3.3. TibA

L'autotransporteur TibA est lui-aussi un homologue d'AIDA-I et possède un domaine passager composé de répétitions de la même séquence consensus de 19 acides aminés. TibA se retrouve plutôt chez des souches d'*E. coli* entérotoxigéniques (ETEC) et semble être associé au fimbriae CFA/I (351). TibA a été découvert dans la souche H10407, qui a été isolée d'une personne atteinte de diarrhée (352, 353). La souche H10407 a la particularité, par rapport aux autres ETEC, d'être capable d'envahir des cellules épithéliales intestinales humaines (354). Cette capacité d'invasion semble provenir de deux locus distincts; les locus d'invasion *A* et *B*. Le locus d'invasion *A* (Tia) exprime une protéine de 25 kDa similaire à la protéine Ail de *Yersinia enterocolitica* (355) et semble être retrouvé chez plusieurs types de souches d'*E. coli* (356). Quant à lui, le locus d'invasion *B* (Tib) permet l'expression d'une molécule de 104 kDa dans la membrane externe qui corrèle avec la capacité d'adhésion et d'invasion de la souche

H10407 et permet à des souches de laboratoire exprimant le locus *tib* d'adhérer et d'envahir des cellules épithéliales en culture (351).

Le locus *tib* est composé de quatre cadres de lecture, *tibDBCA*. En fait, *tibC* codent pour une glycosyltransferase et *tibA* pour un autotransporteur monomérique et les deux gènes forment un opéron similaire à l'opéron *aah-aida* (262). Quand à eux, TibB et TibD semblent être des régulateurs de l'opéron *tibCA* (357). La régulation de TibA n'est pas très bien étudiée. Par contre, une étude cherchant des cibles du régulateur CRP (cAMP regulator protein) chez la souche H10407 a montré que TibA était positivement régulé par CRP, via TibB et TibD (357). CRP est un important régulateur de *E. coli* et est activé par une diminution de la quantité de glucose (358). Chez H10407, TibA est donc exprimé dans des conditions de manque de glucose, conditions retrouvées par exemple en phase stationnaire. Cela suggère donc que TibA est exprimé dans les mêmes conditions qu'AIDA-I.

Finalement, TibA partage plusieurs autres caractéristiques avec AIDA-I, comme la glycosylation et la présence d'une séquence signale longue. Comme AIDA-I, TibA est capable de promouvoir, en plus de l'adhésion et l'invasion de cellules épithéliales, l'auto-agrégation des bactéries et la formation de biofilm (282, 351, 359). Quelques études ont essayé d'identifier un récepteur pour TibA, mais n'ont pas été fructueuses. Des tests d'adhésion, effectués en compétition avec la protéine TibA purifiée, semblent indiquer la présence d'un récepteur spécifique pour TibA (359). De plus, l'expression de TibA chez la souche de laboratoire HB101 permet à la bactérie d'adhérer et d'envahir plusieurs lignées cellulaires, suggérant que le récepteur pour TibA est ubiquitaire (351). Le récepteur de TibA reste toutefois encore inconnu.

3.4. Similarités et différences

Comme il a été mentionné précédemment, AIDA-I, TibA et Ag43 ont été regroupés ensemble pour leur similarité. Tout d'abord, les trois sont des protéines multifonctionnelles capables de médier l'adhésion et l'invasion de cellules épithéliales,

l'auto-agrégation des bactéries ainsi que la formation de biofilm. De plus, les trois SAATs possèdent aussi une organisation similaire (Fig. 11). Ils sont composé (du N-terminal vers le C-terminal) : (i) d'une séquence signal longue, (ii) d'un domaine passager de longueur variable mais bien conservé, (iii) d'une région de jonction de longueur similaire et (iv) d'un tonneau β relativement bien conservé. En fait, le domaine passager des trois SAATs est uniquement composé de répétitions d'une séquence consensus de 19 acides aminés (Fig. 10). Finalement, les trois SAATs peuvent être glycosylés.

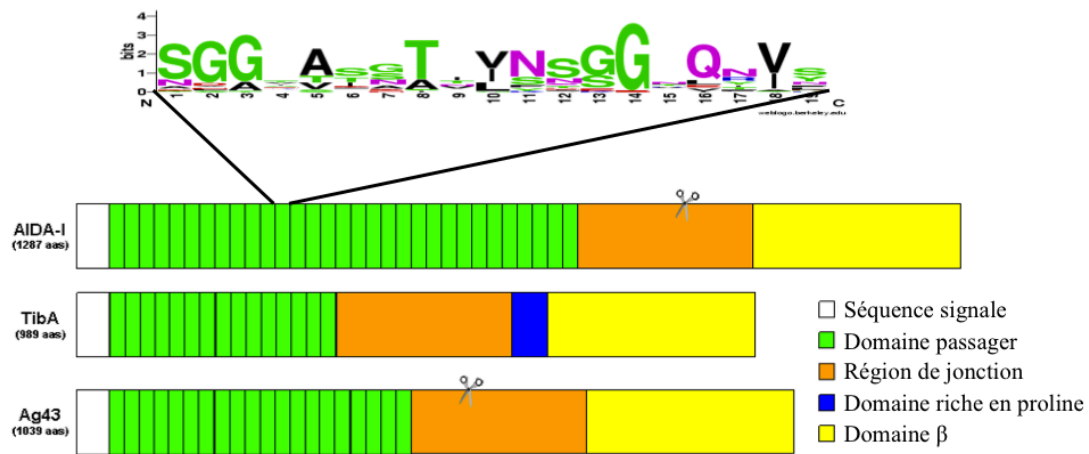


Figure 11 : Similarités et différences entre les SAATs. Schéma représentant l'organisation de AIDA-I, TibA et Ag43. La séquence consensus globale des répétitions de 19 acides aminés retrouvées dans les domaines passagers de AIDA-I, TibA et Ag43 est identifiée au-dessus.

Toutefois, malgré les ressemblances, il existe plusieurs différences entre les SAATs. La différence principale se situe au niveau du nombre de répétitions de la séquence consensus de 19 acides aminés : AIDA-I possède 35 répétitions tandis que TibA et Ag43 sont composés de 16 et 20 répétitions respectivement. Comme chaque répétition forme un tour de l'hélice, cela fait en sorte que les SAATs ont des longueurs qui varient beaucoup. Ensuite, TibA possède un domaine additionnel riche en prolines entre la région de jonction et le tonneau β . De leur côté, AIDA-I et Ag43 sont clivés, bien que le domaine passager reste associé très fortement au domaine β après le clivage.

Aussi, bien qu'une glycosyltransferase soit associée à AIDA-I et à TibA, ce n'est pas le cas pour Ag43. D'ailleurs, AIDA-I et TibA sont toujours retrouvés sous leur forme glycosylée, tandis qu'Ag43 est exprimé sous forme non glycosylée dans la majorité des souches testées. Finalement, bien que les trois SAATs soient impliqués dans la pathogenèse, ils semblent y contribuer de différentes façons. AIDA-I a été découvert en tant qu'adhésine d'une souche d'*E. coli* à adhérence diffuse. TibA, pour sa part, a été découvert en tant que facteur promouvant l'invasion d'une souche ETEC. Finalement, Ag43 est un facteur d'agrégation et a été associé à la formation de biofilm. Par association, on s'est ensuite rendu compte que les trois SAATs étaient capables de promouvoir les mêmes fonctionnalités que leurs homologues. Toutefois, en comparant les SAATs entre eux, on réalise que les SAATs n'ont pas tous le même potentiel pour chacune des fonctionnalités.

4. Problématique et objectif

Toutes ces différences mettent en évidence le fait qu'il reste beaucoup de questions encore sans réponse. D'abord, leur appellation d'autotransporteurs « auto-associatifs » repose sur des suppositions; le mécanisme par lequel les SAATs promouvoient l'auto-agrégation n'a jamais été étudié. On ne sait donc pas si les SAATs sont vraiment auto-associatifs. Dans le cas de l'adhésine Flo de *Saccharomyces cerevisiae*, celle-ci permet l'auto-agrégation spécifique de cellules exprimant Flo par un mécanisme qui n'est pas auto-associatif (316). En fait, l'adhésine Flo reconnaît un composant de la membrane de la cellule voisine et n'interagit donc pas avec elle-même. Ensuite, pourquoi est-ce que la même séquence consensus de 19 acides aminés est conservée chez AIDA-I, TibA et Ag43? Selon les modèles, les SAATs possèdent une architecture similaire à celle de la majorité des autotransporteurs, malgré que ceux-ci ne possèdent pas le motif répété. Il doit donc y avoir une raison, autre que structurale, pour la répétition de cette séquence consensus chez les SAATs. Il y a aussi plusieurs aspects de l'organisation et de la fonctionnalité des SAATs qui demeurent incompris. Par exemple, AIDA-I est clivé, mais le clivage ne semble pas avoir de rôle structural ou fonctionnel, du moins lors d'expériences *in vitro*. De son côté, TibA possède un domaine additionnel riche en prolines. Quel est le rôle de ce domaine? Est-ce qu'il remplace le rôle joué par le clivage chez AIDA-I? Les domaines riches en prolines sont souvent responsables d'interactions protéine-protéine. Toutefois, dans le cas de TibA, sa position, entre la région de jonction et le tonneau β , ne suggère pas qu'il soit responsable de promouvoir l'interaction entre TibA et son récepteur, par exemple. Finalement, quel est le rôle de ces SAATs? Au cours d'une infection, est-ce que c'est la capacité d'adhésion, d'invasion, d'auto-agrégation ou bien de formation un biofilm qui est le plus important pour la bactérie? D'ailleurs, les trois SAATs ont été associés à la pathogenèse d'*E. coli* de différentes façons : par la capacité d'adhésion d'AIDA-I, la capacité d'invasion de TibA et la capacité d'auto-agrégation d'Ag43. Les quelques études portant sur l'expression d'AIDA-I et de TibA suggèrent qu'ils sont exprimés dans des conditions de phase stationnaire ou bien de haute densité de population, des conditions

se rapprochant plutôt de celles retrouvées dans l'environnement. Il est donc possible que les SAATs jouent plutôt un rôle dans l'environnement. Bref, il reste encore beaucoup à comprendre à propos des SAATs avant de savoir comment ils sont utilisés par les bactéries.

Mes travaux de doctorat avaient donc pour but de caractériser l'organisation des SAATs afin de définir leur rôle pour les bactéries. Pour ce faire, mes travaux suivaient deux objectifs : (I) définir l'organisation fonctionnelle des SAATs et (II) déterminer le rôle des SAATs.

Pour répondre au premier objectif, nous avons donc d'abord caractérisé l'autotransporteur TibA, le membre le la famille des SAATs le moins bien étudié. Nous avons effectué une étude structure-fonction pour définir l'organisation de TibA et nous avons étudié le rôle de la glycosylation sur TibA. Ensuite, étant donné le peu de séquences d'*aidA* disponible et l'abondance de souches d'*E. coli* positive pour la présence du gène *aidA* dans le laboratoire du Dr John Fairbrother, nous avons sélectionné plusieurs de ces souches et séquencé leur opéron *aah-aidA*. Notre hypothèse était qu'en définissant l'organisation de TibA et en étudiant la variabilité d'AIDA-I, nous allions identifier les caractéristiques communes et importantes de SAATs, ce qui nous permettrait de déterminer l'étendu réelle et l'importance des SAATs.

Pour répondre au deuxième objectif, en collaboration avec Victoria Girard, nous avons étudié le mécanisme d'auto-agrégation par lequel les SAATs promouvoient l'auto-agrégation des bactéries. Notre hypothèse était que l'auto-agrégation était médiée par une interaction directe protéine-protéine et que ce phénotype était important pour les bactéries.

Les résultats de mes travaux de doctorats m'ont permis de définir la famille des SAATs et d'identifier un rôle pour les SAATs qui n'avaient pas été envisagé auparavant.

Méthodologie et Résultats

Article 1 – Structure-fonction de TibA

Structure-Function Analysis of the TibA Self-Associating Autotransporter Reveals a Modular Organization

Infection and Immunity, Mai 2011; volume 79 (5), p. 1826-32

Avant-propos

De la famille des SAATs, TibA est le membre le moins bien caractérisé. Mon premier objectif était donc de caractériser TibA. Dans un premier temps, nous avons utilisé une approche de mutagenèse aléatoire afin de caractériser l'organisation fonctionnelle de TibA. Notre hypothèse était que l'organisation fonctionnelle de TibA allait nous permettre de définir des caractéristiques communes aux SAATs. Nos résultats nous ont permis de déterminer que le domaine passager de TibA était composé de deux modules: un module d'auto-agrégation en N-terminal suivi d'un module responsable de l'adhésion en C-terminal. Ces résultats montrent que TibA, comme AIDA-I et Ag43, est une protéine modulaire composée de différents modules conférant les différentes fonctionnalités. Nos résultats nous ont permis de comparer TibA avec les autres SAATs et cette comparaison nous a révélé deux choses : (i) les SAATs pourraient avoir évolué par un mécanisme d'acquisition et d'échange de modules et (ii) l'auto-agrégation des bactéries pourraient être le phénotype principal associé à l'expression des SAATs.

Contributions

Contribution pour cet article :

Comme premier auteur et auteur principal de cet article, j'ai effectué toutes les expérimentations et j'ai écrit le manuscrit au complet.

Structure-Function Analysis of the TibA Self-Associating Autotransporter Reveals a Modular Organization

Jean-Philippe Côté and Michael Mourez*

Canada Research Chair on Bacterial Animal Disease, Université de Montréal, Faculté de Médecine Vétérinaire, Saint-Hyacinthe, Québec, Canada

Running Title: Structure-Function of TibA

Abstract

Some enterotoxigenic *Escherichia coli* express the TibA adhesin/invasin, a multifunctional autotransporter that mediates autoaggregation of bacteria, biofilm formation, adhesion to cultured epithelial cells and invasion of these cells. To elucidate the structure-function relationship in TibA, we generated mutants by transposon-based linker scanning mutagenesis and by site-directed mutagenesis. Several insertion mutants had a defect in either adhesion or autoaggregation. Mutants with a defect in autoaggregation were found in the N-terminal half of the extracellular domain while mutants with a defect in adhesion were found in the C-terminal half. Deletion of the putative N-terminal autoaggregation domain abolished autoaggregation of the bacteria, but did not affect adhesion. Deletion of a proline-rich region located at the C-terminus of the extracellular domain abolished the adhesion properties of TibA, but did not affect invasion. This suggests that adhesion and invasion may rely on distinct mechanisms. Thus, our results reveal that TibA possesses a modular organization, with the extracellular domain being separated into an autoaggregation module and an adhesion module.

Introduction

Many *Escherichia coli* strains cause gastrointestinal diseases. Enterotoxigenic *E. coli* (ETEC) is the most common cause of diarrhea mediated by *E. coli* (30). ETEC causes diarrhea by secreting one or more heat-stable or heat-labile enterotoxins. However, the first step in pathogenesis is the attachment to the intestinal epithelium. For this, ETEC possesses mainly fimbrial adhesins but some ETEC strains also possess afimbrial adhesins (44). Furthermore, some ETEC strains, like strain H10407, are able to invade epithelial cells (10). Two different loci in this strain, *tia* (13) and *tib* (11, 27), were shown to mediate invasion (10). The *tib* locus codes for the expression of a 104 kDa surface protein, TibA, that promotes adhesion and invasion (11).

TibA is part of the autotransporter family of proteins (9, 17). Autotransporters are characterized by their organization and their secretion mechanism. They possess an N-

terminal signal sequence that targets the preprotein to the Sec complex and is cleaved after translocation across the inner membrane (4, 34, 42). They have a C-terminal domain that forms a β -barrel inserted into the outer membrane and helps the translocation of the extracellular domain across the outer membrane (25, 33). This passenger domain bears the function of the protein and is usually associated with virulence (16). Finally, at the C-terminus of the passenger domain and leading to the β -barrel, there is a junction region that is important for folding and secretion (32, 36).

Recently, TibA was proposed to be a member of a group called the self-associating autotransporters (SAATs) (22). This group includes two other *E. coli* autotransporters, the adhesin involved in diffuse adherence (AIDA-I) (2) and the aggregation factor Ag43 (15). The three proteins have been grouped together because of their functional similarities. They can all mediate bacterial autoaggregation, biofilm formation, as well as adhesion and invasion of epithelial cells (1, 5, 11, 15, 20, 35, 40, 41). The three proteins also share a peculiar primary structure: the N-terminus of the passenger domain of these three proteins is composed of repeats of the same 19-amino acids consensus sequence. In addition, SAATs can be O-glycosylated and this glycosylation is important for the adhesion properties as well as for the stability and the conformation of the protein (3, 6, 23, 27, 39). In the case of TibA, glycosylation is achieved by TibC, an heptosyltransferase encoded right upstream of *tibA* (28). Despite their similarities, the proteins also have differences: (i) Different numbers of repeats of the consensus sequence; (ii) differences in processing, since both AIDA-I and Ag43 are cleaved while TibA is not (5, 7, 18, 43); and (iii) TibA possesses a unique proline-rich region between the junction domain and the β -barrel (28).

Structure-function relationship studies have been conducted on AIDA-I (8) and Ag43 (21). In both cases, the study revealed that the different functions were not all linked together and that distinct modules in the passenger domain could be associated with different functions. However, there are differences in the organization of AIDA-I and Ag43. The N-terminus of the passenger domain is responsible for adhesion in the case of AIDA-I and for autoaggregation in the case of Ag43. In AIDA-I, autoaggregation seems to be associated with the C-terminus of the passenger domain. Furthermore, a second adhesion domain is present in the C-terminus of the passenger domain of AIDA-I. In Ag43, the C-terminus of the passenger domain is associated with biofilm formation. Thus, despite all the

similarities, there are differences between SAAT proteins that are poorly understood. Therefore, a structure-function relationship in TibA could clarify the organization of this family of virulence factors.

In this study, we gained information on the functionality of the TibA self-associating autotransporter. We have generated several insertion mutants in the extracellular portion of TibA, as well as deletion mutants. We found that the N-terminus of the passenger domain is associated with autoaggregation while the C-terminus of the passenger domain is associated with adhesion. We did not identify mutants that were affected for both adhesion and autoaggregation. However, mutations that affected invasion or biofilm formation also affected either adhesion or autoaggregation. This suggests that while adhesion and autoaggregation are distinct primary functions, invasion and biofilm formation may be secondary and related to these functions.

Material and Methods

Bacterial strains and plasmid

E. coli K-12 strains C600 (New England Biolabs) (F⁻ *thr-1 leuB6 thi-1 lacY1 supE44 rfbD1 fhuA21*) was used in this study. Plasmid pTgH allows expression of TibA and its specific glycosyltransferase. Plasmid pTgH was derived from the pTRC99A vector (Pharmacia Biotech). TibA expressed by pTgH contains a protein tag consisting of six histidine residues at the N-terminus of the passenger domain. The presence of the His tag did not affect the expression and the functionality of TibA (data not shown).

To construct five-amino-acid insertion mutants, the pTgH plasmid was randomly mutagenized using the GPS-LS linker scanning system (New England Biolabs) according to the instructions of the manufacturer. The presence of an insertion in the passenger domain or in the junction region of TibA was assessed by PCR with the following primers: 5'-GTCTGGAATGAATCCACAG and 5'-GTTATCCAGCGTCAATGC. DNA sequencing was used to ascertain the location and integrity of the insertion sequence and that there was no secondary unwanted mutations. The Δ PP and the Δ N deletions were introduced into pTgH by oligonucleotide-directed mutagenesis performed with a QuikChange II site-directed mutagenesis kit (Stratagene) using the primers 5'-

CTGGTATCTGAAGGCTGATACTGGTACATCGTCGTCTCCAGTGCG and 5'-GGCGCATCATCATCATCACCAGTTTGTCTCCAGTGGCGGC, respectively, and according to the instructions of the manufacturer.

Bacterial and cell culture growth conditions

Bacteria containing the different plasmids were grown on Luria-Bertani (LB) agar plates or in liquid LB medium containing 100 $\mu\text{g}\cdot\text{ml}^{-1}$ ampicillin. Bacterial cultures were grown at 30°C and growth was monitored by measuring the optical density at 600nm (OD₆₀₀). At an OD₆₀₀ of 0.8, the cultures were induced with 10 μM isopropyl- β -D-thiogalactopyranoside (IPTG). This low concentration of IPTG was used to limit the toxicity associated with overexpression of TibA and TibC. HEp-2 cells (ATCC CCL-23) were grown at 37°C with 5% CO₂ in Dulbecco's modified Eagle's medium (Gibco) containing 10 mM sodium pyruvate (Sigma), 10% bovine growth serum (HyClone), 2.5 $\mu\text{g}\cdot\text{ml}^{-1}$ amphotericin B (Fungizone), and 100 $\mu\text{g}\cdot\text{ml}^{-1}$ penicillin/streptomycin (Gibco).

SDS-PAGE and immunoblotting

Overnight cultures (5 ml) were grown, normalized, centrifuged for 10 min at 12,000 X g in microcentrifuge tubes, and the pellets were resuspended in 50 μl of Tris-buffered saline (TBS). Whole-cell lysates samples were then diluted in 2X SDS-PAGE loading buffer containing β -mercaptoethanol and denatured by heating at 100°C for 10 min. The samples were then separated by SDS-PAGE on 10% acrylamide gels. The gels were either stained with Coomassie blue or transferred to polyvinylidene fluoride membranes (Millipore). Immunodetection was performed with a serum raised against glycosylated AIDA-I diluted 1:60,000 in blocking buffer (5% skim milk, 50 mM Tris-HCl [pH 7.5], 150 mM NaCl, 0.05% Triton X-100). Alternatively, a His-tag specific antibody (Invitrogen) was used. Horseradish peroxidase-conjugated secondary antibodies (Sigma) were used according to the instructions of the manufacturer. Immune complexes were revealed using a 3,3',5,5'-tetramethylbenzidine solution for membranes (Sigma).

Functional assays

Autoaggregation, biofilm formation, adhesion, and invasion assays were performed as previously described (8). *E. coli* C600 harboring the pTRC99A vector, plasmid pTgH and

plasmid pTgH containing one insertion or a deletion in TibA were induced with 10 μ M IPTG at an OD₆₀₀ of 0.8 and grown overnight. In the autoaggregation assay, the overnight cultures were normalized in 5 ml of LB to an OD₆₀₀ of approximately 1.5 in culture tubes and left standing at 4°C. Samples (100 μ l) were taken 1 cm below the surface at the beginning of the assay and after 120 min, and the OD₆₀₀ of the samples were measured. OD₆₀₀ at 120 min were compared to OD₆₀₀ at the beginning of the assay and results are shown in percentage of initial OD₆₀₀. For TibA Δ PP, samples were collected every 30 min. In the biofilm formation assay, the overnight cultures were normalized in M9 minimal medium and grown, without IPTG, for 24 hours at 30°C in plastic microtiter plates. Biofilms were stained with crystal violet. After washes, the dye was then solubilized with a mixture of ethanol and acetone (80:20), and the absorption at 595 nm of the solution was measured. In the adhesion and invasion assays, the overnight cultures were inoculated onto monolayers of confluent HEp-2 cells in a 24-well plate (approximately 2.5×10^5 cells) using 10^6 colony forming units (CFU) per well. After 3h, the cells were washed with PBS, and the adhering bacteria were recovered with 100 μ l of Triton X-100 (1%), plated, and counted. The total number of bacteria present was also determined and did not vary significantly between mutants and wild-type (data not shown). In the invasion assay, fresh medium containing 100 μ g*ml⁻¹ gentamicin was added after the 3h, and the preparations were incubated for an additional 2h before recovering and plating. All functional assays were performed at least three times in duplicate or triplicate. For each assay, the results obtained with each mutant were compared to those obtained with His-TibA by performing an ANOVA and Dunnett post-tests using Prism 4.0 software (Graphpad Software).

Results

Generation of insertion mutants in TibA

To analyze structure-function relationships in TibA, we used a Tn7-derived transposon system to generate a library of mutants in TibA. The same system was used in a structure-function relationship study of AIDA-I (8). Insertion and excision of this transposon causes a five amino acid insertion in the protein. The mutagenesis was performed on a plasmid that allows the expression of the glycosyltransferase TibC and the

autotransporter TibA under the control of an IPTG-inducible promoter. TibA also possesses a His-tag at the N-terminus of the passenger domain.

We screened for mutants in the repeats of TibA as well as in the N-terminal half of the junction. The procedure yielded 64 insertion mutants. Of these, 29 different mutants were expressed at similar level as His-TibA as assessed by immunoblotting of whole cell lysate with an antibody raised against the His tag or with an antibody raised against glycosylated AIDA-I (data not shown). Our anti-AIDA-I antibody recognizes AIDA-I only when glycosylated (3) and cross-reacts with glycosylated TibA (29). Therefore, we concluded that these 29 mutants were still glycosylated and that these insertions resulted in the expression of stable proteins. We could not detect protein expression for the remaining mutants (data not shown). The mutagenesis procedure creates a stop codon in one of the three possible reading frames of the insertion. Therefore, for the majority of these mutants, lack of expression is due to truncation at the insertion site. Some insertions may also have yielded unstable proteins.

Insertions were spread throughout the repeats, but there were only few insertions in the junction (Fig. 1 and Table I). The junction region is important for the folding of the passenger domain of autotransporters (32, 36). Therefore, most of the insertions in the junction region may have yielded proteins with a folding problem, which would have lead to the degradation of the protein. The same observation was made with AIDA-I, another SAAT; the repeats were permissive for insertions, but outside the repeats, only a few stable insertions were observed (8). Passenger domains of autotransporters are generally rich in β -strands and most of the known structures for passenger domain have a β -helix backbone (31). A model for TibA has been proposed where each 19 amino acid repeat is composed of three β -strands and form one coil of the β -helix (19). Placing the insertions on this model (Fig. 2) shows that the insertions occurred in the loops as well as in the three strands. In the latter case, it is possible that the sequences of the 5-amino acid linker could insert providing an inward facing hydrophobic residue that maintains the integrity of the hydrophobic core. All three strands contained insertions that affected functionality, suggesting that the functions of TibA are not localized on a particular side of the β -helix.

Functionality of the insertion mutants

To determine the effect of the insertions on the functionality of TibA, we tested each mutant for its ability to mediate autoaggregation, to form biofilm, to adhere to cultured epithelial cells and to invade these cells. Figure 3 shows the results for each mutant that had a defect in at least one of the functions. We included one mutant (i1) that had the same functionality as His-TibA as a control. Other mutants that are not shown were as functional as His-TibA in the four assays.

Of the 29 mutants, 17 mutants had a defect in either adhesion or autoaggregation, but no mutants were affected for both phenotypes. All mutants defective for one or more functions were found in the repeats. From these 17 mutants, ten (i14 to i25) had a defect in adhesion and the seven others (i2 to i13) had a defect in autoaggregation. Strikingly, the separation between autoaggregation-defective mutants and adhesion-defective mutants is very clear and lies between mutant i13 and i14. These results suggest that the N-terminal portion of the repeats of TibA is responsible for autoaggregation while the C-terminal part is responsible for adhesion.

Some of these mutants had in addition a defect in invasion or in biofilm formation. However, we did not find mutants that were only affected in invasion or in biofilm formation. Four mutants (i8, i9, i18 and i19) had reduced invasion. It was not possible to associate invasion with a particular region or function (adhesion or autoaggregation), since invasion-defective mutants were found in the N-terminal part associated with autoaggregation (i8, i9) as well as in the C-terminal part associated with adhesion (i18, i19). Interestingly, some mutants deficient for adhesion could still promote invasion (i14-i17 and i22-i25). Adhesion deficient mutants that are still able to invade were also isolated for AIDA-I (8). It suggests that adhesion and invasion are not necessarily related to each other. Biofilm formation and autoaggregation are not strictly linked since many autoaggregation defective mutants still make biofilms. However, it seems it is possible that biofilm formation is to some degree associated with autoaggregation, since 3 of the 4 biofilm-defective mutants (i4, i8, i13 and i22) were also affected in autoaggregation. Since biofilm formation is performed in minimal medium, we ascertained that variations in protein expression levels or growth rates did not account for the defect of these mutants under these conditions (data not shown).

Taken together, our results suggest that the repeats of TibA contain an N-terminal autoaggregation domain and a C-terminal adhesion domain. While it was clear for adhesion and autoaggregation, it is not possible to associate biofilm formation and invasion with a particular region of the repeats. Also, it seems that the functionality of the protein is limited to the repeats and does not extend to the junction region.

Effect of an N-terminal deletion of the repeats

All mutants with a defect in autoaggregation were found in the N-terminal portion of the repeats. In order to confirm that this region contains the autoaggregation domain, we deleted amino acids A55 to Q185 to create TibA Δ N (Fig. 4A). The mutant was correctly expressed and glycosylated, as seen by total protein extract probed with antibodies against the His-tag or recognizing glycosylation (anti-AIDA-I) (Fig. 4B). We then assessed the autoaggregation and adhesion properties of TibA Δ N (Fig. 4C and 4D). While the deletion of the N-terminal half of the passenger domain did not affect the adhesion properties of TibA, TibA Δ N was not able to mediate autoaggregation anymore. This confirms that the domain responsible for autoaggregation is found in the N-terminal part of the repeats of TibA while the adhesion domain is in the C-terminal part of the repeats.

Role of the proline-rich region

One difference between TibA and other SAATs is the presence of a proline-rich region between the junction region and the membrane-embedded domain. To determine the role of this domain in TibA, we deleted it to create TibA Δ PP (Fig. 5A). The mutant was correctly expressed and glycosylated (Fig. 5B). We then tested the functionality of the mutant in autoaggregation, adhesion, invasion and biofilm formation (Fig. 5C-F). TibA Δ PP was not able to mediate adhesion to cultured epithelial cells, but was still able to mediate autoaggregation, biofilm formation and invasion of cultured epithelial cells. TibA Δ PP is another example of a mutant that is defective in adhesion, but is as invasive as His-TibA. Again, this suggests that adhesion and invasion are not necessarily related to each other. The main conclusion of these results is that the proline-rich region is involved in adhesion.

Discussion

In this study, we created different mutations in the self-associating autotransporter TibA to gain information on the functionality of this protein. We first generated 29 insertion mutants in the passenger domain and in the junction region of TibA. Analysis of the functionality of these mutants revealed that the N-terminal portion of the repeats is responsible for autoaggregation while the C-terminal part is associated with adhesion. The autoaggregation domain in Ag43 is also found in the N-terminal portion of the repeats (21) while, in AIDA-I, the N-terminus of the mature protein contains an adhesion domain (8). A region responsible for adhesion was not sought in Ag43. The organization of TibA seems to be closer to the organization of Ag43. Both proteins also have a similar length compared to AIDA-I, which possesses a passenger domain that is twice the size of those of TibA and Ag43 (Fig 6).

Adhesion domains are often located at the N-terminus of adhesins, like AIDA-I, (8) or the trimeric autotransporter YadA from *Yersinia enterocolitica* (38). Having the adhesion domain farther away from the cell surface may facilitate the binding of the adhesin to its receptor. In the case of TibA, the adhesion domain would be closer to the membrane. However, TibA also possesses a proline-rich domain between the extracellular and the membrane-embedded portions. This domain could serve as a spacer so that the passenger domain is farther away from the membrane. Consistent with this hypothesis, deletion of this proline-rich region completely abrogates the ability of TibA to mediate adhesion to epithelial cells. TonB is another example of an *E. coli* protein that harbors a proline-rich region. In the case of TonB, this proline-rich region was shown to adopt an elongated and rigid conformation (12, 24) and its role seems to be limited to providing a physical extension to TonB (26).

In the present study, it is interesting to note that we did not observe mutants that were affected only for invasion or biofilm formation. All mutants affected for these functionalities were also affected in either autoaggregation or adhesion. Similar observations were made in previous studies with AIDA-I (8, 14) and Ag43 (21, 35). This suggests that adhesion and autoaggregation are the main functions of SAATs, whereas biofilm formation and invasion are secondary to these two functions.

In order to invade eukaryotic cells, bacteria need to first adhere to the cells. Therefore, a mutant that is deficient for adhesion is likely to be also deficient for invasion. However, in this study, we isolated several mutants, that were not able to adhere to epithelial cells but could still promote invasion (i16, i17 and TibA Δ PP for instance). All these mutants are still able to promote autoaggregation. Therefore, it is tempting to hypothesize that there is a link between invasion and the two main functions of TibA, adhesion and autoaggregation. For instance, the invasion mediated by TibA could be invasion of bacterial aggregates, as can be seen with *Bartonella hensalae* (37). More work is needed to test this hypothesis.

In conclusion, our study suggests that SAATs evolve by acquisition and/or modification of individual modules that bring new functionalities, either independently or in conjunction with existing modules. In this light, TibA would bear an autoaggregation, an adhesion and a spacer module, along with the junction region and membrane-embedded domain that could be construed as a secretion module (Fig. 7). Other SAATs would stem from a different combination of equivalent modules. This modular aspect is a complexified version of the passenger-translocator organization that is characteristic of all autotransporters, but could represent an evolutionary mechanism that is peculiarly well suited for these proteins.

Acknowledgements

This work was supported by Canadian Institutes for Health Research Grant 84578, funds from the Groupe de Recherche et d'Etudes sur les Maladies Infectieuses du Porc, Canada Research Chair and Canada Foundation for Innovation Grant 201414, and by graduate fellowship to J.-P.C. from the Fonds de Recherche sur la Nature et les Technologies du Québec (FQRNT 126554).

References

1. 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* 6:1539-1546.
2. 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* 57:1506-1511.
3. 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* 40:1403-1413.
4. Brandon, L. D., N. Goehring, A. Janakiraman, A. W. Yan, T. Wu, J. Beckwith, and M. B. Goldberg. 2003. IcsA, a polarly localized autotransporter with an atypical signal peptide, uses the Sec apparatus for secretion, although the Sec apparatus is circumferentially distributed. *Mol Microbiol* 50:45-60.
5. Charbonneau, M. E., F. Berthiaume, and M. Mourez. 2006. Proteolytic processing is not essential for multiple functions of the *Escherichia coli* autotransporter adhesin involved in diffuse adherence (AIDA-I). *J Bacteriol* 188:8504-8512.
6. Charbonneau, M. E., V. Girard, A. Nikolakakis, M. Campos, F. Berthiaume, F. Dumas, F. Lepine, and M. Mourez. 2007. O-linked glycosylation ensures the normal conformation of the autotransporter adhesin involved in diffuse adherence. *J Bacteriol* 189:8880-8889.
7. Charbonneau, M. E., J. Janvare, and M. Mourez. 2009. Autoprocessing of the *Escherichia coli* AIDA-I autotransporter: a new mechanism involving acidic residues in the junction region. *J Biol Chem* 284:17340-17351.
8. Charbonneau, M. E., and M. Mourez. 2007. Functional organization of the autotransporter adhesin involved in diffuse adherence. *J Bacteriol* 189:9020-9029.
9. Dautin, N., and H. D. Bernstein. 2007. Protein secretion in gram-negative bacteria via the autotransporter pathway. *Annu Rev Microbiol* 61:89-112.
10. Elsinghorst, E. A., and D. J. Kopecko. 1992. Molecular cloning of epithelial cell invasion determinants from enterotoxigenic *Escherichia coli*. *Infect Immun* 60:2409-2417.

11. Elsinghorst, E. A., and J. A. Weitz. 1994. Epithelial cell invasion and adherence directed by the enterotoxigenic *Escherichia coli* tib locus is associated with a 104-kilodalton outer membrane protein. *Infect Immun* 62:3463-3471.
12. Evans, J. S., B. A. Levine, I. P. Trayer, C. J. Dorman, and C. F. Higgins. 1986. Sequence-imposed structural constraints in the TonB protein of *E. coli*. *FEBS Lett* 208:211-216.
13. 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* 64:2256-2265.
14. Girard, V., J. P. Cote, M. E. Charbonneau, M. Campos, F. Berthiaume, M. A. Hancock, N. Siddiqui, and M. Mourez. 2010. Conformation change in a self-recognizing autotransporter modulates bacterial cell-cell interaction. *J Biol Chem* 285:10616-10626.
15. 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* 149:115-120.
16. Henderson, I. R., and J. P. Nataro. 2001. Virulence functions of autotransporter proteins. *Infect Immun* 69:1231-1243.
17. 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* 68:692-744.
18. 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* 181:2132-2141.
19. Kajava, A. V., and A. C. Steven. 2006. 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.
20. Kjaergaard, K., M. A. Schembri, C. Ramos, S. Molin, and P. Klemm. 2000. Antigen 43 facilitates formation of multispecies biofilms. *Environ Microbiol* 2:695-702.

21. 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* 51:283-296.
22. Klemm, P., R. M. Vejborg, and O. Sherlock. 2006. Self-associating autotransporters, SAATs: functional and structural similarities. *Int J Med Microbiol* 296:187-195.
23. 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* 412:563-577.
24. Kohler, S. D., A. Weber, S. P. Howard, W. Welte, and M. Drescher. 2010. The proline-rich domain of TonB possesses an extended polyproline II-like conformation of sufficient length to span the periplasm of Gram-negative bacteria. *Protein Sci* 19:625-630.
25. 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* 80:19-34.
26. Larsen, R. A., G. E. Wood, and K. Postle. 1993. The conserved proline-rich motif is not essential for energy transduction by *Escherichia coli* TonB protein. *Mol Microbiol* 10:943-953.
27. Lindenthal, C., and E. A. Elsinghorst. 2001. Enterotoxigenic *Escherichia coli* TibA glycoprotein adheres to human intestine epithelial cells. *Infect Immun* 69:52-57.
28. Lindenthal, C., and E. A. Elsinghorst. 1999. Identification of a glycoprotein produced by enterotoxigenic *Escherichia coli*. *Infect Immun* 67:4084-4091.
29. Moormann, C., I. Benz, and M. A. Schmidt. 2002. Functional substitution of the TibC protein of enterotoxigenic *Escherichia coli* strains for the autotransporter adhesin heptosyltransferase of the AIDA system. *Infect Immun* 70:2264-2270.
30. Nataro, J. P., and J. B. Kaper. 1998. Diarrheagenic *Escherichia coli*. *Clin Microbiol Rev* 11:142-201.

31. Nishimura, K., N. Tajima, Y. H. Yoon, S. Y. Park, and J. R. Tame. 2010. Autotransporter passenger proteins: virulence factors with common structural themes. *J Mol Med* 88:451-458.
32. 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* 47:1367-1383.
33. 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* 23:1257-1266.
34. 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* 281:9038-9048.
35. 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* 299:389-401.
36. Renn, J. P., and P. L. Clark. 2008. A conserved stable core structure in the passenger domain beta-helix of autotransporter virulence proteins. *Biopolymers* 89:420-427.
37. Rhomberg, T. A., M. C. Truttmann, P. Guye, Y. Ellner, and C. Dehio. 2009. A translocated protein of *Bartonella henselae* interferes with endocytic uptake of individual bacteria and triggers uptake of large bacterial aggregates via the invasome. *Cell Microbiol* 11:927-945.
38. Roggenkamp, A., K. Ruckdeschel, L. Leitritz, R. Schmitt, and J. Heesemann. 1996. Deletion of amino acids 29 to 81 in adhesion protein YadA of *Yersinia enterocolitica* serotype O:8 results in selective abrogation of adherence to neutrophils. *Infect Immun* 64:2506-2514.
39. 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* 188:1798-1807.

40. Sherlock, O., M. A. Schembri, A. Reisner, and P. Klemm. 2004. Novel roles for the AIDA adhesin from diarrheagenic *Escherichia coli*: cell aggregation and biofilm formation. *J Bacteriol* 186:8058-8065.
41. 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* 73:1954-1963.
42. Sijbrandi, R., M. L. Urbanus, C. M. ten Hagen-Jongman, H. D. Bernstein, B. Oudega, B. R. Otto, and J. Luirink. 2003. Signal recognition particle (SRP)-mediated targeting and Sec-dependent translocation of an extracellular *Escherichia coli* protein. *J Biol Chem* 278:4654-4659.
43. Suhr, M., I. Benz, and M. A. Schmidt. 1996. Processing of the AIDA-I precursor: removal of AIDA_c and evidence for the outer membrane anchoring as a beta-barrel structure. *Mol Microbiol* 22:31-42.
44. Turner, S. M., A. Scott-Tucker, L. M. Cooper, and I. R. Henderson. 2006. Weapons of mass destruction: virulence factors of the global killer enterotoxigenic *Escherichia coli*. *FEMS Microbiol Lett* 263:10-20.

Figure Legends

Figure 1: Position of the insertion mutants in the passenger domain and the junction region of TibA.

The effects on functionality of the insertions are summarized. Mutants that are affected functionally are shown in black and mutants that did not are shown in gray.

Figure 2: Localization of the insertion mutants in the coil model of TibA.

One repeat of the consensus sequence is predicted to form three β -strands corresponding to one coil of a β -helix. A filled circle represents an isoleucine, a leucine or a valine residue, while “X” represents any residue. Insertion mutants are positioned in this model. Mutants with a defect in auto-aggregation (white box) and mutants with a defect in adhesion (black box) are highlighted. The number as well as the percentage of insertions in each strand is shown below.

Figure 3: Functionality of the insertion mutants

(A) Auto-aggregation assay: *E coli* C600 bearing an empty vector (-) or plasmids allowing expression of His-TibA or the different insertion mutants were normalized to an OD₆₀₀ of 1.5 and left standing. The OD₆₀₀ was measured at the beginning of the assay and after 2h. Results are shown in percentage of initial OD₆₀₀. (B) Adhesion assay: Bacteria were inoculated onto monolayers of confluent Hep-2 cells. After 3h, the adhering bacteria were recovered, plated and counted. Results represent the percentage of adhered bacteria compared to His-TibA. (C) Invasion assay: After adhesion, extracellular bacteria were killed by addition of gentamicin and invaded bacteria were recovered, plated and counted. Results represent the percentage of gentamicin resistant bacteria compared to His-TibA. (D) Biofilm formation assay: Biofilms were stained with crystal violet and the amount of fixed dye is represented as a percentage compared to His-TibA. ANOVA and Dunnett post-tests were used to identify significant differences from WT (white bars; p<0.05).

Figure 4: Effect of the deletion of the N-terminal portion of the passenger domain of TibA.

(A) Schematic representation of TibA and the N-terminal deletion mutant TibA Δ N. (B) Whole-cell lysates of bacteria bearing an empty vector (-), a plasmid allowing the expression of His-TibA or the N-terminal deletion mutant (Δ N) were probed with an antibody raised against glycosylated AIDA-I (upper panel) or against the His-tag (lower panel). (C) Auto-aggregation and (D) adhesion assays were performed as described in Fig.3. ANOVA and Dunnett post-tests were used to distinguish significant (*; p<0.05) from non-significant differences (ns)

Figure 5: Construction and expression of TibA lacking the proline-rich region.

(A) Schematic representation of TibA and the proline-rich region deletion mutant TibA Δ PP. (B) Whole-cell lysates of bacteria bearing an empty vector (-), a plasmid allowing the expression of His-TibA or the proline-rich region deletion mutant (Δ PP) were probed with an antibody raised against glycosylated AIDA-I (upper panel) or against the His-tag (lower panel). (C) Autoaggregation assay, (D) adhesion assay, (E) invasion assay and (F) biofilm formation assay were performed as described in Fig.3 on bacteria bearing

an empty vector (-; circle) or plasmids allowing expression of TibA (WT; square) or TibA deleted for the proline-rich region (Δ PP; triangle). ANOVA and Dunnett post-tests were used to identify significant differences (*; $p < 0.05$) and non-significant differences (ns).

Figure 6: Functional organization of SAATs.

Functional domains found by structure-function studies of AIDA-I (8), TibA (this study) and Ag43 (21) are positioned on a schematic representation of each SAAT. Domains responsible for secretion (blue), autoaggregation (red), adhesion (green) and biofilm formation (orange) are represented. No specific domain for invasion was found in these studies.

Table

Table I: Position and sequence of the insertion mutants

Mutants	Position of the insertion	Inserted sequence
i1	D 58	MFKHD
i2	S 94	CLNTS
i3	N 104	CLNIN
i4	Q 109	LFKHQ
i5	S 119	CLNTS
i6	S 132	CLNTS
i7	N 143	CLNNN
i8	G 145	RLFKQ
i9	Q 147	CLNRQ
i10	G 164	VFKQG
i11	V 168	MFKHV
i12	Q 185	LFKQQ
i13	I 192	CLNSI
i14	K 194	CLNIK
i15	S 227	CLNTS
i16	T 250	CLNTT
i17	I 254	LFKHI
i18	S 263	CLNIS
i19	Y 274	CLNIY
i20	M 297	CLNSM
i21	Q 311	MFKHQ
i22	S 335	CLNIS
i23	A 337	VFKHA
i24	L 346	VFKQL
i25	L 352	CLNNL
i26	D 356	CLNID
i27	V 392	VFKHV
i28	L 435	LFKQL
i29	T 466	CLNST

Figures

Figure 1

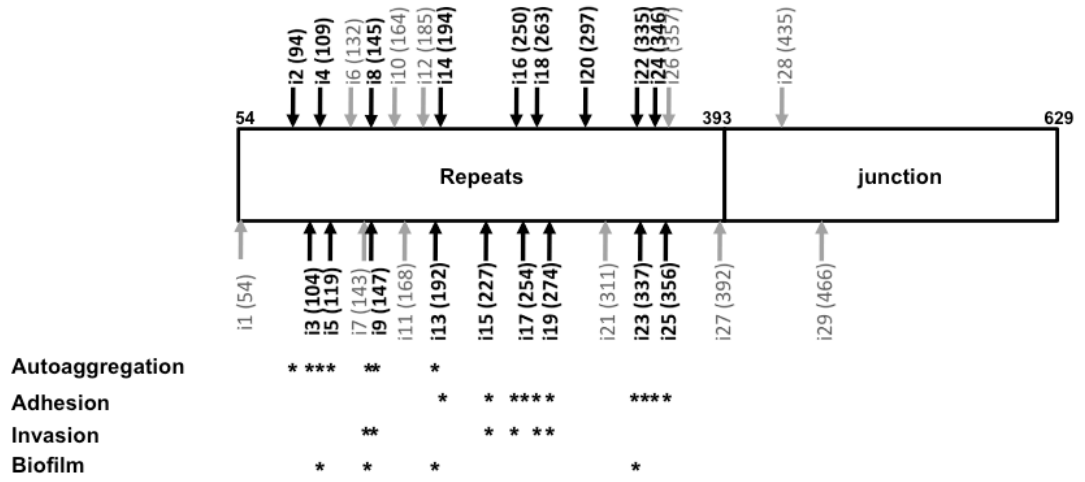
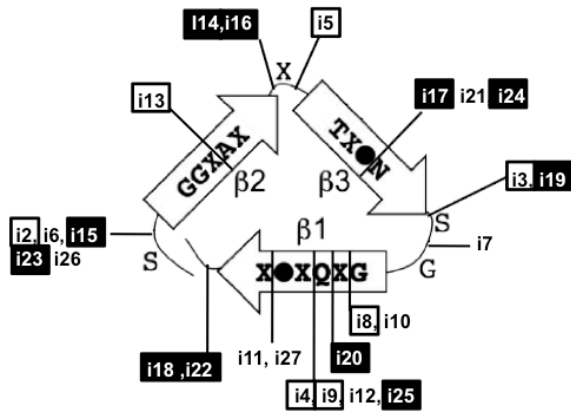


Figure 2



Functionality	$\beta 1$	$\beta 2$	$\beta 3$
Affected	6 (60%)	6 (75%)	5 (71%)
Not affected	4	2	2
Total	10 (40%)	8 (32%)	7 (28%)

Figure 3

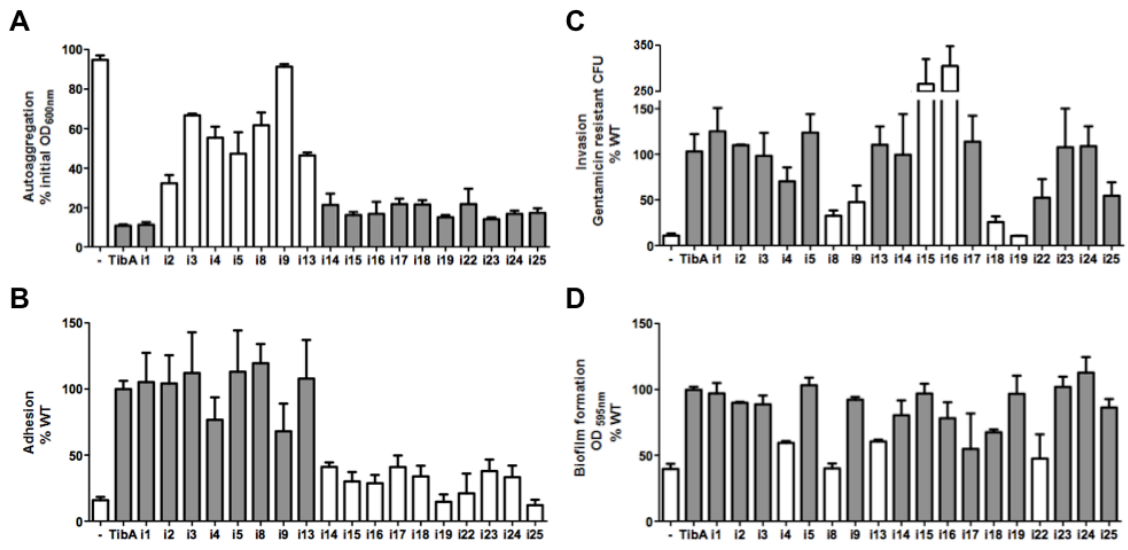


Figure 4

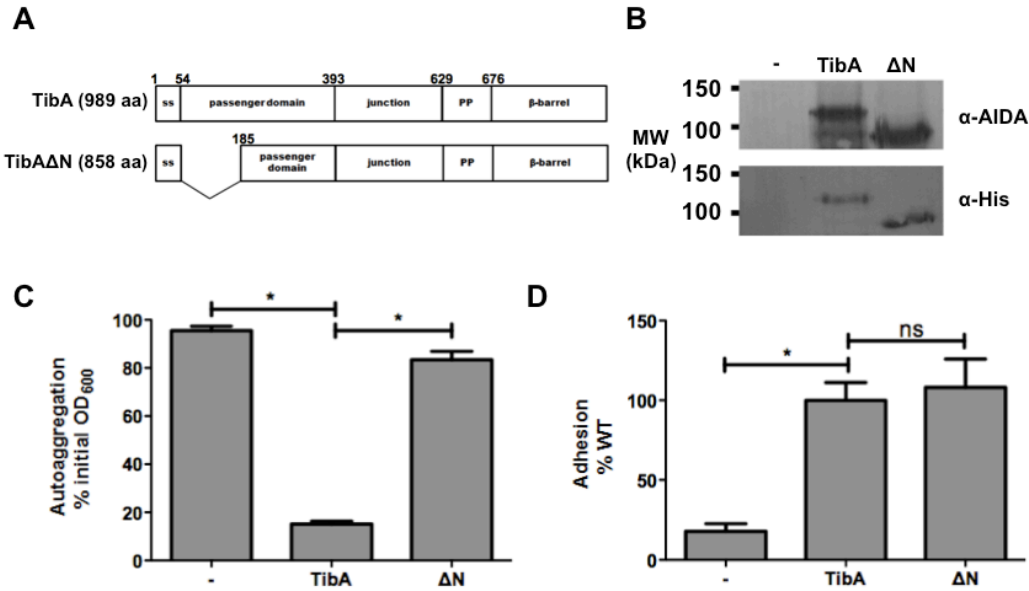


Figure 5

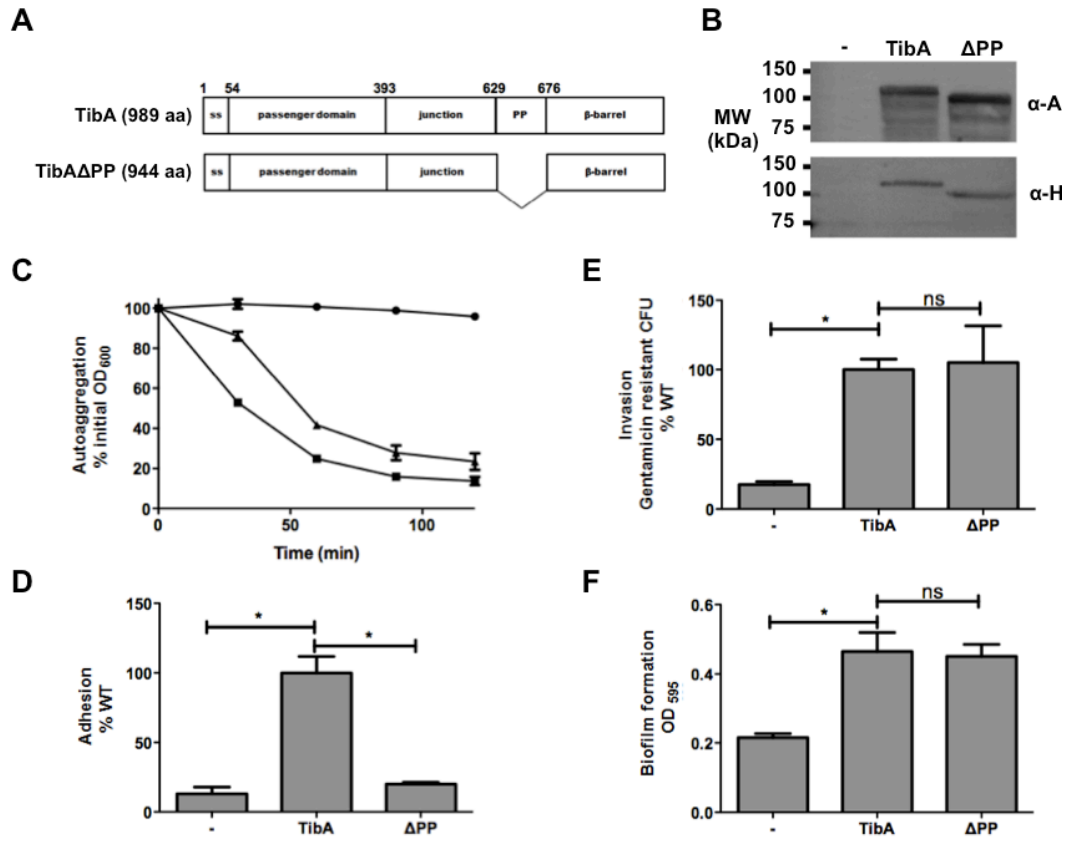
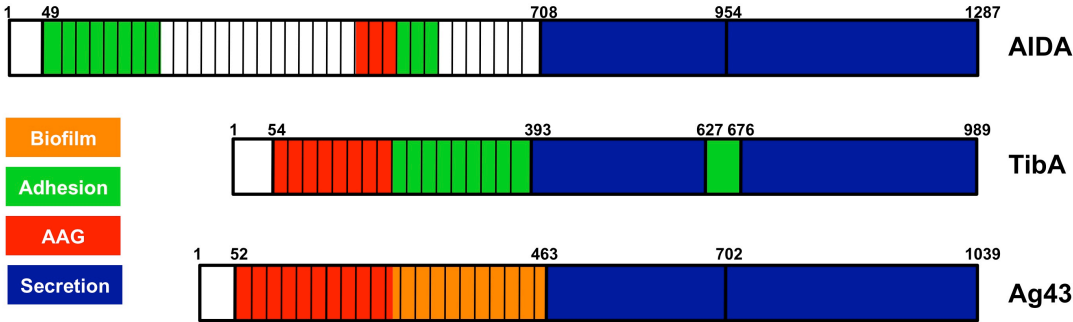


Figure 6



Article 2 – Glycosylation de TibA

Glycosylation of the *Escherichia coli* TibA self-associating autotransporter influences the conformation and the functionality of the protein

PLOS one, sous presse, PONE-D-13-24816R1

Avant-propos

La deuxième partie de la caractérisation de TibA consiste en l'étude de sa glycosylation. Encore une fois, l'effet de la glycosylation a été étudié sur AIDA-I et Ag43, mais pas sur TibA et les conclusions des études sont plutôt contradictoires. Notre hypothèse est qu'en étudiant TibA, une protéine naturellement glycosylée pour laquelle la forme non-glycosylée est assez stable pour être exprimée, nous allons pouvoir déterminer le rôle de la glycosylation des SAATs. Nos résultats démontrent que la glycosylation permet la modulation entre les deux conformations de TibA (voir l'article 4). Lorsque non-glycosylé, TibA se retrouve en conformation oligomérique et est incapable d'adopter sa conformation monomérique, et par le fait même, de permettre l'adhésion aux cellules épithéliales. Ces résultats complètent les études précédentes portant sur la glycosylation des SAATs et permet d'établir que la glycosylation affecte principalement la conformation et la structure des SAATs. L'effet sur la fonctionnalité est donc indirect.

Contributions

Contribution pour cet article :

Comme premier auteur et auteur principal de l'article, j'ai effectué la majorité des expérimentations, à l'exception du dosage du stress cytoplasmique et extracytoplasmique. J'ai analysé les résultats et j'ai écrit le manuscrit.

Glycosylation of the *Escherichia coli* TibA self-associating autotransporter influences the conformation and the functionality of the protein

Jean-Philippe Côté[#], Marie-Ève Charbonneau and Michael Mourez¹

Faculté de Médecine Vétérinaire, Université de Montréal, 3200 Sicotte, Saint-Hyacinthe, Québec J2S 7C6, Canada.

¹ Present address: Sanofi-Aventis, R & D, 195 route d'Espagne, Toulouse 31036, France

Running Title: Glycosylation of TibA

Abstract

The self-associating autotransporters (SAATs) are multifunctional secreted proteins of *Escherichia coli*, comprising the AIDA-I, TibA and Ag43 proteins. One of their characteristics is that they can be glycosylated. Glycosylation of AIDA-I and Ag43 have been investigated, but not that of TibA. It is still not clear whether glycosylation of the SAATs affect their structure or their functionality. Therefore, we have looked at the effects of glycosylation on the TibA adhesin/invasin. TibA is glycosylated by TibC, a specific glycosyltransferase, and the two genes are encoded in an operon. In this study, we have found that the glycosylation of TibA is not limited to the extracellular functional domain, as previously observed with AIDA-I and Ag43. We have determined that unglycosylated TibA is not able to promote the adhesion of bacteria on cultured epithelial cell, even though it is still able to promote invasion, biofilm formation and autoaggregation of bacteria. We have purified the glycosylated and unglycosylated forms of TibA, and determined that TibA is less stable when not glycosylated. We finally observed that glycosylation affects the oligomerisation of TibA and that unglycosylated TibA is locked in a conformation that is not suited for adhesion. Our results suggest that the effect of glycosylation on the functionality of TibA is indirect.

Introduction

Glycosylation is one the most abundant protein modification in eukaryotes and is also an important feature of many bacterial pathogens [1,2]. Many of the bacterial proteins that are glycosylated are secreted or surface exposed and play a role in the interaction of the bacterium with its environment [3]. In some cases, bacteria harbor a complex glycosylation pathway. For instance, in the general N-glycosylation pathway in *Campylobacter jejuni*, up to 10 proteins are involved in the synthesis and transfer of a heptasaccharide to an asparagine residue of the acceptor polypeptide [4]. There are more than 65 different proteins glycosylated by this pathway in *C. jejuni* [5]. Similarly, complex O-glycosylation pathways also exist in bacteria [6]. These pathways often lead to the modification of pilin

or flagellin subunits. Alternatively, there are examples of glycosylation system that are much more simple. These systems are composed of a single glycosyltransferase that glycosylates a single substrate using sugar precursors hijacked from other pathways. The HMW1c glycosyltransferase of *Haemophilus influenzae* is an example of such a simple system: HMW1c uses a sugar precursor from the lipooligosaccharide pathway to glycosylate the HMW1 adhesin [7].

Recently, a new family of these simple O-glycosyltransferases have been described in various proteobacteria [8]. Members of this family include the *Escherichia coli* adhesin associated heptosyltransferase (Aah) and the glycosyltransferase TibC. These two glycosyltransferases function similarly by adding heptoses hijacked from a precursor in the lipopolysaccharide pathway to their substrates, the autotransporters AIDA-I and TibA respectively [9].

Autotransporters are secreted virulence factors of Gram-negative bacteria that possess a modular organization [10,11]. They are composed of (i) an N-terminal signal sequence for inner membrane translocation through the sec translocon, (ii) an extracellular passenger domain, which is the functional domain of the autotransporter [12] and (iii) a C-terminal domain that forms in the outer membrane a β -barrel plugged by an α -helix. Furthermore, between the passenger domain and the β -barrel, there is often a junction region that is extracellular and helps with the secretion and the folding of the passenger domain [13]. In the case of TibA, an additional proline-rich domain separates the typical junction region from the C-terminal domain.

TibA is a member of a sub-family of autotransporters called the self-associating autotransporters (SAATs) [14]. The Adhesin Involved in Diffuse Adherence (AIDA-I) and the aggregation factor Ag43 are also members of the SAATs family. These proteins can be glycosylated [15,16] and are multifunctional, as they mediate the autoaggregation of bacteria, the formation of biofilm and the adhesion and invasion of cultured epithelial cells [17,18,19,20]. They also share sequence similarities; notably, a 19-amino acid sequence that is imperfectly repeated in their passenger domain.

The gene coding for TibA is found in an operon with its glycosyltransferase TibC [21]. *aidA* is also encoded in an operon with its glycosyltransferase *aah* [22]. Aah adds heptose residues to serine or threonine in the passenger domain of AIDA-I and can also

glycosylate Ag43 and TibA, suggesting that Aah and TibC are structurally and functionally similar. In fact, Aah and TibC are functionally interchangeable [23]. Glycosylation of AIDA-I by Aah is highly heterogeneous and no consensus sequence for glycosylation by Aah was found [15]. Instead, it has recently been determined that Aah recognizes a structural motif rather than a sequence [8].

Loss of glycosylation renders AIDA-I or TibA unable to promote adhesion to cultured epithelial cells [15,24]. However, the consequences of glycosylation on the functionality of AIDA-I are likely to be an indirect effect, since when it is not glycosylated, AIDA-I is poorly expressed at the surface of the bacteria. Indeed, loss of the adhesion properties is due to the inability of unglycosylated AIDA-I to adopt a conformation preventing its degradation. An effect of glycosylation on the conformation of Ag43 was also observed, but Ag43 is not associated to a specific glycosyltransferase [25]. However, in these two studies using AIDA-I and Ag43, the authors could not determine if glycosylation also affected the function of the SAAT as well as its conformation, as it is the case for the trimeric autotransporter EmaA [26]. Thus, it is still unclear how and why glycosylation affects the functionality of SAATs.

TibA is an ideal model to study the glycosylation of the SAATs. TibA is associated with its own glycosyltransferase, unlike Ag43; and the unglycosylated form is well expressed at the cell surface and can be purified, as opposed to AIDA-I. In this study, we determined that glycosylation has an impact on the stability and the conformation of TibA. We showed that, when not glycosylated, TibA is locked in an oligomeric conformation that is not suited for adhesion. Therefore, our study completes previous ones and shows that glycosylation of the SAATs allows the modulation between the different conformations of these proteins.

Materials and Methods

Bacterial strains and plasmids

E. coli K-12 strain C600 (New England Biolabs) (F- thr-1 leuB6 thi-1 lacY1 supE44 rfbD1 fhuA21) was used in this study. Plasmid pTgH allows the expression of a his-tagged TibA and its specific glycosyltransferase, TibC [18]. The His-tag was inserted at the N-terminus

of the passenger domain of TibA and does not affect the expression, the glycosylation or the functionality of TibA [18]. Plasmid pTgH C358R was derived from pTgH and contains a mutation at position 358 in TibC that inactivates the enzyme [23]. Plasmid pTngH allows the expression of his-tagged TibA only and was constructed by site-directed deletion of tibC from the plasmid pTgH using the primer Δ tibC (5'-GAATTGGAGCGGATAACAATTTACACAGGAGTAAGCAATGAATAAGGTC-3') and its complement using the Quickchange mutagenesis kit (Stratagene).

Bacterial and mammalian cell growth conditions

Bacteria containing the different plasmids were grown on Luria-Bertani (LB) agar plates or in liquid LB medium containing $100 \mu\text{g}\cdot\text{mL}^{-1}$ ampicillin. Bacterial cultures were grown at 30°C and growth was monitored by measuring the optical density at 600nm (OD_{600}). At an OD_{600} of 0.8, the cultures were induced with $10 \mu\text{M}$ isopropyl- β -D-thiogalactopyranoside (IPTG) and growth was pursued overnight. This low concentration of IPTG was used to limit the toxicity associated with overexpression of TibA and TibC. HEp-2 cells (ATCC CCL-23) were grown at 37°C with 5% CO_2 in Dulbecco's modified Eagle's medium (Gibco) containing 10 mM sodium pyruvate (Sigma), 10% bovine growth serum (HyClone), $2.5 \mu\text{g}\cdot\text{mL}^{-1}$ amphotericin B (Fungizone), and $100 \mu\text{g}\cdot\text{mL}^{-1}$ penicillin/streptomycin (Gibco).

SDS-PAGE and Immunoblotting

Overnight cultures (5 mL) were grown, normalized, centrifuged for 10 min at 12,000 X g in microcentrifuge tubes, and the pellets were resuspended in 50 μL of Tris-buffered saline (TBS; 50 mM Tris-HCl pH 8 and 150 mM NaCl). Whole-cell extracts were then diluted in twice concentrated SDS-PAGE loading buffer containing β -mercaptoethanol and denatured by heating at 100°C for 10 min. The samples were then separated by SDS-PAGE on 10% acrylamide gels. The gels were either stained with Coomassie blue or transferred to polyvinylidene fluoride membranes (Millipore). Immunodetection was performed with a serum raised against glycosylated AIDA-I and cross-reacting with glycosylated TibA diluted 1:50,000 in blocking buffer (5% skim milk, 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.05% Triton X-100). Alternatively, a His-tag specific antibody (Invitrogen) was

used. Horseradish peroxidase-conjugated secondary antibodies (Sigma) were used according to the instructions of the manufacturer. Immune complexes were revealed using a 3,3',5,5'-tetramethylbenzidine solution for membranes (Sigma). Bands were quantified by densitometry using the ImageJ software [27] and were compared to glycosylated TibA.

Protein purification

To purify glycosylated TibA or unglycosylated TibA, one liter of *E. coli* C600 bearing the plasmid pTgH or pTngH, respectively, were grown to an OD₆₀₀ of 0.8 and induced with 10 μM IPTG overnight. Bacteria were harvested and resuspended in 50 mL TBS containing a protease inhibitor cocktail (complete; Roche). Cells were lysed using a French press and an ultrasonicator. Lysates were recovered by a slow speed centrifugation to clear cellular debris and were further centrifuged for 30 min at 250,000 X g. The pellets, containing the membranes, were then resuspended in 25 mL TBS containing 2% Triton X-100 reduced (Sigma) and 10 mM EDTA. The suspensions were incubated for 30 min at 37°C and centrifuged again for 30 min at 250,000 X g. EDTA was removed from the supernatant using PD-10 columns (GE Healthcare) and His-tagged TibA was purified by affinity chromatography using a 1 mL His-Trap HP column (GE Healthcare). Purity of the purified proteins was consistently over 95% as evaluated by Coomassie blue stained SDS-PAGE. Finally, the protein samples were dialyzed to replace the Triton X-100 with 1% n-octyl-β-D-glucoside (Roche).

Mass Spectrometry

Glycosylated his-tagged TibA was purified from *E. coli* C600 as described above and run on an SDS-PAGE 10% acrylamide gel. The protein band corresponding to TibA was cut from the gel and destained with water-sodium bicarbonate buffer and acetonitrile. The protein was reduced with di-thiothreitol and alkylated with iodoacetamide prior to in-gel digestion with trypsin. The tryptic peptides were eluted from the gel with acetonitrile containing 0.1% trifluoroacetic acid. The tryptic peptides were then separated on an Agilent Nanopump system using a C18 ZORBAX trap and a SB-C18 ZORBAX 300 reversed-phase column (Agilent Technologies, Inc.) (150 mm by 75 μm; 3.5 μm particle size). All mass spectra were recorded on a hybrid linear ion trap-triple quadrupole mass spectrometer

(Q-Trap; AB Applied Biosystems) equipped with a nanoelectrospray ionization source. The accumulation of MS-MS data was performed with Analyst software, version 1.4 (AB Applied Biosystems). MASCOT software (Matrix Science, London, United Kingdom) was used to create peak lists from MS and MS-MS raw data.

Functional assays

Autoaggregation, biofilm formation, adhesion and invasion assays were performed as previously described [17]. Briefly, *E. coli* C600 bearing an empty vector, the plasmid pTgH or the plasmid pTngH were induced at an OD₆₀₀ of 0.8 and grown overnight in LB or M9 (Biofilm formation) media. In the autoaggregation assay, the overnight cultures were normalized in 5 mL of LB to an OD₆₀₀ of approximately 1.5 in culture tubes and left standing at 4°C. Samples (100 µL) were taken 1 cm below the surface at the beginning of the assay and after 120 min, and the OD₆₀₀ of the samples were measured. In the biofilm formation assay, the overnight cultures were normalized in M9 minimal medium, grown for 24 hours at 30°C in plastic microtiter plates and the biofilms were stained with crystal violet. After washes, the dye was then solubilized with a mixture of ethanol and acetone (80:20), and the absorption at 595 nm of the solution was measured. In the adhesion and invasion assays, the overnight cultures were inoculated onto monolayers of confluent HEP-2 cells in a 24-well plate (approximately 2.5x10⁵ cells) using 10⁶ colony-forming units (CFU) per well. For adhesion, after 3h, the cells were washed with PBS, and the adhering bacteria were recovered with 100 µL of Triton X-100 (1%), plated, and counted. For invasion, fresh medium containing 100 µg*mL⁻¹ gentamicin was added after the 3h incubation and washing, and the plates were incubated for an additional 2h before recovering and plating.

All functional assays were performed at least three times in duplicate or triplicate. For each assay, the results were compared to those of glycosylated TibA by performing an ANOVA and Dunnett post-tests using Prism 4.0 software (Graphpad Software).

β-Galactosidase reporter assay

The β-galactosidase activity of strains SR1458 [28] and SR1364 [29] transformed with an empty vector, plasmid pTgH, or plasmid pTgHC358R was assessed as described previously

[29], and the results were expressed in Miller units. Statistical comparisons were performed by ANOVA using Prism 4.0 (GraphPad software).

Proteases accessibility assay

Overnight cultures of strain C600 bearing an empty vector, the plasmid pTgH or the plasmid pTngH were normalized and resuspended in TBS in the presence or absence of trypsin (Sigma) or proteinase K (Invitrogen) at a final concentration of 0.5 $\mu\text{g/mL}$, 1 $\mu\text{g/mL}$ or 5 $\mu\text{g/mL}$. After 30 min of incubation on ice, the samples were diluted in twice concentrated SDS-PAGE loading buffer containing β -mercaptoethanol, denatured by heating at 100°C for 10 min, separated by SDS-PAGE and visualized by immunoblotting using the anti-His antibody.

Far-UV Circular Dichroism

Far-UV CD spectra of the proteins were recorded on a spectropolarimeter (Jasco Spectroscopic Co. Ltd.; model J-810) using a 0.1-cm-path length cuvette. The far-UV CD spectra of glycosylated and unglycosylated TibA (500 $\mu\text{g/mL}$ in TBS 1% n-octyl- β -D-glucoside) were recorded between 205 and 260 nm. For each spectrum, ten accumulations were averaged, and the contribution of buffer to the measured ellipticity was subtracted. To study the effect of salts, the protein samples were diluted prior to measurement in TBS containing NaCl in order to give a final NaCl concentration of 500 mM and a final detergent concentration of 0.2-0.3% which is below the critical micelle concentration of n-octyl- β -D-glucoside. To study the effect of temperature, the ellipticity of the proteins was recorded at 218 nm with the temperature varying between 20°C and 80°C at a rate of 5°C/min. Ellipticities were converted to mean residual ellipticities (MRE).

Results

Glycosylation is not required for optimal expression of TibA

Glycosylation is often important for protein stability [30]. This is the case for another SAAT, AIDA-I. The lack of glycosylation of AIDA-I results in a significant decrease in protein expression [15]. Therefore, we first determined the effect of

glycosylation on the expression levels of TibA. To produce glycosylated TibA, we used a plasmid which harbors genes coding for an His-tagged TibA and TibC [18]. To produce unglycosylated TibA, we used a plasmid that expressed His-tagged TibA only. Whole-cell extracts of *E. coli* strain C600 bearing an empty vector, or plasmids allowing the production of glycosylated or unglycosylated TibA were probed with an antibody raised against the His-tag (Fig. 1B). The results showed that the expression level of TibA was similar whether it was glycosylated or not. Therefore, the expression level of TibA is not dependent on its glycosylation status.

The whole cell extracts were also probed with an antibody raised against glycosylated AIDA-I (Fig 1A). The anti-AIDA-I antibody recognizes AIDA-I only when it is glycosylated [9] and cross-reacts with glycosylated TibA [18]. This confirmed that TibA expressed with TibC was glycosylated while TibA expressed alone was not. We have also expressed unglycosylated TibA from a plasmid coding for a mutated and inactive TibC glycosyltransferase [23]. Unglycosylated TibA was also expressed at a similar level than glycosylated TibA when produced from this plasmid (data not shown).

Localization of the glycosylation sites in TibA

We next sought to determine the regions of TibA that were glycosylated. We purified glycosylated his-tagged TibA by affinity chromatography from membrane extracts of *E. coli* strain C600 expressing glycosylated TibA. After SDS-PAGE, the band corresponding to TibA was digested by trypsin within the gel and the resulting peptides were analyzed by electrospray tandem MS/MS [15]. TibA is glycosylated by heptose residues [23]. Therefore, a peptide modified by one or more heptose residues harbors an increase of its mass by a multiple of 192 Da.

Twenty peptides were identified throughout the entire protein and represent 47% of the protein (Fig. 2 and Table 1). Seven of these peptides were found to be glycosylated: their mass presented one to seven addition of 192 Da compared to the theoretical mass. Four of the seven peptides were located in the passenger domain. The remaining three-glycosylated peptides were located in the junction region, showing that glycosylation is not limited to the passenger domain of TibA. As is the case for AIDA-I [15] and Ag43 [25], glycosylation of TibA was heterogenous; one peptide could be found with various amounts

of heptose residues. Unfortunately, we could not determine by this technique onto which amino acid the heptoses were added and, in the case of a peptide modified with multiple heptoses, if the heptoses were on the same residue or on different residues. However, the glycosylated peptides were rich in serines and threonines and did not contain N-glycosylation motifs. Thus, the heptoses were most likely added by O-glycosylation, as it is the case for AIDA-I and Ag43 [15,25].

Glycosylation of TibA is required for adhesion

To determine the effect of glycosylation on the functionality of TibA, we have performed autoaggregation, adhesion, invasion and biofilm formation assays with *E. coli* C600 bearing an empty vector, or plasmids allowing the expression of TibA and TibC or TibA alone (Fig. 3). The results were similar when unglycosylated TibA was produced by the plasmid allowing the expression of TibA and the mutated and inactive TibC (data not shown). Bacteria expressing unglycosylated TibA were able to autoaggregate, form biofilm and invade cultured epithelial cells as well as the bacteria expressing glycosylated TibA, but were impaired in adhesion. Mutants that can invade but are not able to adhere were also isolated in structure-function studies of TibA and AIDA-I [17,18], suggesting that adhesion and invasion are not linked together and rely on distinct mechanism. These results suggest that glycosylation might play a role in the adhesion mediated by TibA.

Unglycosylated TibA causes an extracellular stress

Proteins bearing an abnormal conformation are sensed by specialized stress-sensing system, which leads to the expression of chaperones and proteases in order to correct the situation [31]. To determine the impact of glycosylation on the conformation of TibA, we monitored the cellular stress generated by the expression of glycosylated and unglycosylated TibA using two previously described reporter strains. In strain SR1364, β -galactosidase activity is under the control of the *rpoH* promoter that is induced by a cytoplasmic stress [29]. In strain SR1458, it is under the control of the *degP* promoter, which is induced by an extracytoplasmic stress [28]. Therefore, a cytoplasmic stress will lead to β -galactosidase activity in strain 1364, while an extracytoplasmic stress will lead to β -galactosidase activity in strain 1458. We have allowed the expression of glycosylated or

unglycosylated TibA in strains SR1364 and SR1458 and we have monitored the β -galactosidase activity upon expression of TibA. Glycosylated TibA did not induce the β -galactosidase activity in either strain, suggesting that it does not induce any cytoplasmic or extracytoplasmic stress (Fig. 4). However, unglycosylated TibA induced β -galactosidase activity in strain SR1458, meaning that it induces extracytoplasmic stress and suggesting that it adopts an abnormal conformation. This suggests that unglycosylated and glycosylated TibA harbor different conformations when they reach the extracellular compartment in *E. coli*.

Glycosylation does not affect the final conformation of TibA

Proteins harboring an aberrant conformation are usually more sensitive to proteases [32,33]. Therefore, we have performed limited proteolysis of the bacterial surface proteins in order to determine if unglycosylated TibA is more sensitive to proteolytic degradation than its glycosylated counterpart. We have subjected *E. coli* strain C600 expressing glycosylated TibA or unglycosylated TibA to an increasing amount of trypsin or proteinase K (Fig. 5). We observed that the glycosylation did not affect the sensitivity of TibA to either trypsin or proteinase K. Both proteins were degraded by the same amount of proteases and showed a similar profile of degradation. Since TibA is as well expressed whether it is glycosylated or not, it is possible to purify the unglycosylated form, as opposed to AIDA-I [15]. Therefore, we also performed limited proteolysis on the purified proteins. Again, both proteins showed a similar sensitivity to trypsin and proteinase K (data not shown). Thus, we conclude the final conformations of glycosylated and unglycosylated TibA, as probed by differences in proteolysis sensitivity, do not seem to be different.

We next monitored the conformations of the purified proteins by far-UV circular dichroism. The CD spectra of glycosylated TibA and unglycosylated TibA were similar (Fig. 6A) and showed a single minimum of ellipticity at 218 nm, which is characteristic of β -stranded proteins. This is in agreement with the predicted structure of TibA [34]. This result again suggests that unglycosylated TibA harbors a similar final conformation than glycosylated TibA. To measure the stability of the protein, we followed the thermal denaturation of glycosylated and unglycosylated TibA by measuring the ellipticity at 218 nm. For both glycosylated and unglycosylated TibA, the denaturation followed a simple

two-state model with a denaturation temperature (T_m) of $\sim 65^\circ\text{C}$ (Fig 6B). Right after denaturation, we allowed the protein to refold by decreasing the temperature. We observed that glycosylated TibA was able to refold to a certain extent, but unglycosylated TibA did not refold at all. This suggests that glycosylation affects the dynamics of TibA folding, rather than its final conformation.

Glycosylation allows the modulation between the different conformations of TibA

In the case of AIDA-I, the conformation and oligomerisation of the protein can be modulated by environmental signals, such as bile salts [35]. Furthermore, this change is associated with the functionality of the protein: signals that modulate the conformation and oligomerisation of AIDA-I also modulate AIDA-mediated autoaggregation of bacteria. One such signal is high concentrations of sodium chloride. Since glycosylation seemed to affect the dynamics of TibA folding, we tested the effect of glycosylation on salt-induced conformational changes in TibA.

Previous experiments were realized in the presence of detergent on monomers of TibA because detergent inhibits the oligomerisation of the purified protein (data not shown). As a result, in order to observe the effect of glycosylation on the oligomerisation of TibA, we assessed the conformation of purified glycosylated and unglycosylated TibA in the presence NaCl and in the absence of detergent. To trigger oligomerisation, we diluted the protein sample below the critical micelle concentration (CMC) of n-octyl- β -D-glucoside in the presence or absence of 500 mM NaCl. We observed that addition of salt to glycosylated TibA resulted in a decrease of the ellipticity and a blue-shift in the minimum of ellipticity; two characteristics highlighting conformational changes in the protein (Fig 7A). Furthermore, the addition of salt inhibited the autoaggregation of bacteria expressing glycosylated TibA (Fig. 7C). This suggests that TibA can be found in two different conformations, as is the case for AIDA-I: one conformation that mediates oligomerisation of the protein and autoaggregation of the bacteria and the other that is not competent for oligomerisation and autoaggregation. Interestingly, the addition of salt did not modify the CD spectrum of unglycosylated TibA and did not inhibit the autoaggregation of bacteria expressing unglycosylated TibA (Fig. 7B-C). Thus, when not glycosylated, TibA is only found in the conformation that is competent for oligomerisation and autoaggregation. This

is consistent with the hypothesis that glycosylation of TibA is important for the modulation of its conformation.

Discussion

In this report, we have evaluated the effect of glycosylation on the autotransporter TibA. We have identified glycosylated peptides in TibA and found that the glycosylation was highly heterogenous. Some peptides were modified with various amounts of heptoses and, in some cases, a single peptide could be modified by up to seven heptoses. Unfortunately, in our study, we could not determine the precise sites of glycosylation. However, it is most likely to be O-glycosylation, as for AIDA-I and Ag43 [15,25]. First, Aah, the glycosyltransferase associated with AIDA-I, is similar to TibC and, most likely, uses a similar mechanism of action [8]. In fact, the two glycosyltransferases are interchangeable [23]. Furthermore, most peptides do not contain motifs that are characteristic to N-glycosylation [36]. As in the case of AIDA-I and Ag43, we therefore hypothesize that the glycosylated residues are mainly serines, but could also be threonines [25].

TibA is heavily glycosylated compared to AIDA-I and could possess between 18 and 22 heptose residues. By comparison, 19 glycosylation sites were found in AIDA-I. However, the extracellular portion of AIDA-I, comprising the passenger domain and the junction region, is twice the length of that of TibA. In the case of AIDA-I, glycosylation was only found in the passenger domain [15]. In the case of Ag43, 16 glycosylation sites were found over the ~500 first amino acids, until the position of the cleavage site in the passenger domain [25]. Unfortunately, in that study, the authors could not evaluate whether heptoses could be found beyond this point. From previous reports, it was therefore possible that the glycosylation of SAATs was specific of the imperfect repeats. Our results contradict that hypothesis and are in line with the recent observation that glycosylation is specific for a secondary structure and not a sequence [8]. Interestingly, only the ~600 first amino acids of AIDA-I and TibA were glycosylated. We therefore hypothesize that the SAAT glycosyltransferases glycosylate the first 600 amino acids of the SAAT,

independently of the domain to which they correspond and provided that they harbor the required secondary structure.

Taken together, our observations are consistent with the recently identified mechanism of substrate specificity for Aah [8]. Aah recognizes a structural motif consisting of a strand-loop-strand, with the glycosylated residue at the start of the loop. This rather loose mechanism of action probably applies to TibC and explains the glycosylation independently of any domain. Whether there is a mechanistical reason for the restriction to the first 500-600 amino acids is unclear and is under investigation.

We have also evaluated the effect of glycosylation on the functionality of TibA. We found that the glycosylation is essential for the adhesion properties of TibA, but not the other functions. This observation is in agreement with previous reports [24,37] and also applies to AIDA-I [15]. This raised the possibility that glycosylation conferred lectin-like adhesive properties. However, in the case of AIDA-I, the protein is poorly expressed at the surface of the bacteria when not glycosylated. Therefore, the effect of glycosylation on the adhesion properties of AIDA-I could be an indirect effect caused by the low expression of unglycosylated AIDA-I. Unglycosylated TibA is expressed as well as glycosylated TibA, making it an ideal model to determine the exact effect of glycosylation on the functionality of SAATs. We observed that when unfolded, only purified glycosylated TibA could be refolded. We also observed that when not glycosylated, TibA is not sensitive to salt-induced oligomerisation and conformational changes whether we used purified proteins or proteins expressed in whole bacteria. Lastly, unglycosylated TibA caused a stress when expressed in *E. coli*. Taken together, our results suggest that the dynamics of TibA folding are affected by glycosylation. Thus, it seems that glycosylation gives flexibility to TibA, enabling conformational changes between an oligomeric autoaggregation-competent conformation and a monomeric adhesion-competent conformation. These observations made using a naturally glycosylated autotransporter that we can express in both its glycosylated and unglycosylated forms confirms previous studies performed with AIDA-I and Ag43 [15,25]. Our observations confirm that the primary role of glycosylation is to ensure that the protein adopt a normal conformation, which indirectly affects the functionality of the protein. We therefore hypothesize that for all SAATs, glycosylation specifically affects conformation flexibility rather than allowing lectin-like interactions.

Acknowledgements

We thank Anastasia Nikolakakis and François Lépine for the mass spectrometry experiments. This work was supported by the Canadian Institutes for Health Research and the Natural Sciences and Engineering Research Council, as well as by funds from the Groupe de Recherche sur les Maladies Infectieuses du Porc and the Canada Research Chair program. JPC was supported by a graduate fellowship from Natural Sciences and Engineering Research Council.

References

1. Nothaft H, Szymanski CM (2010) Protein glycosylation in bacteria: sweeter than ever. *Nat Rev Microbiol* 8: 765-778.
2. Schmidt MA, Riley LW, Benz I (2003) Sweet new world: glycoproteins in bacterial pathogens. *Trends Microbiol* 11: 554-561.
3. Szymanski CM, Wren BW (2005) Protein glycosylation in bacterial mucosal pathogens. *Nat Rev Microbiol* 3: 225-237.
4. Szymanski CM, Logan SM, Linton D, Wren BW (2003) *Campylobacter*--a tale of two protein glycosylation systems. *Trends Microbiol* 11: 233-238.
5. Scott NE, Parker BL, Connolly AM, Paulech J, Edwards AV, et al. (2011) Simultaneous glycan-peptide characterization using hydrophilic interaction chromatography and parallel fragmentation by CID, higher energy collisional dissociation, and electron transfer dissociation MS applied to the N-linked glycoproteome of *Campylobacter jejuni*. *Mol Cell Proteomics* 10: M000031-MCP000201.
6. Logan SM (2006) Flagellar glycosylation - a new component of the motility repertoire? *Microbiology* 152: 1249-1262.
7. Grass S, Lichti CF, Townsend RR, Gross J, St Geme JW, 3rd (2010) The *Haemophilus influenzae* HMW1C protein is a glycosyltransferase that transfers hexose residues to asparagine sites in the HMW1 adhesin. *PLoS Pathog* 6: e1000919.

8. Charbonneau ME, Cote JP, Haurat MF, Reiz B, Crepin S, et al. (2012) A structural motif is the recognition site for a new family of bacterial protein O-glycosyltransferases. *Mol Microbiol* 83: 894-907.
9. Benz I, Schmidt MA (2001) Glycosylation with heptose residues mediated by the aah gene product is essential for adherence of the AIDA-I adhesin. *Mol Microbiol* 40: 1403-1413.
10. Dautin N, Bernstein HD (2007) Protein secretion in gram-negative bacteria via the autotransporter pathway. *Annu Rev Microbiol* 61: 89-112.
11. Leyton DL, Rossiter AE, Henderson IR (2012) From self sufficiency to dependence: mechanisms and factors important for autotransporter biogenesis. *Nat Rev Microbiol* 10: 213-225.
12. Nishimura K, Tajima N, Yoon YH, Park SY, Tame JR (2010) Autotransporter passenger proteins: virulence factors with common structural themes. *J Mol Med* 88: 451-458.
13. Oliver DC, Huang G, Nodel E, Pleasance S, Fernandez RC (2003) A conserved region within the *Bordetella pertussis* autotransporter BrkA is necessary for folding of its passenger domain. *Mol Microbiol* 47: 1367-1383.
14. Klemm P, Vejborg RM, Sherlock O (2006) Self-associating autotransporters, SAATs: functional and structural similarities. *Int J Med Microbiol* 296: 187-195.
15. Charbonneau ME, Girard V, Nikolakakis A, Campos M, Berthiaume F, et al. (2007) O-linked glycosylation ensures the normal conformation of the autotransporter adhesin involved in diffuse adherence. *J Bacteriol* 189: 8880-8889.
16. Sherlock O, Dobrindt U, Jensen JB, Munk Vejborg R, Klemm P (2006) Glycosylation of the self-recognizing *Escherichia coli* Ag43 autotransporter protein. *J Bacteriol* 188: 1798-1807.
17. Charbonneau ME, Mourez M (2007) Functional organization of the autotransporter adhesin involved in diffuse adherence. *J Bacteriol* 189: 9020-9029.
18. Cote JP, Mourez M (2011) Structure-function analysis of the TibA self-associating autotransporter reveals a modular organization. *Infect Immun* 79: 1826-1832.

19. Klemm P, Hjerrild L, Gjermansen M, Schembri MA (2004) Structure-function analysis of the self-recognizing Antigen 43 autotransporter protein from *Escherichia coli*. *Mol Microbiol* 51: 283-296.
20. Reidl S, Lehmann A, Schiller R, Salam Khan A, Dobrindt U (2009) 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.
21. Lindenthal C, Elsinghorst EA (1999) Identification of a glycoprotein produced by enterotoxigenic *Escherichia coli*. *Infect Immun* 67: 4084-4091.
22. Berthiaume F, Leblond MF, Harel J, Mourez M (2010) Growth-phase-dependent expression of the operon coding for the glycosylated autotransporter adhesin AIDA-I of pathogenic *Escherichia coli*. *FEMS Microbiol Lett* 311: 176-184.
23. Moormann C, Benz I, Schmidt MA (2002) Functional substitution of the TibC protein of enterotoxigenic *Escherichia coli* strains for the autotransporter adhesin heptosyltransferase of the AIDA system. *Infect Immun* 70: 2264-2270.
24. Lindenthal C, Elsinghorst EA (2001) Enterotoxigenic *Escherichia coli* TibA glycoprotein adheres to human intestine epithelial cells. *Infect Immun* 69: 52-57.
25. Knudsen SK, Stensballe A, Franzmann M, Westergaard UB, Otzen DE (2008) Effect of glycosylation on the extracellular domain of the Ag43 bacterial autotransporter: enhanced stability and reduced cellular aggregation. *Biochem J* 412: 563-577.
26. Tang G, Ruiz T, Mintz KP (2012) O-polysaccharide glycosylation is required for stability and function of the collagen adhesin EmaA of *Aggregatibacter actinomycetemcomitans*. *Infect Immun* 80: 2868-2877.
27. Schneider CA, Rasband WS, Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* 9: 671-675.
28. Raina S, Missiakas D, Georgopoulos C (1995) The rpoE gene encoding the sigma E (sigma 24) heat shock sigma factor of *Escherichia coli*. *EMBO J* 14: 1043-1055.
29. Mourez M, Skouloubris S, Betton JM, Dassa E (1997) Heat shock induction by a misassembled cytoplasmic membrane protein complex in *Escherichia coli*. *Mol Microbiol* 26: 821-831.
30. Wyss DF, Wagner G (1996) The structural role of sugars in glycoproteins. *Curr Opin Biotechnol* 7: 409-416.

31. Missiakas D, Mayer MP, Lemaire M, Georgopoulos C, Raina S (1997) Modulation of the *Escherichia coli* sigmaE (RpoE) heat-shock transcription-factor activity by the RseA, RseB and RseC proteins. *Mol Microbiol* 24: 355-371.
32. Mogensen JE, Tapadar D, Schmidt MA, Otzen DE (2005) Barriers to folding of the transmembrane domain of the *Escherichia coli* autotransporter adhesin involved in diffuse adherence. *Biochemistry* 44: 4533-4545.
33. Fontana A, de Laureto PP, Spolaore B, Frare E, Picotti P, et al. (2004) Probing protein structure by limited proteolysis. *Acta Biochim Pol* 51: 299-321.
34. Kajava AV, Steven AC (2006) Beta-rolls, beta-helices, and other beta-solenoid proteins. *Adv Protein Chem* 73: 55-96.
35. Girard V, Cote JP, Charbonneau ME, Campos M, Berthiaume F, et al. (2010) Conformation change in a self-recognizing autotransporter modulates bacterial cell-cell interaction. *J Biol Chem* 285: 10616-10626.
36. Kowarik M, Young NM, Numao S, Schulz BL, Hug I, et al. (2006) Definition of the bacterial N-glycosylation site consensus sequence. *EMBO J* 25: 1957-1966.
37. Sherlock O, Vejborg RM, Klemm P (2005) The TibA adhesin/invasin from enterotoxigenic *Escherichia coli* is self recognizing and induces bacterial aggregation and biofilm formation. *Infect Immun* 73: 1954-1963.

Figure legends

Figure 1: Effect of glycosylation on the expression level of TibA

Whole cell extracts of *E. coli* strain C600 bearing an empty vector (-), or a plasmid allowing the expression of glycosylated TibA (TibA/TibC) or unglycosylated TibA (TibA) were separated by SDS-PAGE and revealed by immunoblotting with an anti-AIDA-I antibody (A) or an anti-His antibody (B). The intensity of the bands probed with the α -His antibody was quantified using ImageJ and values were compared to glycosylated TibA (lower panel). Experiments were done five times and ANOVA and Dunnett post-tests were used to identify significant (*; $p < 0.05$) and non-significant (ns) differences with glycosylated TibA.

Figure 2: Localization of the peptides in TibA

Glycosylated (filled circle) and unglycosylated (empty circle) peptides from Table 1 are localized on a schematic representation, drawn to scale, of TibA highlighting the signal sequence (residues 1 to 54), the passenger domain (residues 54 to 393), the junction region (residues 393 to 627), the proline-rich domain (residues 627 to 677) and the membrane-embedded domain (residues 677 to 989).

Figure 3: Effect of glycosylation on the functionality of TibA

(A) Autoaggregation assay: *E. coli* C600 bearing an empty vector (-), a plasmid allowing the expression of glycosylated TibA (TibA/TibC) or unglycosylated TibA (TibA) were normalized to an OD₆₀₀ of 1.5 and left standing. OD₆₀₀ at the top of the culture was measured at the beginning of the assay and after 2h. Results are shown in percentage of initial OD₆₀₀. (B) Adhesion assay: Bacteria were inoculated onto monolayers of confluent Hep-2 cells. After 3h, the adhering bacteria were recovered, plated and counted. Results represent the percentage of adhered bacteria compared to glycosylated TibA. (C) Invasion assay: After adhesion, extracellular bacteria were killed by addition of gentamicin and invaded bacteria were recovered, plated and counted. Results represent the percentage of gentamicin resistant bacteria compared to glycosylated TibA. (D) Biofilm formation assay: Biofilms were stained with crystal violet and the amount of fixed dye was determined by measuring OD₅₉₅. Experiments were done three times in duplicate and ANOVA and Dunnett post-tests were used to identify significant (*; p<0.05) and non-significant (ns) differences with glycosylated TibA.

Figure 4: Effect of glycosylation on the generation of cytoplasmic and extracytoplasmic stress

Reporter *E. coli* strains SR1364 (A) and SR1458 (B) were transformed with an empty vector (-), a plasmid allowing the expression of glycosylated TibA (TibA/TibC) or unglycosylated TibA (TibA/TibC_{C358R}) and the β-galactosidase activity was measured. Experiments were done three times in duplicate and ANOVA and Dunnett post-tests were used to identify significant (*; p<0.05) and non-significant (ns) differences with glycosylated TibA.

Figure 5: Effect of glycosylation on protease susceptibility

E. coli strain C600 bearing an empty vector (-), or a plasmid allowing the expression of glycosylated TibA (TibA/TibC) or unglycosylated TibA (TibA) were pelleted and resuspended in TBS in the presence or absence of trypsin or proteinase K (0.5 µg/ml, 1 µg/mL and 5 µg/mL). After 30 min, whole cell extracts were prepared and revealed by immunoblotting with an anti-His antibody. MW, molecular weight.

Fig. 6 : Effect of glycosylation on purified TibA

TibA was purified from *E. coli* strain C600 bearing plasmids allowing expression of glycosylated and unglycosylated TibA and solubilized in TBS 1% n-octyl-β-D-glucoside. (A) Far-UV CD spectrum of glycosylated TibA (black) and unglycosylated TibA (red) are shown between 205 and 260 nm. (B) Proteins were also subjected to a thermal denaturation followed by a renaturation monitored by far-UV CD. Temperature was increased from 25°C and 80°C at a rate of 5°C per minute and the ellipticities at 218 nm were recorded for glycosylated TibA (black) and unglycosylated TibA (red). Right after the denaturation, temperature was decreased from 80°C to 25°C at a rate of 5°C per minute and the ellipticities were recorded at 218 nm for glycosylated TibA (gray) and unglycosylated TibA (orange).

Figure 7: Effect of glycosylation on the modulation of conformation by salt

Far-UV CD spectrum of glycosylated TibA (A) and unglycosylated TibA (B) are shown between 205 and 260 nm. The purified proteins were solubilized in TBS 1% n-octyl-β-D-glucoside (bOG). Prior to CD measurement, the proteins were diluted to a final bOG concentration of 0.3%, below the CMC, in TBS (black line) or in TBS 500 mM NaCl (dotted line). (C) *E. coli* strain C600 bearing an empty vector (-), or a plasmid allowing the expression of glycosylated TibA (TibA/TibC) or unglycosylated TibA (TibA) were resuspended in TBS (black bars) or in TBS 500 mM NaCl (white bars) and autoaggregation assays were performed.

Table

Table I: Glycosylated and unglycosylated peptides in TibA

Glycosylated peptides				
Sequence	Theoretical mass	# of heptoses	Observed mass	Score ^a
a Q ₁₄₇ TVFSGGSAMGTIINSGGDQYVISGGSATSASVTSGAR ₁₈₄	3579.85	5	4538.97	52
		6	4730.09	45
		7	4923.64	66
b T ₂₃₄ TINSGGGMYLYGGSATGTTSIYNGGR ₂₆₀	2555.76	1	2747.65	37
		2	2939.51	39
c Q ₂₆₁ YVSSGGSATNTTVYSGGR ₂₇₉	1891.97	3	2468.36	35
d S ₃₇₆ GGVLYGTTTTLTK ₃₈₉	1412.56	0	1411.74	53
		1	1603.65	55
e F ₄₃₂ SGLLSQDGGIFLQSGGAMTMDALQAK ₄₅₈	2744.13	1	2935.29	68
f A ₄₅₉ NVTTSQGTTLTLDNGTILTGNVAGDSTGAGDMAVK ₄₉₄	3452.75	6	4604.48	36
g N ₅₇₆ TGLEPVSAGAPLQVVQTGGGDAFTLK ₆₀₃	2699.01	0	2697.07	36
		1	2889.65	75
Unglycosylated peptides				
Q ₂₀₄ YVRDGGSATDTVLNNTGR ₂₂₂	2024.13		2023.23	43
G ₄₉₅ ASVWHLDDGSTV GALTLDN GTVDFRPSTTTR ₅₂₆	3347.60		3346.64	47
G ₆₀₄ GKVDAGTWEYGLSK ₆₁₈	1567.72		1566.80	66
E ₆₁₉ NTNWYLK ₆₂₇	1067.17		1067.40	39
S ₆₈₉ VDAVLGMATAPAYVFENSELDNLR ₇₁₂	2553.87		2553.26	56
S ₇₉₀ NVDSSGGGLYATWFDNDGYVVDGVLK ₈₁₆	2900.06		2899.47	66
G ₈₃₆ DYSQNGFGGSLEAGR ₈₅₁	1614.65		1613.94	92
T ₈₅₂ FSLNENAWAQPYYR ₈₆₆	1795.97		1795.79	93
L ₉₀₁ GMTLDVAGK ₉₀₉	1004.21		1004.61	55
E ₉₁₀ VKPYLSAAVSHEFSDNNK ₉₂₉	2135.32		2135.18	65
I ₉₃₂ NDTYDFRNDISGTTGK ₉₄₈	1917.02		1917.25	35
Y ₉₄₉ GLGVNAQLTPNAGVWAEAR ₉₆₈	2087.32		2086.94	92
Y ₉₆₉ ENKQTESPITGGVGFR ₉₈₆	1940.10		1940.12	62

^a A score >35 indicate identity or extensive homology (probability of a random match, p<0.05).

Figures

Figure 1

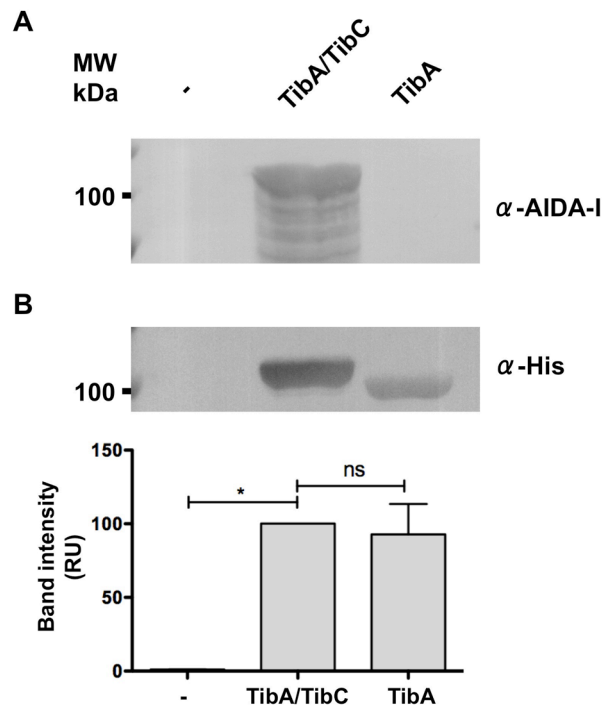


Figure 2

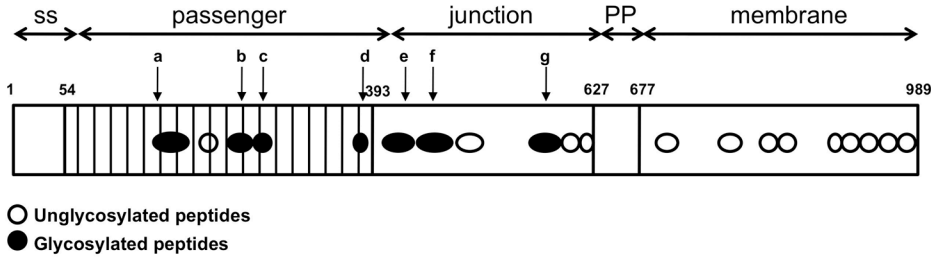


Figure 3

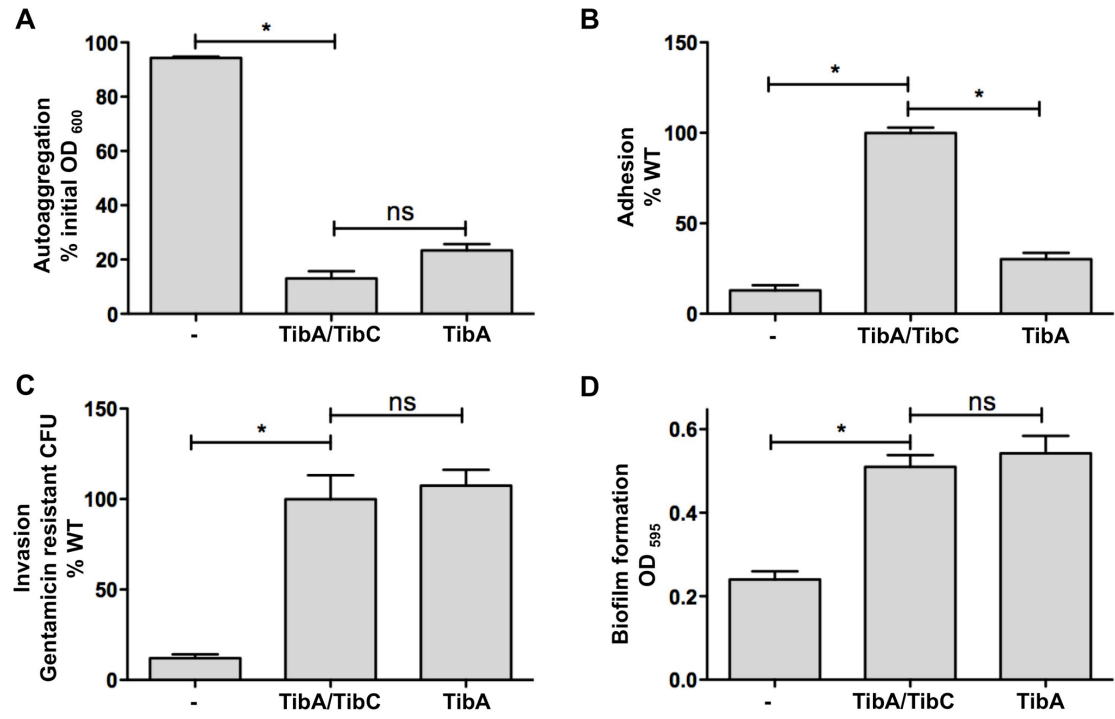


Figure 4

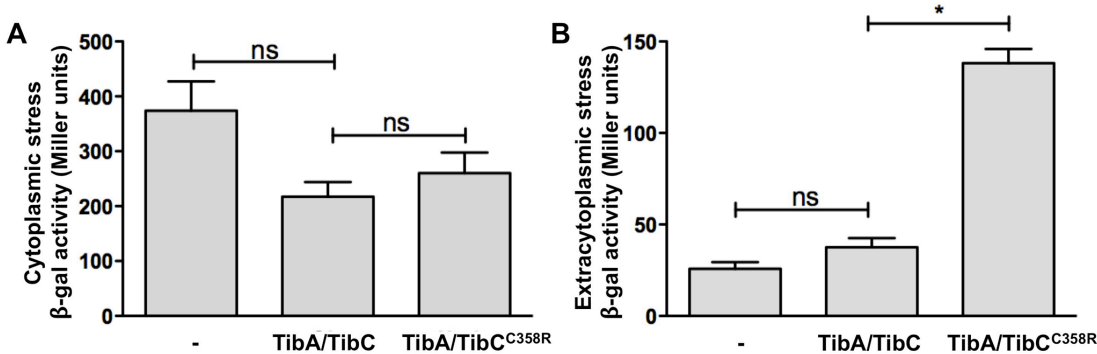


Figure 5

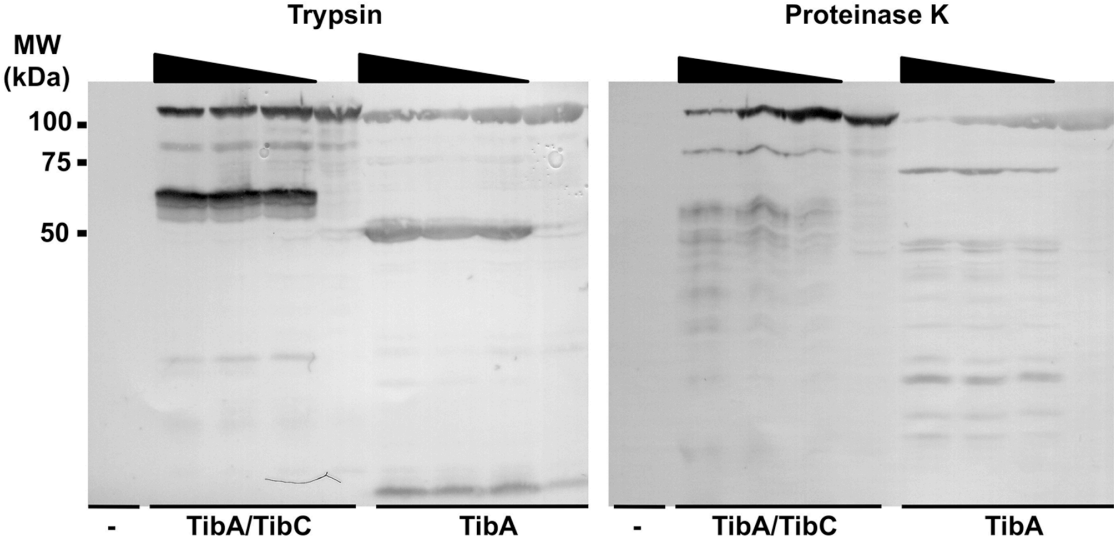


Figure 6

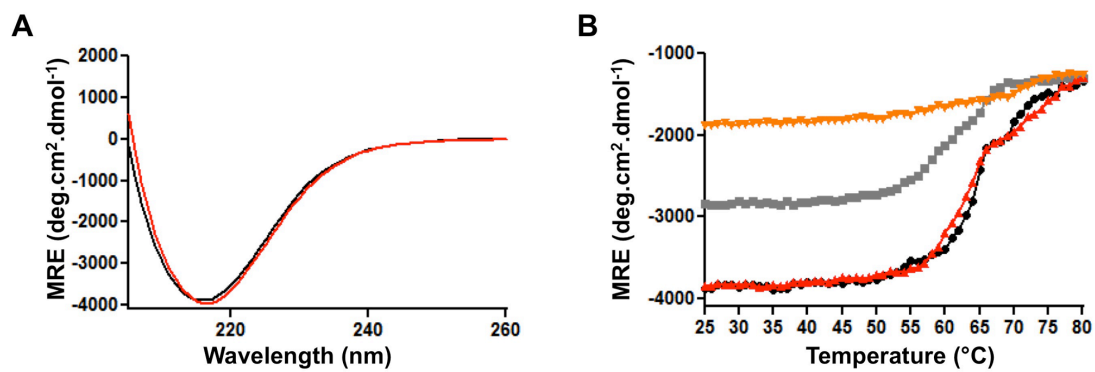
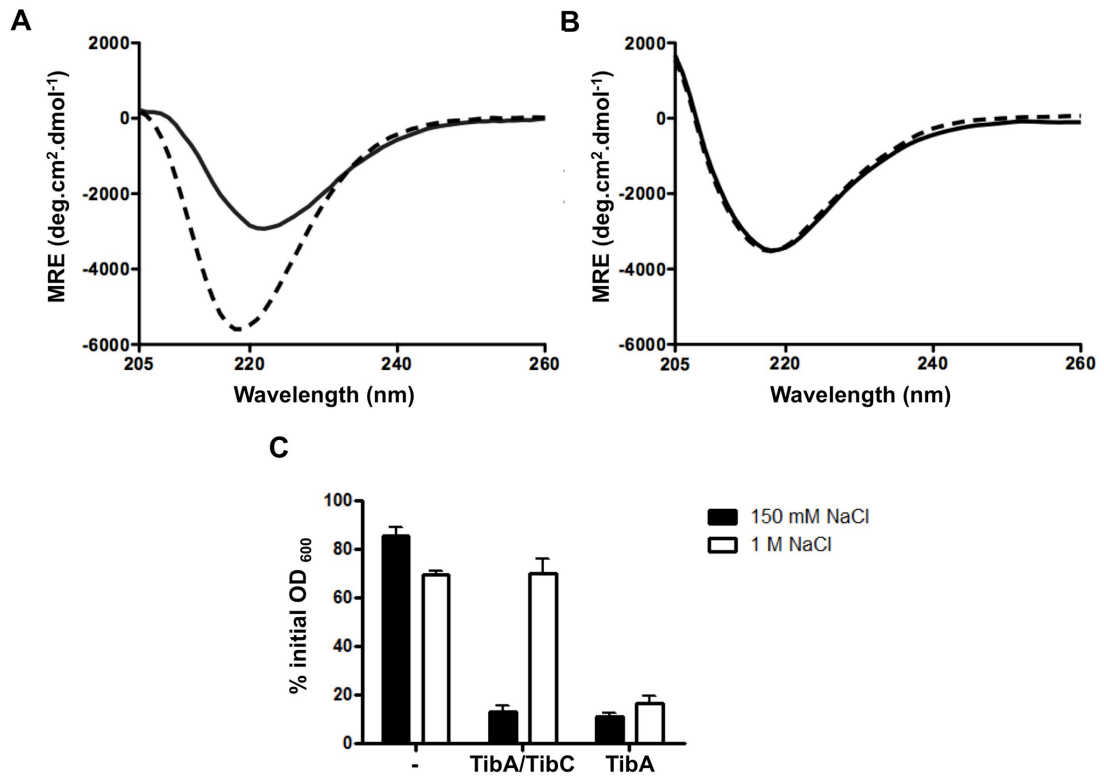


Figure 7



Article 3 – Allèles d’AIDA-I

Identification and evolution mechanism of new alleles coding for the AIDA-I autotransporter of porcine pathogenic *Escherichia coli*

Applied and Environmental Microbiology, Juillet 2012; volume 18 (13), p. 4597-4605

Avant-propos

L’article 1 met en évidence un possible mécanisme d’évolution des SAATs par acquisition et échange de modules. Nous avons donc voulu confirmer cette hypothèse en séquençant l’opéron *aah-aidA* de plusieurs souches d’*E. coli*. Par le fait même, ces nouvelles séquences d’AIDA-I vont compléter le manque de séquences disponibles (seulement 2 séquences d’*aidA* connues) et nous permettre de voir la variabilité qu’il y a dans le gène *aidA*. Nos résultats montrent la présence de trois allèles différents de *aah-aidA* qui sont associés à différents profils de facteurs de virulence. De plus, notre analyse montre qu’*aidA* est bel et bien construit par blocs de séquence et montre des échanges de blocs complets de séquence entre les différents allèles, ce qui suggère que les SAATs pourraient bien avoir évolué par acquisition et échange de modules. Ces résultats nous ont permis de déterminer les caractéristiques communes aux SAATs qui ont été utilisées plus loin pour identifier de nouveaux membres de la famille des SAATs (article 5).

Contributions

Contribution pour cet article :

Comme premier auteur, j’ai effectué toutes les analyses et presque toutes les expérimentations, à l’exception des PFGE et des hybridations d’ADN. J’ai aussi écrit le manuscrit.

Identification and evolution mechanism of new alleles coding for the AIDA-I autotransporter of porcine pathogenic *Escherichia coli*

Jean-Philippe Côté¹, Frédéric Berthiaume¹, Sébastien Houle², John M. Fairbrother¹, Charles M. Dozois² and Michael Mourez^{1,#}

¹ Faculté de Médecine Vétérinaire, Université de Montréal, Saint-Hyacinthe, Québec, Canada,

² INRS-Institut Armand-Frappier, Laval, Québec, Canada

Running Title: Alleles of *aidA* autotransporter

Abstract

Autotransporters are a large family of virulence factors of Gram-negative bacterial pathogens. The autotransporter Adhesin Involved in Diffuse Adherence (AIDA-I) is an outer membrane protein of *Escherichia coli*, which allows binding to epithelial cells as well as autoaggregation of bacteria. AIDA-I is glycosylated by a specific heptosyltransferase encoded by the *aah* gene that forms an operon with the *aidA* gene. *aidA* is highly prevalent in strains causing disease in pigs. Nevertheless, there are only two published whole-length sequences for this gene. In this study, we sequenced the *aah* and *aidA* genes of 24 *aidA*-positive porcine strains harboring distinct virulence factor profiles. We compared the obtained sequences and performed phylogenetic and pulse-field electrophoresis analyses. Our results suggest that there are at least 3 different alleles for *aidA*, which are associated with distinct virulence factor profiles. The genes are found on high molecular weight plasmids and seem to evolve via shuffling mechanisms, with one of the sequences showing evidence of genetic recombination. Our work suggests that genetic plasticity allows the evolution of *aah-aidA* alleles that are selected during pathogenesis.

Introduction

Escherichia coli is an important cause of severe diarrhea resulting in mortality and morbidity in humans and animals worldwide (26). Adhesion is often a critical first step in the infection process, and pathogenic *E. coli* strains employ a vast array of fimbrial or afimbrial adhesins (27). The adhesin involved in diffuse adherence (AIDA-I) is one such adhesin and was first identified in a strain isolated from a case of infantile diarrhea (1). However, AIDA-I is rarely found in *E. coli* strains causing disease in humans (18, 29). In contrast, it is frequently found in *E. coli* strains causing post-weaning diarrhea or edema disease in pigs (24, 28, 29, 36, 37). Indeed, experimental infections have shown that AIDA-I is important for the development of disease in pigs (31).

AIDA-I belongs to the family of autotransporters, outer membrane proteins of Gram-negative bacteria characterized by their modular organization and secretion

mechanism (12, 16). Autotransporters comprise: (i) an N-terminal signal sequence for Sec-dependent translocation across the inner membrane: (ii) an extracellular central 'passenger' domain: and (iii) a C-terminal domain embedded in the outer membrane. In several autotransporters, a conserved region between the passenger domain and the membrane-embedded domain forms a junction, or 'stable core', which is important for secretion (30, 32).

The N-terminal part of the passenger domain of AIDA-I is unusual: It is formed by multiple repetitions of a 19-amino acid consensus sequence. This characteristic is shared by two other *E. coli* autotransporters, the adhesin/invasin TibA (23) and the autoaggregation factor Ag43 (15). With AIDA-I, these proteins have been called self-associating autotransporters (SAATs), as they can mediate bacterial auto-aggregation through self-association (14, 22).

Two main alleles of *agn43* code for proteins with differences in the number of consensus sequence repetitions (33, 35). An allele of *tibA* with a different number of consensus sequence repetitions was also recently identified (19). Allelic variations in *agn43* have been associated with functional differences and distinct phylogenetic groups (21, 33). This suggests that there is an evolution mechanism allowing the appearance of variations in the number of repetitions of the consensus sequence, but such a mechanism remains to be understood. Similar studies examining putative alleles of *aidA* are lacking. The *aidA* sequences from two strains, the human pathogenic strain 2787 (1) and the porcine pathogenic strain E393 (37), are known. Thus, the question of the existence of *aidA* alleles is still unresolved.

In porcine pathogenic strains of *E. coli*, *aidA* has been found in isolates associated with different pathogenicity mechanisms, corresponding to its association with distinct virulence factors. AIDA-I has been associated with the STb heat-stable toxin of enterotoxigenic *E. coli* (ETEC) (13, 34). It has also been associated with Stx2 toxin and F18 fimbriae of Shiga-toxin producing *E. coli* (STEC), or with F18 alone (non pathogenic) (29). Lastly, AIDA-I is present in some strains of enteropathogenic *E. coli* (EPEC), which are characterized by attaching/effacing lesions caused by intimate adhesion mediated in part by the adhesin intimin (6). These strains cause various diseases in pigs, including

edema disease and post-weaning diarrhea. Thus, it is possible that *aida*, like *agn43*, has functionally distinct alleles.

In this study, we have sequenced the *aah-aida* operon from 24 porcine pathogenic *E. coli* strains in order to investigate the genetic organization and the allelic distribution of *aah-aida*. By comparing these sequences, along with the ones previously known, we found that there are three different alleles for *aida* and that the sequences are distributed, in part, according to the virulence factor profiles. Our data also reveal insights into the functionality of AIDA-I and the evolution mechanism of the *aida* alleles.

Materials and Methods

Bacterial strains and growth conditions

E. coli strains 2787 and PD20 and 23 other porcine pathogenic *E. coli* strains isolated from pigs presenting disease in Québec in 2007 were provided by the reference Laboratory for *E. coli* (Faculty of veterinary medicine, University of Montreal). The presence of genes encoding for virulence factors was tested by conventional PCR and by colony hybridization. Bacteria were grown on Luria-Bertani (LB) agar plates or in liquid LB broth at 37°C overnight. A PD20 Δ *aida* strain was previously generated by allelic exchange (31). Antimicrobial susceptibility was tested using the disk diffusion assay (Kirby-Bauer). The following 8 antimicrobials were used: Ampicillin (AMP: 10 μ g), Apramycin (AM:15 μ g), Ceftiofur (TIO: 30 μ g): Florfenicol (FLR: 30 μ g): Neomycin (NEO: 30 μ g): Spectinomycin (SPCM: 100 μ g): Tetracycline (TET: 30 μ g) and Trimethoprim-sulfamethoxazole (SXT:25 μ g). Breakpoints were those recommended by the CLSI (17).

E. coli strain C600 (New England biolabs: F⁻, *thr1*, *leuB6*, *thi1*, *lacY1*, *supE44*, *rfaD1*, *fhuA21*) was used for phenotypic assays. Bacteria containing the different plasmids were grown at 30°C on LB agar plates or in liquid LB medium containing 100 μ g/ml ampicillin and 50 μ g/ml chloramphenicol. Growth was followed by measuring turbidity as the optical density at 600 nm (OD_{600 nm}). At an OD_{600 nm} of 0.8, cultures were induced with 10 μ M IPTG. This low concentration of IPTG was used to limit toxicity associated with overexpression of AIDA-I and/or Aah.

Sequencing of the *aah/aidA* genes

DNA was extracted using the DNeasy Blood and Tissue kit (Qiagen) according to the instructions of the manufacturer. Following extraction, *aah*, *aidA* and the region upstream of *aah* were amplified by PCR and sequenced using the primers listed in Table 1. The resulting sequences were deposited in GenBank under accession numbers: HQ641298 – HQ641310 and JQ044400 – JQ044410

Phylogenetic analysis

Sequence alignments were made with the Geneious Pro 5.0.4 software (Biomatters Ltd.) using the BLOSUM matrix 62. Phylogenetic trees were generated using the neighbor-joining method and the Jukes-Cantor model. The silent variations, conservative variations and non-conservative variations were designated by comparison with the sequence from strain 2787.

SDS-PAGE and immunoblotting

Overnight cultures (5 ml) were adjusted to identical optical density measured at 600 nm, centrifuged for 10 min at 12,000 X g in microcentrifuge tubes, and the pellets were resuspended in 50 µl of tris-buffered saline (TBS). Suspensions were then diluted in 2X SDS-PAGE loading buffer containing β-mercaptoethanol, denatured by heating at 100°C for 10 min and separated by SDS-PAGE on 10% acrylamide gels. The gels were transferred to polyvinylidene fluoride membranes (Millipore). Immunodetection was performed with a serum raised against glycosylated AIDA-I (7) diluted 1:60,000 in blocking buffer (5% skim milk, 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.05% Triton X-100). A goat anti-rabbit horseradish peroxidase (HRP)-conjugated antibody (Sigma) was used as a secondary antibody according to the instructions of the manufacturer. Immune complexes were revealed using a 3,3',5,5'-tetramethylbenzidine solution for membranes (Sigma).

Plasmid extraction and Pulse field gel electrophoresis

High molecular weight plasmids were extracted following the procedure of Kado and Liu (20). Overnight LB cultures of the tested strains were collected and the pellets suspended in E buffer (40 mM Tris, 2 mM EDTA, pH 7.8). Bacteria were incubated in lysis buffer (50 mM Tris, 3% SDS, pH12.6) for one hour at 56°C. A crude phenol:chloroform mix (1:1)

was added and the solution emulsified by inverting. The emulsion was broken by centrifugation and the upper aqueous phase transferred to a microfuge tube. Electrophoresis was performed on 1 % agarose gels in 0.5 x TBE buffer on a CHEF-Mapper XA PFGE system (Bio-Rad) for 20 hours at 14°C, with run conditions of 6 V/cm, a pulse angle of 120° and pulse times from 0.47 to 10.29 s.

DNA hybridization

Southern hybridization was performed using standard methods. Briefly, high molecular weight plasmids were transferred from gel by a modified downward alkaline capillary transfer method (11). The gel was stained with a 1 µg/ml ethidium bromide solution for 30 minutes and the DNA nicked with the use of a GS Gene linker UV chamber (Bio-Rad) at 60 mJ. The gel was then destained for 20 minutes and equilibrated for 15 minutes in 0.4 N NaOH - 1.5 M NaCl. The gel was then put onto a positively charged nylon membrane (Boehringer Mannheim) and transfer proceeded for 40 hours. The membrane was finally neutralized in 0.5 M Tris-HCl pH 7 for 5 min, rinsed briefly in 2x SSC (0.3 M NaCl, 30 mM sodium citrate) and dried on blotting paper.

The membranes were rehydrated in DIG Easy Hyb buffer (Roche) at 42°C for 1 hour. Hybridization was performed with DIG-coupled probes, generated by PCR using the DIG DNA Labeling Mix (Roche) and corresponding to the 683 nucleotides at the N terminus of *aida*, to the *estB* gene or to the 230 nucleotides at the N-terminus of *fedA* (29), the gene coding for the main subunit of the F18 fimbriae (Table 1). Hybridization was performed with the denatured probes (10 min at 100°C) diluted 1:1,000 in DIG Easy Hyb at 42°C for 18 hours with gentle agitation. The probe was then washed with 2x SSC, 0.1 % SDS and revealed with an anti-DIG antibody conjugated to horseradish peroxidase (Anti-DIG-POD: Roche) diluted 1:1,000 in blocking buffer. Immune complexes were revealed using a 3,3',5,5'-tetramethylbenzidine solution for membranes (Sigma).

Cloning of AIDA-I alleles

The *aida* gene from strains 2787, 14326 (STb) and 17530 (F18) were amplified by PCR using the primers AIDA_F and AIDA_R (Table 1) using the Expand Long Template PCR system (Roche) according to the instructions of the manufacturer. The PCR products were

then cloned in the pTRC99A vector (Pharmacia Biotech) using *NcoI* and *XbaI*. Constructions were verified by restriction mapping and sequencing.

Phenotypic assays

Autoaggregation, biofilm formation, and adhesion assays were performed as previously described (10). *E. coli* C600 harboring a plasmid allowing the expression of Aah (8) and the pTRC99A vector or the plasmids allowing the expression of the different alleles of AIDA-I were induced with 10 μ M IPTG at an OD₆₀₀ of 0.8 and grown overnight in LB (autoaggregation and adhesion assays) or M9 (biofilm assay) medium. In the autoaggregation assay, the overnight cultures were normalized in 5 ml of LB to an OD₆₀₀ of approximately 1.5 in culture tubes and left standing at 4°C. Samples (100 μ l) were taken 1 cm below the surface at the beginning of the assay and after 120 min, and the OD₆₀₀ of the samples were measured. OD₆₀₀ at 120 min were compared to OD₆₀₀ at the beginning of the assay and results are shown in percentage of initial OD₆₀₀. In the biofilm formation assay, the overnight cultures were normalized in M9 minimal medium and grown for 24 hours at 30°C in plastic microtiter plates. Biofilms were stained with crystal violet. After washes, the dye was then solubilized with a mixture of ethanol and acetone (4:1), and the absorption at 595 nm of the solution was measured. In the adhesion assay, the overnight cultures were inoculated onto monolayers of confluent HEp-2 cells in a 24-well plate (approximately 2.5×10^5 cells) using 10^6 colony-forming units (CFU) per well. After 3h, the cells were washed with PBS, and the adhering bacteria were recovered with 100 μ l of Triton X-100 (1%), plated, and counted.

All functional assays were performed at least three times in duplicate or triplicate. For each assay, the results obtained were compared to the 2787 allele by performing an ANOVA and Dunnett post-tests using Prism 4.0 software (Graphpad Software).

Results and Discussion

Assembling a collection of *aidA*+ strains with differing virulence gene profiles

We selected 24 strains of *E. coli* isolated from pigs and harboring the *aidA* gene, as demonstrated by PCR. Our selection included the previously described porcine pathogenic strain PD20 (28). The presence of the various virulence genes was determined at the

Reference Laboratory for *E. coli* of the Faculty of Veterinary Medicine, Université de Montréal (Table 2). These strains represent the 4 distinct virulence factor profiles and pathogenicity mechanisms described above: (i) F18 alone (non pathogenic): (ii) F18, Stx2 and the thermostable toxin East1 (STEC): (iii) the intimin Eae (EPEC): or (iv) STb with or without East1 toxins (ETEC). The two strains from which *aidA* sequences are known are ETEC, with the profiles STb-East1 (E393) or East1 (2787). Other similarities and differences exist between the 24 porcine strains and the previously studied ones: All newly selected strains were isolated in Canada, whereas 2787 was isolated in Europe and E393 was from China. In addition, all strains were recovered from pigs, except for strain 2787, which was recovered from a child. To assess the diversity of our strains, we also determined their antimicrobial resistance profiles (Table 2). Different resistance profiles were observed, showing that the strains were not identical clones.

Expression of AIDA-I

We first looked at the expression of AIDA-I from the 24 porcine strains by blotting total protein extracts and probing with an antibody raised against AIDA-I (Fig. 1). We included in the blot total protein extracts from strain 2787 as a positive control and from PD20 where *aidA* had been deleted as a negative control (PD20 Δ *aidA*)(31). All strains expressed AIDA-I at a similar level, except for strain 15194. In this strain, AIDA-I is not well expressed. There is no difference in the sequence of the promoter region, *aah* or *aidA* that could explain this lowered expression (see below). It is therefore possible that the *aah-aidA* operon is regulated differently in this strain.

AIDA-I is matured by self-cleavage (9) and glycosylated by the heptosyltransferase Aah (2). Our antibody specifically recognizes the glycosylated AIDA-I fragment (8). Thus, our results show that both Aah and AIDA-I are expressed, consistent with the fact that genes coding for these proteins form an operon (3, 4). Furthermore, the polypeptide for each of the strains had the same apparent size, which suggests that there was no important variation in the number of consensus sequence repetitions.

Building sequence alignments and phylogenetic trees

From the 24 porcine strains, we amplified by PCR and sequenced *aah*, *aidA* and the promoter region upstream of *aah*. We aligned the resulting sequences of *aah* and *aidA*,

along with the sequences from 2787 and E393. All *aah* and *aidA* sequences had the same length, confirming that there was no variation in the number of repetitions in *aidA*, and were identical over 96% and 93.2% of the gene respectively. In Figure 2, we report all the nucleotide positions for which there were variations amongst the sequences, and only these positions. Variations in the part of the sequences corresponding to *aah* and the first half of *aidA* reveal two distinct groups: the sequences from 2787 and E393 are similar and differ from those of all the other strains. Variations in the second half of *aidA* (corresponding to the junction region and membrane-embedded domain) reveal three groups: the sequences of 2787 and E393 are still similar and different from those of the other strains. However, the latter separated into two groups: one consisting of the strains with genes coding for the F18 fimbriae and the other consisting of the strains with genes coding for the STb toxin or intimin (Eae). The presence of three distinct groups of sequences were further evidenced by building phylogenetic trees with the *aah* and *aidA* sequences (Fig. 3). Again, the 2787/E393 strains clustered away from all the other strains, which split into the F18 and STb/Eae groups. The sequence from one strain, 14466, clustered with the F18 or the STb/Eae group depending on whether it was the *aah* or *aidA* sequence that was considered in the analysis (Fig. 3). This will be discussed below.

Lastly, we amplified the promoter region upstream of *aah*, using primers based on the known sequence from 2787 (4). Sequencing of the PCR fragments indicated a difference between strain 2787 and all the other strains (Fig. 4). For instance, one of the two directly repeated sequences upstream of the *aah* Shine-Dalgarno sequence is absent in all strains but 2787 (Fig. 4A). The sequences of the promoter also highlight some differences between the STb and the F18 strains (Fig. 4B). Of note, there is an insertion of 3 nucleotides (TAT) near the -10 region of the promoter in the STb and some Eae strains. The transcription start of P149 in these strains might therefore be slightly different. The consequence of these changes should be investigated in future experiments.

Taken together, our results suggest that one determinant for the sequence variations is the presence of specific virulence genes, namely those encoding F18 and STb/Eae. Since the receptor for F18 is expressed only in older age pigs, strains possessing F18 usually infect older animals (25), whereas strains producing STb or intimin may be present in younger pigs. Thus, strains expressing these virulence factors might segregate, which could result in

sequence differences. In addition, the similarity of sequences from strains expressing STb or Eae is surprising, since they correspond to different pathogenicity mechanisms, those of ETEC and EPEC respectively.

Additional factors play an important role, however, as the sequence from E393 was very different from that of strains with the same STb/East1 profile. Such a difference could be due to the geographical origin of the strains: since the sequences from 2787 and E393 are highly similar, it is possible that strains from Eurasia differ from those isolated in North America. More sequences from eurasian strains are needed to validate this hypothesis. Lastly, the host species origin does not seem to be a factor for sequence variability, since E393 and 2787 have similar sequences but were isolated from human and pig, respectively.

Mapping the positions for which there are genetic variations

We observed a bias when we averaged over a window of 20 nucleotides the number of positions for which there are variations amongst the *aah-aidA* sequences. Far more variations clustered in the region coding for the junction region of AIDA-I and two smaller regions in the C terminus of the passenger domain (Fig. 5A). These regions correspond to blocks of sequence that are different between alleles. These blocks are readily identifiable when observing the alignment of the nucleotides for which there are variations. In fact, the block corresponding to the junction region extends to include the sequence coding for the whole membrane-embedded domain (Fig. 2). This suggests that the sequence of *aidA* has evolved by acquisition of blocks of nucleotides, in addition to the accumulation of point mutations. It is striking to note in particular that the *aidA* sequences from the STb/Eae and F18 groups have a nearly identical passenger domain but distinct junction and membrane-embedded domains. A mechanism involving modular changes could also contribute to the evolution of other autotransporters genes. Indeed, domain shuffling has been implicated in the evolution of some autotransporters, such as EspP (5) and Ag43. For the latter, there are two main alleles of *agn43* that differ in length, one allele having lost the C-terminal part of the passenger domain (35). The “full-length” allele is further subdivided into three different sub-families. Indeed, by looking at the polymorphisms of these sub-families, blocks of sequence that have been exchanged are also easily identifiable (Fig S1).

We then distinguished variations resulting in conservative or silent mutations, versus variations resulting in non-conservative mutations. Positions resulting in silent and

conservative mutations are spread throughout the entire sequence but some regions of the gene encoding AIDA-I clearly lack variations resulting in non-conservative mutations (Fig. 5B). In particular, the sequences coding for the C-terminal membrane-embedded domain and the two previously identified adhesion regions of the passenger domain (10) were highly conserved. This is not surprising, since these domains of the protein play key roles in biogenesis and function, respectively. It is surprising, however, that there are many variations resulting in non-conservative mutations in the junction region, since the junction region is known to be important for translocation and folding of the passenger domain (30, 32) and one would expect it to be more sensitive to mutations.

Localization of the *aah-aidA* genes on large plasmids

The evolution of three *aah-aidA* alleles could be due to the localization of these genes on specific and different genetic elements. Indeed, *aidA* has been found on large plasmids also coding for the F18 fimbriae (24). It is possible that the *aidA* allele corresponding to the STb/Eae group is located on a different plasmid than that corresponding to the F18 group. Using pulse field gel electrophoresis, we investigated the presence of plasmids extracted from a subset of 13 strains, including strain 2787 (Fig. 6). All strains contained large plasmids, although there was no common pattern of plasmids corresponding to *aidA* alleles in the F18 or STb/Eae groups. This heterogeneity confirms that none of the strains were identical clones, in agreement with the phenotypical diversity we noted when selecting the strains (Table 2). We used an *aidA*-specific probe to determine in each strain which plasmid harbored the *aah-aidA* operon (Fig. 7A). We observed that the *aah-aidA* operon is present on large plasmids that vary in sizes. In a previous study it was also reported that large plasmids extracted from strains expressing F18 and harboring *aidA* had various sizes (24). Thus, the sizes of plasmids harboring *aidA* do not seem to correlate with specific *aidA* alleles.

We further used *estB*- and *fedA*-specific probes to determine if *aidA* co-localized on plasmids encoding these other virulence factors (Fig. 7). In every strain, *aidA* was found on the same plasmid as the genes coding for, either, STb or the F18 fimbriae. This suggests that *aidA* is physically linked to these other virulence factor genes.

Overall, our results suggest that *aidA* alleles are not harbored by specific plasmids.

Therefore, the association of different alleles with distinct virulence factors might stem

from the fact that *aidA* genes are readily moving between plasmids and/or that the plasmids harboring *aidA* are easily gaining or losing genes.

Fortuitously, the sequences we obtained illustrate this plasticity. Indeed, the sequence of *aah* from strain 14466 clusters with the STb/Eae group whereas the sequence of *aidA* from this strain clusters with the F18 group (Fig. 3). This suggests that there was a recombination event between the alleles of the two groups. When representing only the variations specific for the F18 and STb/Eae groups (Fig. 8), it becomes evident that the recombination occurred between nucleotides 1119 and 1380 of the *aah-aidA* operon, which corresponds to the limit between *aah* and *aidA*.

Functional properties of the AIDA-I variants

The three alleles of *aidA* possess a similar passenger domain. In order to determine if they are also functionally similar, we have cloned one representative for each allele. We previously cloned and characterized the functionality of *aidA* from strain 2787 (10). We further cloned *aidA* from strains 14326 (STb/Eae group) and 17530 (F18 group) into the same expression vector. Glycosylation is important for the stability of AIDA-I (8). Therefore, we have co-transformed the plasmids allowing the expression of the different alleles of AIDA-I with a plasmid allowing the expression of Aah from strain 2787. The three alleles were expressed and glycosylated properly, as seen by total protein extracts immunoblotted with an antibody raised against glycosylated AIDA-I (Fig. 9A). We then tested the constructs for their ability to mediate adhesion to cultured epithelial cells, autoaggregation of bacteria and formation of biofilm. The three alleles mediated the same level of adhesion, autoaggregation and formation of biofilm (Fig 9B-D). This is not surprising since the sequences indicated that the passenger domain is well conserved, especially for the parts of the passenger domain that are associated with the functionality of the protein (Fig. 5).

In conclusion, we propose the existence of at least three different alleles of the *aah-aidA* operon: one group corresponds to strains 2787/E393, which could reflect their eurasian geographical origin. The two remaining groups are divided according to their virulence gene profiles: the presence of the F18-encoding genes versus the presence of the STb toxin- or the Eae adhesin-encoding genes. The evolution of these alleles seems to rely

on a high genetic plasticity involving mutations as well as shuffling of whole blocks of sequence and recombination. Sequence variations could then become fixed during segregation of strains, through self-interaction of the new alleles and/or other factors. Such a mechanism allowing a quick apparition and spread of new alleles of a self-associating surface protein could be an important mechanism to permit recognition of 'self' in bacteria. In that respect, association of alleles with specific virulence factors could indicate the emergence of strains with a specific pathogenic process. This idea remains to be tested.

Acknowledgements

This work was supported by the Canadian Institutes for Health Research (MM) and the Natural Sciences and Engineering Research Council (CMD), as well as by funds from the Groupe de Recherche sur les Maladies Infectieuses du Porc and the Canada Research Chair program (MM and CMD). JPC was supported by a graduate fellowship from Natural Sciences and Engineering Research Council.

References

1. 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* 57:1506-1511.
2. 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* 40:1403-1413.
3. Benz, I., T. van Alen, J. Bolte, M. E. Wormann, and M. A. Schmidt. 2010. Modulation of transcription and characterization of the promoter organization of the autotransporter adhesin heptosyltransferase and the autotransporter adhesin AIDA-I. *Microbiology* 156:1155-1166.
4. Berthiaume, F., M. F. Leblond, J. Harel, and M. Mourez. 2010. Growth-phase-dependent expression of the operon coding for the glycosylated autotransporter adhesin AIDA-I of pathogenic *Escherichia coli*. *FEMS Microbiol Lett* 311:176-184.

5. Brockmeyer, J., M. Bielaszewska, A. Fruth, M. L. Bonn, A. Mellmann, H. U. Humpf, and H. Karch. 2007. Subtypes of the plasmid-encoded serine protease EspP in Shiga toxin-producing *Escherichia coli*: distribution, secretion, and proteolytic activity. *Appl Environ Microbiol* 73:6351-6359.
6. Bruant, G., Y. Zhang, P. Garneau, J. Wong, C. Laing, J. M. Fairbrother, V. P. Gannon, and J. Harel. 2009. Two distinct groups of porcine enteropathogenic *Escherichia coli* strains of serogroup O45 are revealed by comparative genomic hybridization and virulence gene microarray. *BMC Genomics* 10:402.
7. Charbonneau, M. E., F. Berthiaume, and M. Mourez. 2006. Proteolytic processing is not essential for multiple functions of the *Escherichia coli* autotransporter adhesin involved in diffuse adherence (AIDA-I). *J Bacteriol* 188:8504-8512.
8. Charbonneau, M. E., V. Girard, A. Nikolakakis, M. Campos, F. Berthiaume, F. Dumas, F. Lepine, and M. Mourez. 2007. O-linked glycosylation ensures the normal conformation of the autotransporter adhesin involved in diffuse adherence. *J Bacteriol* 189:8880-8889.
9. Charbonneau, M. E., J. Janvore, and M. Mourez. 2009. Autoprocessing of the *Escherichia coli* AIDA-I autotransporter: a new mechanism involving acidic residues in the junction region. *J Biol Chem* 284:17340-17351.
10. Charbonneau, M. E., and M. Mourez. 2007. Functional organization of the autotransporter adhesin involved in diffuse adherence. *J Bacteriol* 189:9020-9029.
11. Chomczynski, P. 1992. One-hour downward alkaline capillary transfer for blotting of DNA and RNA. *Anal Biochem* 201:134-139.
12. Dautin, N., and H. D. Bernstein. 2007. Protein secretion in gram-negative bacteria via the autotransporter pathway. *Annu Rev Microbiol* 61:89-112.
13. Dubreuil, J. D. 2010. STb and AIDA-I: the missing link? *Crit Rev Microbiol* 36:212-220.
14. Girard, V., J. P. Cote, M. E. Charbonneau, M. Campos, F. Berthiaume, M. A. Hancock, N. Siddiqui, and M. Mourez. 2010. Conformation change in a self-recognizing autotransporter modulates bacterial cell-cell interaction. *J Biol Chem* 285:10616-10626.

15. 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* 149:115-120.
16. 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* 68:692-744.
17. Clinical and Laboratory Standard Institute. 2008. Performance Standards for Antimicrobial Disk and Dilution Susceptibility Tests for Bacteria Isolated From Animals; Approved Standard. CLSI document M31-A3, 3 ed. Clinical and Laboratory Standard Institute (CLSI).
18. Jallat, C., V. Livrelli, A. Darfeuille-Michaud, C. Rich, and B. Joly. 1993. *Escherichia coli* strains involved in diarrhea in France: high prevalence and heterogeneity of diffusely adhering strains. *J Clin Microbiol* 31:2031-2037.
19. Johnson, T. J., C. DebRoy, S. Belton, M. L. Williams, M. Lawrence, L. K. Nolan, and J. L. Thorsness. 2010. Pyrosequencing of the Vir plasmid of necrotoxicogenic *Escherichia coli*. *Vet Microbiol* 144:100-109.
20. Kado, C. I., and S. T. Liu. 1981. Rapid procedure for detection and isolation of large and small plasmids. *J Bacteriol* 145:1365-1373.
21. 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* 51:283-296.
22. Klemm, P., R. M. Vejborg, and O. Sherlock. 2006. Self-associating autotransporters, SAATs: functional and structural similarities. *Int J Med Microbiol* 296:187-195.
23. Lindenthal, C., and E. A. Elsinghorst. 1999. Identification of a glycoprotein produced by enterotoxigenic *Escherichia coli*. *Infect Immun* 67:4084-4091.
24. Mainil, J. G., E. Jacquemin, P. Pohl, A. Kaeckenbeeck, and I. Benz. 2002. DNA sequences coding for the F18 fimbriae and AIDA adhesin are localised on the same plasmid in *Escherichia coli* isolates from piglets. *Vet Microbiol* 86:303-311.

25. Nagy, B., T. A. Casey, S. C. Whipp, and H. W. Moon. 1992. Susceptibility of porcine intestine to pilus-mediated adhesion by some isolates of piliated enterotoxigenic *Escherichia coli* increases with age. *Infect Immun* 60:1285-1294.
26. Nataro, J. P., and J. B. Kaper. 1998. Diarrheagenic *Escherichia coli*. *Clin Microbiol Rev* 11:142-201.
27. Nataro, J. P., J. B. Kaper, R. Robins-Browne, V. Prado, P. Vial, and M. M. Levine. 1987. Patterns of adherence of diarrheagenic *Escherichia coli* to HEP-2 cells. *Pediatr Infect Dis J* 6:829-831.
28. Ngeleka, M., J. Pritchard, G. Appleyard, D. M. Middleton, and J. M. Fairbrother. 2003. Isolation and association of *Escherichia coli* AIDA-I/STb, rather than EAST1 pathotype, with diarrhea in piglets and antibiotic sensitivity of isolates. *J Vet Diagn Invest* 15:242-252.
29. 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* 8:143-149.
30. 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* 47:1367-1383.
31. Ravi, M., M. Ngeleka, S. H. Kim, C. Gyles, F. Berthiaume, M. Mourez, D. Middleton, and E. Simko. 2007. Contribution of AIDA-I to the pathogenicity of a porcine diarrheagenic *Escherichia coli* and to intestinal colonization through biofilm formation in pigs. *Vet Microbiol* 120:308-319.
32. Renn, J. P., and P. L. Clark. 2008. A conserved stable core structure in the passenger domain beta-helix of autotransporter virulence proteins. *Biopolymers* 89:420-427.
33. 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* 73:1553-1562.
34. Taillon, C., E. Nadeau, M. Mourez, and J. D. Dubreuil. 2008. Heterogeneity of *Escherichia coli* STb enterotoxin isolated from diseased pigs. *J Med Microbiol* 57:887-890.

35. van der Woude, M. W., and I. R. Henderson. 2008. Regulation and function of Ag43 (flu). *Annu Rev Microbiol* 62:153-169.
36. Zhang, W., M. Zhao, L. Ruesch, A. Omot, and D. Francis. 2007. Prevalence of virulence genes in *Escherichia coli* strains recently isolated from young pigs with diarrhea in the US. *Vet Microbiol* 123:145-152.
37. Zhao, L., X. Chen, X. Xu, G. Song, and X. Liu. 2009. Analysis of the AIDA-I gene sequence and prevalence in *Escherichia coli* isolates from pigs with post-weaning diarrhoea and oedema disease. *Vet J* 180:124-129.

Figure Legends

Figure 1: Expression of AIDA-I

Whole-cell lysates from normalized overnight cultures of the 24 porcine isolates were separated by SDS-PAGE and probed by immunoblotting with an antibody raised against glycosylated AIDA-I. Whole-cell lysates from strains 2787 and PD20 Δ *aidA* were added as a positive and a negative control respectively.

Figure 2: Polymorphism in the *aah/aidA* operon

The sequence of the operons were aligned using the BLOSUM 62 matrix. Shown here are only the nucleotides showing polymorphisms. We omitted all the nucleotides that were identical for all the sequences in order to see the pattern of polymorphisms. Position 1176 corresponds to the end of *aah* and the start of *aidA*.

Figure 3: Phylogenetic analysis of *aah* and *aidA*

aah (A) and *aidA* (B) were aligned separately and a phylogenetic tree was build from these alignments using the neighbor-joining method.

Figure 4: Promoter region upstream of *aah*

The sequences of the region upstream of *aah* were aligned to show (A) the direct repeat right in front of *aah* and (B) the two promoters at position -128 and -149.

Figure 5: Regions of variability and conservation in *aah/aidA*

Polymorphisms (from Fig. 2) were placed at their actual position in the operon. (A) The number of variations was averaged over 20 nucleotides along the sequence and this average number of variations was plotted for each position. Over the length of *aah/aidA*, there are 3.39 variations/20 nucleotides. (B) Variations were separated in three categories: silent variations (blue), conservative variations (red) and non-conservative variations (green). To determine in which category a polymorphism should be placed, we compared it to the sequence from strain 2787. (C) Schematic representation showing the organization of AIDA-I: the signal sequence (ss): the passenger domain with adhesion domains in gray: the junction region and the membrane-embedded domain (MED).

Figure 6: Localization of *aidA* on large plasmids

Plasmids from 12 of the porcine isolates as well as from strain 2787 were extracted and separated by pulse field gel electrophoresis and revealed by ethidium bromide staining. A plasmid extract from strain χ 7122 was used as a molecular weight control.

Figure 7: Colocalization of *aidA* and other virulence factor

Plasmids from 11 of the porcine isolates as well as from strain 2787 were extracted and separated by pulse field gel electrophoresis, transferred to a nitrocellulose membrane and probed with a DIG-coupled oligonucleotide specific for *aida* (A), *fedA* (B) or *estB* (C). A plasmid extract from strain χ 7122 was used as a negative control.

Figure 8: Recombination of the sequence from strain 14466

The first half of the polymorphisms defining the STb group and the F18 group were used to analyze the position where the recombination of strain 14466 occurred. From the alignment in Fig. 2, we selected only the polymorphisms specifically defining the STb group and F18 group. We omitted strains 2787 and E393 because they form a separate group. We also omitted polymorphisms that were specific for only one sequence.

Figure 9: Functionality of the three alleles

E. coli C600 bearing a plasmid allowing the expression of Aah from strain 2787 and an empty vector (-), a plasmid allowing the expression of AIDA-I from strain 2787, 14326 (STb) or 17530 (F18) were probed with an antibody raised against glycosylated AIDA-I (A). The bacteria were also assessed for their ability to mediate autoaggregation (B), adhesion to Hep-2 cells (C) and formation of biofilm (D). The functional assays were made in triplicate. ANOVA and Dunnett post-tests were used to identify significant differences from the 2787 allele (*: $p < 0.05$).

Tables

Table I: Primers used in this study

Name	Sequence (5' → 3')	Function
AIDA_F	TAATAACCATGGGTAACGCCTACAGTATCATATGG	Amplification and cloning of aidA
AIDA_R	TTATCTAGATTATCAGAAGCTGTATTTTATCC	
Aah_F	ATGACTTTCTTATCACCACC	Amplification of aah
Aah_R	AATGATCCTGTCATGATTGG	
Promo_F	CTCCGTATGTAACGCTGTG	Amplification of the region upstream of aah
Promo_R	GGTATCTCTGGTGGTGATAAG	
Seq_aidA1	ATGGGTAAGGCCTACAGTATCATATGG	Sequencing of aidA
Seq_aidA2	GCGCTAATGCCACTGAGAC	
Seq_aidA3	CTGTGGCGAGGTATCTGG	
Seq_aidA4	CAGAGTTGTCAGGTAGTGG	
Seq_aidA5	TATGTCAATGAAGATGGCAG	
Seq_aidA6	GCTTCTGCAGGTGGGGG	
Seq_aidA7	GGTTTTCCCACCATTATTGACG	
Seq_aah1	CGGATAACCAGCCTGTGG	Sequencing of aah
Seq_aah2	CCACTTCAGCCCACCCC	
Seq_prom	CTCCGTATGTAACGCTGTG	Sequencing of the region upstream of aah
AIDA_1	ATGGGTAAGGCCTACAGTATCATATGG	
DIG-AIDA-R	GTGGCATTAGCGCCAGCAG	
F18_Probe_F	GTGAAAAGACTAGTGTTTATTTTC	Generation of the DIG-coupled probes
F18_Probe_R	CTTGTAAGTAACCGCGTAAGC	
STb_Probe_F	ATGAAAAAGAATATCGCATTCTTC	
STb_Probe_R	TTAGCATCCTTTTGCTGC	

Table II: Strains used in this study

Strain	Virulence factor profile ¹	Host	Location	Antimicrobial resistance ²	Reference
14430	AIDA: F18	Pig	Canada	AMP: AM: TET	This study
15139	AIDA: F18	Pig	Canada	AMP: TET: SXT	This study
15194	AIDA: F18	Pig	Canada	AMP: AM: FLR: TET	This study
17530	AIDA: F18	Pig	Canada	AM: NEO	This study
15334	AIDA: F18	Pig	Canada	AMP: FLR: TET	This study
15372	AIDA: F18	Pig	Canada	AMP: FLR: NEO: TET	This study
14596	Stx2: East1: AIDA: F18	Pig	Canada	AMP: NEO: TET: SXT	This study
14603	Stx2: East1: AIDA: F18	Pig	Canada	AMP: TET: SXT	This study
15478	Stx2: East1: AIDA: F18	Pig	Canada	AMP: FLR: TET	This study
14377	Stx2: East1: AIDA: F18	Pig	Canada	AMP: FLR: TET: SXT	This study
15742	Stx2: East1: AIDA: F18	Pig	Canada	AMP: NEO: TET: SXT	This study
15941	Stx2: East1: AIDA: F18	Pig	Canada	AMP: FLR: TET	This study
14748	Eae: AIDA	Pig	Canada	AMP: TET	This study
14503	Eae: AIDA	Pig	Canada	TET	This study
2010	Eae: AIDA	Pig	Canada	AMP: TIO: FLR: TET	This study
2020	Eae: AIDA	Pig	Canada	FLR: TET	This study
2034	Eae: AIDA	Pig	Canada	AMP: FLR: TET	This study
14466	STb: AIDA: East1	Pig	Canada	AMP: TET	This study
14326	STb: AIDA: East1	Pig	Canada	AMP: AM: SPCM: TET: SXT	This study
14707	STb: AIDA: East1	Pig	Canada	FLR: SPCM: TET	This study
14675	STb: AIDA: East1	Pig	Canada	AMP: SPCM: TET: SXT	This study
15146	STb: AIDA: East1	Pig	Canada	AMP: AM: NEO: TET: SXT	This study
14187	STb: AIDA: East1	Pig	Canada	AMP: TET: SXT	This study
2787	East1: AIDA	Human	Germany	-	(1)
PD20	STb: AIDA	Pig	Canada	-	(28)
E393	STb: East1: AIDA	Pig	China	ND ³	(37)

¹ Strains tested by colony hybridization using radioactively labeled (P32) DNA probes for the genes encoding virulence factors: STa, STb, LT, EAST1, Stx1, Stx2, CNF, F4, F17, F18, F5, F6, F41, P, Eae, AIDA, Paa, AFA, Tsh, Aerobactin

² Antimicrobials tested: Ampicillin (AMP), Apramycin (AM), Ceftiofur (TIO): Florfenicol (FLR): Neomycin (NEO): Spectinomycin (SPCM): Tetracycline (TET) and Trimetoprim-sulfamethoxazole (SXT)

³ No data is available for the antimicrobial resistance of this strain

Figures

Fig. 1

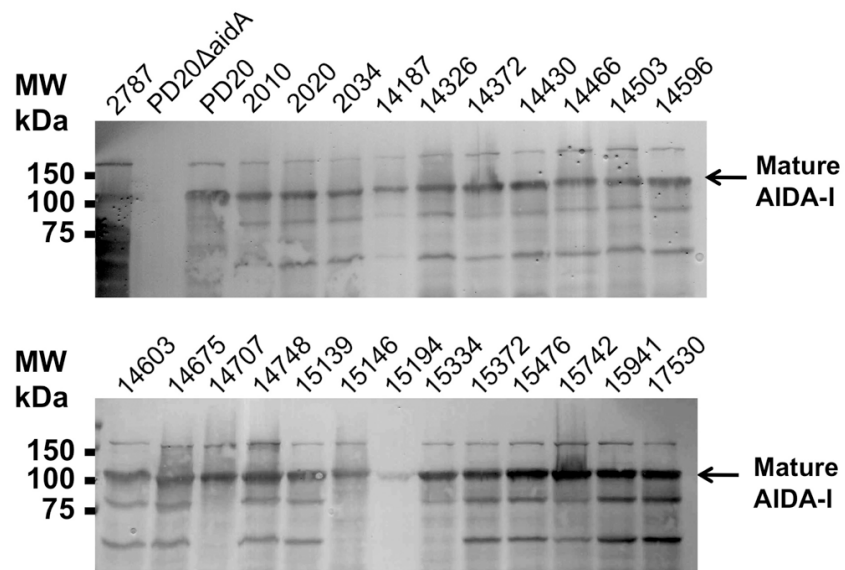


Fig. 3

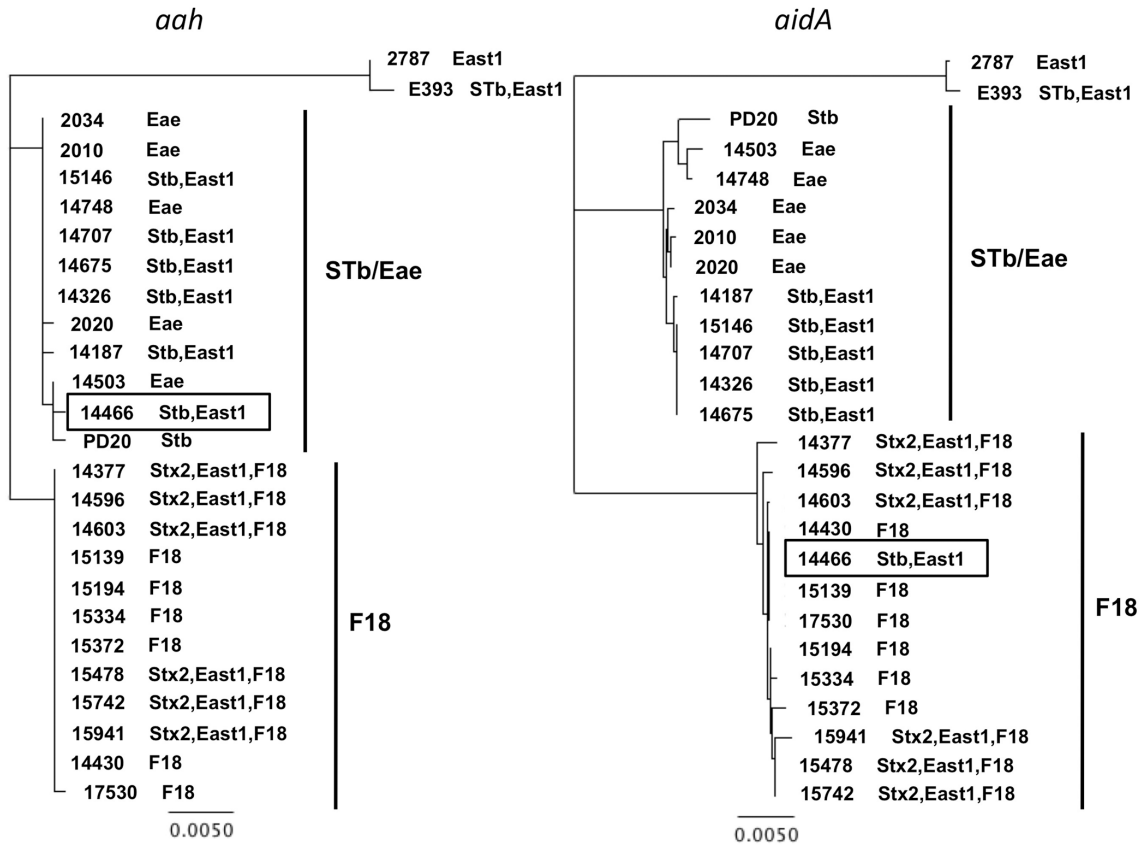


Fig. 5

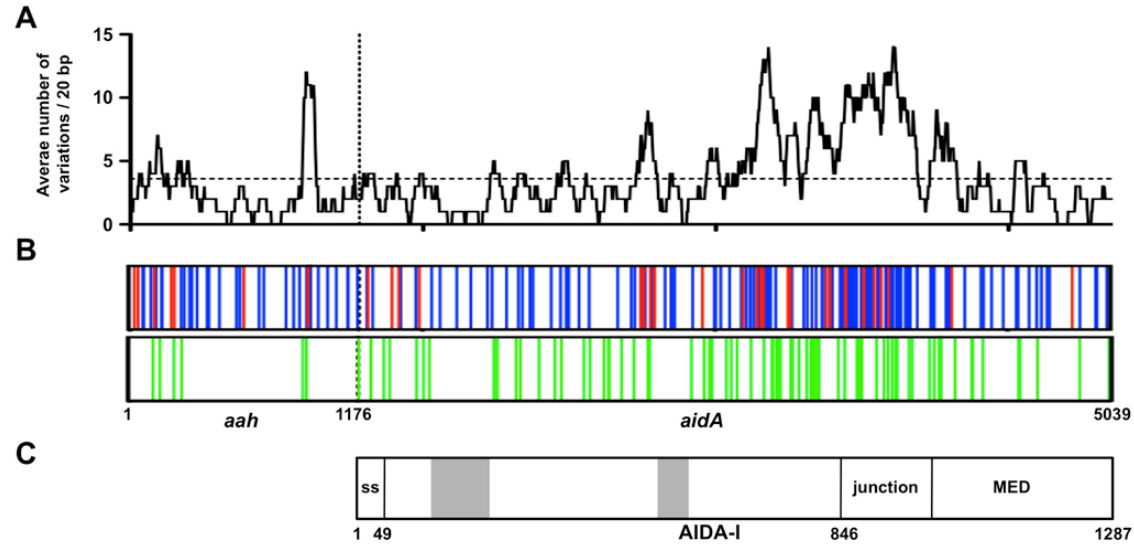


Fig. 6

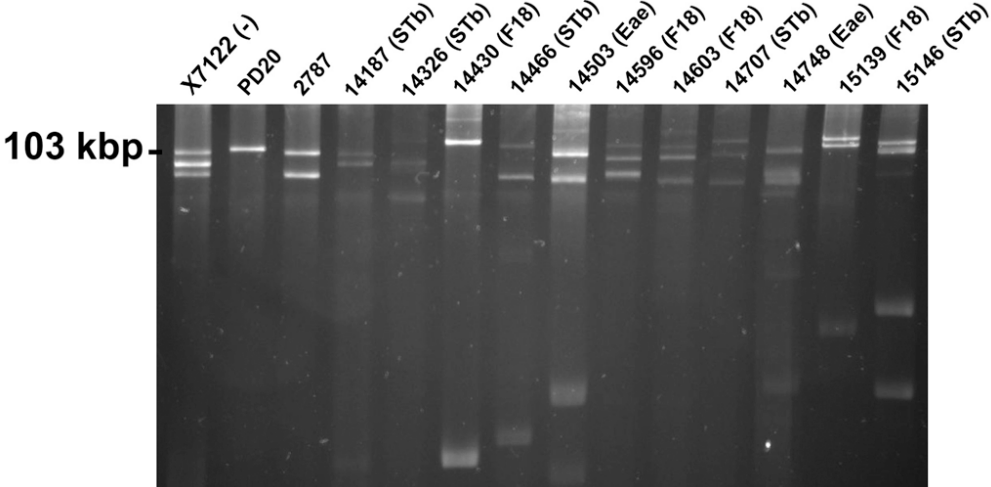


Fig. 7

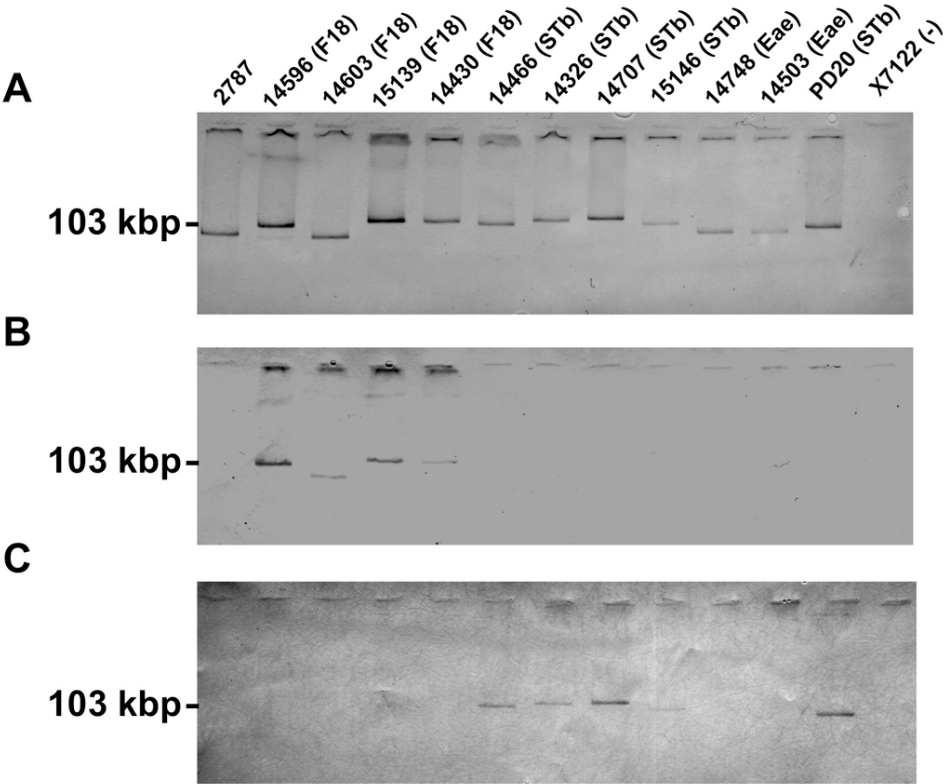


Fig. 8

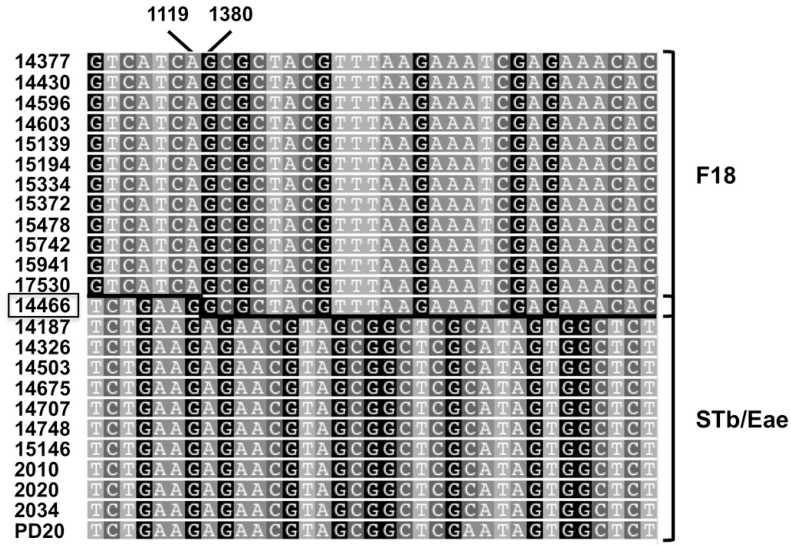
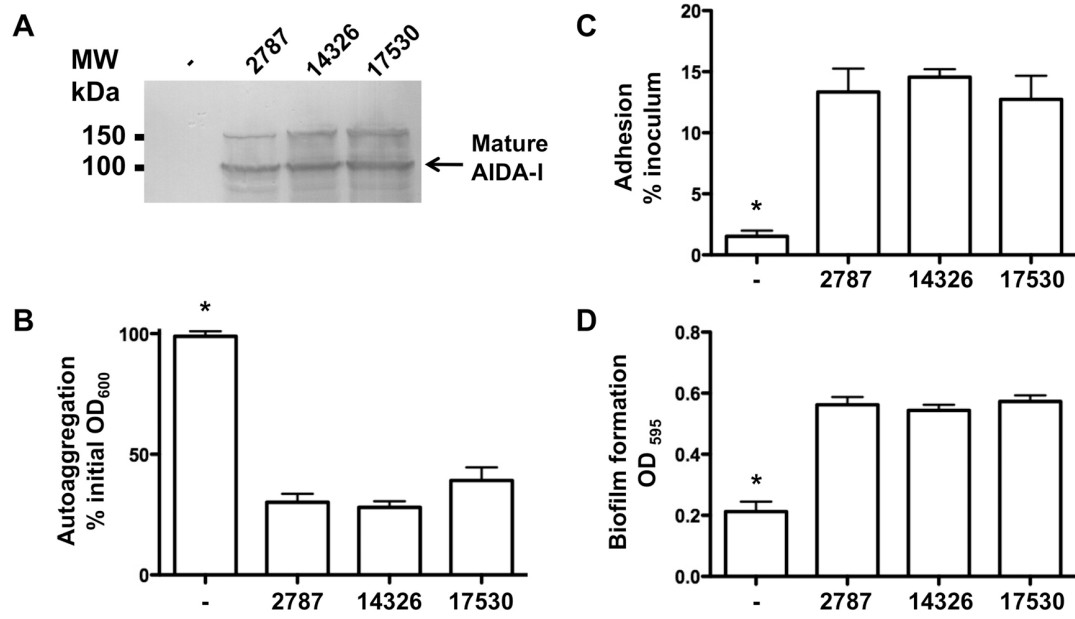


Fig. 9



Supplementary information

Fig. S1

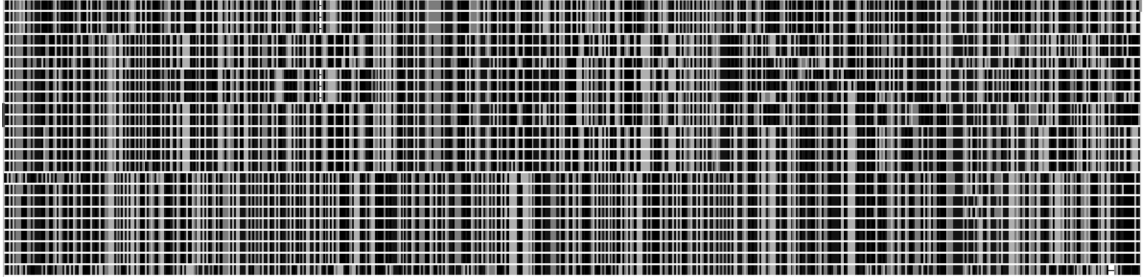


Figure S1: Polymorphism of the *agn43* gene

We aligned the sequence of 24 *agn43* “full-length” genes with the Geneious Pro 5.0.4 software (Biomatters Ltd.) using the BLOSUM matrix 62. Shown here are only the nucleotides showing polymorphisms. We omitted all the nucleotides that were identical for all the sequences in order to see the pattern of polymorphisms. The sequences used are those from the reference numbers: AE014075, AJ586887, CP001671, AJ586888, CU982162, FN554766, CU928145, CU928163, CP000247, X16664, CP001383, AE005674, AE014073, AF200692, AF233271, AF233272, AP009048, AP010953, CP000948, CP001396, CP001637, U00096 and AJ303141.

Article 4 – Oligomérisation d'AIDA-I

Conformation Change in a Self-recognizing Autotransporter Modulates Bacterial Cell-Cell Interaction

Journal of Biological Chemistry, Avril 2010; volume 285 (14), p. 10616-10626

Avant-propos

L'article 1 met aussi en évidence l'importance de l'auto-agrégation dans la fonctionnalité des SAATs. Nous avons donc étudié le mécanisme par lequel les SAATs permettent l'auto-agrégation des bactéries. Il a été proposé que les SAATs permettent l'auto-agrégation des bactéries par un mécanisme d'auto-association. Or, il n'existe aucune preuve de ce mécanisme auto-associatif. Nous avons donc utilisé AIDA-I comme modèle pour étudier le mécanisme d'auto-agrégation des SAATs. Nos résultats confirment qu'il s'agit bel et bien d'un mécanisme auto-associatif et que l'interaction d'AIDA-I avec lui-même en est une de haute affinité. Finalement, nous avons observé que AIDA-I pouvait osciller entre deux conformations selon la concentration en sel et en déoxycholate. Ces résultats confirment en plus l'importance de l'auto-agrégation comme phénotype principal associé à l'expression d'un SAAT.

Notre étude du mécanisme d'auto-agrégation se poursuit dans la section « Résultats supplémentaires » présentée en annexe de cette thèse.

Contributions

Contribution pour cet article :

Comme co-premier auteur, j'ai participé à la rédaction de l'article et à l'analyse des résultats. Côté expérimentation, j'ai effectué la reconstitution de l'auto-agrégation sur les billes et les expériences analysant l'effet du sel et du déoxycholate sur l'auto-agrégation.

Conformation Change in a Self-recognizing Autotransporter Modulates Bacterial Cell-Cell Interaction

Victoria Girard^{*1}, Jean-Philippe Côté^{*1}, Marie-Ève Charbonneau¹, Manuel Campos^{1,4}, Frédéric Berthiaume¹, Mark A. Hancock², Nadeem Siddiqui³ and Michael Mourez¹

* Authors contributed equally to the work

¹ Canada Research Chair on Bacterial Animal Diseases, Université de Montréal, Faculté de Médecine Vétérinaire, Saint-Hyacinthe, Québec, Canada

² Sheldon Biotechnology Centre, McGill University, Montréal, Québec, Canada

³ Institut de Recherche en Immunologie et Cancer, Université de Montréal, Montréal, Québec, Canada

⁴ Present address: Institut Pasteur, CNRS URA2172, 25 rue du Docteur Roux; 75015 Paris, France.

Running title: Characterizing AIDA-I autotransporter self-interaction

Abstract

Bacteria mostly live as multicellular communities, although they are unicellular organisms, yet the mechanisms that tie individual bacteria together are often poorly understood. The adhesin involved in diffuse adherence (AIDA-I) is an adhesin of diarrheagenic *Escherichia coli* strains. AIDA-I also mediates bacterial auto-aggregation and biofilm formation and thus could be important for the organization of communities of pathogens. Using purified protein and whole bacteria, we provide direct evidence that AIDA-I promotes auto-aggregation by interacting with itself. Using various biophysical and biochemical techniques, we observed a conformational change in the protein during AIDA-AIDA interactions, strengthening the notion that this is a highly specific interaction. The self-association of AIDA-I is of high affinity but can be modulated by sodium chloride. We observe that a bile salt, sodium deoxycholate, also prevents AIDA-I oligomerization and bacterial auto-aggregation. Thus, we propose that AIDA-I, and most likely other similar autotransporters such as antigen 43 (Ag43) and TibA, organize bacterial communities of pathogens through a self-recognition mechanism that is sensitive to the environment. This could permit bacteria to switch between multicellular and unicellular lifestyles to complete their infection.

Introduction

Many Gram-positive and Gram-negative bacterial species, including *Escherichia coli* (1), *Bordetella pertussis* (2), *Staphylococcus aureus* (3), or *Streptococcus pyogenes* (4), possess the ability to auto-aggregate. Auto-aggregation can be visualized microscopically by the clumping of bacterial cells and, in some cases, macroscopically by the settling of static cultures. It is thought that bacterial aggregates can enhance colonization, participate in biofilm formation, and confer resistance to host defenses (4-7). In addition, cell-cell aggregation is a characteristic of many multicellular behaviors such as swarming (8). Auto-aggregation is usually conferred by surface-exposed adhesins like curli (9), pili/fimbriae (1, 10, 11), or autotransporters (12-14). How bacteria achieve and control their multicellular

lifestyle has recently become a major research area. With a few notable exceptions (15), however, little is known on the mechanisms of action of the molecules that promote tight association between identical bacterial cells. A recently described family of *E. coli* outer membrane proteins called self-associating autotransporters (SAAT) provides a desirable model system to study these interactions (6).

SAAT are involved in adhesion to epithelial cells, biofilm formation, and bacterial auto-aggregation (6). To date, this family comprises three proteins: the adhesin involved in diffuse adherence (AIDA-I) (16, 17), the auto-aggregation factor Ag43 (18, 19), and the TibA adhesin/invasin (20, 21). AIDA-I was originally discovered in an *E. coli* strain isolated from a case of infantile diarrhea (16). Since then it has been associated with a high percentage of pathogenic strains of *E. coli* involved in neonatal and post-weaning diarrhea in piglets (22-26). Experimental infections have confirmed the importance of AIDA-I in the pathogenesis of AIDA-expressing strains (27). Like all autotransporters (28), AIDA-I comprises an N-terminal signal sequence, a large extracellular domain, and a C-terminal membrane-embedded domain (see Fig. 1A). The extracellular domain of AIDA-I is mainly composed of imperfect repeats of a 19-amino acid sequence that is characteristic of SAAT (6). AIDA-I is cleaved after insertion into the outer membrane (29). The extracellular fragment mature AIDA-I and the membrane-embedded fragment AIDA_c remain strongly associated despite cleavage (30). Mature AIDA-I can be released from the bacterial surface by a brief heat treatment (16), which denatures mature AIDA-I (31).

Based on observation of the aggregation of fluorescently labeled bacteria and other evidence, it was hypothesized that SAAT-mediated auto-aggregation is caused by direct SAAT-SAAT interaction (17, 18, 21). But direct biochemical evidence of this interaction is lacking. This is problematic because interaction of a surface adhesin from one cell with a different surface molecule of an adjacent cell can still permit specific auto-aggregation. For instance, yeast cells expressing the FLO1 adhesin selectively auto-aggregate, yet FLO1 does not bind to itself but rather to the cell wall of neighboring cells (32). It is therefore important to clearly ascertain whether SAAT are truly “self-associating” and to characterize the interaction these proteins can engage. For instance, whether environmental cues affect these interactions is unknown.

Here we provide a detailed characterization of the auto-aggregation mediated by AIDA-I. We first confirmed that the extracellular domain of AIDA-I mediates auto-aggregation. Using purified protein or whole bacteria, we then directly showed that AIDA-I interacts with itself, giving the first unequivocal proof that AIDA-I is self-associating. Finally, and unexpectedly, we demonstrated that AIDA-AIDA interactions and bacterial auto-aggregation can be modulated through induction of reversible conformational change in AIDA-I. This suggests that SAAT allow self-recognition of a pathogenic strain, but in a specific and highly controllable fashion, which could be a way to quickly adapt to a new environment once in the host.

Experimental procedures

Bacterial Strains and Growth Conditions

The *E. coli* K-12 strain C600 (F- thr-1 leuB6 thi-1 lacY1 supE44 rfbD1 fhuA21) obtained from New England Biolabs was used in this study and does not contain the locus coding for AIDA-I. The plasmid pAgH has been described previously (31). The plasmid expresses AIDA-I with a six-histidine and glycine tag (HisG) in the N-terminal part of mature AIDA-I and Aah, the specific glycosyltransferase of AIDA-I. Bacterial cultures were grown at 30°C on LB agar plates or in LB medium containing 100 µg/ml of ampicillin. The expression of AIDA-I was induced overnight when bacteria reached an OD_{600 nm} of 0.8 with 10 µM of isopropyl-β-D-thiogalactopyranoside, unless indicated otherwise. This low level of induction results in the production of an amount of AIDA-I that is slightly less than the maximal amount produced in the wild-type strain 2787, as checked by quantitative PCR and Western blotting (data not shown).

Purification of Proteins and Preparation of Cellular Lysates and Membrane Fractions

The purification of His-tagged AIDA-I was performed essentially as previously described (33). Briefly, 1 liter of culture of C600 expressing glycosylated AIDA-I, either wild-type or mutated, was grown to an OD_{600nm} of 0.8 and induced overnight with 10 µM isopropyl-β-D-thiogalactopyranoside. The cells were harvested and resuspended in 40 ml of Tris-buffered saline (TBS; 50 mM Tris-HCl, pH 8, 150 mM NaCl) containing lysozyme (50

$\mu\text{g/ml}$ final concentration) and EDTA (10 mM final concentration). The cells were lysed using a French press and an ultrasonic processor. After a low speed centrifugation to clear cellular debris, the total lysates were recovered. The lysates were centrifuged for 50 min in an ultracentrifuge at 250,000 X g. The pellets containing bacterial membranes were resuspended in TBS. For AIDA-I purification purposes, Triton X-100 (1% final concentration) was added to the membrane fractions, which were incubated for 1 h on ice and centrifuged again for 50 min in an ultracentrifuge. The supernatant, containing the HisG-tagged AIDA-I, was purified by immobilized metal affinity chromatography using a His Trap HP column (Amersham Biosciences). The purity of the purified protein was evaluated by SDS-PAGE followed by Coomassie Blue staining and was consistently over 95%. The concentration of the protein was measured with the BCA protein assay kit (Pierce). The purified protein was dialyzed in TBS overnight prior to use in all assays to eliminate all detergent. To purify heat-extracted protein, 1 liter from overnight culture of C600 expressing pAgH was grown as above. The cells were harvested and resuspended in 15 ml of 10 mM sodium phosphate buffer, pH 7, and heated at 60 °C for 20 min. The cells were centrifuged for 15 min at 16,000 X g, and the supernatant, containing the HisG-tagged heat-extracted AIDA-I, was purified as described before (31).

Auto-aggregation Assays

The assays were performed as previously described (31). Overnight cultures of C600 harboring an empty vector or the plasmid pAgH were normalized in 1.5 ml of LB to an OD_{600nm} of 1 to 1.5 and left standing at 4°C in deep well 96-well plates. Samples were taken at the surface every 30 min, and the OD_{600nm} was measured. Alternatively, 20 μl of each suspension was spotted on a microscope slide and incubated at 4°C, and aggregates of bacteria were examined after 3 h of incubation with a phase contrast microscope. For the inhibition experiments, purified protein in TBS, cell lysates, or membrane preparations were added to the wells at the start of the assay. To study the effect of salt on autoaggregation, the cultures were harvested and resuspended in TBS at different NaCl or sodium deoxycholate concentrations. In dose-response curves, the concentration required to achieve 50% inhibition was obtained by nonlinear regression fitting to a sigmoidal dose-

response curve, using Prism 4.0 (GraphPad Software). All other statistical comparisons (one-way and two-way ANOVA, Student's t test) were performed using Prism 4.0.

Bead Aggregation Assays

Purified AIDA-I (0.15 mg/ml) was incubated for 4 h at room temperature with 0.8 µg of sulfate-modified polystyrene green fluorescent beads (average size, 1 µm; Sigma) in 500 µl of TBS. Control beads were prepared by incubation in TBS. After incubation, the beads were centrifuged, washed with TBS containing 1% (w/v) bovine serum albumin, and resuspended in 500 µl of TBS, 1% bovine serum albumin. For sedimentation assays, the beads were diluted four times in TBS in a final volume of 500 µl and left standing at room temperature overnight. In some experiments 1 M NaCl or 0.1% sodium deoxycholate (final concentrations) was added. The samples were taken at the surface at the beginning and end of the assay, and fluorescence at 515 nm (excitation at 485 nm) was measured with a fluorometer (Jasco Spectroscopic Co. Ltd.; model FMO-427). The samples were taken at the bottom of the same overnight solutions and visualized by fluorescence microscopy with a Leica DMI4000B microscope (Meyer Instruments) equipped with epifluorescence and UV excitation modules (Chroma Technology Corporation). Alternatively, the beads were diluted five times in TBS in a final volume of 500 µl and left standing for 6 h at room temperature before flow cytometry acquisition and analysis using Cellquest Pro software on a FACScalibur instrument (BD Biosciences).

Surface Plasmon Resonance

Using a Biacore 3000 instrument (GE Healthcare), SPR experiments were performed on research grade CM5 sensor chips at 25 °C using filtered (0.2 µm) and degassed HBS-EP (10 mM HEPES, pH 7.4, 150 mM NaCl, 3 mM EDTA, 0.005% (v/v) Tween 20) for the immobilization procedure. Immobilized AIDA-I surfaces (1000-6000 RU) were prepared using the Biacore amine coupling kit as recommended; corresponding bovine serum albumin reference surfaces were prepared in the similar manner. To assess binding specificity, the kinetic analyses were performed in which AIDA-I (0-1000 nM), diluted in TBS containing 2% (w/v) bovine serum albumin, was injected over immobilized surface at 30 µl/min using KINJECT mode (5 min of association and 8 min of dissociation). In

additional trials, AIDA-I (200 nM) was heated at 60 °C for 20 min prior to injection, or AIDA-I (150 nM) was injected in the presence of increasing NaCl concentrations (150, 300, and 500 mM). To assess binding affinity, equilibrium analyses were performed in which AIDA-I was injected in a similar manner at 10 μ l/min (15 min of association and 8 min of dissociation). Steady-state binding responses were plotted as a function of AIDA-I concentration and then subjected to nonlinear regression analysis (GraphPad Prism software). For all SPR assays, the surfaces were regenerated between sample injections using 60-s pulses of 1 M NaCl at 50 μ l/min. All of the binding data are representative of duplicate injections acquired from at least three independent trials.

Far-UV CD and Intrinsic Fluorescence Spectroscopy

Far-UV CD spectra of the protein were recorded on a spectropolarimeter (Jasco Spectroscopic Co. Ltd.; model J-810) using a 0.1-cm-path length cuvette. The far-UV CD spectra of AIDA-I (125–200 μ g/ml in TBS) were recorded between 200 and 260 nm. For each spectrum, five to ten accumulations were averaged, and the contribution of buffer to the measured ellipticity was subtracted. To study the effect of salts, the protein samples were diluted in TBS with different concentrations of NaCl or sodium deoxycholate prior to measurement. Ellipticities were converted to mean residual ellipticities. Intrinsic fluorescence was recorded in the same spectropolarimeter equipped with a fluorescence emission monochromator (Jasco Spectroscopic Co. Ltd.; model FMO-427). The spectra were recorded between 300 and 400 nm after excitation at 290 nm. Five accumulations were averaged, and the contribution of the buffer was subtracted.

Chemical Cross-linking

Overnight cultures of C600 harboring the different plasmids were diluted 1:100 in fresh medium and grown to an OD_{600nm} of 0.4, at which point the expression of AIDA-I was induced for 3 h with 10 μ M isopropyl- β -D-thiogalactopyranoside. The cultures were normalized to the same OD_{600nm} of 1 and washed in phosphate-buffered saline. The samples (25 μ l) were incubated for 30 min at room temperature with different concentrations of bis[sulfosuccinimidyl] suberate (BS³; Pierce), which is a membrane-impermeable NHS-ester reagent reacting with amino groups to form stable amide bonds. The best reactions

were found to occur at 300 μM BS³. The reaction was stopped by incubating for an additional 15 min with 50 mM Tris-HCl, pH 7.5. A similar reaction was conducted with purified proteins at a concentration of 0.2 mg/ml in 1 mM Tris-HCl, pH 8, 150 mM NaCl. The samples were diluted in 2 X SDS-PAGE loading buffer containing β -mercaptoethanol and denatured by heating at 100°C for 10 min. The samples were then separated by SDS-PAGE at 4°C on 7% acrylamide gels and analyzed by immunoblotting. Immunodetection was performed with a serum raised against heat-extracted mature AIDA-I diluted 1:50,000 in blocking buffer (5% skim milk, 50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.05% Triton X-100). A goat anti-rabbit horseradish peroxidase-conjugated antibody (Sigma) was used as a secondary antibody according to the instructions of the manufacturer. Immune complexes were revealed using a 3,3',5,5'-tetramethylbenzidine solution for membranes (Sigma).

Analytical Ultracentrifugation

Purified AIDA-I in TBS was diluted in TBS alone, in TBS containing NaCl so as to reach a final concentration of 1 M, or in TBS containing sodium deoxycholate so as to reach a final concentration of 0.5%. All of the samples contained \sim 0.69 mg/ml of protein, which gave an absorbance of \sim 0.7 at 280 nm. Sedimentation velocity experiments were performed on the samples using a Beckman XL-I analytical ultracentrifuge. The reference and sample sectors of a 1.2-cm-path length double-sector ultracentrifuge cell were filled with 400 μl of the appropriate buffer and 390 μl of protein, respectively. Sedimentation velocity runs were performed at either 20,000 or 30,000 rpm, with absorbance scans monitored at 280 nm in 10-min intervals over a total spin time of 4-6 h at 20°C. Molecular mass, buffer viscosity, density, and hydration of the protein were calculated using the software SEDNTERP. Molecular mass and hydration of the protein were calculated assuming 20 heptose residues/molecule. Sedimentation velocity data were analyzed by the computer program SEDFIT (34). Sedimentation profiles were obtained by fitting the absorbance data to the continuous $c(S)$ model.

Results

Evidence That AIDA-I Mediates Auto-aggregation via Direct Self-interaction

As shown in previous reports (6, 17, 31), we observed that bacteria from standing cultures of *E. coli* expressing AIDA-I form large aggregates of cells that can be observed by phase contrast microscopy (supplemental Fig. S1). This was not the case with bacteria bearing a control empty plasmid. These aggregates settle at the bottom of culture tubes, which can be easily followed with a spectrophotometer by measuring optical density at the top of the bacterial cultures (supplemental Fig. S1 and Fig. 1B).

We isolated in a previous study (33) a mutant of AIDA-I, I24, bearing an insertion of a 5-amino acid sequence at amino acid 574 in the C-terminal region of the imperfect repeats of AIDA-I (Fig. 1A). As observed before, this mutant is unable to mediate auto-aggregation (Fig. 1B). When we mixed equal amounts of bacteria expressing wild-type AIDA-I and bacteria expressing I24, we observed an intermediate amount of settling. Furthermore, at the top of the cultures, there were fewer bacteria expressing WT compared with bacteria expressing I24 (Fig. 1C). Thus, AIDA-I is able to mediate auto-aggregation preferentially between bacteria expressing the wild-type protein and exclude the bacteria expressing I24. This argues that AIDA-I functions by interacting with itself, directly or indirectly.

We decided to exclude the possibility that AIDA-I engage another molecule to mediate auto-aggregation using another approach. As indicated above, strain-specific auto-aggregation of yeast mediated by the FLO1 adhesin involves interaction of FLO1 with neighboring cell wall molecules (32). We therefore added different subcellular extracts of bacteria harboring an empty vector or expressing AIDA-I. We observed that auto-aggregation is totally inhibited by the addition of a lysate or a membrane extract of bacteria expressing AIDA-I but, by contrast, the addition of an identical amount of cell lysate or a membrane extract of bacteria harboring an empty vector had no significant effect (Fig. 2A). This experiment suggested that AIDA-I does not mediate auto-aggregation by interacting with another cell surface molecule.

We then added purified whole AIDA-I to standing cultures of bacterial cells expressing AIDA-I. As shown in Fig. 2B, we observed that the protein can inhibit auto-

aggregation in a dose-dependent and saturable manner with 50% inhibition (IC_{50}) reached at a concentration of ~ 34 nM. Such a low concentration strongly suggests that this effect is specific. By contrast, neither the purified I24 mutant nor the heat-extracted portion of the extracellular domain of AIDA-I were able to prevent auto-aggregation (Fig. 2C). This is consistent with the inability of I24 to mediate auto-aggregation even with wild-type AIDA-I and with the fact that the heat-extracted extracellular domain of AIDA-I is denatured (31). In addition, the fact that heat-denatured AIDA-I did not prevent auto-aggregation suggests that the inhibition is not due to a putative heat-resistant contaminant in our protein preparation.

To finally prove that AIDA-I mediates auto-aggregation by direct self-interaction, we coated fluorescent polystyrene beads with purified protein. As shown in Fig. 3A, we observed that beads coated with wild-type AIDA-I could auto-aggregate, whereas control empty beads or beads coated with the purified I24 mutant did not. This phenomenon could be quantified using a sedimentation assay similar to the one used with whole bacteria (Fig. 3B) or by flow cytometry analysis of particle fluorescence and scattering (Fig. 3C). These assays clearly show the magnitude of the aggregation mediated solely by AIDA-I, and the I24 control ascertains that this is a specific effect. In conclusion, taken together our results represent, to our knowledge, the first direct proof that SAAT-mediated auto-aggregation is due to direct self-interaction.

Direct Interaction of the Extracellular Domain of AIDA-I in Vitro

To directly demonstrate that AIDA-I can associate with itself, we used SPR. Reference and purified AIDA-I surfaces were immobilized on an SPR chip. A series of soluble AIDA-I concentrations were then titrated over the prepared surfaces in tandem. Typical kinetic analyses exhibited specific, dose-dependent binding of soluble AIDA-I to immobilized AIDA-I (Fig. 4A). No detectable binding responses were observed, however, when AIDA-I was heated at 60°C for 20 min prior to injection. This result of heat treatment, known to denature the mature portion of AIDA-I but not AIDA_c (31), strongly suggests that the binding observed by SPR involves interactions between the mature AIDA-I extracellular fragments. To assess binding affinity, the extremely slow off rates observed precluded fitting of the kinetic titrations to a simple 1:1 binding model (i.e. simultaneous on

and off rate determinations). Alternatively, equilibrium analyses were performed in which the steady-state binding responses were plotted as a function of AIDA-I concentration to determine the overall equilibrium dissociation constant, K_D (Fig. 4B). The apparent K_D predicts an interaction of high affinity (48 ± 10 nM) and was not dependent on the level of protein immobilized on the chip. The latter confirms that the SPR signal is specific. Additional SPR experiments with different chips, buffers, and conditions yielded highly similar results (supplemental Fig. S2), strengthening the notion that AIDA-AIDA interaction is specific. Interestingly, the K_D of the interaction is in excellent agreement with the IC_{50} determined in our inhibition experiment with purified AIDA-I (34 nM; Fig. 2B). Our results therefore show that AIDA-I can interact with itself in vitro with a high affinity and, to our knowledge, constitutes the first direct observation of a SAAT-SAAT interaction.

Observation of AIDA-I Oligomers in Vitro and in Whole Bacteria

To characterize the oligomers made by AIDA-I, we used analytical ultracentrifugation (AUC). The sedimentation profile of the purified protein showed two discrete peaks (Table 1 and supplemental Figs. S3 and S4). When the same experiment was performed with the purified I24 mutant, we observed that the protein was monomeric, with a sedimentation coefficient similar to the one of the wild-type monomer. This is consistent with the fact that I24 does not allow auto-aggregation and thus most likely remains monomeric. The computed molecular masses of the first peak observed with wild-type AIDA-I and of the peak of I24 are consistent with the predicted molecular mass of AIDA-I. However, the computed molecular mass of the second peak observed with wild-type AIDA-I is too small for a dimer. There is little doubt that this second peak comes from a dimeric species, however. The aberrant computed mass could be explained if the two species have different shapes and therefore different frictional coefficients, whereas the fitting program uses the same coefficient for both to make its calculations. When constraining the molecular masses of the two species as a monomer and a dimer of AIDA-I, a solution could be found with two different frictional coefficients (Table 1). This shows that the peaks could be interpreted as a monomer and a dimer if these two species have different conformations. More evidence confirming this hypothesis is presented below. All

of the calculated frictional coefficients indicate that the overall shape of the AIDA-I protein is elongated. This is in agreement with structural prediction (35) and the known structures of other extracellular domains of autotransporters, which mainly consists of elongated β -helices (36). Overall, this experiment confirms that AIDA-I is able to oligomerize in solution. It is intriguing to see more of the monomer form of AIDA-I than the dimer form, despite the protein concentration being far above the K_D measured by SPR. This could be explained if dimers themselves were associating into a range of higher order oligomers, which would be too big and heterogeneous to be observed in this AUC experiment. Evidence for this is provided below.

To test for direct interactions between AIDA-I molecules in whole bacteria, we performed cell surface cross-linking experiments with BS³, a membrane-impermeable NHS-ester reagent bridging primary amino groups of proteins. We added a low amount of BS³ to cultures of bacteria expressing AIDA-I. By immunoblotting with an antibody raised against mature AIDA-I, we observed the 100- and 150-kDa bands of mature extracellular AIDA-I and pro-protein, respectively (Fig. 5A). In addition, we observed bands of ~200 and 300 kDa as well as fainter bands of even higher molecular masses. These bands could correspond to dimers and higher order oligomers of mature extracellular domains of AIDA-I. None of these oligomers can be observed when bacteria expressed the I24 mutant, showing that the cross-links observed with the wild-type protein are indeed specific and that AIDA-I interacts with itself in whole bacteria through its mature AIDA-I portion. Similar patterns of cross-linking were observed with purified proteins in solution (Fig. 5B) and confirm that AIDA-I is making similar higher order oligomers in solution. As stated above, this would explain the unexpected monomer:dimer ratio observed during sedimentation in analytical ultracentrifugation. In summary, these results characterize AIDA-I interactions with itself, both in vitro and in vivo, and show that they involve heterogeneous oligomers.

Modulation of AIDA-AIDA Interaction and Bacterial Auto-aggregation

Because ionic interactions are often important in protein-protein interactions, we assessed AIDA-AIDA interactions at different NaCl concentrations. Typical SPR binding profiles showed that AIDA-AIDA interaction is salt-sensitive, and the association

decreased dramatically with increasing NaCl concentrations (Fig. 6A). The AUC sedimentation profile of purified AIDA-I in the presence of 1 M NaCl also showed that AIDA-I is a single species (Table 1). In addition, the presence of 1 M NaCl prevented the aggregation of beads coated with AIDA-I (Fig. 3, A and B). Together, these experiments show that high salt concentrations disrupt AIDA-AIDA interactions. We investigated the effect of ionic strength on the settling of cultures of bacterial cells expressing AIDA-I by resuspending bacteria in buffers at different NaCl concentrations (Fig. 6B). We observed a significant decrease in the settling of bacterial cells at 300 mM NaCl and no aggregation at 500 mM NaCl. These results correlate well with the *in vitro* measurements of AIDA-AIDA interactions, thereby reinforcing the notion that direct AIDA-AIDA interactions mediate auto-aggregation. As previously noted (31), the extracellular domain of AIDA-I is extremely tightly associated to AIDAc and did not dissociate even in the presence of up to 2 M NaCl (data not shown). In summary, these results suggest that electrostatic interactions are involved in the self-association of AIDA-I *in vitro* and in whole bacteria.

Modulation of AIDA-AIDA Interaction Involves a Reversible Conformation Change

As indicated above, AUC suggested that monomeric and oligomeric AIDA-I have different conformations. We performed far-UV CD spectroscopy on purified AIDA-I at different ionic strengths to directly assess the influence of oligomerization on the conformation of AIDA-I. The CD spectrum of AIDA-I shows a clear single minimum of ellipticity at 218 nm (Fig. 7A), a characteristic of β -strands, in agreement with the predicted structure of AIDA-I and our previous observations (31, 34). Upon increase of NaCl concentration, the CD spectrum of AIDA-I changed dramatically, with an important decrease in ellipticity and the minimum of ellipticity shifting toward a lower wavelength. This result signals important conformational changes. Measuring the intrinsic fluorescence of purified AIDA-I in the presence of various concentrations of NaCl showed that high salt concentrations also caused a dramatic increase in intrinsic fluorescence (Fig. 7B), another indicator of conformational changes. Interestingly, I24 had the same CD and intrinsic fluorescence spectra at low and high NaCl concentrations, and these spectra were similar to the ones of AIDA-I in the presence of high salt concentration (Fig. 7, C and D). This is consistent with the fact that I24 is unable to mediate auto-aggregation, does not form

complexes by chemical cross-linking, and is monomeric as seen in AUC. Taken together, these experiments suggest that the conformational change observed with increasing NaCl concentration corresponds to a transition toward the conformation of a monomeric protein and that the monomer and oligomers have different conformations. The difference in frictional coefficients between the two species observed in AUC is also in agreement with these results. Strikingly, regardless of the method, we observed that these conformational changes are completely reversible and the protein can repeatedly switch back and forth between conformations, strengthening the notion that these changes are very specific.

Bile Salt Can Modulate AIDA-AIDA Interaction and Auto-aggregation

We wondered whether another physiologically relevant condition could alter AIDA-AIDA interaction and bacterial auto-aggregation. Because AIDA-I is expressed in diarrheagenic strains of *E. coli*, it is bound to encounter the bile secreted in the intestine (37, 38). We therefore tested the effects of sodium deoxycholate, a typical bile salt, on AIDA-I conformation and oligomerization. Strikingly, we observed that bacteria expressing AIDA-I could not auto-aggregate when incubated in the presence of sodium deoxycholate (Fig. 8A). The sedimentation of AIDA-I observed by AUC in the presence of deoxycholate showed that it is monomeric (Table 1). The CD spectrum of AIDA-I in these conditions (Fig. 8B) was also similar to the one observed in monomeric AIDA-I. Lastly, aggregation of beads coated with AIDA-I was also disrupted by sodium deoxycholate (Fig. 3, A and B). Taken together, these results show that a physiological cue is able to prevent AIDA-AIDA interaction as well as auto-aggregation.

Discussion

The autotransporter adhesins of *E. coli* AIDA-I, TibA and Ag43, have been grouped into the family of self-associating autotransporters (6). As the name implies, preliminary evidence suggested that these proteins mediate bacterial auto-aggregation through self-interaction (17, 21, 39). The assumption that a SAAT interacts with itself was never actually directly tested, however. In this study we bring proof that AIDA-I does indeed interact with itself, using *in vitro* and *in vivo* approaches. Our experiments also show that

this self-interaction is directly responsible for auto-aggregation. In particular, we observed that purified protein could trigger auto-aggregation of beads and compete with auto-aggregation of bacteria. In addition we observed that high salt concentrations were able to disrupt AIDA-AIDA interaction as well as auto-aggregation. The latter result suggests that charged or polar residues contribute to AIDA-AIDA interactions. A similar effect of high salt concentrations was observed on bacterial aggregation mediated by Ag43 (18), consistent with the idea that the self-interaction mechanism might be common in all SAAT. Interestingly, we observed that AIDA-I interactions result in the formation of a range of oligomers beyond dimers. This suggests that AIDA-AIDA interactions involve proteins on the same bacterium as well as proteins on adjacent bacteria. Additional experiments suggest that this is indeed the case (Cote JP et al., unpublished results).

We observed that high salt concentrations also caused a substantial change in AIDA-I conformation, suggesting that the AIDA-I protein has significantly different conformations depending on its oligomeric state. Large conformational changes are also observed in the mammalian cell adhesion surface proteins cadherins (40), allowing them to play their role as fine-tuned and modulable organizers of all tissues. Their conformations are calcium-dependent, and the conformational changes act as a switch modulating the engagement of cell-cell interactions. We observed that sodium deoxycholate, a bile salt, was able to disrupt auto-aggregation and AIDA-AIDA interaction, as well as change AIDA-I conformation. Because the osmolality of the intestine is usually high (~300 mM) and in the range of the NaCl concentration we tested and because the concentration of bile salt we tested is also in the physiological range, it is tempting to speculate that AIDA-I is ideally suited to allow bacteria to sense the intestinal environment and modulate cell-cell interactions. Similarly, zinc released from immune cells was recently hypothesized to induce conformational changes and auto-association of Staphylococcal adhesins (15). It should also be noted that the osmolality and bile salt content of the intestine are not constant spatially and temporally. Thus, environmental regulation of AIDA-I self-interaction could help bacteria to adapt to a specific niche and/or a specific time. In this respect it is interesting to note that diarrheagenic strains of *E. coli* bearing AIDA-I are most often isolated in cases of diarrhea occurring post-weaning, a situation where the animals are stressed and undergo significant changes in intestinal homeostasis.

It has been previously observed that some enteropathogenic and enteroaggregative strains of *E. coli* are able to make bacterial aggregates through the expression of pili or fimbriae and that these aggregates are important for pathogenesis (11, 41). The roles these bacterial aggregates could play are not precisely known, but they could enhance colonization, participate in biofilm formation, and/or confer resistance to host defenses. Additionally, auto-aggregation could be construed as a way to discriminate “self” from “nonself.” This distinction could be advantageous by ensuring that pathogens only associate with their own kin, maximizing their infectivity by increasing their local density as well as by optimizing their pathogenic potential in bringing together their virulence factors. In this view, AIDA-I would represent a peculiarly simple and versatile example of such a tag defining self. In the previous examples of pathogenic *E. coli* forming aggregates described above, a dedicated protein can disperse the aggregates, and this dispersion was also shown to be required for efficient pathogenesis (11, 40). Although the reason why dispersion is needed is unclear, it was hypothesized that disaggregation is required to establish new foci of infection (11) or penetrate the mucous layer of the intestinal epithelium (40). Similarly, but in a considerably simpler and faster manner, bile salts could disperse AIDA-mediated aggregates and allow efficient infection. In addition, SAAT, like other autotransporters promoting auto-aggregation (13, 14), are multifunctional proteins and can mediate adhesion to host cells (6). Regulation of self-interaction could also act as a switch mechanism between these multiple activities.

In conclusion, we have shown here that AIDA-I can directly mediate auto-aggregation through self-interaction and that the environment could potentially modulate this interaction through specific conformation change. We thus propose that AIDA-I and similar other autotransporters such as Ag43 and TibA represent a versatile and simple family of tunable bacterial adhesion molecules enabling specific social behavior at an appropriate time. Further work on this type of adhesins is warranted to thoroughly test and pursue the implications of this hypothesis.

Acknowledgements

We thank M. Ramjeet, J.-M. Ghigo and C. Beloin for critical reading of the manuscript and insightful comments. This work was supported by financial contributions from the

Canadian Institutes for Health Research (CIHR grant 84578), the Groupe de Recherche et d'Études sur les Maladies Infectieuses du Porc (GREMIP), and the Canada Research Chair and Canada Foundation for Innovation programs (grant 201414). Sheldon Biotechnology Centre is supported by a Research Resource Grant from CIHR. The Analytical Ultra Centrifugation experiments were completed at the Biophysical Platform at IRIC.

References

1. 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* 60:2297-2304.
2. Menozzi, F. D., P. E. Boucher, G. Riveau, C. Gantiez, and C. Locht. 1994. Surface-associated filamentous hemagglutinin induces autoagglutination of *Bordetella pertussis*. *Infect Immun* 62:4261-4269.
3. McDevitt, D., P. Francois, P. Vaudaux, and T. J. Foster. 1994. Molecular characterization of the clumping factor (fibrinogen receptor) of *Staphylococcus aureus*. *Mol Microbiol* 11:237-248.
4. Frick, I. M., M. Morgelin, and L. Bjorck. 2000. Virulent aggregates of *Streptococcus pyogenes* are generated by homophilic protein-protein interactions. *Mol Microbiol* 37:1232-1247.
5. Berge, A., B. M. Kihlberg, A. G. Sjöholm, and L. Bjorck. 1997. *Streptococcal* protein H forms soluble complement-activating complexes with IgG, but inhibits complement activation by IgG-coated targets. *J Biol Chem* 272:20774-20781.
6. Klemm, P., R. M. Vejborg, and O. Sherlock. 2006. Self-associating autotransporters, SAATs: functional and structural similarities. *Int J Med Microbiol* 296:187-195.
7. Ochiai, K., T. Kurita-Ochiai, Y. Kamino, and T. Ikeda. 1993. Effect of co-aggregation on the pathogenicity of oral bacteria. *J Med Microbiol* 39:183-190.

8. Verstraeten, N., K. Braeken, B. Debkumari, M. Fauvart, J. Fransaer, J. Vermant, and J. Michiels. 2008. Living on a surface: swarming and biofilm formation. *Trends Microbiol* 16:496-506.
9. Hammar, M., Z. Bian, and S. Normark. 1996. Nucleator-dependent intercellular assembly of adhesive curli organelles in *Escherichia coli*. *Proc Natl Acad Sci U S A* 93:6562-6566.
10. Schembri, M. A., G. Christiansen, and P. Klemm. 2001. FimH-mediated autoaggregation of *Escherichia coli*. *Mol Microbiol* 41:1419-1430.
11. 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* 280:2114-2118.
12. Girard, V., and M. Mourez. 2006. Adhesion mediated by autotransporters of Gram-negative bacteria: structural and functional features. *Res Microbiol* 157:407-416.
13. Alamuri, P., and H. L. Mobley. 2008. A novel autotransporter of uropathogenic *Proteus mirabilis* is both a cytotoxin and an agglutinin. *Mol Microbiol* 68:997-1017.
14. Felek, S., M. B. Lawrenz, and E. S. Krukonis. 2008. The *Yersinia pestis* autotransporter YapC mediates host cell binding, autoaggregation and biofilm formation. *Microbiology* 154:1802-1812.
15. Conrady, D. G., C. C. Brescia, K. Horii, A. A. Weiss, D. J. Hasset, and A. B. Herr. 2008. A zinc-dependent adhesion module is responsible for intercellular adhesion in staphylococcal biofilms. *Proc Natl Acad Sci U S A* 105:19456-19461.
16. 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* 57:1506-1511.
17. Sherlock, O., M. A. Schembri, A. Reisner, and P. Klemm. 2004. Novel roles for the AIDA adhesin from diarrheagenic *Escherichia coli*: cell aggregation and biofilm formation. *J Bacteriol* 186:8058-8065.
18. 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* 51:283-296.

19. 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* 149:115-120.
20. Elsinghorst, E. A., and J. A. Weitz. 1994. Epithelial cell invasion and adherence directed by the enterotoxigenic *Escherichia coli* tib locus is associated with a 104-kilodalton outer membrane protein. *Infect Immun* 62:3463-3471.
21. 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* 73:1954-1963.
22. 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* 15:378-381.
23. Mainil, J. G., E. Jacquemin, P. Pohl, A. Kaeckenbeeck, and I. Benz. 2002. DNA sequences coding for the F18 fimbriae and AIDA adhesin are localised on the same plasmid in *Escherichia coli* isolates from piglets. *Vet Microbiol* 86:303-311.
24. Ngeleka, M., J. Pritchard, G. Appleyard, D. M. Middleton, and J. M. Fairbrother. 2003. Isolation and association of *Escherichia coli* AIDA-I/STb, rather than EAST1 pathotype, with diarrhea in piglets and antibiotic sensitivity of isolates. *J Vet Diagn Invest* 15:242-252.
25. 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* 8:143-149.
26. Zhang, W., M. Zhao, L. Ruesch, A. Omot, and D. Francis. 2007. Prevalence of virulence genes in *Escherichia coli* strains recently isolated from young pigs with diarrhea in the US. *Vet Microbiol* 123:145-152.
27. Ravi, M., M. Ngeleka, S. H. Kim, C. Gyles, F. Berthiaume, M. Mourez, D. Middleton, and E. Simko. 2007. Contribution of AIDA-I to the pathogenicity of a porcine diarrheagenic *Escherichia coli* and to intestinal colonization through biofilm formation in pigs. *Vet Microbiol* 120:308-319.

28. 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* 68:692-744.
29. Suhr, M., I. Benz, and M. A. Schmidt. 1996. Processing of the AIDA-I precursor: removal of AIDA_c and evidence for the outer membrane anchoring as a beta-barrel structure. *Mol Microbiol* 22:31-42.
30. Charbonneau, M. E., J. Janvore, and M. Mourez. 2009. Autoprocessing of the *Escherichia coli* AIDA-I autotransporter: a new mechanism involving acidic residues in the junction region. *J Biol Chem* 284:17340-17351.
31. Charbonneau, M. E., F. Berthiaume, and M. Mourez. 2006. Proteolytic processing is not essential for multiple functions of the *Escherichia coli* autotransporter adhesin involved in diffuse adherence (AIDA-I). *J Bacteriol* 188:8504-8512.
32. Smukalla, S., M. Caldara, N. Pochet, A. Beauvais, S. Guadagnini, C. Yan, M. D. Vinces, A. Jansen, M. C. Prevost, J. P. Latge, G. R. Fink, K. R. Foster, and K. J. Verstrepen. 2008. FLO1 is a variable green beard gene that drives biofilm-like cooperation in budding yeast. *Cell* 135:726-737.
33. Charbonneau, M. E., and M. Mourez. 2007. Functional organization of the autotransporter adhesin involved in diffuse adherence. *J Bacteriol* 189:9020-9029.
34. Lebowitz, J., M. S. Lewis, and P. Schuck. 2002. Modern analytical ultracentrifugation in protein science: a tutorial review. *Protein Sci* 11:2067-2079.
35. Kajava, A. V., and A. C. Steven. 2006. 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.
36. 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* 103:4918-4923.
37. Begley, M., C. G. Gahan, and C. Hill. 2005. The interaction between bacteria and bile. *FEMS Microbiol Rev* 29:625-651.

38. Malik-Kale, P., C. T. Parker, and M. E. Konkel. 2008. Culture of *Campylobacter jejuni* with sodium deoxycholate induces virulence gene expression. *J Bacteriol* 190:2286-2297.
39. 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* 188:1798-1807.
40. Leckband, D., and A. Prakasam. 2006. Mechanism and dynamics of cadherin adhesion. *Annu Rev Biomed Eng* 8:259-287.
41. Sheikh, J., J. R. Czczulin, 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* 110:1329-1337.

Figure legends

Figure 1: Auto-aggregation of bacterial cells expressing AIDA-I (A) schematic representation of AIDA-I and the location of an auto-aggregation defective mutant, I24. The white box represents the signal sequence, the light gray boxes represent the imperfect 19-amino acids repeats, and the dark box represents the membrane-embedded fragment, AIDA_c. (B and C) auto-aggregation in mixes of bacteria expressing AIDA-I or I24. Overnight cultures of bacteria expressing wild-type AIDA-I (WT) or the I24 mutant were left standing in culture tubes for 180 min. The cultures have resistance against ampicillin (AmpⁱR) or against ampicillin and chloramphenicol (AmpⁱR/Cml^rR), respectively. In (B), the OD_{600nm} values at the top of the cultures were compared with the values at the beginning of the assay. The columns were compared with the WT control by performing a one-way ANOVA and Dunnett's post tests. In (C), the colony forming units (cfu) of bacteria bearing the different combinations of markers were evaluated, and a ratio was calculated. The ratios were compared by a Student's t test and were found to be statistically different. *, statistical significance (p < 0.05); NS, not significant. The assays were performed three times in duplicate.

Figure 2: Inhibition of auto-aggregation by purified AIDA-I and not by cellular extracts (A) cultures bearing an empty vector (-) or expressing AIDA-I (WT) were incubated in deep well 96-wells plates with or without total lysates (L) or membrane extracts (M) prepared from cultures harboring the empty vector or expressing AIDA-I and left standing for 90 min. The columns were compared with the wild-type control by performing a one-way ANOVA and Dunnett's post tests. (B) auto-aggregation assays of cultures of bacteria expressing AIDA-I were performed with various concentrations of purified AIDA-I present. The points were fitted by nonlinear regression to a sigmoidal dose-response. (C) normalized cultures of bacteria harboring an empty vector or expressing AIDA-I were incubated with or without 200 nM wild-type protein, mutated protein I24, or heat-extracted purified protein (HE). The columns were compared with the wild-type vector control by performing a one-way ANOVA and Dunnett's post tests. *, statistical significance ($p < 0.05$); *NS*, not significant.

Figure 3: Auto-aggregation of fluorescent beads mediated by purified AIDA-I (A) solutions containing empty fluorescent polystyrene beads (Control), AIDA-coupled beads (AIDA), and I24-coupled beads (I24) were left standing overnight, and the samples were taken at the bottom of the tube and examined by fluorescent microscopy. The experiments were performed with AIDA-coupled beads in the presence of 1 M NaCl or 0.1% sodium deoxycholate (DOC). (B) Samples were taken on top of the solutions containing the beads at the beginning of the assay (white bars) and after overnight incubation (black bars), and fluorescence at 515 nm was measured. (C) The solutions containing empty beads (black) or AIDA-coupled beads (red) solutions were left standing at room temperature for 6 h and then analyzed by flow cytometry. All of the experiments were repeated three times; typical results are shown.

Figure 4: *In vitro* evidence of AIDA-AIDA interactions (A) representative SPR for kinetic titrations of soluble AIDA-I (0-1000 nM) binding to 6,700 RU amine-coupled AIDA-I at 30 μ l/min. The dashed line indicates 200 nM injection of heat-treated AIDA-I (60°C for 20 min). (B) representative SPR for equilibrium titrations of soluble AIDA-I (0-1000 nM) binding to immobilized AIDA-I at 10 μ l/min. The symbols represent RU at

equilibrium (Req), and the lines represent nonlinear regression analysis. The average of three independent trials yielded apparent K_D constants of 36 nM (top curve, 6700 RU AIDA-I), 59 nM (middle curve, 2500 RU AIDA-I), and 52 nM (bottom curve, 1000 RU AIDA-I).

Figure 5: Chemical cross-linking of AIDA-I (A) cultures of bacteria expressing wild-type AIDA-I (WT) or the I24 mutant were incubated with or without 300 μ M BS³ cross-linker for 30 min. The samples were diluted in 2 X SDS-PAGE loading buffer and immunoblotted with an antibody directed against the extracellular part of AIDA-I. (B) purified proteins (0.2 mg/ml in 1 mM Tris, pH 8, 150 mM NaCl) were cross-linked in the same conditions, and the samples were treated by SDS-PAGE and immunoblotting similarly. pro., pro-protein; mat., mature extracellular AIDA-I.

Figure 6: Modulation of AIDA-AIDA interaction and bacterial auto-aggregation (A) kinetic SPR analysis in which purified AIDA-I (150 nM) was injected in the presence of increasing NaCl concentration. The experiments were performed three times. The average responses observed were normalized to the response observed at 150 mM NaCl. The columns were compared by performing a one-way ANOVA and Dunnett's post tests. (B) overnight cultures of bacteria expressing AIDA-I (WT) or harboring an empty vector were resuspended in 50 mM Tris-HCl pH 8 buffers at different ionic strengths and were left standing at 4°C. The OD_{600nm} at the top of the culture was measured at the end of the assay after 180 min and compared with the value at the beginning of the assay. The assays were performed three times in duplicate, and the columns were compared by a two-way ANOVA and Bonferroni post-tests. *, statistical significance ($p < 0.05$); NS, not significant.

Figure 7: Conformation change in AIDA-I (A) far-UV CD spectra of purified wild-type AIDA-I when sodium chloride was added to the buffer to yield a final concentration of 150, 330, 660, or 1000 mM. The signal was converted to mean residual ellipticity (MRE). (B) intrinsic fluorescence of purified AIDA-I after excitation at 290 nm was recorded in the same buffers and reported in arbitrary units (a.u.). (C and D) far-UV CD spectra (C) and

intrinsic fluorescence (D) of the purified I24 mutant were obtained as in A and B, in the presence of 150 mM or 1 M NaCl.

Figure 8: Effect of bile salt on AIDA-AIDA interaction and auto-aggregation (A)

overnight cultures of bacteria expressing AIDA-I (WT) or harboring an empty vector were resuspended in TBS with or without 0.1% sodium deoxycholate and were left standing at 4°C. The OD_{600nm} at the top of the culture was measured at the end of the assay after 180 min and compared with the value at the beginning of the assay. The assays were performed three times in duplicate, and columns were compared by a two-way ANOVA and Bonferroni post-tests. *, statistical significance ($p < 0.05$); *NS*, not significant. (B) far-UV CD spectrum of purified AIDA-I was recorded in TBS without (solid line) or with (dashed line) 0.1% sodium deoxycholate (DOC).

Table**Table I: Analytical ultracentrifugation of purified AIDA-I**

Sedimentation velocity experiments were performed with purified wild-type or mutant AIDA-I diluted in 50 mM Tris-HCl, pH 8, containing 150 mM NaCl, 1 M NaCl, or 150 mM NaCl and 0.5% sodium deoxycholate. Sedimentation (c) and frictional (f/f_0) coefficients and molecular mass were then determined by fitting the data from the sedimentation profiles, and the root mean square deviation (RMSD) from the best solution is indicated.

Protein	Condition	c (S)	Molecular mass	f/f_0	RMSD
			<i>kDa</i>		
Wild-type	Peak 1 with 150 mM NaCl	5.65 ± 0.234	133 ± 10	1.95^a	0.00434
	Peak 2 with 150 mM NaCl	7.27 ± 0.36	189 ± 12	1.95^a	0.00434
	1 M NaCl	4.375 ± 0.17	124 ± 8	1.77	0.00561
	0.5% sodium deoxycholate	5.85 ± 0.2	135 ± 8	1.66	0.00542
I24	150 mM NaCl	5.13 ± 0.25	133 ± 10	1.89	0.00537

^a This is the average frictional coefficient for the two species assumed to have a similar shape; if the two species are constrained to be a monomer and a dimer of AIDA-I during the fitting, then two separate frictional coefficient can be calculated and are 1.69 for peak 1 and 2.08 for peak 2.

Figures

Figure 1

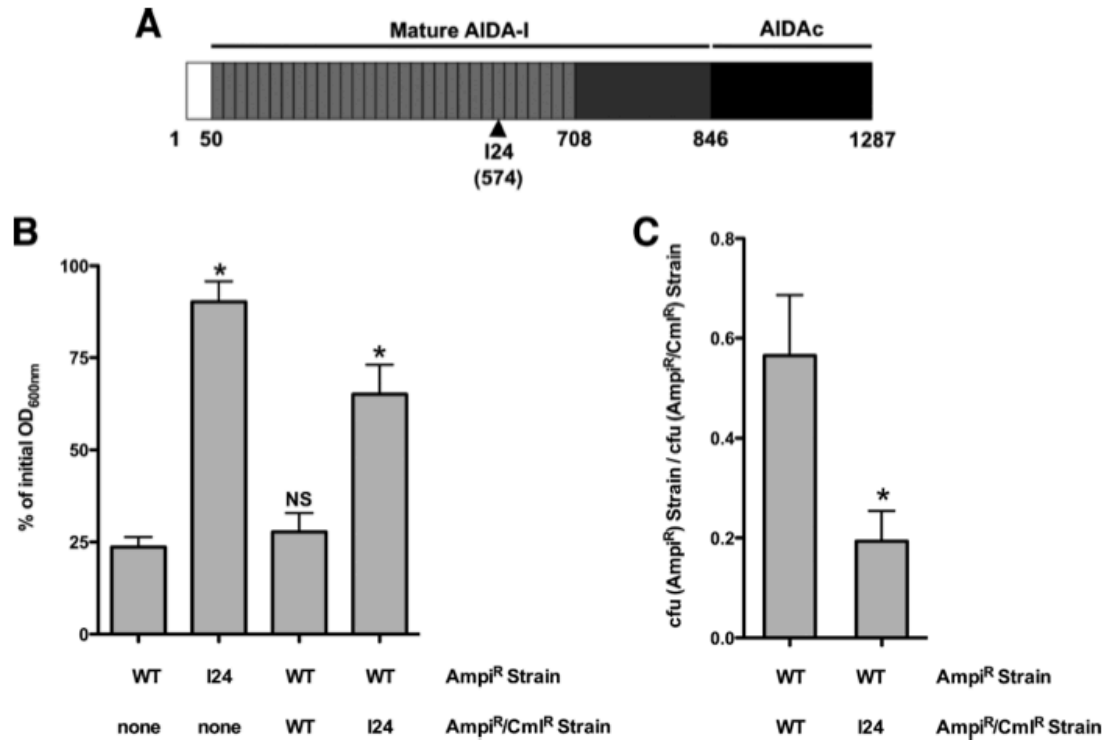


Figure 2

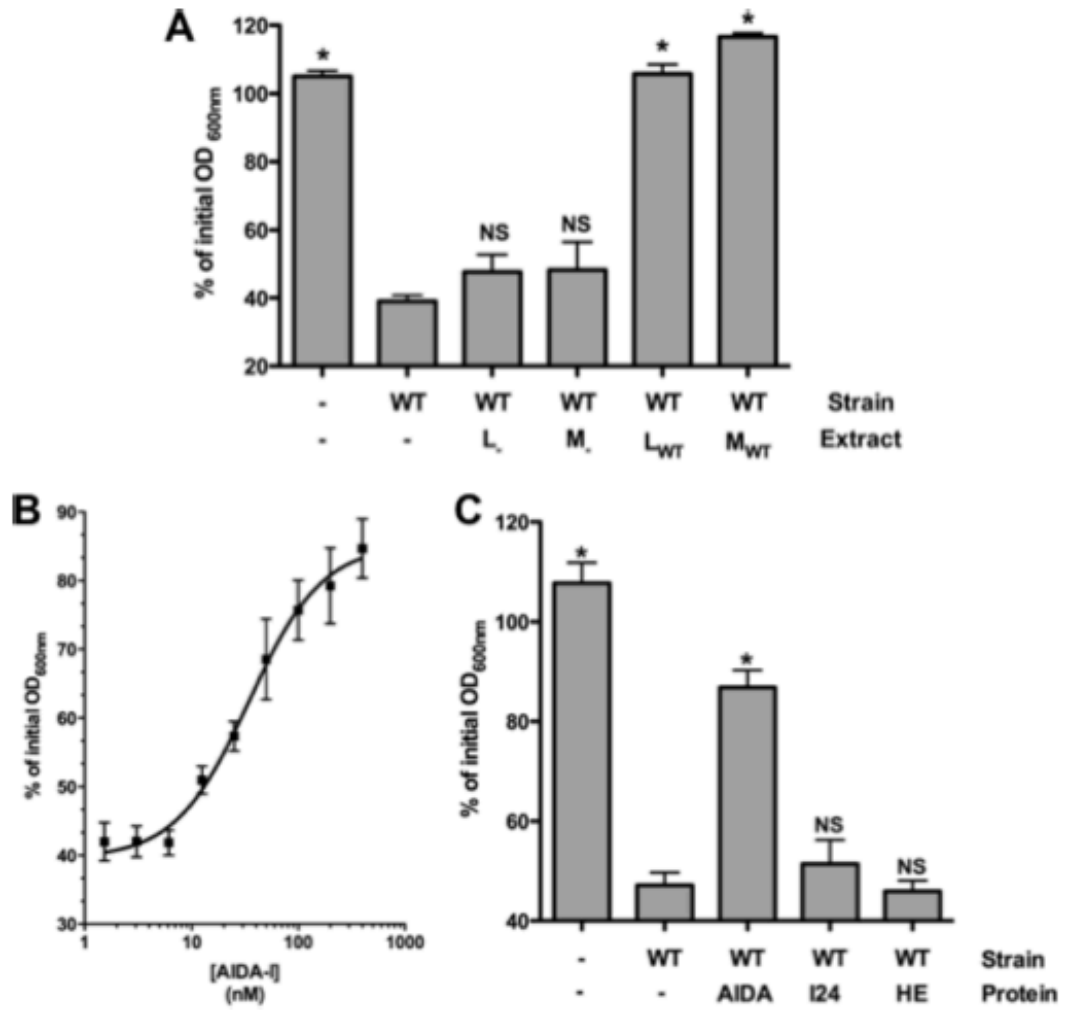


Figure 3

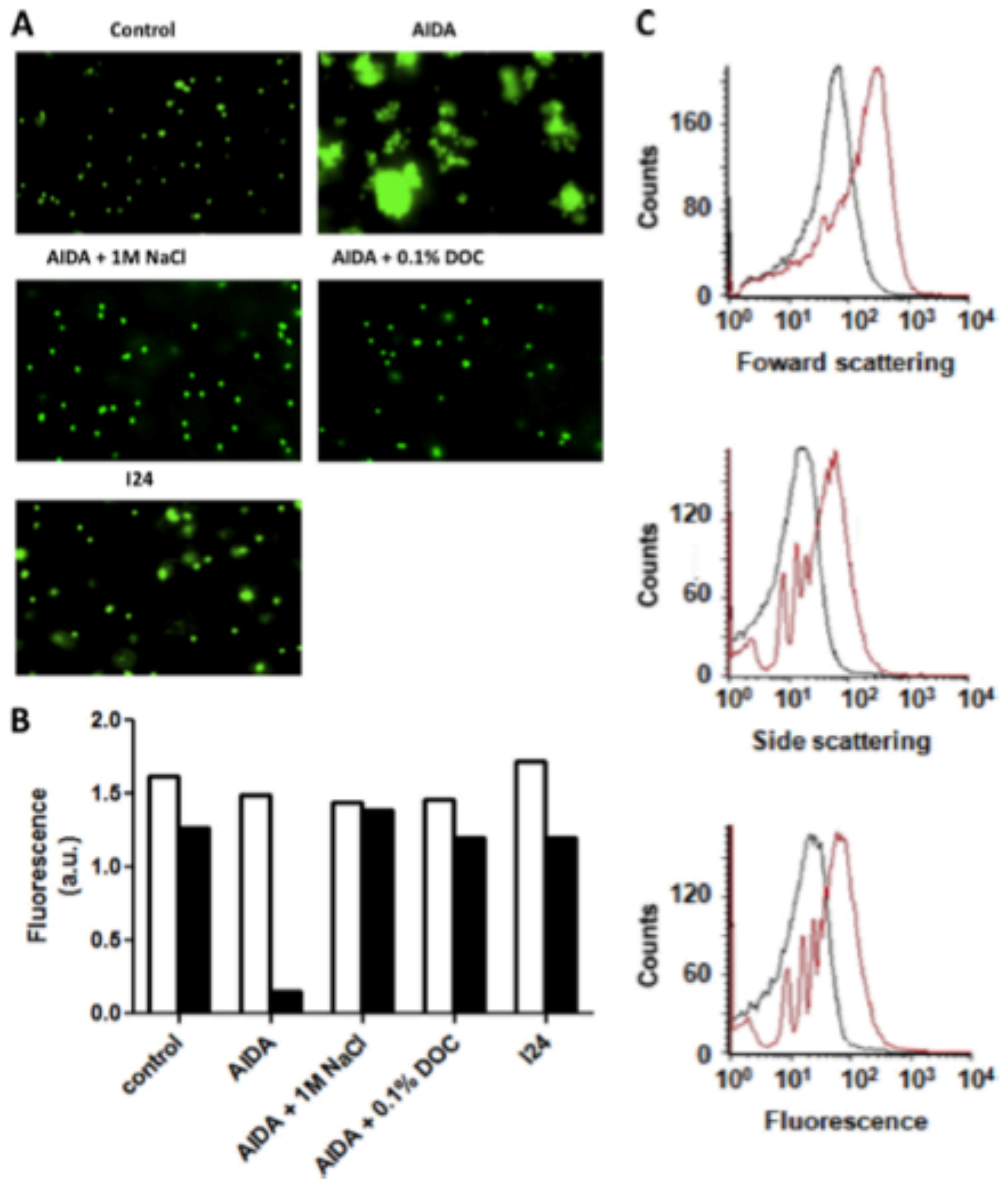


Figure 4

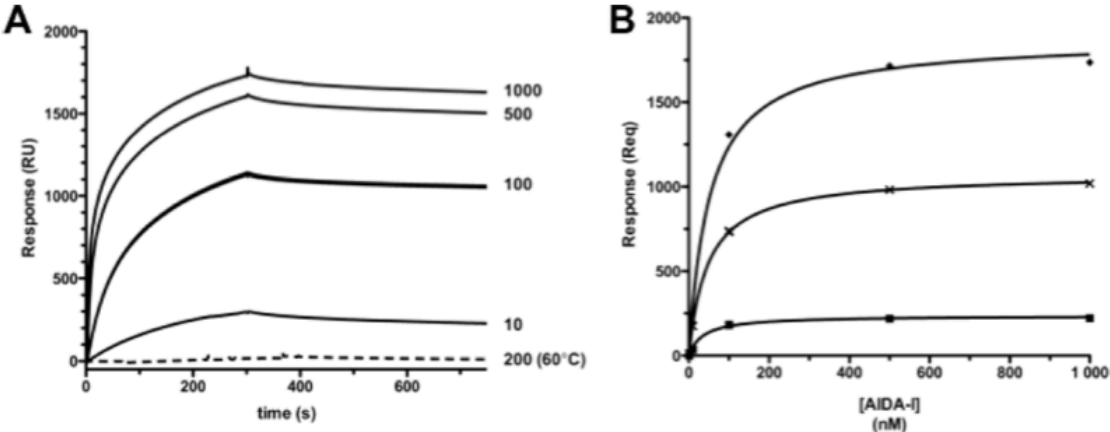


Figure 5

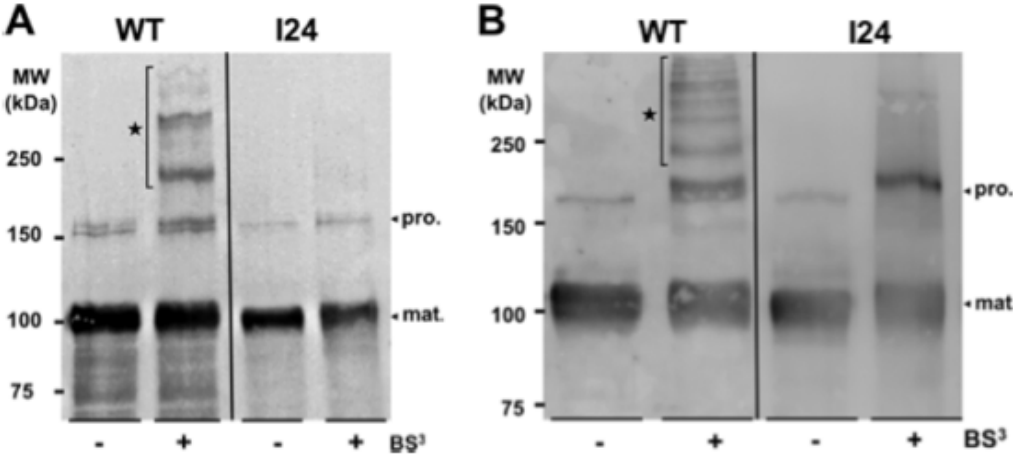


Figure 6

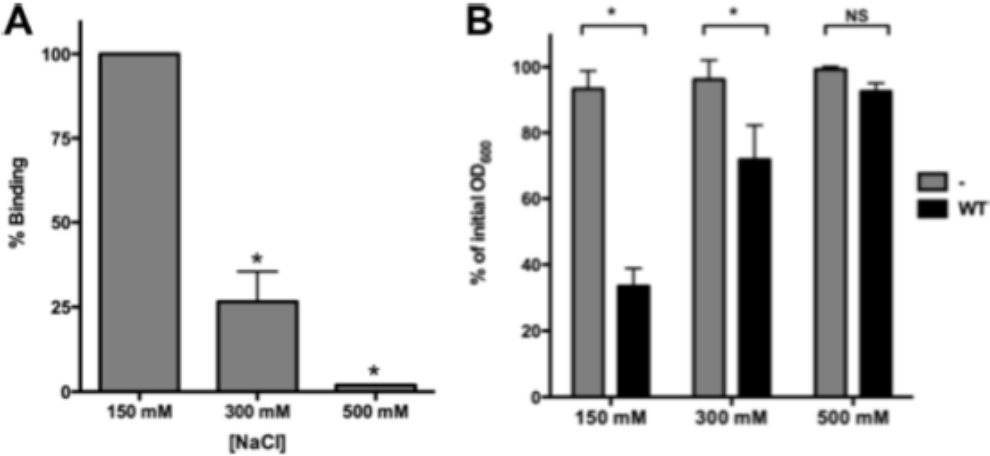


Figure 7

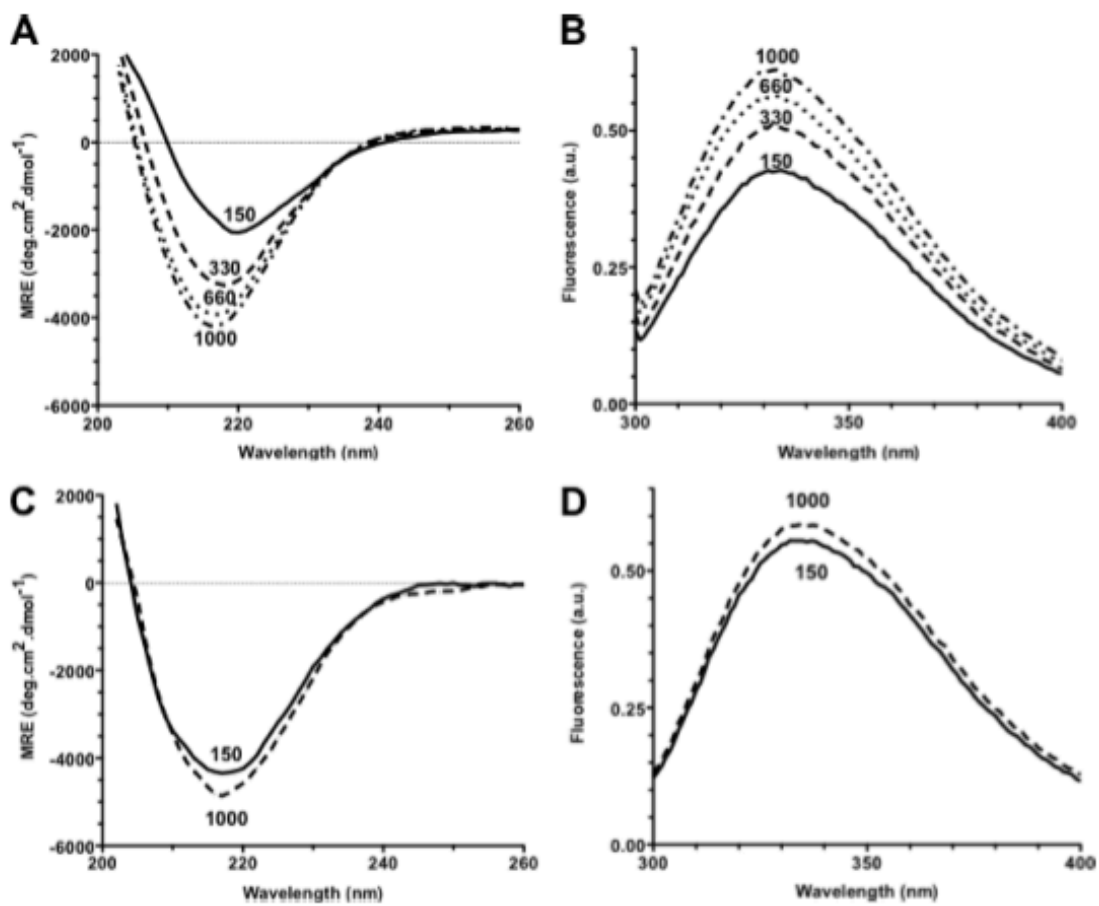
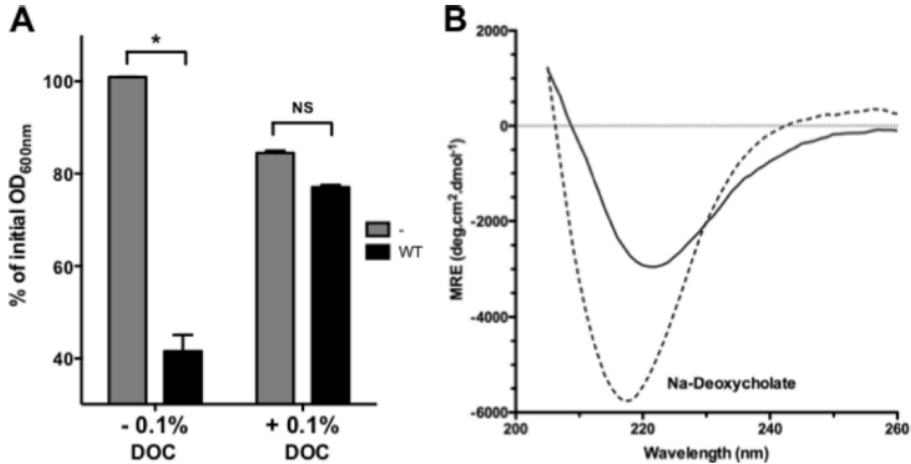
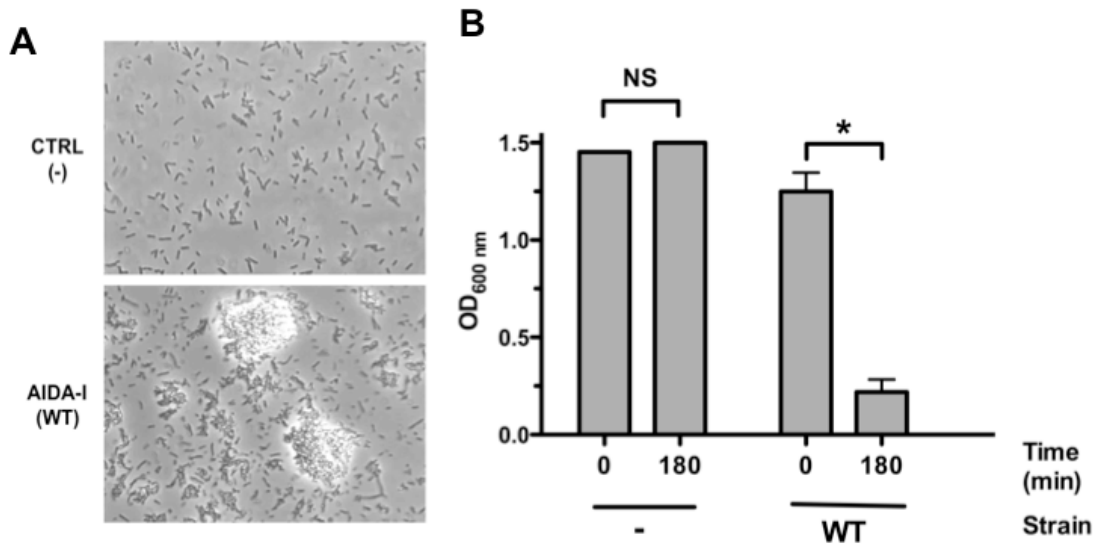


Figure 8



Supplementary information

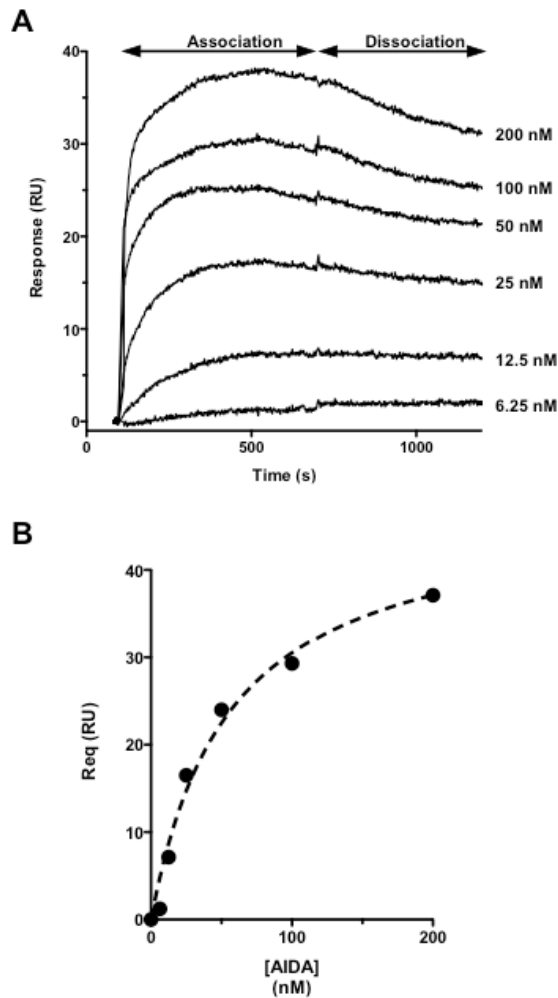
Fig. S1



Supplementary Fig. 1: Auto-aggregation of bacteria expressing AIDA-I.

(A) Samples of overnight cultures of a strain of *Escherichia coli* bearing an empty vector (CTRL, -) or a plasmid expressing AIDA-I (WT) were spotted on a microscope slide, incubated at 4°C and aggregates of bacteria were examined after 180 min of incubation with a phase-contrast microscope. (B) The same overnight cultures were left standing at 4°C and OD_{600nm} at the top of the culture were measured at the beginning of the assay and after 180 min. The assays were performed three times in duplicates and columns were compared by a two-way ANOVA and Bonferroni post-tests (* indicates statistical significance, $p < 0.05$; NS, not significant).

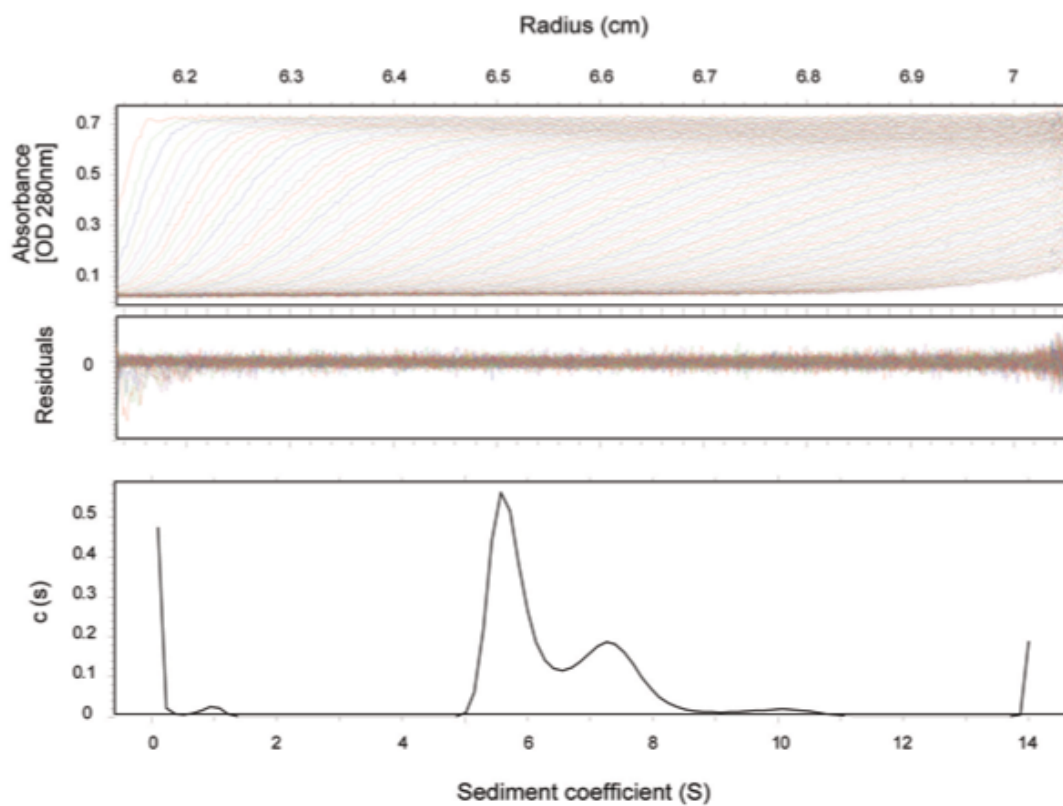
Fig. S2



Supplementary Fig. 2: AIDA-AIDA interactions monitored by Surface Plasmon Resonance (SPR) on a very low-density AIDA-I surface.

(A) SPR analysis was performed on BIACore CM4 chips on which ~ 250 RU of purified AIDA-I was immobilized by amine-coupling. Soluble AIDA-I (0 - 200 nM) was injected in HBS-ET buffer (10 mM Hepes pH 7.4, 150 mM NaCl, 3 mM EDTA, 0.0005% Tween-20) at 50 μ l/min. (B) The steady-state amounts bound (Req) were plotted against the concentration of AIDA-I and analyzed according to the steady-state affinity model with the BIAevaluation software in order to calculate an equilibrium dissociation constant. This analysis yielded a K_D of ~56 nM.

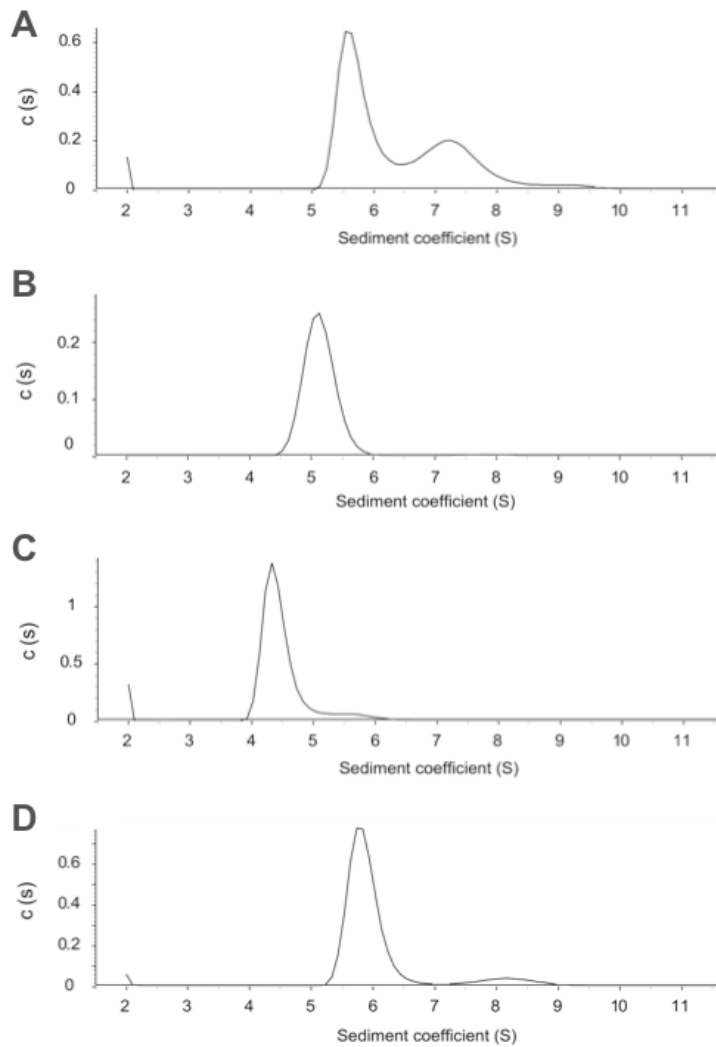
Fig. S3



Supplementary Fig. 3: Sedimentation profile of purified wild-type AIDA-I in TBS observed by AUC.

In this representative profile, 5 μM of purified wild-type AIDA-I in 50 mM Tris-HCl pH 8, 150 mM NaCl was sedimented in an analytical ultracentrifuge at 30,000 rpm, with absorbance scans monitored at 280 nm in 10 min intervals over a total spin time of 4-6 hours at 20°C. The top panel shows the absorbance at 280 nm that was observed. The middle and lower panels show the residuals and results, respectively, of the fitting of the experiment to the sedimentation of theoretical species of the indicated sedimentation coefficient.

Fig. S4



Supplementary Fig. 4: Sedimentation profiles observed by AUC.

Sedimentation coefficients observed from sedimentation velocity experiments performed as in Supplementary Fig. 3 with (A) wild-type AIDA-I in 50 mM Tris-HCl pH 8, 150 mM NaCl; (B) I24 in 50 mM Tris-HCl pH 8, 150 mM NaCl; (C) wild-type AIDA-I in 50 mM Tris-HCl pH 8, 1 M NaCl or (D) wild-type AIDA-I in 50 mM Tris-HCl pH 8, 150 mM NaCl, 0.5% sodium deoxycholate.

Article 5 – SAATs : molécules « cadherine-like »

Self-associating autotransporters as cadherin-like molecules involved in recognition of self

Manuscrit en preparation pour *Proceedings of the National Academy of Sciences*

Avant-propos

Précédemment (article 4 + résultats supplémentaires en annexe), nous avons observé que les SAATs utilisaient un mécanisme d'interaction similaire aux cadhérines. Comme les cadhérines interagissent préférentiellement avec elles-mêmes plutôt qu'avec un autre type de cadhérine, nous avons donc voulu voir si les SAATs étaient impliqués dans des interactions homophiliques. Nos résultats démontrent que les SAATs interagissent spécifiquement avec eux-mêmes, mais pas avec un autre SAAT. De ce fait, les SAATs permettent donc de discriminer entre le soi et le non-soi chez les bactéries et nos résultats suggèrent que les SAATs jouent un rôle dans les interactions sociales des bactéries ainsi que dans la pathogenèse.

Contributions

Contribution pour cet article :

En tant que premier auteur, j'ai effectué la majorité des expérimentations, j'ai analysé les résultats et j'ai écrit le manuscrit. Les expériences de « Single molecule force spectroscopy » ont été effectuées par Jumey Kochu et Dr Louis Cuccia, tandis que les expériences d'adhésion en flux ont été effectuées avec l'aide de Magali Soyer et Dr Guillaume Duménil.

Self-associating autotransporters as cadherin-like molecules involved in recognition of self

Jean-Philippe Côté^{1,#}, Magali Soyer², Jumey Kochuparampil³, Louis A. Cuccia³, Guillaume Duménil² and Michael Mourez¹

¹ Université de Montréal, Faculté de Médecine Vétérinaire, 3200 Sicotte, St.-Hyacinthe, J2S 7C6, Québec, Canada

² INSERM, U970, Paris Cardiovascular Research Center PARCC, 75737, Paris cedex 15, France

³ Department of Chemistry and Biochemistry, Concordia University, 7141 Sherbrooke St. West, H4B 1R6, Québec, Canada

Short Title: Self-recognition mediated by autotransporters

Abstract

Bacteria mostly live in community and, accordingly they have developed many mechanisms to communicate and interact with their environment. The self-associating autotransporters (SAATs) of *Escherichia coli* are a family of outer membrane proteins that mediate close contact between bacterial cells, which lead to bacterial auto-aggregation. The SAAT family is composed of the autotransporters AIDA-I, TibA and Ag43. The three SAATs are similar and are composed of repetitions of the same 19 amino acid consensus sequence in their passenger domain. Furthermore, the mechanism of interaction used by SAATs is similar to that of eukaryotic cadherins. Therefore, we have evaluated if SAATs showed a similar specificity of interaction than cadherins towards homophilic interactions. We found by beads aggregation assays, ELISA and single molecule force spectroscopy that AIDA-I interact with itself, but not with TibA, showing that SAATs are involved in homophilic interactions. We have determined that this specificity of interaction relies on both the number of repetitions in the passenger domain and on the subtle sequence difference in the imperfect portion of the repeats. We finally observed that this specificity of interaction is used to discriminate self from non-self and that the auto-aggregation leads to increased level of adhesion. Our results therefore suggest that SAATs are a family of cadherin-like molecules that promote recognition of self in bacteria. This could then lead to an increase in colonization or, alternatively, such an ability could be used by bacteria to organize microbial community.

Introduction

Bacteria are social organisms that live in community (1). Accordingly, bacteria have developed many mechanisms to communicate with each other and to sense their environment. Many of these systems involve large surface structures that allow bacteria to directly interact with their neighbors, whether they are host or other bacterial cells (2). Interaction with other bacteria is important as it can lead to aggregation and formation of biofilm and is performed by various surface structures, including fimbriae, curli, and

secretion systems (3). In most of these cases, the interactions involve very long structures at the surface of the bacteria. On the other hand, the shorter size of autotransporters could provide a more intimate contact between bacterial cells.

Autotransporters are part of the type V secretion system (4). They are modular proteins that do not require a specific secretion system for their translocation across the cell envelope. Instead, they only use the general secretion machineries of bacteria. They are composed of an N-terminal SEC-dependant signal sequence for translocation of the inner membrane and a C-terminal β -barrel that is inserted in the outer membrane by the BAM complex (5). Between the two, there is a functional domain, the passenger domain, that normally forms a β -helical structure in the extracellular space and is usually associated with virulence (6). In some cases, the passenger domain is cleaved and released from the cell surface. However, in others, the passenger domain stay associated at the cell surface. The latter is the case, for instance, for the adhesin involved in diffuse adherence (AIDA-I) (7).

AIDA-I was discovered in the diffusely adhering *Escherichia coli* strain 2787, but is mostly found in porcine pathogenic strains (8, 9). AIDA-I is a very versatile autotransporter capable of promoting adhesion and invasion of epithelial cells as well as promoting the autoaggregation of bacteria and the formation of biofilm (10). AIDA-I has recently been grouped with two other *E. coli* autotransporter, the adhesin/invasin TibA and the autoaggregation factor Ag43, in a family called the self-associating autotransporters (SAATs) (11). These three proteins were grouped together because they were functionally similar. However, they also share important sequence similarities. Notably, their passenger domain is composed of repeats of a 19-amino acid consensus sequence. Like other passenger domains, the passenger domain of SAATs folds into a β -helix, with each repeat forming one turn of the β -helix (12). The main difference between the three SAATs is the number of repeats in the passenger domain and therefore the length of the proteins.

Other autotransporters are also called self-associating autotransporters, even though they do not harbor the 19-amino acid repetition, because they also promote the autoaggregation of bacteria (13). These autotransporters possess a “SAAT domain” that, even without the 19-amino acid consensus sequence, is structurally similar and forms a β -helix composed of 3 β -strands that is responsible for bacterial autoaggregation. The crystal structure of such an autotransporter, the autotransporter Hap from *Haemophilus influenzae*,

recently established how two “SAAT domains” interact together to promote bacterial aggregation (14). The structure showed that one turn of the β -helix interacts with one face of the other and that the interaction is the result of the stacking of weak van der Waals forces.

Additionally, we have recently determined the mechanism of AIDA-I-mediated autoaggregation of bacteria (15). This mechanism involved direct AIDA-I/AIDA-I interactions that are associated with conformational changes in the protein and was reminiscent of the mechanism used by cadherin molecules (16). Cadherins are eukaryotic molecules that are responsible for cell-cell adhesion and the organization of tissues (16) and are principally involved in homophilic interactions (17). Therefore, it is possible that SAATs are cadherin-like molecules that also mediate homophilic interactions.

Here, we have evaluated the specificity of interaction between different SAATs and found that the autoaggregation mediated by SAATs is highly specific. We observed that one SAAT could interact with itself but not with another SAAT, even though the proteins are very similar. Our results suggest that the SAATs may be cadherin-like molecules that are involved in recognition of self. We further evaluated the role of this specificity of interaction and found that interaction with self increased the colonization potential of bacteria expressing one SAAT.

Results and discussion

SAAT-SAAT interactions are specific

To evaluate if SAATs are involved in homophilic interactions, we have purified the autotransporters AIDA-I and TibA and we have reconstituted autoaggregation on fluorescent beads (15). His-tagged AIDA-I and TibA were purified by affinity chromatography from *E. coli* strain C600 and coated onto red and green fluorescent beads. The beads were then mixed and left standing for 6h without shaking in order for the beads to aggregate together. After incubation, the suspensions were visualized by fluorescence microscopy (Fig. 1A). When mixed together, red and green beads coated with the same purified proteins, AIDA-I/AIDA-I or TibA/TibA, formed mixed aggregates. In contrast, when we mixed AIDA-I-coated beads with TibA-coated beads, the beads formed

homogenous aggregates. We also performed an ELISA assay where we coated 96-well plates with purified proteins and we probed the protein-protein interactions using the fluorescent beads (Fig. 1B-C). The AIDA-I beads (red lines) interacted only in the wells coated with AIDA-I (Fig. 1B) whereas the TibA beads (blue lines) interacted only in the wells coated with TibA (Fig. 1C).

We further evaluated the protein-protein interactions by single-molecule force spectroscopy (18). In this approach, the tip of a cantilever containing purified proteins gets in contact with a protein coated-slide in order to promote protein-protein interactions. After the interaction, the cantilever is withdrawn from the slide until the interaction is disrupted, and the rupture force is measured. The rupture force of the AIDA-I/AIDA-I interactions was 52 ± 13 pN and the rupture force of the TibA/TibA interactions was 64 ± 14 pN (Fig. S2). These interactions are relatively strong, although not as strong as the FimH/mannose interactions, that are disrupted by a force of >140 pN (19). The rupture forces of the AIDA-I/AIDA-I and TibA/TibA interactions were more comparable to the rupture force of 33 pN needed to disrupt the interaction between two VE-cadherin molecules (20). On the other hand, when we tested the interaction of the heterologous protein pairs, i.e. the interaction of AIDA-I with TibA in both possible orientations, we could not detect any specific interactions.

Taken together, these results show that AIDA-I interacts with itself but not with TibA, suggesting that SAATs mediate homophilic interactions. This specificity towards homophilic interactions could be used as a mechanism of recognition of self by bacteria in order to choose their interacting partners, as cadherins dictate how eukaryotic cells interact with each other. Interestingly, many molecules that promote the aggregation of cells are formed by a repetitive structure. Examples include, besides SAATs and cadherins, the Ser-rich proteins of streptococci (21), CsaA of *Dyctiostellium discoideum* (22) and the Flo1 adhesin of *Saccharomyces cerevisiae* (23). The example of Flo1 is interesting because the repeated structure of the *flo* gene allowed for the evolution of the gene and gave rise to a complete family of variants (24). Each Flo variant varies in length and shows a different specificity of interaction; organisms expressing one Flo variant aggregate only with others that expressed the same Flo variant.

Members of the SAATs family possess a different number of a 19-amino acid imperfect repeat in their passenger domain. Moreover, recent studies with AIDA-I and TibA have highlighted a mechanism using domain shuffling for the evolution of the SAATs (9, 25). Therefore, the three SAATs might stem from the evolution of one primary SAAT, as it is the case for Flo1, and there might be a complete SAAT family. Indeed, we found many autotransporters containing the same 19-amino acid imperfect repeat in their passenger domain (Table 1). These new SAATs were not limited to *E. coli* or γ -proteobacteria and were even not limited to pathogens. The presence of SAAT in non-pathogenic bacteria, such as *Oxalobacter formigenes*, suggests that these autotransporters are not limited to a role in pathogenesis. Thus, a role in the recognition of self could explain the presence of SAATs in non-pathogenic bacteria. Furthermore, all these autotransporters share the same organization, but differ mainly by the number of repeats of the 19-amino acid consensus sequence in their passenger domain. Therefore, SAATs may be a family of molecules of different length that use a cadherin-like mechanism to promote cell-cell interactions.

Length and sequence are both determinant of specificity

Since the main difference between the SAATs is the number of repeats of the 19-amino acid consensus sequence in the passenger domain, we hypothesized that this could be a determinant of the specificity. In order to test this hypothesis, we used an N-terminal truncated variant of AIDA-I, AIDA-I Δ N (10). The deletion removed an adhesion domain and did not change the autoaggregation properties of AIDA-I. As a result, this mutant is well suited to evaluate the role of length in the specificity of interaction since it harbors a similar number of repeats than TibA. AIDA-I Δ N was purified, adsorbed on the fluorescent beads and we performed our interaction assays.

In the autoaggregation assay, the AIDA-I Δ N beads interacted with themselves and still interacted with the AIDA-I beads, but not with the TibA beads (Fig. 2A). This suggested that the sequence was the main determinant of specificity. However, in the ELISA assay, a difference between the interaction of WT AIDA-I and AIDA-I Δ N was observed. In wells coated with AIDA-I Δ N (Fig. 2B), the interaction of the WT AIDA-I beads was not as strong as that of the AIDA-I Δ N beads (black line). A similar observation

was made when using wells coated with WT AIDA-I (Fig. S1). These results suggest that both the sequence and the length of the passenger domain are important determinants of specificity, since AIDA-I Δ N did not interact with TibA and has less affinity for WT AIDA-I.

The recent crystal structure of Hap has revealed the oligomerisation mechanism of the “SAAT domain” (14). In the crystal structure, the edge of one “SAAT domain” interacts with the face of the other and the interaction is carried out by weak van der Waals forces from self-complementary surfaces. The strong overall interaction is created by the stacking of these forces that comes from the β -helical structure. The crystal structure highlights the importance of complementary surface since the interaction was strongly sensitive to steric effect. Furthermore, in the crystal structure, the interaction between two “SAAT domains” occurred throughout the entire domain. For the SAATs, the whole passenger domain is a “SAAT domain”, which suggests that the interaction occur throughout the whole passenger domain. Two identical “SAAT domain” most likely maximize the interaction potential by correctly positioning the self-complementary surfaces. Therefore, the crystal structure of Hap explains the importance of both the sequence and the number of repeats on the specificity towards homophilic interactions.

Specificity of SAAT interaction promotes the segregation of similar bacteria

AIDA-I and TibA are found in different strains of *E. coli*. AIDA-I is mostly found in porcine strain associated with post-weaning diarrhea and oedema disease, such as PD20 (9). TibA was discovered in the human enterotoxigenic *E. coli* (ETEC) strain H10407 and is found in other human ETEC strains (26). Therefore, we have used H10407 as a WT context for TibA and the ETEC strain PD20 as a WT context for AIDA-I (27). In these strains, AIDA-I and TibA are regulated differently and the regulation mechanisms are not fully understood (28, 29). To ensure that AIDA-I and TibA are expressed at similar levels, we have expressed AIDA-I and TibA in these strains from an IPTG-inducible promoter. In addition, to visualize autoaggregation, bacteria also expressed either GFP or mCherry. Bacteria were mixed together, left standing and visualized by fluorescence microscopy (Fig. 3). When two cultures of *E. coli* strain PD20 expressing respectively GFP and mCherry were mixed together, the GFP-expressing bacteria aggregated with the mCherry-

expressing bacteria. The same observation was made when two H10407 strains were mixed together. However, when we mixed *E. coli* strain PD20 with strain H10407, the two ETEC strains segregated from each other and formed homogenous aggregates. This suggests that the specificity of interactions allows bacteria to interact only with their own kin and that SAATs may be involved in recognition of self.

AIDA-I and TibA are expressed by different strains of *E. coli*. AIDA-I is mainly expressed in strains that infect pigs while TibA is present in strains that infect humans (9, 26). Since these strains infect different hosts, they need not to interact together. Therefore, the recognition of self by SAATs might be used by bacteria to segregate from other bacteria that colonize different niche. Interestingly, all *E. coli* strains possess an allele of Ag43 or an allele of its close homologue Cah. In fact, Cah is often viewed as an allele of Ag43, but most likely is a deletion variant of Ag43 (Fig. S3) (30). Again, Cah and Ag43 are not found in the same strains: Cah is found in intestinal strains while Ag43 is found in other extra-intestinal strains, like the uropathogenic *E. coli* (UPEC). It is therefore possible that Cah has evolved from Ag43 in order to segregate *E. coli* strains that colonize different niche.

Autoaggregation increases adhesion of bacteria to epithelial cells

This mechanism of the recognition of self by SAATs could also serve for pathogenesis. The preference towards homophilic interactions could be used by bacteria in order to interact only with their own kin, which would give these bacteria a significant advantage in survival and colonization. In fact, autoaggregation of bacteria often have an effect on the adhesion potential of these bacteria, especially under shear stress conditions (31). In the intestine, shear stress values can go up to 1 dynes/cm² (32). Besides promoting autoaggregation of bacteria, AIDA-I and TibA are two potent adhesins of *E. coli* (33, 34). Therefore, we have evaluated the effect of autoaggregation on the adhesion of bacteria to epithelial cells under shear stress conditions.

The bacteria were allowed to adhere onto Hep-2 cells in a laminar flow chamber. The shear stress value during the adhesion was 0.1 dynes/cm², which is similar to the shear stress encountered in the gut (32). After adhesion, unadhered bacteria were washed by increasing the flow. We have first looked at the adhesion of *E. coli* C600 expressing TibA by time-lapse video microscopy. In some cases, as for *Neisseria meningitidis*, only

individual bacterial adhere to epithelial cells (31). In our case, aggregates were resistant to shear stress (Fig. S4) and bacteria in aggregates adhered as well as individual bacteria. Furthermore, once adhered, these aggregates were highly dynamic; the aggregates expanded and we observed smaller aggregates that detached and adhered elsewhere onto cells or other aggregates (Fig. 4 and Movie S1). Here, images were taken when no bacteria were added to the system; newly adhered bacteria come from aggregates that have detached from the cells downstream. Therefore, it seemed that aggregation could help adhesion under shear stress conditions.

To test the hypothesis that aggregation could help adhesion, we have used a mutant of TibA, the mutant i3, that no longer promote the autoaggregation of bacteria (25). In static conditions, *E. coli* strain C600 expressing this mutant adhered as well as bacteria expressing WT TibA (25). However, under shear stress conditions, the adhesion mediated by the i3 mutant was dramatically reduced compared to WT TibA (Fig. S5). This result suggested that autoaggregation of bacteria is important for adhesion under shear stress conditions.

To further evaluate this effect of autoaggregation on adhesion, *E. coli* strain C600 expressing AIDA-I or TibA were allowed to aggregate or not and the bacteria were allowed to adhere onto Hep-2 cells in a laminar flow chamber. When the bacteria expressing either AIDA-I or TibA were allowed to aggregate, their adhesion level was significantly increased compared to the adhesion of individual bacteria (Fig. 5A). This was especially true when adhesion time was prolonged. The adhesion level of individual bacteria was the same independently of the adhesion time. However, adhesion of aggregated bacteria increased when adhesion time was increased. It seemed that the number of adhesion event was similar for individual and aggregated bacteria, but, since the aggregates were resistant to high shear stress levels (Fig. S4), there were more adhered bacteria overall when bacteria were in aggregates (Fig. 5B). Furthermore, the specificity of interaction was conserved during adhesion under shear stress conditions (Fig S6). These results suggest that autoaggregation of bacteria by the SAATs increase the adhesion capacity of these bacteria under shear stress conditions.

Taken together, our results suggest that SAATs are a family of cadherin-like molecules that may be used by bacteria to discriminate between self and non-self, even below the species level. Such an ability to discriminate self from non-self is rarely evoked in bacteria, but could play a role in pathogenesis as well as in the organization of multicellular communities. In fact, the organization of bacterial communities is similar to that of eukaryotic tissues. While eukaryotes secrete an extracellular matrix composed of proteoglycans, adhesion proteins and collagen, bacteria produce a matrix composed of polysaccharides, proteins and DNA that fulfills the same role. In this context, eukaryotic cells use cadherins to mediate cell-cell interactions. SAATs could be the cadherin-like molecules used by bacteria to promote cell-cell interaction in communities.

Material and Methods

Bacterial strains, cell line and growth conditions

E. coli K-12 strains C600 (New England Biolabs) (F- thr-1 leuB6 thi-1 lacY1 supE44 rfbD1 fhuA21) and the ETEC strains PD20 (27) and H10407 (26) were used in this study. Bacteria were transformed with plasmids pAgH (35), pTgH (25), pAgH-deltaN (10) or pTgH-i3 (25) that allow the expression of glycosylated his-tagged AIDA-I, TibA, AIDA-IΔN and the i3 mutant respectively. The bacteria were also transformed with plasmids allowing the expression of GFP or mCherry. Bacteria containing the different plasmids were grown on Luria-Bertani (LB) agar plates or in liquid LB medium containing 100 $\mu\text{g}\cdot\text{ml}^{-1}$ ampicillin and 50 $\mu\text{g}\cdot\text{ml}^{-1}$ chloramphenicol. Bacterial cultures were grown at 30°C and the optical density at 600nm (OD_{600}) was measure to monitor growth. At an OD_{600} of 0.8, the cultures were induced with 10 μM isopropyl- β -D-thiogalactopyranoside (IPTG) and growth was pursued overnight. This low concentration of IPTG was used to limit the toxicity associated with overexpression of AIDA-I, TibA and their glycosyltransferase. HEp-2 cells (ATCC CCL-23) were grown at 37°C with 5% CO₂ in Dulbecco's modified Eagle's medium (Gibco) containing 10 mM sodium pyruvate (Sigma), 10% bovine growth serum (HyClone), 2.5 $\mu\text{g}\cdot\text{ml}^{-1}$ amphotericin B (Fungizone), and 100 $\mu\text{g}\cdot\text{ml}^{-1}$ penicillin/streptomycin (Gibco).

Protein purification

To purify the His-tagged proteins, one liter of *E coli* C600 bearing the plasmid pAgH, pTgH or pAgH-deltaN were grown to an OD₆₀₀ of 0.8 and induced with 10 μM IPTG overnight. Bacteria were harvested, resuspended in 50 mL TBS containing a protease inhibitor cocktail (complete; Roche) and the cells were lysed using a French press and an ultrasonicator. Lysates were recovered by a slow speed centrifugation to clear cellular debris and were further centrifuged for 30 min at 250,000 X g. Pellets were resuspended in 25 mL of TBS containing either 1% Triton X-100 reduced (Sigma) [for AIDA-I and AIDA-ΔN] or 2% Triton X-100 reduced (Sigma) and 10 mM EDTA [for TibA]. Suspensions were incubated for 30 min at 37°C and centrifuged again for 30 min at 250,000 X g. For TibA, EDTA was removed from the supernatant using PD-10 columns (GE Healthcare). The His-tagged proteins were further purified from the supernatants by affinity chromatography using 1 mL His-Trap HP columns (GE Healthcare). Purity of the purified proteins was consistently over 95% as evaluated by Coomassie blue stained SDS-PAGE. Finally, the protein samples were dialyzed to replace the Triton X-100 by 1% n-octyl-β-D-glucoside (Roche).

Autoaggregation assay and modified ELISA

Purified proteins (1.5 mg/mL) were adsorbed onto 1 μm fluorescent polystyrene beads (FluoSpheres sulfate microspheres; Invitrogen) according to the instructions of the manufacturer. Beads were stored at a concentration of 0.2% in TBS 1% BSA. For the autoaggregation assay, the beads were diluted to a concentration of 0.01% in TBS and mixed. The mixtures were left standing at room temperature for 6h and visualized using a DMI 4000B reverse fluorescence microscope (Leica Microsystems Inc.). Images were further analyzed with ImageJ (36) and Mander's coefficients were calculated using the JACoP plugin (37). Mander's coefficients represent the percentage of colocalization for each color: $M_R = \sum R_{coloc} / \sum R_{total}$ and $M_G = \sum G_{coloc} / \sum G_{total}$.

Alternatively, the beads were used in a modified ELISA assay. Purified proteins (500 nM) were coated in 96-well plates overnight and blocked with TBS 1% BSA for 2h at room temperature. To evaluate the protein-protein interactions, AIDA-I and TibA-coated green fluorescent beads were added to the wells at different concentrations (0.01% to 0.2%) in

TBS for 4h. Wells were washed 5X with TBS and the bound beads were resuspended with TBS 1M NaCl. High salt concentrations disrupt the protein-protein interactions (15). Fluorescence was then measured using a Synergy HT fluorescence plate reader (BioTek). The fluorescence emission was registered at 528 nm (bandpass 20 nm) after excitation at 485 nm (bandpass 20 nm).

For the autoaggregation of bacteria, overnight cultures were normalized in 5 ml of LB to an OD600 of 1.0 in culture tubes. The initial aggregates in the suspension were disrupted by repeated passages through 27G syringes and the suspensions were mixed in a 1:1 ratio. After 2h incubation at 4°C without shaking, the mixtures were analyzed by fluorescence microscopy.

Bioinformatic analysis

The sequence of AIDA-I, TibA and Ag43 were blasted against the non-redundant protein database. In order to find other autotransporters bearing the same 19-amino acid consensus sequence, we used PHI-Blast using the consensus sequence XGGXAXXTX(V/I/L)NXGGXQX(V/I/L)X where X is any amino acid. 19-amino acids repetitions were confirmed visually and autotransporters that do not harbor these repetitions were discarded. Finally, the repeats for each SAATs identified were aligned in order to create the consensus sequence (Weblogo) (38).

Single molecule force spectroscopy

Single molecule force spectroscopy experiments were performed with silicon nitride triangular microlever probes MLCT-AUNM (Veeco/Bruker Probes). The probe surfaces were aminated with a 2% solution of VECTABOND™ reagent (Vector Laboratories), in acetone for 10 minutes via amino silane coupling. The aminated probes were then immersed in chloroform solution containing NHS-PEG6000-NHS (2 mg/mL) and CH3O-PEG750-NHS (250 mg/mL) (Rapp Polymere GmbH) with 0.5% triethylamine (Sigma) for 1 hour. Finally, probes were coupled with AIDA-I or TibA (0.1 mg protein/ml TBS 1M NaCl) for 1 h. TBS 1M NaCl was used at this point because the proteins are monomeric at high salt concentration (15). Functionalized tips were either used immediately or stored for no more than 36 hours at 4 °C in TBS 1M NaCl buffer. Prior to their use, the tips were

rinsed TBS 150 mM NaCl. Alternatively, the surfaces were prepared by the adsorption 10 μ L of AIDA-I or TibA (0.1 mg/mL) in TBS 1M NaCl onto freshly cleaved mica for 30 minutes at RT and excess of proteins was removed by rinsing with TBS.

The force measurements were obtained using an MFP-1D (Asylum Research) at RT. The tips were allowed to interact with the surface for 1 s with an applied force of 100-300 pN. Over 5000 force-distance curves were acquired for each protein pairs with a tip velocity of 500 nm/s. The force-distance curves were analyzed by a custom program written in IGOR Pro 6 (Wavemetrics) which calculated rupture forces (39). The most probable rupture force for each loading rate was determined by careful data analysis. Each force-distance curve was first screened for unbinding events. The retraction force-distance curves having single unbinding events were analyzed on the basis of the Extended-Freely jointed chain model. The most probable rupture forces were obtained from the Gaussian fit to the force histogram using OriginPro9 software.

Adhesion assay

Adhesion assays were done under flow conditions (40). Hep2 cells were grown into ibiTreat μ -slide VI 0.4 (Ibidi). The Ibidi slide was placed directly under a flow generating a shear stress of 0.2 dynes*cm⁻². Overnight bacterial cultures (10⁸ bacteria) were allowed to adhere onto the cells for 30 min to 2h at 37°C and unbound bacteria were washed by increasing the flow. Adhesion was then visualized using a DMI 4000B reverse fluorescence microscope (Leica Microsystems Inc.). Total fluorescence by images was quantify using ImageJ (36).

Acknowledgements

This work was supported by the Canadian Institutes for Health Research and the Natural Sciences and Engineering Research Council, as well as by funds from the Groupe de Recherche sur les Maladies Infectieuses du Porc and the Canada Research Chair program. JPC was supported by a graduate fellowship from Natural Sciences and Engineering Research Council.

References

1. Nadell CD, Xavier JB, & Foster KR (2009) The sociobiology of biofilms. *FEMS Microbiol Rev* 33(1):206-224.
2. Hayes CS, Aoki SK, & Low DA (2010) Bacterial contact-dependent delivery systems. *Annu Rev Genet* 44:71-90.
3. Fronzes R, Remaut H, & Waksman G (2008) Architectures and biogenesis of non-flagellar protein appendages in Gram-negative bacteria. *EMBO J* 27(17):2271-2280.
4. Dautin N & Bernstein HD (2007) Protein secretion in gram-negative bacteria via the autotransporter pathway. *Annu Rev Microbiol* 61:89-112.
5. Leyton DL, Rossiter AE, & Henderson IR (2012) From self sufficiency to dependence: mechanisms and factors important for autotransporter biogenesis. *Nat Rev Microbiol* 10(3):213-225.
6. Nishimura K, Tajima N, Yoon YH, Park SY, & Tame JR (2010) Autotransporter passenger proteins: virulence factors with common structural themes. *J Mol Med* 88(5):451-458.
7. Charbonneau ME, et al. (2012) A structural motif is the recognition site for a new family of bacterial protein O-glycosyltransferases. *Mol Microbiol* 83(5):894-907.
8. Benz I & Schmidt MA (1989) Cloning and expression of an adhesin (AIDA-I) involved in diffuse adherence of enteropathogenic *Escherichia coli*. *Infect Immun* 57(5):1506-1511.
9. Cote JP, et al. (2012) Identification and Mechanism of Evolution of New Alleles Coding for the AIDA-I Autotransporter of Porcine Pathogenic *Escherichia coli*. *Appl Environ Microbiol* 78(13):4597-4605.
10. Charbonneau ME & Mourez M (2007) Functional organization of the autotransporter adhesin involved in diffuse adherence. *J Bacteriol* 189(24):9020-9029.
11. Klemm P, Vejborg RM, & Sherlock O (2006) Self-associating autotransporters, SAATs: functional and structural similarities. *Int J Med Microbiol* 296(4-5):187-195.
12. Kajava AV & Steven AC (2006) The turn of the screw: variations of the abundant beta-solenoid motif in passenger domains of Type V secretory proteins. *J Struct Biol* 155(2):306-315.

13. Wells TJ, Totsika M, & Schembri MA (2010) Autotransporters of *Escherichia coli*: a sequence-based characterization. *Microbiology* 156(Pt 8):2459-2469.
14. Meng G, Spahich N, Kenjale R, Waksman G, & St Geme JW, 3rd (2011) Crystal structure of the *Haemophilus influenzae* Hap adhesin reveals an intercellular oligomerization mechanism for bacterial aggregation. *EMBO J* 30(18):3864-3874.
15. Girard V, et al. (2010) Conformation change in a self-recognizing autotransporter modulates bacterial cell-cell interaction. *J Biol Chem* 285(14):10616-10626.
16. Leckband D & Prakasam A (2006) Mechanism and dynamics of cadherin adhesion. *Annu Rev Biomed Eng* 8:259-287.
17. Katsamba P, et al. (2009) Linking molecular affinity and cellular specificity in cadherin-mediated adhesion. *Proc Natl Acad Sci U S A* 106(28):11594-11599.
18. Weisel JW, Shuman H, & Litvinov RI (2003) Protein-protein unbinding induced by force: single-molecule studies. *Curr Opin Struct Biol* 13(2):227-235.
19. Yakovenko O, et al. (2008) FimH forms catch bonds that are enhanced by mechanical force due to allosteric regulation. *J Biol Chem* 283(17):11596-11605.
20. Baumgartner W, et al. (2000) Cadherin interaction probed by atomic force microscopy. *Proc Natl Acad Sci U S A* 97(8):4005-4010.
21. Zhou M & Wu H (2009) Glycosylation and biogenesis of a family of serine-rich bacterial adhesins. *Microbiology* 155(Pt 2):317-327.
22. Queller DC, Ponte E, Bozzaro S, & Strassmann JE (2003) Single-gene greenbeard effects in the social amoeba *Dictyostelium discoideum*. *Science* 299(5603):105-106.
23. Smukalla S, et al. (2008) FLO1 is a variable green beard gene that drives biofilm-like cooperation in budding yeast. *Cell* 135(4):726-737.
24. Goossens K & Willaert R (2010) Flocculation protein structure and cell-cell adhesion mechanism in *Saccharomyces cerevisiae*. *Biotechnol Lett* 32(11):1571-1585.
25. Cote JP & Mourez M (2011) Structure-function analysis of the TibA self-associating autotransporter reveals a modular organization. *Infect Immun* 79(5):1826-1832.
26. Elsinghorst EA & Kopecko DJ (1992) Molecular cloning of epithelial cell invasion determinants from enterotoxigenic *Escherichia coli*. *Infect Immun* 60(6):2409-2417.

27. Ravi M, et al. (2007) Contribution of AIDA-I to the pathogenicity of a porcine diarrheagenic *Escherichia coli* and to intestinal colonization through biofilm formation in pigs. *Vet Microbiol* 120(3-4):308-319.
28. Berthiaume F, Leblond MF, Harel J, & Mourez M (2010) Growth-phase-dependent expression of the operon coding for the glycosylated autotransporter adhesin AIDA-I of pathogenic *Escherichia coli*. *FEMS Microbiol Lett* 311(2):176-184.
29. Espert SM, Elsinghorst EA, & Munson GP (2011) The tib Adherence Locus of Enterotoxigenic *Escherichia coli* Is Regulated by Cyclic AMP Receptor Protein. *J Bacteriol* 193(6):1369-1376.
30. van der Woude MW & Henderson IR (2008) Regulation and function of Ag43 (flu). *Annu Rev Microbiol* 62:153-169.
31. Mikaty G, et al. (2009) Extracellular bacterial pathogen induces host cell surface reorganization to resist shear stress. *PLoS Pathog* 5(2):e1000314.
32. Tchesnokova V, et al. (2010) Shear-enhanced binding of intestinal colonization factor antigen I of enterotoxigenic *Escherichia coli*. *Mol Microbiol* 76(2):489-502.
33. Benz I & Schmidt MA (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.
34. Lindenthal C & Elsinghorst EA (2001) Enterotoxigenic *Escherichia coli* TibA glycoprotein adheres to human intestine epithelial cells. *Infect Immun* 69(1):52-57.
35. Charbonneau ME, Berthiaume F, & Mourez M (2006) Proteolytic processing is not essential for multiple functions of the *Escherichia coli* autotransporter adhesin involved in diffuse adherence (AIDA-I). *J Bacteriol* 188(24):8504-8512.
36. Schneider CA, Rasband WS, & Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* 9(7):671-675.
37. Bolte S & Cordelieres FP (2006) A guided tour into subcellular colocalization analysis in light microscopy. *J Microsc* 224(Pt 3):213-232.
38. Crooks GE, Hon G, Chandonia JM, & Brenner SE (2004) WebLogo: a sequence logo generator. *Genome Res* 14(6):1188-1190.
39. Yermolenko IS, et al. (2010) Origin of the nonadhesive properties of fibrinogen matrices probed by force spectroscopy. *Langmuir* 26(22):17269-17277.

40. Soyer M & Dumenil G (2012) A laminar-flow chamber assay for measuring bacterial adhesion under shear stress. *Methods Mol Biol* 799:185-195.

Figure legends

Figure 1: SAAT-SAAT interactions are specific

Purified AIDA-I and TibA were adsorbed on fluorescent polystyrene beads. (A) Beads autoaggregation assay: Protein-coated beads were mixed, left standing without shaking and analyzed by fluorescence microscopy. The experiment was conducted at least 3 times and images represent typical observations. Bars represent the Mander's coefficient mean, the percentage of colocalization, of 15 different images. (B and C) ELISA assay: 96-well plates were coated overnight with 500 nM of (B) purified AIDA-I or (C) purified TibA. To probe the interactions, AIDA-I (red lines) and TibA (blue lines) -coated green fluorescent beads were added to the wells at different concentrations and the fluorescence was measured at 528 nm using a fluorescence plate reader. Results represent three experiments performed in duplicates.

Figure 2: Effect of length on the specificity

The interaction of AIDA-I Δ N with AIDA-I and TibA was evaluated by the beads autoaggregation assay (A) and by the ELISA assay (B). In the ELISA assay, the wells were coated with 500 nM of AIDA-I Δ N and the protein-protein interactions were probed with green fluorescent beads coated with AIDA-I (red line), TibA (blue line) and AIDA-I Δ N (black line).

Figure 3: SAAT specificity of interaction allows the segregation of two similar *E. coli* Enterotoxigenic *E. coli* (ETEC) strains PD20 and H10407 expressing AIDA-I and TibA respectively and either GFP or mCherry were mixed together, left standing without shaking and analyzed by fluorescence microscopy. The experiment was conducted 3 times and images represent typical observations.

Figure 4: Adhesion of bacteria in aggregates on Hep-2 cells

Time-lapse image of the adhesion of *E. coli* strain C600 expressing TibA in aggregates on Hep-2 cells. Bacteria were allowed to adhere to Hep-2 cells during 30 min under a shear stress of 0.1 dyn/cm² and unadhered bacteria were washed. At this point, images were taken each 5 min during 2 h, starting at t=0. Images highlight the detachment (A) and the adhesion (B) of an aggregate of bacteria expressing TibA. At the time the pictures were taken, no new bacteria were added to the system.

Figure 5: Autoaggregation by the SAATs increases adhesion of bacteria

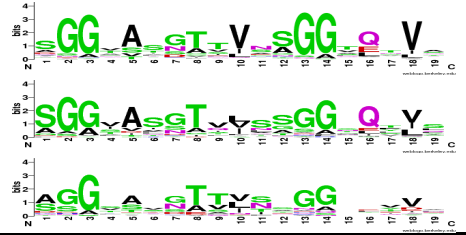
Escherichia coli strain C600 bearing an empty vector (-) or a plasmid allowing the expression of AIDA-I (red lines) or TibA (blue lines) were allowed to adhere onto Hep-2 cells under shear stress conditions. (A) Bacteria were loaded on the cells either as aggregates (full lines) or as single cells (dotted lines). Adhesion was visualized by fluorescence microscopy and is represented by the total fluorescence per image from three different experiments. (B) Typical images of the adhesion on Hep-2 cells of aggregated (left) and individual (right) *E. coli* strain C600 bacteria expressing TibA.

Table

Table I: Putative SAATs bearing the 19 amino-acid consensus sequence repeats

SAAT	Organism	Length / # of repeats	Consensus sequence
Ag43	<i>E. coli</i> / <i>Shigella</i> (γ)	1039 / 20	
Cah	<i>E. coli</i> / <i>Shigella</i> (γ)	949 / 16	
YejO	<i>E. coli</i> / <i>Shigella</i> (γ)	863 / 13	
YpjA	<i>E. coli</i> / <i>Shigella</i> (γ)	1527 / 47	
AIDA-I	<i>E. coli</i> (γ)	1287 / 35	
TibA	<i>E. coli</i> (γ)	989 / 16	
TibA ZP_09335089	<i>Citrobacter freundii</i> (γ)	963 / 16	
YP_003033977 (TibA-like)	<i>E. coli</i> (γ)	1135 / 27	
ESA_02084	<i>Cronobacter sakazakii</i> (γ)	2356 / 93	
SARI_04626	<i>Salmonella enterica</i> (γ)	916 / 14	
YP_003368457	<i>Citrobacter rodentium</i> (γ)	1593 / 44	
Cah YP_003368339	<i>Citrobacter rodentium</i> (γ)	948 / 16	
ZP_02902636	<i>Escherichia albertii</i> (γ)	877 / 14	
ZP_05727242	<i>Pantoea sp.</i> (γ)	1408 / 39	
YP_003334434	<i>Dickeya dadantii</i> (γ)	1075 / 16	
YP_002795894	<i>Laribacter hongkongensis</i> (β)	1023 / 24	
ZP_06468800	<i>Burkholderia sp.</i> (β)	1527 / 44	
ZP_04578582	<i>Oxalobacter formigenes</i> (β)	1641 / 55	
ZP_04576972	<i>Oxalobacter formigenes</i> (β)	1414 / 42	

YP_001110606	<i>Burkholderia vietnamiensis</i> (β)	1981 / 71
YP_745424	<i>Granulibacter bethesdensis</i> (α)	1266 / 34
YP_190953	<i>Gluconobacter oxydans</i> (α)	1103 / 28



Figures

Figure 1

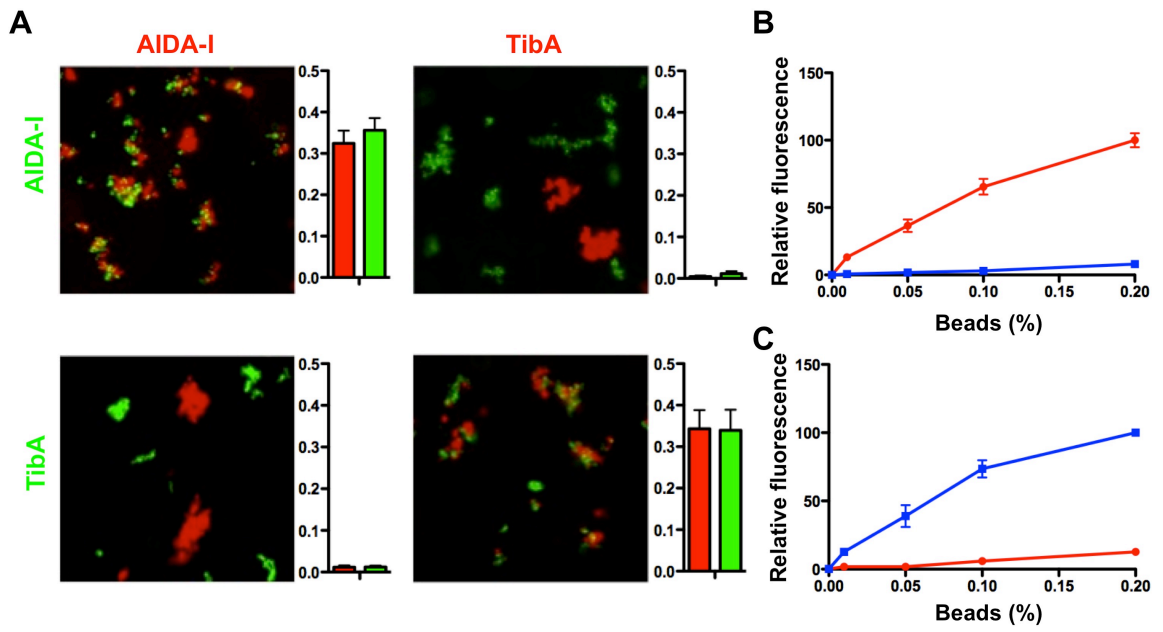


Figure 2

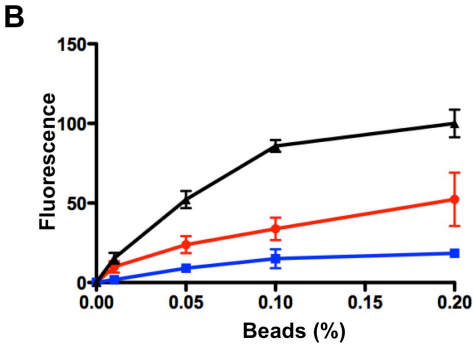
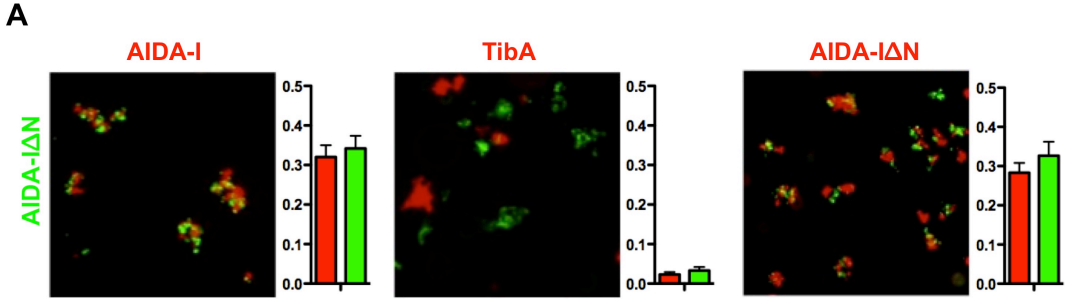


Figure 3

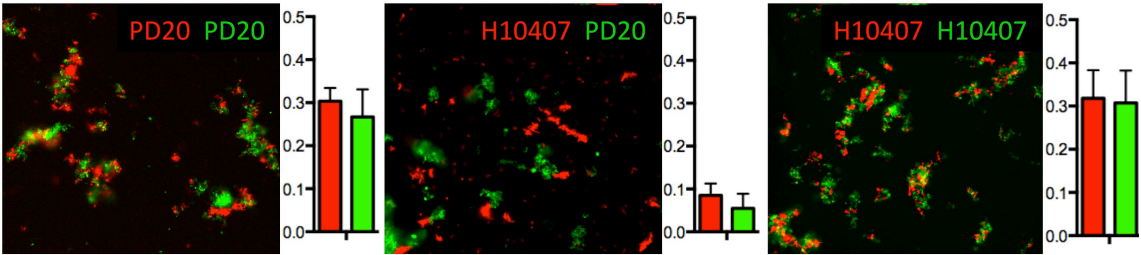


Figure 4

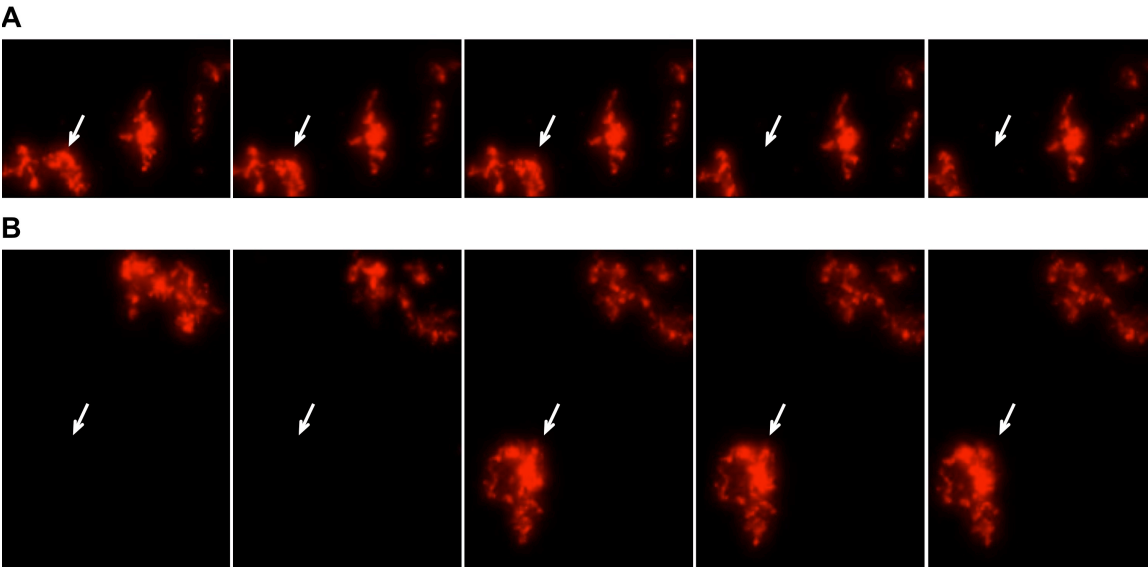
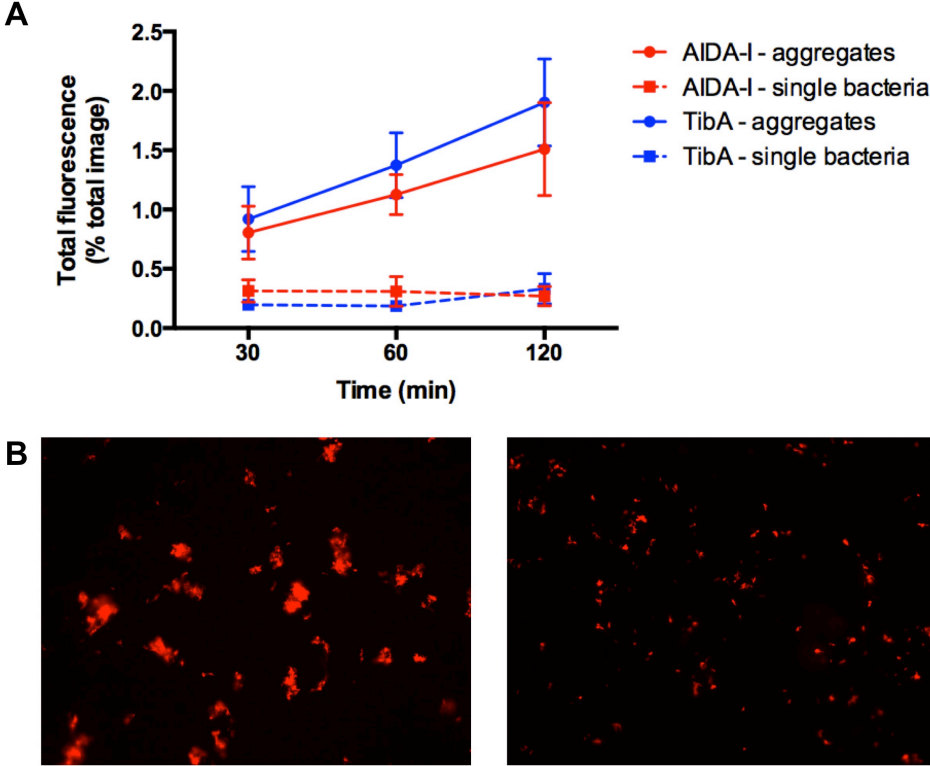


Figure 5



Supplementary informations

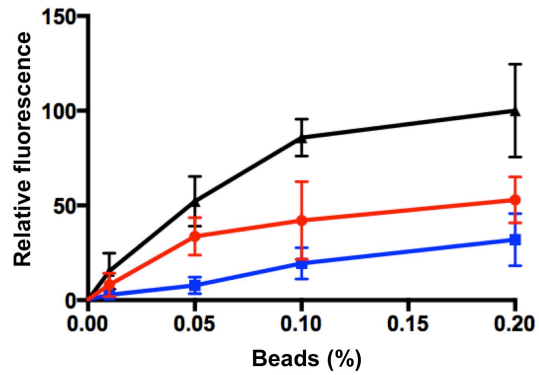


Figure S1: Interaction between AIDA-I, TibA and AIDA-IΔN by ELISA assay

96-well plates were coated overnight with 500 nM of purified AIDA-I (red), purified TibA (blue) or the purified deletion mutant, AIDA-IΔN (black). To probe the interactions, AIDA-IΔN-coated green fluorescent beads were added to the wells at different concentrations and the fluorescence was measured at 528 nm using a fluorescence plate reader. Results represent three experiments performed in duplicates.

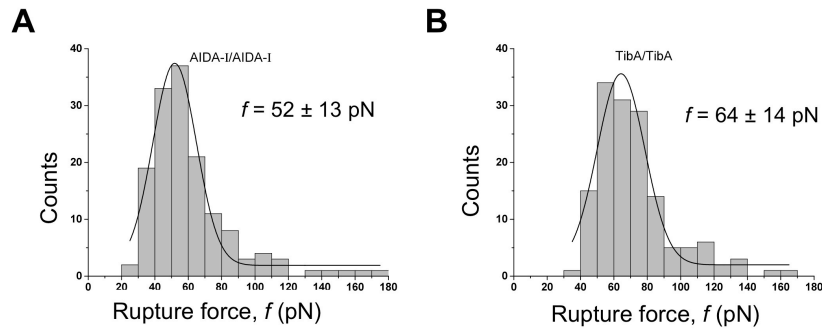


Figure S2: Rupture force of the AIDA-I/AIDA-I and TibA/TibA interaction probed by single-molecule force spectroscopy

All specific interaction events for the AIDA-I/AIDA-I (A) and TibA/TibA (B) pairs were plotted according to the rupture force. The most probable rupture force was obtained from the Gaussian fit to the rupture force histogram

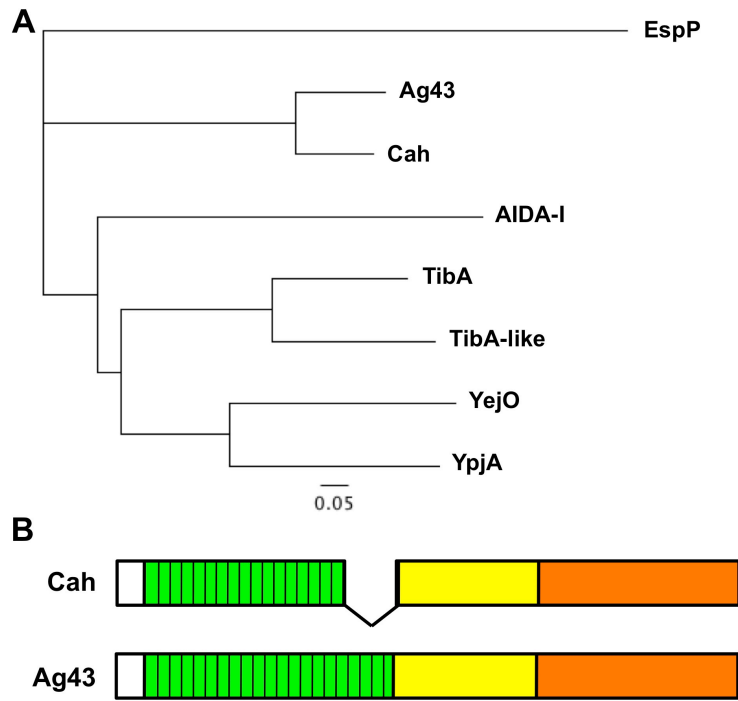


Figure S3: Cah is a deletion variant of Ag43

(A) SAATs from *E. coli* were aligned and a phylogenetic tree was built from these alignments using the neighbor-joining method with Geneious Pro. The sequence of the autotransporter EspP, from *E. coli* O157:H7 was used as an out-group. (B) Schematic representation of Cah and Ag43.

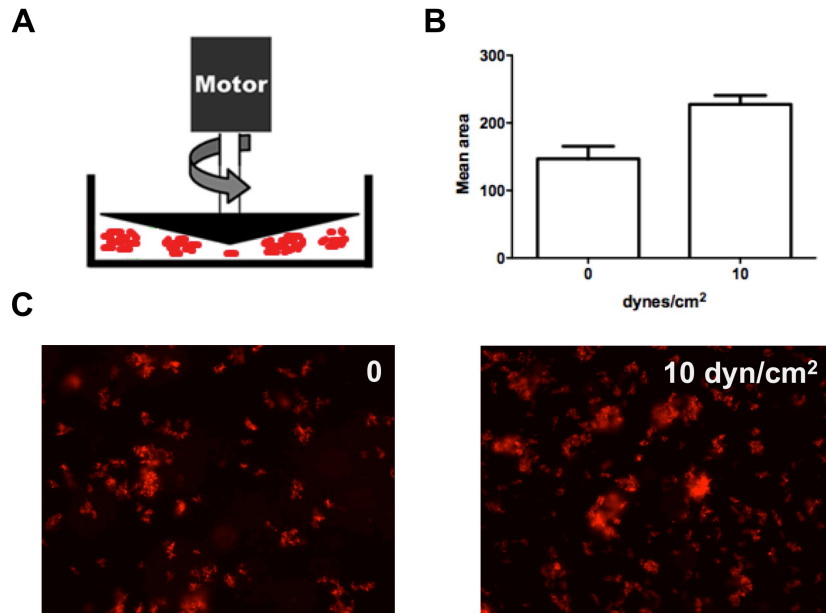


Figure S4: Aggregates are not sensitive to shear stress

(A) Suspensions of *E. coli* strain C600 expressing TibA were submitted to shear stress in a cone and plate device and were analyzed by fluorescence microscopy. (B) The size of aggregates in the suspension were measured by automated image analysis using the ImageJ software. (C) Representative images of the bacterial suspensions before and after a shear stress of 10 dynes/cm²

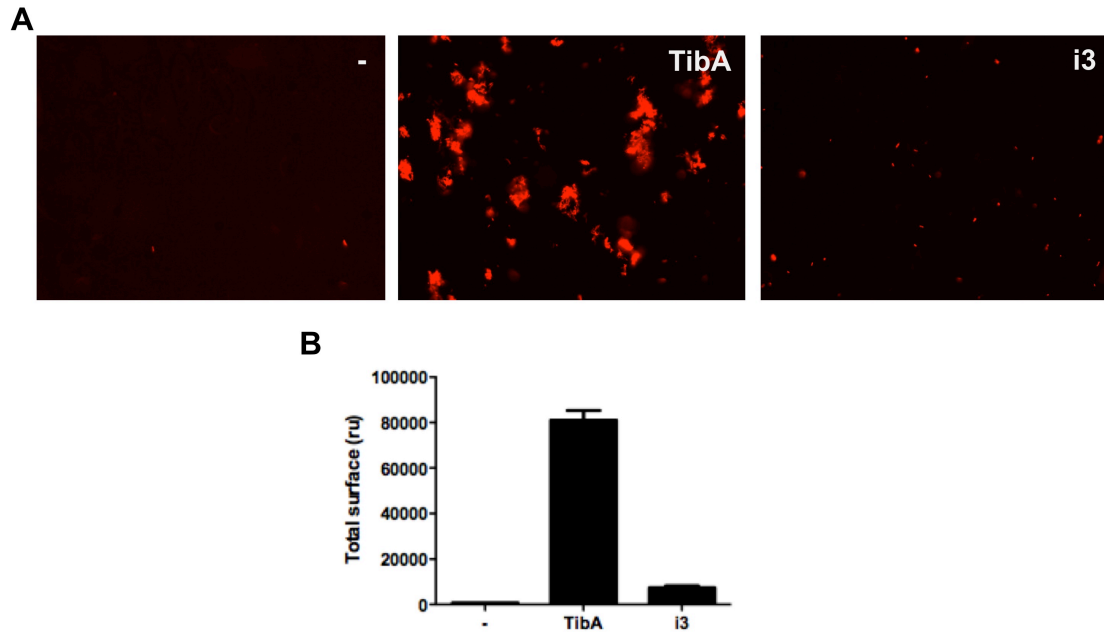


Figure S5: i3 mutant does not adhere under shear stress conditions

(A) *E. coli* strain C600 bearing an empty vector (-) or a plasmid allowing the expressing of TibA or the mutant i3 were allowed to adhere to Hep-2 cells under a shear stress of 0.1 dynes/cm² and visualized by fluorescence microscopy. (B) The total fluorescence per image was measured by automated image analysis using the ImageJ software.

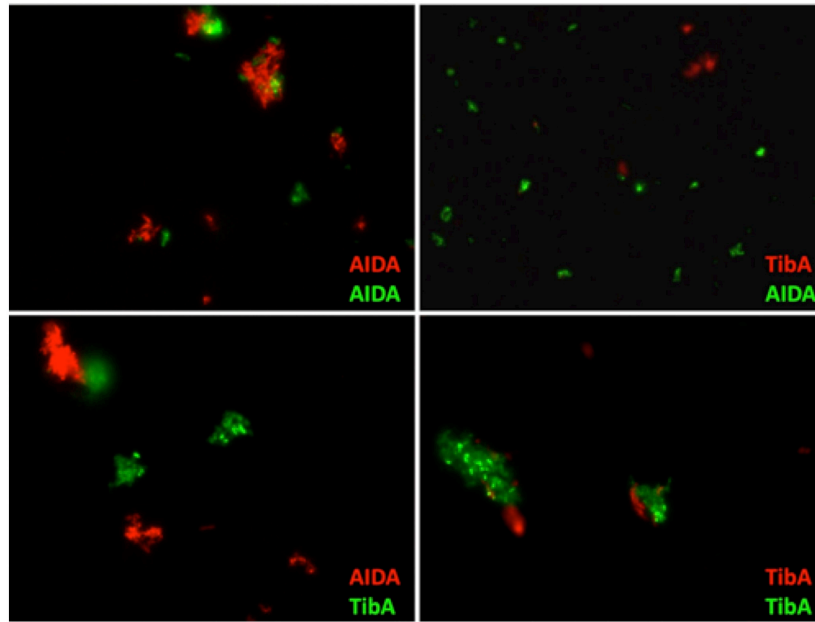


Figure S6: Specificity of interaction during adhesion under shear stress conditions

E. coli strain C600 expressing AIDA-I or TibA and either GFP or mCherry were allowed to adhere to Hep-2 cells under a shear stress of 0.1 dynes/cm² and visualized by fluorescence microscopy.

Discussion

Le but de cette thèse était de caractériser les autotransporteurs auto-associatifs (SAATs) d'*E. coli* afin de définir leur rôle. La famille des SAATs est formée par les autotransporteurs AIDA-I, TibA et Ag43 (6). Ces trois autotransporteurs sont des protéines multifonctionnelles capables de promouvoir l'adhésion et l'invasion de cellules épithéliales ainsi que l'auto-agrégation des bactéries et la formation de biofilm. À cause de ces phénotypes, les SAATs sont considérés comme étant d'importants facteurs de virulence d'*E. coli*. Par ailleurs, lors d'une infection expérimentale de porcelets par la souche PD20, qui contient les facteurs de virulence AIDA-I et STb, la présence de AIDA-I était nécessaire au développement de diarrhée chez les porcelets (331). Toutefois, le rôle exact des SAATs au cours du cycle de vie des bactéries est toujours inconnu. Dans un contexte d'infection, on ne sait toujours pas si les SAATs servent aux bactéries en tant qu'adhésine, en tant que facteur d'agrégation ou bien en tant que facteur permettant la formation de biofilm par exemple.

AIDA-I et Ag43 ont été amplement étudiés. Par contre, il y a peu d'informations à propos de TibA. De plus, il existe plusieurs différences dans l'organisation, la fonctionnalité et la biogenèse d'AIDA-I et d'Ag43 et ces différences ne sont pas très bien comprises. Une étude approfondie de TibA permettrait donc possiblement de comprendre et d'expliquer les différences qui existent entre les SAATs. Par ailleurs, TibA est lui-même assez différent d'AIDA-I et d'Ag43. Par exemple, contrairement à ses deux homologues, TibA ne subit pas de clivage une fois rendu à la surface de la bactérie et possède un domaine additionnel riche en prolines. La première étape en vue d'une caractérisation des SAATs est donc d'abord de définir l'organisation et la fonctionnalité de TibA.

Caractérisation de TibA

TibA, comme plusieurs autres autotransporteurs et comme ses homologues AIDA-I et Ag43, est une protéine multifonctionnelle. AIDA-I et Ag43 sont des protéines modulaires qui sont composées de différents modules; chaque module étant responsable d'un phénotype différent (313, 346). Toutefois, malgré qu'ils partagent les mêmes répétitions dans le domaine passager, il existe des différences importantes dans

l'organisation d'AIDA-I et d'Ag43. Par exemple, AIDA-I possède deux domaines d'adhésion, dont un présent à l'extrémité en N-terminal du domaine passager. Par contre, aucun domaine d'adhésion n'a été identifié dans Ag43 et on retrouve plutôt un domaine responsable de l'auto-agrégation dans la partie N-terminale. Donc, afin de mieux comprendre ce qui caractérise les SAATs, nous avons effectué une étude structure-fonction sur TibA à l'aide d'une méthode de mutagenèse aléatoire par insertion de transposon (article 1), comme cela a été fait pour AIDA-I (313). Les résultats ainsi acquis avec TibA nous ont permis de comparer les trois SAATs ensemble et de mieux comprendre les différences qui existent entre eux.

Tout d'abord, nos résultats démontrent que le domaine passager de TibA est séparé en deux modules : un module N-terminal responsable du phénotype d'auto-agrégation et un module C-terminal responsable du phénotype d'adhésion. Remarquablement, tous les mutants montrant un défaut fonctionnel avaient soit un défaut d'auto-agrégation ou bien un défaut d'adhésion. Par ailleurs, la séparation entre les domaines d'adhésion et d'auto-agrégation est très claire ; tous les mutants avec un défaut d'auto-agrégation sont retrouvés entre les acides aminés 92 et 193 et tous ceux avec un défaut d'adhésion sont compris entre les résidus 193 et 352. Nous n'avons toutefois pas réussi à identifier de domaines précis pour les phénotypes d'invasion ou de formation de biofilm. En fait, tous les mutants avec un défaut d'invasion et de formation de biofilm avaient aussi soit un défaut d'adhésion ou d'auto-agrégation et ces mutants étaient distribués à travers tout le domaine passager. Ces résultats suggèrent donc que les phénotypes principaux associés à l'expression de TibA sont l'adhésion aux cellules épithéliales et l'auto-agrégation des bactéries.

L'invasion et la formation de biofilm pourraient donc être des phénotypes secondaires qui découlent de l'adhésion et de l'auto-agrégation. Par exemple, l'expression de TibA permet à une souche *E. coli* K12 de former un biofilm sur des plaques de polystyrène, malgré l'absence de gènes normalement associés à la formation de biofilm. Il est donc possible que TibA entraîne la formation de biofilm uniquement grâce à sa capacité d'adhérer au polystyrène et sa capacité d'auto-agrégation. Le phénotype d'invasion pourrait aussi découler des deux phénotypes principaux. Il est intéressant de noter que l'observation

au microscope de l'invasion de bactéries exprimant AIDA-I a permis de constater que les bactéries résistantes à la gentamycine étaient principalement retrouvées en agrégats (Mélicca René, résultats non publiés). Il est donc possible que l'adhésion d'agrégats de bactéries entraîne l'invasion subséquente des cellules épithéliales. Certaines bactéries comme *Bartonella henselae* envahissent les cellules endothéliales surtout lorsque les bactéries sont en agrégats (360). Il est aussi intéressant de constater que certains variants de TibA, comme le mutants Δ PP et la forme non glycosylée de TibA (article 5), montrent une forte capacité d'invasion malgré une baisse significative de la capacité d'adhésion. Ceci suggère que les phénotypes d'adhésion et d'invasion ne sont pas liés ensemble. De plus, il pourrait y avoir une corrélation entre l'invasion et l'auto-agrégation puisque TibA Δ PP et TibA non glycosylé sont fortement auto-agrégatifs. Il serait donc intéressant de déterminer le mécanisme utilisé par TibA et les autres SAATs pour envahir les cellules épithéliales. Il serait intéressant de voir si l'auto-agrégation a un rôle dans l'invasion et ce qu'il advient des bactéries une fois à l'intérieur des cellules.

Habituellement les domaines d'adhésion sont retrouvés dans la région d'une protéine la plus éloignée de la membrane, ce qui les rend accessible pour une interaction avec leur récepteur (361). Or dans le cas de TibA, le domaine d'adhésion est retrouvé dans la partie C-terminale du domaine passager. Il est intéressant aussi de noter que la délétion du domaine riche en prolines a un effet sur l'adhésion. Les régions riches en prolines sont souvent impliquées dans les interactions protéine-protéine et beaucoup d'interactions récepteur-ligand reposent sur des domaines riches en prolines (362). Il est donc possible que ce domaine riche en prolines interagisse directement avec le récepteur de TibA à la surface des cellules. Toutefois, il est situé entre le domaine β et la région de jonction et est donc probablement masqué par la structure de la région de jonction. Par ailleurs, les domaines riches en prolines adoptent souvent une conformation étendue, à cause de la rigidité des résidus prolines (363). Celui de TibA est composé de huit répétitions de la séquence de 5 acides aminés GNPDA et ce type particulier de répétition adopte généralement une conformation étendue et flexible (364). Le domaine riche en prolines de TibA pourrait donc servir d'extension, permettant d'éloigner le domaine passager de la membrane de la bactérie et, ainsi, d'exposer le domaine d'adhésion.

Alternativement, le phénotype associé à la délétion du domaine riche en prolines est similaire à celui de la forme non glycosylée de TibA. Dans les deux cas, la protéine n'est plus capable de promouvoir l'adhésion aux cellules épithéliales, mais permet toujours l'invasion, l'auto-agrégation et la formation de biofilm. En fait, la forme non glycosylée de TibA est bloquée dans sa conformation oligomérique d'auto-agrégation (voir plus loin). La glycosylation pourrait donner une certaine flexibilité à la protéine permettant le passage de la conformation oligomérique d'auto-agrégation à la conformation monomérique d'adhésion. Il est donc possible que TibA Δ PP soit aussi bloqué dans cette conformation et que le domaine riche en prolines soit responsable de cette flexibilité dans la protéine. Nous n'avons toutefois pas testé la conformation de TibA Δ PP à différentes concentrations de sel pour voir si la protéine était capable de changer de conformation ou non.

Lorsque l'on compare les trois SAATs ensembles (article 1; Fig. 6), il devient évident qu'il s'agit de trois protéines modulaires, bien que leur organisation soit différente. Il est intéressant de constater que malgré, une séquence et une structure uniforme, c'est-à-dire une hélice β où chaque tour est constitué d'une répétition de la même séquence consensus de 19 acides aminés, chaque domaine passager est composé de différents modules. Cela suggère que la structure en hélice β n'est peut-être pas aussi rigide qu'on l'imagine. D'ailleurs, la structure du domaine passager de VacA de *Helicobacter pylori* montre un domaine passager formé de modules où chaque module est incliné un par rapport à l'autre de façon à ce que l'hélice β n'est pas continue et uniforme (215). En fait, il s'agit exactement de l'observation que nous avons faite avec AIDA-I en microscopie électronique (résultats supplémentaires). AIDA-I adopte une structure plutôt courbée et montre un arrangement de « beads on a string ». Cet arrangement est donc possiblement le reflet des différents modules présents dans la protéine et montre que l'hélice β n'est pas complètement uniforme. Ces différents modules nous montrent aussi qu'il est difficile de prédire le phénotype associé à un domaine passager seulement en regardant la séquence, étant donné que, dans le cas des SAATs, une structure et une séquence similaire donne lieu à des modules associés à des phénotypes différents.

Malgré les différences d'organisation entre les trois SAATs, quelques points communs ressortent de ces études de structure-fonction (article 1; Fig. 6). Premièrement, il

est d'ailleurs intéressant de constater que les trois SAATs sont composés de modules différents. Il semble donc que AIDA-I, TibA et Ag43, malgré une séquence commune, représentent chacun un assemblage distinct de différents modules. Cela met l'accent sur l'étonnante capacité du domaine passager à acquérir de nouvelles fonctions et suggère que les SAATs ont évolué par un mécanisme d'échange et d'acquisition de nouveaux modules. Par exemple, le domaine passager d'AIDA-I ressemble à celui de TibA à l'exception qu'un domaine d'adhésion a été ajouté en N-terminal. D'un autre côté, TibA diffère d'Ag43 par le remplacement du domaine « biofilm » en un domaine d'adhésion, couplé à l'acquisition d'un deuxième domaine (le domaine riche en prolines) aidant l'adhésion. Deuxièmement, la seule fonctionnalité commune aux trois SAATs, selon les études de structure-fonction, est l'auto-agrégation. Il est donc possible que le rôle principal des SAATs soit de permettre l'auto-agrégation des bactéries, alors que les autres fonctionnalités soient le résultat de l'acquisition de domaines spécifiques. Ces deux points, l'évolution des SAATs et le phénotype d'auto-agrégation, seront discutés en détail dans les prochaines sections.

Glycosylation de TibA

La deuxième partie de la caractérisation de TibA concerne sa glycosylation. Il est connu que la glycosylation affecte la capacité de TibA de promouvoir l'adhésion des bactéries (359). Le rôle exact de la glycosylation sur la fonctionnalité de TibA demeure toutefois incertain. Dans le cas des autres SAATs AIDA-I et Ag43, la glycosylation a un effet sur la stabilité de la protéine (279, 280). Pour AIDA-I, lorsque l'on empêche la glycosylation, AIDA-I est moins stable, est très peu exprimé à la surface des bactéries et est donc incapable de promouvoir l'adhésion des bactéries à des cellules épithéliales en culture (279). Un effet de la glycosylation sur la stabilité d'Ag43 a aussi été observé (280). Toutefois, dans la majorité des cas, Ag43 n'est pas glycosylé dans un contexte sauvage. Seules quelques souches d'*E. coli* expriment Ag43 sous une forme glycosylée, mais le facteur qui permet la glycosylation d'Ag43 dans un contexte sauvage reste toujours indéterminé (263, 278). Il est intéressant de constater toutefois que l'expression artificielle d'une forme glycosylée d'Ag43 augmente sa capacité à médier l'adhésion des bactéries à des cellules épithéliales (278). L'effet de la glycosylation sur la stabilité des SAATs est

plutôt bien documenté. Par contre, on ne sait toujours pas si l'effet de la glycosylation sur la fonctionnalité est bien réel. Pour l'instant, on assume que, comme la glycosylation sert à la stabilité de la protéine, l'effet sur la fonctionnalité est plutôt indirect. Il est toutefois possible que la glycosylation ait autant un effet sur la structure de la protéine que sur sa fonctionnalité, comme c'est le cas pour l'autotransporteur trimérique EmaA de *Aggregatibacter actinomycetemcomitans* (283).

L'intérêt de TibA par rapport à AIDA-I et à Ag43 est que la glycosylation de TibA se fait dans un contexte sauvage, avec sa propre glycosyltransferase, et que la forme non-glycosylée de TibA est aussi bien exprimée. Nos résultats ont confirmé que la glycosylation affecte la capacité de TibA de promouvoir l'adhésion sans influencer sa capacité d'invasion, de formation de biofilm et d'auto-agrégation (article 5). De plus, comme pour les autres SAATs, nous avons observé que la glycosylation augmentait la stabilité de la protéine. Par contre, nous avons aussi observé que la forme non glycosylée de TibA adopte sa conformation oligomérique peu importe les conditions, comparativement à la forme glycosylée qui oscille entre une conformation oligomérique et une conformation monomérique selon la concentration en sel. D'ailleurs, comme l'ajout de sel n'entraîne pas un changement de conformation pour TibA non glycosylé, il n'est pas étonnant que l'ajout de sel n'inhibe pas l'auto-agrégation de bactéries exprimant TibA non glycosylé. Il semble donc que lorsque non glycosylé, TibA est pris dans sa conformation oligomérique et que cette conformation oligomérique ne soit pas compatible avec l'adhésion. Donc, pour les SAATs, nous croyons que, plutôt que d'attribuer des propriétés « lectine » à la protéine, la glycosylation augmente la stabilité de la protéine et confère la flexibilité nécessaire à la protéine pour osciller entre ses deux conformations possibles.

Il est intéressant de constater que l'effet de la glycosylation n'est pas aussi drastique sur TibA comparativement à AIDA-I; la forme non glycosylée de AIDA-I est très sensible aux protéases et n'est pas bien exprimé à la surface des bactéries (279). Dans le même ordre d'idée, Ag43 n'a pas besoin d'être glycosylé pour être bien exprimé. AIDA-I est beaucoup plus grand que TibA et Ag43 et possède environ deux fois plus de répétitions de la séquence consensus dans le domaine passager. Il est donc tentant de penser que cette

différence de longueur explique la différence entre l'importance de la glycosylation sur la stabilité et l'expression de AIDA-I et TibA. Plusieurs des nouveaux SAATs identifiés sont aussi associés à une glycosyltransférase de la même famille que Aah et TibC (276). Remarquablement, à l'exception de YpjA, toutes les SAATs ayant 1100 acides aminés ou plus sont associés à une glycosyltransférase, tandis que les SAATs plus petits, à quelques exceptions près, ne sont pas associés à une glycosyltransférase. Il est donc attrayant de croire que les SAATs plus grands nécessitent la glycosylation afin d'adopter une conformation normale résistante à la dégradation. Cette hypothèse implique donc que le rôle premier de la glycosylation des SAATs est de stabiliser la protéine. De plus, nos résultats suggèrent que lorsque non glycosylés, les SAATs adoptent une conformation compatible avec l'auto-agrégation, mais pas avec l'adhésion. Cela suggère donc que le rôle premier des SAATs pourrait donc être de permettre l'interaction et les contacts bactérie-bactérie. Alternativement, il est possible que les SAATs ayant acquis un domaine d'adhésion ait aussi acquis une glycosyltransférase.

Finalement, au cours de notre étude de la glycosylation de TibA, nous avons aussi identifié les peptides modifiés par l'ajout d'un ou plusieurs heptoses pour déterminer à quel endroit de la protéine TibA était glycosylé. Nous avons identifié 7 peptides modifiés dans TibA. Il est intéressant de noter que les peptides glycosylés ne sont pas limités au domaine passager, mais sont aussi retrouvés dans la région de jonction. Il s'agit du premier exemple où la glycosylation des SAATs n'est pas limitée aux répétitions dans le domaine passager (278-280). En fait, comme AIDA-I et Ag43 sont clivés et qu'il est facile d'extraire le domaine passager seul par extraction à la chaleur, la recherche de peptides glycosylés avait été faite dans le domaine passager seulement. Récemment, il a été montré que Aah reconnaissait un motif structural plutôt qu'une séquence (276). Cet élément structural consiste en un motif brin β -boucle-brin β , où le résidu glycosylé est présent au début de la boucle. TibC est fonctionnellement interchangeable avec Aah et utilise fort probablement le même mécanisme d'action (276, 277). La glycosylation observée dans la région de jonction est compatible avec ce modèle, puisque la région de jonction a une structure de deux feuillets β face-à-face et consiste en un enchaînement d'un motif brin β -boucle-brin β .

Évolution des SAATs

Le premier point soulevé par la comparaison des SAATs est leur organisation modulaire. Comme mentionné précédemment, en regardant l'organisation fonctionnelle des trois SAATs, on remarque qu'il s'agit de trois assemblages distincts de modules et que les différences représentent des échanges ou bien des additions de modules. Cela suggère donc que l'évolution des SAATs s'est faite par un remaniement de domaines (« domaine shuffling »), plutôt que par une simple accumulation de mutations ponctuelles. Nous avons d'ailleurs vu ses échanges de modules en regardant plusieurs séquences de l'opéron *aah-aidA* (article 2). En analysant les polymorphismes entre plusieurs séquences de l'opéron *aah-aidA*, nous avons remarqué que ceux-ci se regroupaient par blocs (Article 2; Fig. 2). De plus, les séquences *d'aah-aidA* sont regroupées en trois allèles. De façon intéressante, les allèles sont d'abord séparés selon la provenance géographique : un premier allèle correspond aux deux souches provenant de l'Europe et de l'Asie et les deux autres allèles correspondent aux souches provenant de l'Amérique. Il serait intéressant d'obtenir plus de séquences provenant de l'Europe et de l'Asie afin de confirmer la première séparation entre les allèles selon l'origine géographique. Ensuite, une deuxième séparation se fait selon le profil de facteurs de virulence, alors que les deux allèles américains sont séparés selon la présence du fimbriae F18 et de la toxine STb ou de l'intimine Eae (allèle F18 et allèle STb/Eae). D'ailleurs, le gène codant pour AIDA-I semble se retrouver sur un plasmide qui contient aussi soit le fimbriae F18 ou bien la toxine STb, ce qui suggère que *aidA* est lié à d'autres facteurs de virulence. (Article 2; Fig. 7).

La différence entre les allèles d'AIDA-I se retrouve donc au niveau de blocs complets de séquence et il semble que les blocs soient échangeables d'un allèle à un autre. Par exemple, l'allèle associé au fimbriae F18 présente certains blocs qui proviennent de l'allèle Europe/Asie, tandis que d'autres proviennent de l'allèle STb/Eae (Article 2; Fig. 2). Une des séquences que nous avons identifiée montre d'ailleurs un événement de recombinaison entre les deux allèles américains, confirmant les échanges des modules entre les allèles (Article 2; Fig. 8). Cette construction par blocs semble aussi être partagé par les autres SAAT puisque les blocs de séquences sont aussi visibles pour Ag43 (Article 2; Fig.

S1). Il semble donc que les SAATs aient bel et bien évolué par un mécanisme d'échange et d'acquisition de domaines.

L'idée de l'échange de domaines (« domain shuffling ») a souvent été évoquée en parlant de l'évolution des autotransporteurs, quoique en général on parle de l'échange du domaine passager complet ou bien du domaine β (365-367). Par exemple, l'autotransporteur IgA1 protéase de *Neisseria meningitidis* et les SPATEs des *Enterobacteriaceae* partagent un domaine passager similaire, malgré des domaines β très différents, suggérant que le même domaine passager s'est greffé sur des domaines β différents (100). Ce phénomène avait aussi été observé pour le gène *lav* de *Haemophilus influenzae* qui est formé par la fusion de deux gènes aux contenus GC complètement différents et donc d'origines différentes (368). En fait, en remontant à l'origine des autotransporteurs, on imagine bien qu'un autotransporteur provient de la fusion de deux gènes : un domaine fonctionnel, devenu le domaine passager, et un tonneau β (100, 366). Dans le cas d'AIDA-I, toutefois, l'organisation modulaire est plus complexe que l'organisation normale d'un autotransporteur (séquence signale, domaine passager, région de jonction et domaine β), alors que le domaine passager est lui-même composé de plusieurs modules. AIDA-I représente ainsi un des seuls exemples, avec EspP (367), pour lequel il a été démontré que ce mécanisme d'échange de domaines était à l'origine de l'évolution d'un autotransporteur.

Les échanges de modules observés pour AIDA-I sont probablement à l'origine des différents SAATs. Comme mentionné précédemment, il est possible que la différence entre Ag43 et TibA soit l'échange du domaine biofilm en domaine d'adhésion, alors que la différence entre AIDA-I et TibA est l'ajout d'un domaine d'adhésion en N-terminal. Il est donc possible qu'il y ait plus de SAATs qu'AIDA-I, TibA et Ag43. Nous avons donc cherché d'autres autotransporteurs possédant des répétitions de la même séquence consensus de 19 acides aminés et nous avons identifié plusieurs SAATs potentiels (Article 4; Table I). Tous ces autotransporteurs possèdent des répétitions de la séquence consensus de 19 acides aminés et montrent une organisation similaire à AIDA-I, TibA et Ag43, avec un domaine β et une région de jonction similaires.

La principale différence entre tous ces SAATs est la longueur de la protéine, qui reflète le nombre de répétition de la séquence consensus. Le nombre de répétition varie énormément entre les SAATs, allant de 13 répétitions pour YejO jusqu'à 93 répétitions pour le SAAT de *Cronobacter sakazakii*. En fait, une telle séquence répétée est bien adaptée pour une évolution rapide, puisque les séquences répétées sont reconnues pour être sensibles à des événements de duplication ou de délétion (369). D'ailleurs la différence principale entre Ag43 et Cah est la délétion de la portion C-terminale du domaine passager chez Cah (Article 4; Fig. S3). De plus, le domaine passager du SAAT (YP_003368457) de *Citrobacter rodentium* est formé de répétitions d'un bloc de séquence, lui-même composé de quelques répétitions de la séquence de 19 acides aminés (Fig. 1).

```

186 VNGSASNMLLENGGFLSWL NGHQATNTTINSGGYLSVY GDGSAVDITQNSGGAISDTSATLSGNTNINGSFSI 258
389 AGGSASNMLLENGGYLNVN SGHQATNTTINSGGGLRVS GGGTAVDITQNSGGAISADTSATLSGNTNINGSFSI 461
462 ANGSASNMLLENGGSLYVN SGHQATNTTINSGGGLRVS GGGTAVDITQNSGGAIDNTIYATLSGNTNINGSFSI 534
535 ANGSASNMLLENGGYLYVD GGHQATNTTINSGGILSVS GGGTAVDITQNSGGVIDNTIYATLSGNTNINGSFSI 607

```

Figure 1 : Duplications dans le domaine passager du SAAT de *C. rodentium*

Alignement des répétitions d'une séquence (correspondant à trois répétitions de 19 acides aminés + une boucle longue) provenant du domaine passager du SAAT (YP_003368457) de *Citrobacter rodentium*. Les résidus en vert sont conservés dans la majorité des répétitions.

Nous avons d'abord identifié plusieurs autres SAATs chez *E. coli*. La majorité de ces SAATs nouvellement identifiés n'ont pas été étudiés, à l'exception de Cah (299) et YpjA (300). Cah a été identifié comme étant un homologue d'Ag43 liant le calcium et est fonctionnellement similaire à Ag43 (299). YpjA quant à lui a été renommé EhaD et semble promouvoir la formation de biofilm (300). Dans certains cas, ces SAATs sont spécifiques à certaines souches, comme AIDA-I et TibA, alors que dans d'autres, ils sont retrouvés dans la majorité des souches d'*E. coli*, comme YejO, YpjA, Cah et Ag43. D'ailleurs, les SAATs qui sont le plus souvent retrouvés dans les génomes d'*E. coli* sont aussi partagés par d'autres espèces bactériennes relativement similaire, comme *Shigella* et *Citrobacter*. Par ailleurs, le cas d'Ag43 et Cah est particulièrement intéressant. Plusieurs groupes identifient Cah comme étant un allèle d'Ag43 (348, 349). Cependant, comme la différence entre les SAATs se situe principalement au niveau du nombre de répétition, il est selon moi pertinent de séparer Cah et Ag43. De plus, la distribution entre Cah et Ag43 chez *E. coli* est plutôt

intéressante; Cah est retrouvé dans toutes les souches intestinales tandis que Ag43 est retrouvé chez les autres types d'*E. coli*, notamment les *E. coli* uropathogènes (UPEC), suggérant qu'il s'agit bien de deux SAATs distincts.

Finalement, les SAATs sont retrouvés chez des α -, β - et γ -protéobactéries. Plusieurs SAATs sont retrouvés dans des espèces bactériennes qui sont pathogènes pour l'humain ou les animaux, comme *E. coli*, *Shigella* et *Cronobacter sakazakii*, alors que d'autres sont retrouvés chez des pathogènes de plantes, comme *Pantoea* sp. et *Dickeya dadantii*. Il est toutefois intéressant de noter que les SAATs ne sont pas limités aux bactéries pathogènes, mais sont aussi retrouvés chez des bactéries non-pathogènes, telles que *Oxalobacter formigenes* et *Gluconobacter oxydans*. *O. formigenes* est un symbionte du gros intestin de plusieurs vertébrés et dégrade les cristaux d'oxalate de calcium (370). De son côté, *G. oxydans* se retrouve dans l'environnement dans des endroits riches en sucres et a d'ailleurs une incomparable capacité à dégrader les sucres et les produits dérivés des sucres (371). Finalement, chez *E. coli*, même les souches commensales et de laboratoire, codent pour des SAATs. Ces exemples suggèrent donc que les SAATs ne sont pas limités à un rôle dans l'intestin, ni à un rôle dans la virulence.

Mécanisme d'auto-agrégation

Comme il a été mentionné précédemment, le phénotype d'auto-agrégation semble être le point commun des SAATs. En 2006, Per Klemm et collaborateurs ont proposé de regrouper AIDA-I, TibA et Ag43 ensemble sur la base que ces trois autotransporteurs promouvaient l'auto-agrégation des bactéries par un mécanisme d'auto-association (6). Ils ont alors nommé la famille composée d'AIDA-I, TibA et Ag43, les autotransporteurs auto-associatifs d'*Escherichia coli*. Toutefois, le mécanisme d'auto-association est purement spéculatif et repose sur la simple observation que l'expression d'AIDA-I, de TibA ou d'Ag43 dans une souche *E. coli* de laboratoire (K12) entraînait l'agrégation spécifique des bactéries exprimant la protéine. Ainsi, comme le phénotype d'agrégation était causé par la seule expression d'un des SAAT et qu'il n'y avait pas d'interactions avec les bactéries « vides », ils en ont déduit qu'il s'agissait d'un mécanisme d'auto-association. Or, il existe

d'autres mécanismes qui pourraient permettre cette auto-agrégation spécifique des bactéries. Il est possible par exemple qu'AIDA-I interagisse avec un facteur accessoire présent dans le milieu extracellulaire et que l'interaction de plusieurs molécules d'AIDA-I avec une molécule de ce facteur accessoire entraîne l'auto-agrégation des bactéries. Une autre possibilité est qu'AIDA-I reconnaisse une autre molécule à la surface de la bactérie adjacente et que c'est cette interaction qui permette l'agrégation. Par exemple, l'adhésine Flo1 de *Saccharomyces cerevisiae* permet l'agrégation spécifique des cellules exprimant Flo1 via une interaction entre l'adhésine et un composant de la paroi de l'autre cellule (316). C'est l'interaction réciproque par des molécules Flo1 des deux cellules qui rend l'agrégation spécifique. Nous avons donc voulu déterminer le mécanisme par lequel les SAATs entraînent l'auto-agrégation des bactéries et nous avons utilisé AIDA-I comme modèle étant donné la disponibilité d'outils pour étudier l'auto-agrégation, comme le mutant non agrégatif i24 (313).

La première observation que nous avons faite en étudiant le mécanisme d'auto-agrégation est qu'AIDA-I interagissait bel et bien avec lui-même pour permettre l'auto-agrégation des bactéries. D'abord, la protéine AIDA-I purifiée, mais pas le mutant i24, permet d'inhiber l'auto-agrégation des bactéries exprimant AIDA-I et permet l'auto-agrégation de billes fluorescentes (article 3). Ces résultats suggèrent donc que la protéine AIDA-I seule est suffisante pour permettre l'auto-agrégation des bactéries et donc qu'il n'y a pas de recrutement d'une molécule accessoire. Nous avons aussi vu que la protéine AIDA-I purifiée interagissait avec elle-même par différentes techniques biochimiques, comme le pontage chimique, l'ultracentrifugation analytique et le SPR. Il s'agit d'une interaction de forte affinité, avec un K_D de 45 nM. Nous avons finalement observé par pontage chimique extracellulaire que cette interaction d'AIDA-I avec elle-même se faisait aussi à la surface des bactéries. Ces résultats constituent la première preuve de l'interaction directe de la protéine AIDA-I avec elle-même et que l'auto-agrégation se produit grâce un mécanisme d'auto-association. Au moment de notre démonstration des interactions AIDA-I/AIDA-I, une interaction directe comme celle-là avait été démontré pour très peu de facteurs permettant l'auto-agrégation des bactéries. La protéine H de *Streptococcus*

pyogenes et l'hémagglutinine FHA de *Bordetella pertussis* constituaient les deux autres exemples (317, 372).

Dans un deuxième temps, nous avons observé qu'AIDA-I pouvait former des oligomères de grandes tailles, bien plus importants que des dimères. D'ailleurs, en microscopie électronique, nous avons observé des oligomères contenant jusqu'à 12 molécules d'AIDA-I (résultats supplémentaires). Nous avons ensuite déterminé que AIDA-I pouvait interagir avec d'autres molécules d'AIDA-I en *cis*, dans une orientation « côte-à-côte », ou bien en *trans*, dans une orientation « face-à-face » (résultats supplémentaires). Le modèle qui en découle est donc qu'AIDA-I forme d'abord des dimères en *cis* à la surface d'une même bactérie et que ce sont ces dimères qui interagissent alors en *trans* pour promouvoir l'auto-agrégation des bactéries (résultats supplémentaires; Fig. 8). Nous n'avons malheureusement pas pu déterminer le rôle exact des dimères en *cis*. Toutefois, comme l'auto-agrégation de bactéries exprimant AIDA-I n'est inhibée par la protéine purifiée que lorsque celle-ci peut oligomériser (en absence de détergent), cela suggère que l'oligomérisation d'AIDA-I sur une même bactérie est préalable à l'auto-agrégation des bactéries. De plus, des expériences de pontages chimiques fait sur bactéries, avec un agent pontant incapable de traverser la membrane des bactéries, montrent que l'interaction entre les bactéries se fait via les dimères en *cis* (Victoria Girard, résultats non publiés).

La présence de dimères en *cis* contraste avec le modèle d'interaction proposé par Meng et collaborateurs suivant la structure cristallographique de l'autotransporteur Hap de *Haemophilus influenzae* (214). Hap permet l'auto-agrégation des bactéries grâce à un domaine « SAAT » (214, 373, 374). La structure de Hap montre un dimère en *trans*, suggérant qu'une molécule de Hap à la surface d'une bactérie interagit avec une deuxième molécule à la surface d'une autre bactérie (Fig. 2A). Le modèle ne contient donc pas d'interactions en *cis*. Il est à noter que Hap possède un domaine supplémentaire conférant l'activité sérine protéase. Il s'agit d'un imposant domaine globulaire qui ressort de l'hélice β et il est donc possible que ce domaine empêche les interactions en *cis* par encombrement stérique. Toutefois, remarquablement, le modèle d'interaction prévoit que les oligomères de Hap peuvent grossir par l'ajout d'autres molécules. L'interaction entre deux molécules de

Hap se fait via la face F2 d'une molécule et la boucle F1/F2 de l'autre molécule, ce qui veut dire que chaque molécule de Hap peut interagir avec deux autres molécules présentes sur l'autre bactérie (Fig. 2B). Les oligomères peuvent donc devenir très importants. Donc, bien qu'il n'y ait pas d'interactions en *cis*, il semble que l'oligomérisation de Hap provoque aussi le recrutement de plusieurs molécules de Hap au même endroit sur une même bactérie (Fig. 2A).

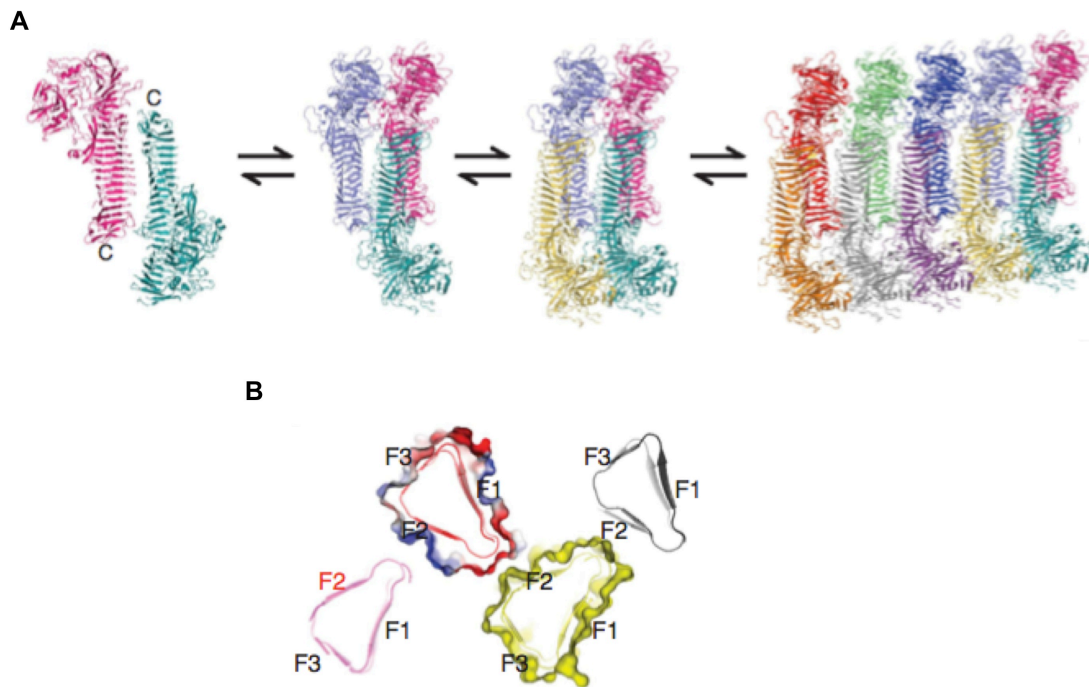


Figure 2 : Modèle d'interaction de l'autotransporteur Hap [figure adaptée de la référence (214)] (A) Une molécule de Hap interagit avec une deuxième molécule en *trans*. Le dimère peut ensuite recruter un autre monomère de Hap, via une nouvelle interaction en *trans*, et ainsi de suite jusqu'à la formation de très gros oligomères. (B) L'interaction entre deux molécules de Hap se fait via la face F2 de l'hélice β d'un monomère et la boucle F1/F2. Comme la surface d'interaction n'est pas la même pour les deux molécules, une deuxième interaction pour chaque molécule est donc possible.

Des interactions en *cis* ont cependant été observées pour d'autres autotransporteurs. Par exemple, l'autotransporteur IcsA de *Shigella flexneri* oligomérisé dans la membrane d'une même bactérie (375). IcsA est retrouvé dans la famille AIDA dans les classifications des autotransporteurs et possède aussi un domaine « SAAT » (100, 214). IcsA interagit avec la molécule N-WASP (début de la cascade de polymérisation de l'actine) des cellules

eucaryotes et entraîne la formation d'une queue d'actine à un pôle de la bactérie, ce qui permet la motilité intra- et inter-cellulaire (376). L'oligomérisation de IcsA est importante pour ce processus en permettant le recrutement de plusieurs molécules N-WASP au même endroit (377). L'oligomérisation semble d'ailleurs être un thème récurrent pour les molécules provoquant l'activation de N-WASP (378). À priori, comme pour AIDA-I, le domaine passager de IcsA ne semble pas contenir de domaine globulaire empêchant les interactions en *cis*. L'exemple d'IcsA suggère donc que l'oligomérisation d'AIDA-I en *cis* est possible.

Finalement, nous avons aussi observé que l'oligomérisation d'AIDA-I entraînait un changement de conformation dans la protéine (article 3). Nous avons observé que l'ajout de sel et de déoxycholate (DOC), deux conditions qui inhibent l'oligomérisation d'AIDA-I ainsi que l'auto-agrégation des bactéries, provoquait d'importants changements de conformation dans la protéine AIDA-I purifiée, qui correspondent à un gain de structure secondaire. Remarquablement, le mutant i24 adopte toujours la même conformation peu importe les conditions et celle-ci correspond à la conformation plus structurée qu'adopte AIDA-I dans des conditions où la protéine est monomérique.

Il semble aussi que la conformation adoptée par AIDA-I corrèle avec sa fonctionnalité. La conformation oligomérique est évidemment responsable de l'auto-agrégation, alors que la conformation monomérique plus structurée serait responsable de l'adhésion. Deux observations suggèrent que la conformation monomérique est responsable de l'adhésion. Premièrement, le mutant i24 capable d'adhésion, mais pas d'auto-agrégation, adopte toujours cette conformation. Deuxièmement, TibA oscille aussi entre une conformation monomérique et une conformation oligomérique. Toutefois, la forme non glycosylée de TibA, qui est capable d'auto-agrégation, mais pas d'adhésion, n'est pas capable d'adopter sa conformation monomérique. Ceci suggère donc que c'est la conformation monomérique qui est responsable du phénotype d'adhésion. Ainsi, il semble qu'AIDA-I puisse adopter une conformation oligomérique responsable de l'auto-agrégation des bactéries ou bien une conformation monomérique responsable du phénotype d'adhésion selon les conditions.

Par ailleurs, l'effet du sel et du DOC, un composant des sels biliaires, indiquent qu'AIDA-I est sensible à des facteurs environnementaux puisque le DOC et d'importantes concentrations de sel sont présents dans l'intestin. Le modèle probable est donc que les agrégats de bactéries sont dispersés en arrivant dans l'intestin, laissant la possibilité aux bactéries seules d'adhérer aux cellules épithéliales. Ces résultats suggèrent donc que l'auto-agrégation des bactéries a plutôt une importance dans l'environnement. D'ailleurs, AIDA-I et son homologue TibA semblent être exprimés dans des conditions similaires à celles retrouvées dans l'environnement (337, 357). Comme mentionné plus tôt, TibA subit le même changement de conformation et semble utiliser le même mécanisme d'interaction qu'AIDA-I, ce qui suggère donc que le mécanisme d'interaction est commun pour tous les SAATs. Il existe d'autres molécules permettant l'adhésion aux cellules ou à d'autres composants et qui subissent des changements de conformation lors de l'interaction. Un exemple intéressant est celui des cadhérines chez les cellules eucaryotes.

Les cadhérines sont les molécules responsables des interactions cellules-cellules chez les eucaryotes et entraînent la séparation des différentes couches de cellules dans les tissus (379). Ce sont des glycoprotéines membranaires dépendantes du calcium qui possèdent un important domaine extracellulaire. Remarquablement, comme pour les SAATs, le domaine extracellulaire des cadhérines est composé d'unité répétée. L'unité de base commune des cadhérines est un domaine cadhérine appelé domaine EC. Les cadhérines classiques possèdent 5 répétitions de ce domaine EC, mais le nombre de domaine EC peut varier pour d'autres types de cadhérine (Fig. 3A). La séquence des domaines EC est considérablement variable, mais ils adoptent toujours une structure conservée de « β -sandwich » similaire à un domaine Ig (immunoglobuline) (380). La jonction entre les domaines EC fixe plusieurs ions Ca^{2+} , ce qui induit un important changement de conformation dans la protéine (381). La fixation du calcium rigidifie le domaine extracellulaire, stabilisant la protéine et lui conférant ainsi une conformation compatible avec son oligomérisation. Donc, les cadhérines, comme les SAATs, réagissent à leur environnement (calcium pour les cadhérines et sel/DOC pour les SAATs) par un changement de conformation qui permet de moduler les interactions cellules-cellules.

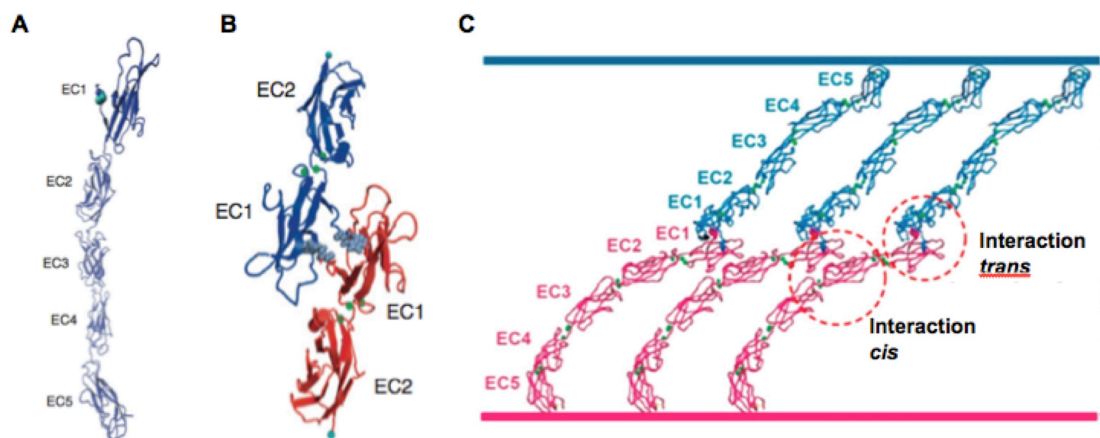


Figure 3 : Structure et mécanisme d'interaction des cadhérines [figure adaptée des références (382) et (383)] (A) Structure du domaine extracellulaire d'une cadhérine classique. (B) Structure montrant le point d'interaction en *trans* (échange du brin N-terminal) entre deux cadhérines. (C) Modèle d'interaction en *cis* et en *trans* des cadhérines.

La comparaison entre les cadhérines et les SAATs ne s'arrête pas là. En effet, les cadhérines utilisent un mécanisme d'interaction semblable à celui des SAATs, c'est-à-dire un mécanisme impliquant des interactions en *cis* et en *trans* (384). Plusieurs groupes ont observés la formation de dimères en *cis* ainsi que l'oligomérisation en *trans*, mais il n'est toujours pas clair si les interactions en *cis* sont toujours présentes lorsqu'il y a oligomérisation en *trans* (384-389). La raison de l'incertitude quant à la présence des dimères en *cis* lors de l'oligomérisation en *trans* vient du fait que certains groupes ont vu que les deux types d'interactions utilisaient le même site de liaison (388, 389). Le domaine EC1, le plus éloigné de la membrane, est le point de contact initial et principal pour les interactions en *trans* (Fig. 3B). Les domaines EC1 de deux cadhérines en *trans* participe à un échange de leur brin β terminal et un seul résidu, le Trp2, semble crucial pour l'interaction (380). Or, d'après certains groupes, ce tryptophane semble aussi être impliqué dans les interactions *cis*, suggérant que les interactions *cis* ne peuvent pas avoir lieu en même temps que les interactions *trans* (389). Toutefois, l'interaction entre cadhérines n'est pas seulement située au niveau du domaine EC1 et plusieurs études ont subséquemment montré qu'ils pouvaient y avoir un chevauchement complet des domaines extracellulaires de cadhérines qui interagissent en *trans* (390-392). Maintenant, plusieurs groupes suggèrent que les interactions en *cis* et en *trans* ne sont pas mutuellement exclusives (380, 393, 394).

Dans tous les cas, tout le monde s'entend sur l'importance des interactions en *cis* afin de solidifier l'interaction entre les deux cellules (Fig. 3C). En fait, l'interaction entre les monomères en *trans* n'est pas assez forte pour expliquer un contact aussi solide que celui observé entre deux cellules (383). Toutefois, en regroupant plusieurs cadhérines au même endroit de la cellule, l'interaction devient alors assez forte. La coopérativité entre les deux types d'interactions est donc crucial pour la formation d'un contact solide entre deux cellules (383, 394).

Il existe donc plusieurs points communs entre les SAATs et les cadhérines. D'abord, ce sont deux types de glycoprotéines membranaires ayant un large domaine extracellulaire formé d'unités répétées et permettant les interactions cellules-cellules. Ensuite, la fonction des SAATs et des cadhérines est régulée par des facteurs environnementaux, le sel et le DOC ainsi que le calcium respectivement, et pour ce faire, les protéines subissent un changement de conformation. Finalement, ces contacts cellules-cellules sont dépendants de la formation d'un important complexe de protéines impliquant des interactions en *cis* et des interactions en *trans*. Il est donc possible que les SAATs soient des molécules « cadhérine-like » chez les bactéries et qu'ils soient responsables des contacts entre bactéries au sein des communautés microbiennes. Une autre caractéristique des cadhérines est qu'elles interagissent préférentiellement de façon homophile (395).

Nous avons observé le même type de spécificité pour les SAATs (Article 5). AIDA-I interagit spécifiquement avec lui-même, mais pas avec TibA. Cette spécificité est basée sur le nombre de répétitions dans le domaine passager ainsi que sur les différences de séquences dans la partie moins conservée des répétitions. Remarquablement, lorsque l'on mélange des bactéries similaires exprimant AIDA-I ou TibA, les agrégats bactériens se séparent les uns des autres de telle sorte que les bactéries exprimant AIDA-I interagissent ensemble et les bactéries exprimant TibA interagissent ensemble. Ces résultats suggèrent que les SAATs permettent de discriminer le soi du non-soi chez les bactéries et ce, chez la même espèce bactérienne. Il est intéressant de constater que plusieurs SAATs sont retrouvés dans des bactéries particulières. Par exemple, AIDA-I est retrouvé surtout chez des *E. coli* qui infectent le porc tandis que TibA est surtout retrouvé chez des souches

humaines. Il est donc très sensé que des bactéries exprimant AIDA-I et TibA n'interagissent pas ensemble étant donné qu'elles colonisent des hôtes complètement différents. Un autre exemple intéressant est celui d'Ag43 et de Cah. Comme mentionné précédemment, Cah est retrouvé chez toutes les souches intestinales de *E. coli* tandis que Ag43 se retrouve chez les autres, comme les UPEC (Article 4; Fig. S3). De plus, Cah et Ag43 sont très similaires et la principale différence est la délétion de la partie C-terminale du domaine passager chez Cah. Il est donc tentant de supposer que Cah a évolué à partir d'Ag43 afin de faire la discrimination entre les souches d'*E. coli* qui doivent se rendre à l'intestin et les autres.

Notre hypothèse est donc que les SAATs représentent une famille de protéines permettant la reconnaissance de soi chez les bactéries. La reconnaissance de soi est rarement évoquée chez les bactéries et il existe très peu de systèmes chez les bactéries qui permettent la reconnaissance de soi. Les autres exemples sont les systèmes Ids chez *Proteus mirabilis* et TraAB chez *Myxococcus xanthus* (396, 397). *Proteus mirabilis* est capable de se déplacer par un processus de « swarming ». Lorsque des « swarms » provenant de deux souches différentes se rencontrent, ils ne se mélangent pas. Ce phénomène repose sur un locus de quelques gènes, appelé Ids pour « Identification of self » car il permet la discrimination entre le soi et le non-soi (396). Ce locus a été identifié par la suite comme étant un système de sécrétion de type 6, mais les effecteurs responsables du phénotype de reconnaissance de soi n'ont toujours pas été identifiés (398). Il ne semble toutefois pas qu'il s'agisse d'un système toxine-antitoxine, comme les effecteurs Tse chez *Pseudomonas aeruginosa* (187), car il n'y a pas de cellules mortes au point de rencontre des deux « swarms ». La reconnaissance de soi chez *Proteus mirabilis* pourrait ainsi avoir un rôle pour la colonisation en conservant une population clonale.

Chez *Myxococcus xanthus*, il a été récemment montré que deux souches pouvaient s'échanger des lipoprotéines de membranes externes, rendant ainsi la population plus apte à répondre à des changements de conditions (399). Cet échange est rendu possible grâce aux protéines TraA des deux bactéries qui interagissent ensemble et provoquent la fusion des membranes (397). Or, les deux bactéries doivent exprimer la même copie de TraA pour que l'échange soit possible. Il semble donc que TraA, en plus de participer dans le mécanisme

d'échange, discrimine les bactéries qui peuvent participer dans cet échange, suggérant que TraA soit une molécule de reconnaissance du soi.

Dans ces deux exemples, la reconnaissance du soi se produit dans un contexte bien précis. Toutefois, un système de reconnaissance de soi comme celui des SAATs qui entraîne l'agrégation de bactéries exprimant le même SAAT ne semble pas avoir une fonction aussi précise. Nous avons observé que l'auto-agrégation des bactéries entraînait une augmentation de la colonisation, puisque l'adhésion des bactéries est significativement augmentée lorsque celles-ci sont en agrégats plutôt que seules (article 4). La reconnaissance de soi pourrait donc aider la colonisation des bactéries. Ce résultat est toutefois contradictoire avec l'effet du sel et du DOC que nous avons observé. Précédemment, nous avons observé que d'importantes concentrations de sel et la présence de DOC inhibaient l'auto-agrégation des bactéries. Or, ceci représente les conditions présentes dans l'intestin, ce qui suggère que les bactéries ne sont plus en agrégats une fois dans l'intestin. Toutefois, lorsqu'un agrégat atteint l'intestin, il est possible que l'effet du sel et du DOC se limitent aux bactéries, et aux molécules d'AIDA-I, qui se retrouvent à l'extérieur de l'agrégat. Les bactéries seraient alors toujours en agrégats et les bactéries retrouvées à l'extérieur des agrégats seraient alors prêtes pour l'adhésion.

Globalement, ces résultats suggèrent que les SAATs ont une importance dans l'environnement plutôt que dans l'hôte lors d'une infection. La présence de SAATs dans des bactéries environnementales supporte cette hypothèse. De plus, AIDA-I et TibA semblent être exprimés lorsque les bactéries sont dans une phase stationnaire où les nutriments sont présents en quantité limitée (337, 357). Comme les cadhérines qui permettent la cohésion des tissus chez les eucaryotes, les SAATs pourrait donc organiser les communautés bactériennes, comme les biofilms. Un tel mécanisme d'interaction et de reconnaissance de soi serait avantageux pour les bactéries en biofilm en permettant un contact étroit entre les bactéries et en permettant aux bactéries de choisir avec qui former un biofilm.

D'ailleurs, les communautés bactériennes ne sont pas si différentes des tissus eucaryotes. Dans les tissus eucaryotes, les cellules sont attachées solidement les unes aux autres par des cadhérines et sont entourés d'une matrice extracellulaire (MEC) avec laquelle les cellules interagissent abondamment, via des protéines appelées intégrines (400). Cette MEC est composée de protéoglycans, de collagène et de protéines et joue un rôle dans la structure des tissus ainsi que dans l'attachement, la migration et le fonctionnement cellulaire. Les protéoglycans sont des chaînes de polysaccharides couplées à une protéine centrale qui occupent l'espace extracellulaire (401). Les protéoglycans sont hydrophiles et adoptent une conformation étendue. Il en résulte la formation d'un gel hydraté qui donne une structure aux tissus et leur confère une résistance à la compression. De plus, plusieurs protéines interagissent avec les protéoglycans, faisant en sorte que plusieurs molécules soient retenues dans les tissus. Le collagène est une protéine qui s'associe en fibrilles et sert d'élément structural à la matrice extracellulaire en plus de réguler l'attachement cellulaire. De plus, le collagène interagit avec l'élastine, ce qui confère l'élasticité aux tissus. Finalement, la MEC est aussi composée de plusieurs protéines, comme la fibronectine et la laminine, qui permettent l'attachement et le fonctionnement cellulaire.

De leur côté, les bactéries ont des éléments qui remplacent chacun des composés de la MEC, dans la matrice d'un biofilm par exemple. Ainsi, la matrice d'un biofilm est composée de polysaccharides, d'ADN extracellulaire et de protéines, ce qui lui permet de maintenir sa structure, sa cohésion et sa fonction. Les polysaccharides constituent un réseau complexe formé de plusieurs longues chaînes linéaires et branchées et jouent un rôle similaire aux protéoglycans de la MEC (318). De son côté, l'ADN extracellulaire sert de composant structural et procure une résistance mécanique au biofilm en plus de promouvoir l'attachement initial des bactéries dans le biofilm (402). Finalement, la matrice du biofilm est composée de plusieurs protéines extracellulaires ayant un rôle structural, par exemple en faisant le lien entre les bactéries et les autres composants de la matrice, ou bien un rôle fonctionnel, par exemple en procurant des nutriments aux bactéries dans le biofilm. Dans ce contexte, les SAATs pourraient agir à titre molécules « cadhérine-like » en permettant l'interaction entre les bactéries dans le biofilm.

Un modèle alternatif est que les SAATs pourraient représenter des gènes de barbe verte (Green Beard genes). Le terme « green beard » a été introduit par Richard Dawkins dans son livre de 1976, *The Selfish Gene*, d'après une idée de William D. Hamilton (403). Le terme décrit des gènes qui procurent un comportement altruiste à un individu. Le premier véritable exemple est celui des fourmis de feu (fire ants) où les fourmis ouvrières possédant un allèle particulier du gène GP-9 reconnaissent les reines possédant ce même allèle par leur phéromone et tuent les autres reines qui ne possèdent pas cet allèle particulier (404). Les gènes de barbe verte n'entrent pas très bien dans la théorie de l'évolution et la théorie du gène égoïste a alors été proposée. Les gènes de barbes vertes sont donc des gènes qui veulent promouvoir la survie des organismes possédant le même gène. Un gène de barbe verte est un gène qui produit une molécule qui est reconnaissable, qui reconnaît cette même molécule chez les autres individus et qui traite ces individus préférentiellement (405) (Fig. 4A).

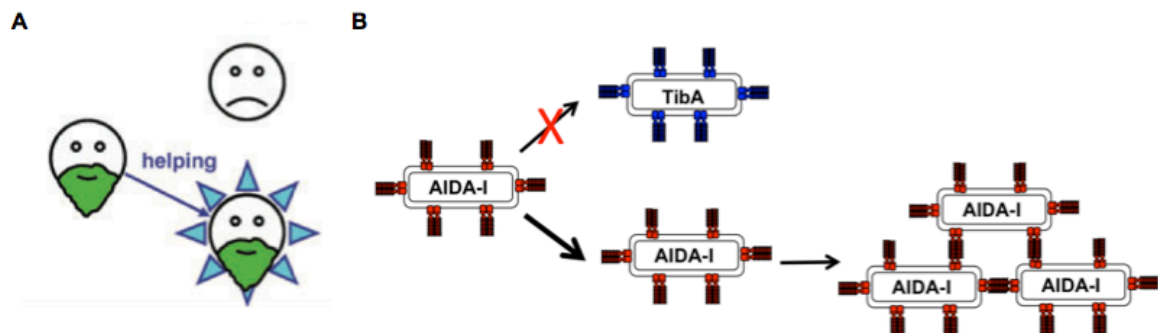


Figure 4 : Les SAATs comme gènes de barbe verte [figure adaptée de la référence (405)] (A) Les gènes de barbes vertes sont des gènes qui : (i) produisent un trait reconnaissable, (ii) reconnaissent ce trait chez d'autres individus et (iii) traitent préférentiellement ces individus. (B) Le gène codant pour AIDA-I : (i) produit un autotransporteur à la surface des bactéries, (ii) qui interagit spécifiquement avec lui-même, mais pas avec son homologue TibA et (iii) il en résulte un agrégat de bactéries exprimant AIDA-I, ce qui augmente la colonisation et la survie des bactéries.

Tout cela est souvent réalisé par deux gènes qui sont très fortement liés entre eux; un gène qui produit la molécule de reconnaissance et un deuxième gène qui produit l'effet bénéfique. En général, l'expression de gènes de barbes vertes a donc un coût énergétique

important. Cela provoque le développement de tricheurs qui profitent du bénéfice apporté à la population sans toutefois déboursier le coût associé à sa production (production de la molécule de reconnaissance mais pas du bien commun). Comme les tricheurs profitent de l'effet bénéfique sans déboursier le coût de production, ce sont plutôt les tricheurs qui prennent le dessus dans la population. À cause de cela, l'effet « barbe verte » ne semble pas être très intéressant et peu de gènes de barbes vertes ont été identifiés. L'association des deux effets (la reconnaissance et l'effet bénéfique) en un seul gène est donc un moyen très efficace pour empêcher le développement de tricheur. D'ailleurs, deux exemples où un seul gène est responsable de l'effet « barbe verte » sont particulièrement intéressants. Il s'agit des gènes *csA* de *Dictyostelium discoideum* (406) et *flo* de *Saccharomyces cerevisiae* (316). Comme les SAATs, CsaA et Flo1 sont des adhésines qui permettent l'agrégation spécifique des cellules exprimant ces mêmes gènes. Dans tous les cas, l'agrégation spécifique des cellules exprimant le gène de barbe verte améliore la survie de ces cellules. AIDA-I et les autres SAATs pourraient donc être une famille de gènes de barbe verte (et autres couleurs) puisqu'ils (i) produisent des molécules de reconnaissance à la surface des bactéries, (ii) qui interagissent spécifiquement avec elle-mêmes, (iii) provoquant la formation d'un agrégat et favorisant ainsi la colonisation et la survie de ces bactéries (Fig. 4B). Si AIDA-I était bel et bien un gène de barbe verte, nous devrions observer une augmentation des bactéries possédant le gène *aida*. De façon intéressante, sur les fermes porcines du Québec, en 2012, il y a une augmentation de la prévalence du gène *aida* chez les souches qui infectent le porc (John M. Fairbrother, communication personnelle), suggérant que AIDA-I donnent un avantage de colonisation et de survie aux bactéries.

Conclusion

Les résultats présentés dans cette thèse permettent de définir beaucoup mieux les autotransporteurs auto-associatifs (SAATs). Nous avons d'abord mis en évidence l'organisation modulaire des SAATs, où chaque SAATs représente un assemblage distinct de différents modules. Conséquemment, le domaine passager adopte une structure en hélice β qui est beaucoup moins droite et uniforme que les modèles le suggèrent. Les modules présents dans les SAATs semblent être facilement échangeables, ce qui suggère que les SAATs ont évolué par acquisition et échange de modules. Conséquemment, les SAATs forment une grande famille d'autotransporteurs et sont présents chez différents types de protéobactéries.

Nous résultats ont aussi mis en évidence l'importance de l'auto-agrégation pour la fonctionnalité des SAATs. Nous suggérons donc que les SAATs ont pour rôle principal de permettre l'auto-agrégation des bactéries. Dans certains cas, AIDA-I par exemple, l'acquisition de domaines spécifiques peut alors permettre de nouvelles fonctionnalités. Il serait donc intéressant de vérifier si les SAATs présents dans les autres bactéries sont capable de promouvoir l'auto-agrégation des bactéries. Il serait aussi intéressant de voir si la capacité d'auto-agrégation est importante pour les bactéries, en testant un mutant non agrégatifs de AIDA-I dans un modèle *in vivo*.

De plus, nos résultats ont permis de faire un parallèle entre les SAATs et les cadhérines eucaryotes. Les cadhérines sont les molécules utilisées par les eucaryotes pour organiser les différents tissus en permettant les interactions cellule-cellule ainsi que les interactions entre les cellules et la matrice extracellulaire. Nous suggérons donc que les SAATs sont des molécules « cadhérines-like » qui organise les communautés bactériennes. Il serait donc intéressant d'étudier les SAATs dans un contexte de biofilm afin de voir si la présence des SAATs joue un rôle sur l'organisation des bactéries dans un biofilm, autant dans le cadre des interactions entre les bactéries que dans les interactions entre les bactéries et la matrice du biofilm.

Nous suggérons donc que les SAATs ne sont peut-être pas des facteurs de virulence au sens stricte, mais sont avant tout des déterminants sociaux pour les bactéries. Nos résultats suggèrent que les SAATs sont des molécules « cadhérines-like » qui permettent la reconnaissance de soi chez les bactéries et donc la discrimination du soi et du non soi. Il est donc aussi possible que les SAATs soient des gènes de barbe verte, c'est-à-dire des gènes égocentriques, qui, par une interaction spécifique avec eux-mêmes, promouvoient la survie des bactéries exprimant le même SAAT.

Références

1. **Nadell, C. D., J. B. Xavier, and K. R. Foster.** 2009. The sociobiology of biofilms. *FEMS Microbiol Rev* **33**:206-224.
2. **Waksman, G.** 2012. Bacterial secretion comes of age. *Philos Trans R Soc Lond B Biol Sci* **367**:1014-1015.
3. **Dautin, N., and H. D. Bernstein.** 2007. Protein secretion in gram-negative bacteria via the autotransporter pathway. *Annu Rev Microbiol* **61**:89-112.
4. **Leyton, D. L., A. E. Rossiter, and I. R. Henderson.** 2012. From self sufficiency to dependence: mechanisms and factors important for autotransporter biogenesis. *Nat Rev Microbiol* **10**:213-225.
5. **Nishimura, K., N. Tajima, Y. H. Yoon, S. Y. Park, and J. R. Tame.** 2010. Autotransporter passenger proteins: virulence factors with common structural themes. *J Mol Med* **88**:451-458.
6. **Klemm, P., R. M. Vejborg, and O. Sherlock.** 2006. Self-associating autotransporters, SAATs: functional and structural similarities. *Int J Med Microbiol* **296**:187-195.
7. **Nataro, J. P., and J. B. Kaper.** 1998. Diarrheagenic *Escherichia coli*. *Clin Microbiol Rev* **11**:142-201.
8. **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**:139-145.
9. **Finlay, B. B., and S. Falkow.** 1997. Common themes in microbial pathogenicity revisited. *Microbiol Mol Biol Rev* **61**:136-169.
10. **Henderson, I. R., J. P. Nataro, J. B. Kaper, T. F. Meyer, S. K. Farrand, D. L. Burns, B. B. Finlay, and J. W. St Geme, 3rd.** 2000. Renaming protein secretion in the gram-negative bacteria. *Trends Microbiol* **8**:352.
11. **Salmond, G. P., and P. J. Reeves.** 1993. Membrane traffic wardens and protein secretion in gram-negative bacteria. *Trends Biochem Sci* **18**:7-12.
12. **Waksman, G., and S. J. Hultgren.** 2009. Structural biology of the chaperone-usher pathway of pilus biogenesis. *Nat Rev Microbiol* **7**:765-774.
13. **DiGiuseppe Champion, P. A., and J. S. Cox.** 2007. Protein secretion systems in *Mycobacteria*. *Cell Microbiol* **9**:1376-1384.
14. **Desvaux, M., N. J. Parham, A. Scott-Tucker, and I. R. Henderson.** 2004. The general secretory pathway: a general misnomer? *Trends Microbiol* **12**:306-309.
15. **Bingle, L. E., C. M. Bailey, and M. J. Pallen.** 2008. Type VI secretion: a beginner's guide. *Curr Opin Microbiol* **11**:3-8.
16. **Natale, P., T. Bruser, and A. J. Driessen.** 2008. Sec- and Tat-mediated protein secretion across the bacterial cytoplasmic membrane--distinct translocases and mechanisms. *Biochim Biophys Acta* **1778**:1735-1756.
17. **Papanikou, E., S. Karamanou, and A. Economou.** 2007. Bacterial protein secretion through the translocase nanomachine. *Nat Rev Microbiol* **5**:839-851.
18. **Chatzi, K. E., M. F. Sardis, S. Karamanou, and A. Economou.** 2013. Breaking on through to the other side: protein export through the bacterial Sec system. *Biochem J* **449**:25-37.

19. **Cannon, K. S., E. Or, W. M. Clemons, Jr., Y. Shibata, and T. A. Rapoport.** 2005. Disulfide bridge formation between SecY and a translocating polypeptide localizes the translocation pore to the center of SecY. *J Cell Biol* **169**:219-225.
20. **Matsuyama, S., J. Akimaru, and S. Mizushima.** 1990. SecE-dependent overproduction of SecY in Escherichia coli. Evidence for interaction between two components of the secretory machinery. *FEBS Lett* **269**:96-100.
21. **Brundage, L., J. P. Hendrick, E. Schiebel, A. J. Driessen, and W. Wickner.** 1990. The purified E. coli integral membrane protein SecY/E is sufficient for reconstitution of SecA-dependent precursor protein translocation. *Cell* **62**:649-657.
22. **Van den Berg, B., W. M. Clemons, Jr., I. Collinson, Y. Modis, E. Hartmann, S. C. Harrison, and T. A. Rapoport.** 2004. X-ray structure of a protein-conducting channel. *Nature* **427**:36-44.
23. **Gumbart, J., and K. Schulten.** 2006. Molecular dynamics studies of the archaeal translocon. *Biophys J* **90**:2356-2367.
24. **Tian, P., and I. Andricioaei.** 2006. Size, motion, and function of the SecY translocon revealed by molecular dynamics simulations with virtual probes. *Biophys J* **90**:2718-2730.
25. **Bonardi, F., E. Halza, M. Walko, F. Du Plessis, N. Nouwen, B. L. Feringa, and A. J. Driessen.** 2011. Probing the SecYEG translocation pore size with preproteins conjugated with sizable rigid spherical molecules. *Proc Natl Acad Sci U S A* **108**:7775-7780.
26. **Zimmer, J., Y. Nam, and T. A. Rapoport.** 2008. Structure of a complex of the ATPase SecA and the protein-translocation channel. *Nature* **455**:936-943.
27. **Egea, P. F., and R. M. Stroud.** 2010. Lateral opening of a translocon upon entry of protein suggests the mechanism of insertion into membranes. *Proc Natl Acad Sci U S A* **107**:17182-17187.
28. **Scheuring, J., N. Braun, L. Nothdurft, M. Stumpf, A. K. Veenendaal, S. Kol, C. van der Does, A. J. Driessen, and S. Weinkauf.** 2005. The oligomeric distribution of SecYEG is altered by SecA and translocation ligands. *J Mol Biol* **354**:258-271.
29. **Deville, K., V. A. Gold, A. Robson, S. Whitehouse, R. B. Sessions, S. A. Baldwin, S. E. Radford, and I. Collinson.** 2011. The oligomeric state and arrangement of the active bacterial translocon. *J Biol Chem* **286**:4659-4669.
30. **Nishiyama, K., T. Suzuki, and H. Tokuda.** 1996. Inversion of the membrane topology of SecG coupled with SecA-dependent preprotein translocation. *Cell* **85**:71-81.
31. **Kusters, I., and A. J. Driessen.** 2011. SecA, a remarkable nanomachine. *Cell Mol Life Sci* **68**:2053-2066.
32. **Sianidis, G., S. Karamanou, E. Vrontou, K. Boulias, K. Repanas, N. Kyrpides, A. S. Politou, and A. Economou.** 2001. Cross-talk between catalytic and regulatory elements in a DEAD motor domain is essential for SecA function. *EMBO J* **20**:961-970.
33. **Gouridis, G., S. Karamanou, I. Gelis, C. G. Kalodimos, and A. Economou.** 2009. Signal peptides are allosteric activators of the protein translocase. *Nature* **462**:363-367.
34. **Eser, M., and M. Ehrmann.** 2003. SecA-dependent quality control of intracellular protein localization. *Proc Natl Acad Sci U S A* **100**:13231-13234.

35. **Lycklama, A. N. J. A., and A. J. Driessen.** 2012. The bacterial Sec-translocase: structure and mechanism. *Philos Trans R Soc Lond B Biol Sci* **367**:1016-1028.
36. **Gardel, C., K. Johnson, A. Jacq, and J. Beckwith.** 1990. The secD locus of *E. coli* codes for two membrane proteins required for protein export. *EMBO J* **9**:4205-4206.
37. **Matsuyama, S., Y. Fujita, and S. Mizushima.** 1993. SecD is involved in the release of translocated secretory proteins from the cytoplasmic membrane of *Escherichia coli*. *EMBO J* **12**:265-270.
38. **Tsukazaki, T., H. Mori, Y. Echizen, R. Ishitani, S. Fukai, T. Tanaka, A. Perederina, D. G. Vassylyev, T. Kohno, A. D. Maturana, K. Ito, and O. Nureki.** 2011. Structure and function of a membrane component SecDF that enhances protein export. *Nature* **474**:235-238.
39. **Samuelson, J. C., M. Chen, F. Jiang, I. Moller, M. Wiedmann, A. Kuhn, G. J. Phillips, and R. E. Dalbey.** 2000. YidC mediates membrane protein insertion in bacteria. *Nature* **406**:637-641.
40. **Nouwen, N., and A. J. Driessen.** 2002. SecDFyajC forms a heterotetrameric complex with YidC. *Mol Microbiol* **44**:1397-1405.
41. **Paetzel, M., A. Karla, N. C. Strynadka, and R. E. Dalbey.** 2002. Signal peptidases. *Chem Rev* **102**:4549-4580.
42. **Hegde, R. S., and H. D. Bernstein.** 2006. The surprising complexity of signal sequences. *Trends Biochem Sci* **31**:563-571.
43. **Schatz, P. J., and J. Beckwith.** 1990. Genetic analysis of protein export in *Escherichia coli*. *Annu Rev Genet* **24**:215-248.
44. **Eisner, G., H. G. Koch, K. Beck, J. Brunner, and M. Muller.** 2003. Ligand crowding at a nascent signal sequence. *J Cell Biol* **163**:35-44.
45. **Poritz, M. A., H. D. Bernstein, K. Strub, D. Zopf, H. Wilhelm, and P. Walter.** 1990. An *E. coli* ribonucleoprotein containing 4.5S RNA resembles mammalian signal recognition particle. *Science* **250**:1111-1117.
46. **Kaiser, C. M., H. C. Chang, V. R. Agashe, S. K. Lakshminpathy, S. A. Etchells, M. Hayer-Hartl, F. U. Hartl, and J. M. Barral.** 2006. Real-time observation of trigger factor function on translating ribosomes. *Nature* **444**:455-460.
47. **Hoffmann, A., B. Bukau, and G. Kramer.** 2010. Structure and function of the molecular chaperone Trigger Factor. *Biochim Biophys Acta* **1803**:650-661.
48. **Buskiewicz, I., E. Deuerling, S. Q. Gu, J. Jockel, M. V. Rodnina, B. Bukau, and W. Wintermeyer.** 2004. Trigger factor binds to ribosome-signal-recognition particle (SRP) complexes and is excluded by binding of the SRP receptor. *Proc Natl Acad Sci U S A* **101**:7902-7906.
49. **Gierasch, L. M.** 1989. Signal sequences. *Biochemistry* **28**:923-930.
50. **Ulbrandt, N. D., J. A. Newitt, and H. D. Bernstein.** 1997. The *E. coli* signal recognition particle is required for the insertion of a subset of inner membrane proteins. *Cell* **88**:187-196.
51. **Patzelt, H., S. Rudiger, D. Brehmer, G. Kramer, S. Vorderwulbecke, E. Schaffitzel, A. Waitz, T. Hesterkamp, L. Dong, J. Schneider-Mergener, B. Bukau, and E. Deuerling.** 2001. Binding specificity of *Escherichia coli* trigger factor. *Proc Natl Acad Sci U S A* **98**:14244-14249.
52. **Driessen, A. J., E. H. Manting, and C. van der Does.** 2001. The structural basis of protein targeting and translocation in bacteria. *Nat Struct Biol* **8**:492-498.

53. **Grudnik, P., G. Bange, and I. Sinning.** 2009. Protein targeting by the signal recognition particle. *Biol Chem* **390**:775-782.
54. **Bechtluft, P., R. G. van Leeuwen, M. Tyreman, D. Tomkiewicz, N. Nouwen, H. L. Tepper, A. J. Driessen, and S. J. Tans.** 2007. Direct observation of chaperone-induced changes in a protein folding pathway. *Science* **318**:1458-1461.
55. **Qi, H. Y., J. B. Hyndman, and H. D. Bernstein.** 2002. DnaK promotes the selective export of outer membrane protein precursors in SecA-deficient *Escherichia coli*. *J Biol Chem* **277**:51077-51083.
56. **Sakr, S., A. M. Cirinesi, R. S. Ullers, F. Schwager, C. Georgopoulos, and P. Genevoux.** 2010. Lon protease quality control of presecretory proteins in *Escherichia coli* and its dependence on the SecB and DnaJ (Hsp40) chaperones. *J Biol Chem* **285**:23506-23514.
57. **Fekkes, P., C. van der Does, and A. J. Driessen.** 1997. The molecular chaperone SecB is released from the carboxy-terminus of SecA during initiation of precursor protein translocation. *EMBO J* **16**:6105-6113.
58. **Economou, A., J. A. Pogliano, J. Beckwith, D. B. Oliver, and W. Wickner.** 1995. SecA membrane cycling at SecYEG is driven by distinct ATP binding and hydrolysis events and is regulated by SecD and SecF. *Cell* **83**:1171-1181.
59. **du Plessis, D. J., N. Nouwen, and A. J. Driessen.** 2011. The Sec translocase. *Biochim Biophys Acta* **1808**:851-865.
60. **Palmer, T., and B. C. Berks.** 2012. The twin-arginine translocation (Tat) protein export pathway. *Nat Rev Microbiol* **10**:483-496.
61. **De Buck, E., E. Lammertyn, and J. Anne.** 2008. The importance of the twin-arginine translocation pathway for bacterial virulence. *Trends Microbiol* **16**:442-453.
62. **Widdick, D. A., K. Dilks, G. Chandra, A. Bottrill, M. Naldrett, M. Pohlschroder, and T. Palmer.** 2006. The twin-arginine translocation pathway is a major route of protein export in *Streptomyces coelicolor*. *Proc Natl Acad Sci U S A* **103**:17927-17932.
63. **Rose, R. W., T. Bruser, J. C. Kissinger, and M. Pohlschroder.** 2002. Adaptation of protein secretion to extremely high-salt conditions by extensive use of the twin-arginine translocation pathway. *Mol Microbiol* **45**:943-950.
64. **Cristobal, S., J. W. de Gier, H. Nielsen, and G. von Heijne.** 1999. Competition between Sec- and TAT-dependent protein translocation in *Escherichia coli*. *EMBO J* **18**:2982-2990.
65. **Stanley, N. R., T. Palmer, and B. C. Berks.** 2000. The twin arginine consensus motif of Tat signal peptides is involved in Sec-independent protein targeting in *Escherichia coli*. *J Biol Chem* **275**:11591-11596.
66. **Berks, B. C., F. Sargent, and T. Palmer.** 2000. The Tat protein export pathway. *Mol Microbiol* **35**:260-274.
67. **Graubner, W., A. Schierhorn, and T. Bruser.** 2007. DnaK plays a pivotal role in Tat targeting of CueO and functions beside SlyD as a general Tat signal binding chaperone. *J Biol Chem* **282**:7116-7124.
68. **Pommier, J., V. Mejean, G. Giordano, and C. Iobbi-Nivol.** 1998. TorD, a cytoplasmic chaperone that interacts with the unfolded trimethylamine N-oxide reductase enzyme (TorA) in *Escherichia coli*. *J Biol Chem* **273**:16615-16620.

69. **Alami, M., I. Luke, S. Deitermann, G. Eisner, H. G. Koch, J. Brunner, and M. Muller.** 2003. Differential interactions between a twin-arginine signal peptide and its translocase in *Escherichia coli*. *Mol Cell* **12**:937-946.
70. **Mori, H., and K. Cline.** 2002. A twin arginine signal peptide and the pH gradient trigger reversible assembly of the thylakoid [Δ]pH/Tat translocase. *J Cell Biol* **157**:205-210.
71. **Gohlke, U., L. Pullan, C. A. McDevitt, I. Porcelli, E. de Leeuw, T. Palmer, H. R. Saibil, and B. C. Berks.** 2005. The TatA component of the twin-arginine protein transport system forms channel complexes of variable diameter. *Proc Natl Acad Sci U S A* **102**:10482-10486.
72. **Luke, I., J. I. Handford, T. Palmer, and F. Sargent.** 2009. Proteolytic processing of *Escherichia coli* twin-arginine signal peptides by LepB. *Arch Microbiol* **191**:919-925.
73. **Cline, K., and M. McCaffery.** 2007. Evidence for a dynamic and transient pathway through the TAT protein transport machinery. *EMBO J* **26**:3039-3049.
74. **Barnett, J. P., R. T. Eijlander, O. P. Kuipers, and C. Robinson.** 2008. A minimal Tat system from a gram-positive organism: a bifunctional TatA subunit participates in discrete TatAC and TatA complexes. *J Biol Chem* **283**:2534-2542.
75. **Sklar, J. G., T. Wu, D. Kahne, and T. J. Silhavy.** 2007. Defining the roles of the periplasmic chaperones SurA, Skp, and DegP in *Escherichia coli*. *Genes Dev* **21**:2473-2484.
76. **Paschen, S. A., T. Waizenegger, T. Stan, M. Preuss, M. Cyrklaff, K. Hell, D. Rapaport, and W. Neupert.** 2003. Evolutionary conservation of biogenesis of beta-barrel membrane proteins. *Nature* **426**:862-866.
77. **Voulhoux, R., M. P. Bos, J. Geurtsen, M. Mols, and J. Tommassen.** 2003. Role of a highly conserved bacterial protein in outer membrane protein assembly. *Science* **299**:262-265.
78. **Kim, S., J. C. Malinverni, P. Sliz, T. J. Silhavy, S. C. Harrison, and D. Kahne.** 2007. Structure and function of an essential component of the outer membrane protein assembly machine. *Science* **317**:961-964.
79. **Knowles, T. J., M. Jeeves, S. Bobat, F. Dancea, D. McClelland, T. Palmer, M. Overduin, and I. R. Henderson.** 2008. Fold and function of polypeptide transport-associated domains responsible for delivering unfolded proteins to membranes. *Mol Microbiol* **68**:1216-1227.
80. **Robert, V., E. B. Volokhina, F. Senf, M. P. Bos, P. Van Gelder, and J. Tommassen.** 2006. Assembly factor Omp85 recognizes its outer membrane protein substrates by a species-specific C-terminal motif. *PLoS Biol* **4**:e377.
81. **Bos, M. P., V. Robert, and J. Tommassen.** 2007. Functioning of outer membrane protein assembly factor Omp85 requires a single POTRA domain. *EMBO Rep* **8**:1149-1154.
82. **Wu, T., J. Malinverni, N. Ruiz, S. Kim, T. J. Silhavy, and D. Kahne.** 2005. Identification of a multicomponent complex required for outer membrane biogenesis in *Escherichia coli*. *Cell* **121**:235-245.
83. **Gatsos, X., A. J. Perry, K. Anwari, P. Dolezal, P. P. Wolyneec, V. A. Likic, A. W. Purcell, S. K. Buchanan, and T. Lithgow.** 2008. Protein secretion and outer membrane assembly in Alphaproteobacteria. *FEMS Microbiol Rev* **32**:995-1009.

84. **Chan, N. C., V. A. Likic, R. F. Waller, T. D. Mulhern, and T. Lithgow.** 2006. The C-terminal TPR domain of Tom70 defines a family of mitochondrial protein import receptors found only in animals and fungi. *J Mol Biol* **358**:1010-1022.
85. **Ruiz, N., D. Kahne, and T. J. Silhavy.** 2006. Advances in understanding bacterial outer-membrane biogenesis. *Nat Rev Microbiol* **4**:57-66.
86. **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**:206-214.
87. **Krojer, T., J. Sawa, E. Schafer, H. R. Saibil, M. Ehrmann, and T. Clausen.** 2008. Structural basis for the regulated protease and chaperone function of DegP. *Nature* **453**:885-890.
88. **Jiang, J., X. Zhang, Y. Chen, Y. Wu, Z. H. Zhou, Z. Chang, and S. F. Sui.** 2008. Activation of DegP chaperone-protease via formation of large cage-like oligomers upon binding to substrate proteins. *Proc Natl Acad Sci U S A* **105**:11939-11944.
89. **Korotkov, K. V., M. Sandkvist, and W. G. Hol.** 2012. The type II secretion system: biogenesis, molecular architecture and mechanism. *Nat Rev Microbiol* **10**:336-351.
90. **Pugsley, A. P., O. Francetic, O. M. Possot, N. Sauvonnet, and K. R. Hardie.** 1997. Recent progress and future directions in studies of the main terminal branch of the general secretory pathway in Gram-negative bacteria--a review. *Gene* **192**:13-19.
91. **Hirst, T. R., J. Sanchez, J. B. Kaper, S. J. Hardy, and J. Holmgren.** 1984. Mechanism of toxin secretion by *Vibrio cholerae* investigated in strains harboring plasmids that encode heat-labile enterotoxins of *Escherichia coli*. *Proc Natl Acad Sci U S A* **81**:7752-7756.
92. **d'Enfert, C., C. Chapon, and A. P. Pugsley.** 1987. Export and secretion of the lipoprotein pullulanase by *Klebsiella pneumoniae*. *Mol Microbiol* **1**:107-116.
93. **Reichow, S. L., K. V. Korotkov, W. G. Hol, and T. Gonen.** 2010. Structure of the cholera toxin secretion channel in its closed state. *Nat Struct Mol Biol* **17**:1226-1232.
94. **Guilvout, I., M. Chami, A. Engel, A. P. Pugsley, and N. Bayan.** 2006. Bacterial outer membrane secretin PulD assembles and inserts into the inner membrane in the absence of its pilotin. *EMBO J* **25**:5241-5249.
95. **Peabody, C. R., Y. J. Chung, M. R. Yen, D. Vidal-Ingigliardi, A. P. Pugsley, and M. H. Saier, Jr.** 2003. Type II protein secretion and its relationship to bacterial type IV pili and archaeal flagella. *Microbiology* **149**:3051-3072.
96. **Burrows, L. L.** 2005. Weapons of mass retraction. *Mol Microbiol* **57**:878-888.
97. **Filloux, A.** 2010. Secretion signal and protein targeting in bacteria: a biological puzzle. *J Bacteriol* **192**:3847-3849.
98. **Douzi, B., G. Ball, C. Cambillau, M. Tegoni, and R. Voulhoux.** 2011. Deciphering the Xcp *Pseudomonas aeruginosa* type II secretion machinery through multiple interactions with substrates. *J Biol Chem* **286**:40792-40801.
99. **Reichow, S. L., K. V. Korotkov, M. Gonen, J. Sun, J. R. Delarosa, W. G. Hol, and T. Gonen.** 2011. The binding of cholera toxin to the periplasmic vestibule of the type II secretion channel. *Channels (Austin)* **5**:215-218.

100. **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**:692-744.
101. **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**:1088-1101.
102. **Grijpstra, J., J. Arenas, L. Rutten, and J. Tommassen.** 2013. Autotransporter secretion: varying on a theme. *Res Microbiol* **164**:562-82.
103. **Jacob-Dubuisson, F., C. Locht, and R. Antoine.** 2001. Two-partner secretion in Gram-negative bacteria: a thrifty, specific pathway for large virulence proteins. *Mol Microbiol* **40**:306-313.
104. **Jacob-Dubuisson, F., R. Fernandez, and L. Coutte.** 2004. Protein secretion through autotransporter and two-partner pathways. *Biochim Biophys Acta* **1694**:235-257.
105. **Mazar, J., and P. A. Cotter.** 2006. Topology and maturation of filamentous haemagglutinin suggest a new model for two-partner secretion. *Mol Microbiol* **62**:641-654.
106. **St Geme, J. W., 3rd, and H. J. Yeo.** 2009. A prototype two-partner secretion pathway: the *Haemophilus influenzae* HMW1 and HMW2 adhesin systems. *Trends Microbiol* **17**:355-360.
107. **Jacob-Dubuisson, F., J. Guerin, S. Baelen, and B. Clantin.** 2013. Two-partner secretion: as simple as it sounds? *Res Microbiol* **164**:583-95.
108. **Kajava, A. V., and A. C. Steven.** 2006. 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.
109. **Delattre, A. S., N. Saint, B. Clantin, E. Willery, G. Lippens, C. Locht, V. Villeret, and F. Jacob-Dubuisson.** 2011. Substrate recognition by the POTRA domains of TpsB transporter FhaC. *Mol Microbiol* **81**:99-112.
110. **Grass, S., and J. W. St Geme, 3rd.** 2000. Maturation and secretion of the non-typable *Haemophilus influenzae* HMW1 adhesin: roles of the N-terminal and C-terminal domains. *Mol Microbiol* **36**:55-67.
111. **Jacob-Dubuisson, F., V. Villeret, B. Clantin, A. S. Delattre, and N. Saint.** 2009. First structural insights into the TpsB/Omp85 superfamily. *Biol Chem* **390**:675-684.
112. **Clantin, B., A. S. Delattre, P. Rucktooa, N. Saint, A. C. Meli, C. Locht, F. Jacob-Dubuisson, and V. Villeret.** 2007. Structure of the membrane protein FhaC: a member of the Omp85-TpsB transporter superfamily. *Science* **317**:957-961.
113. **Surana, N. K., A. Z. Buscher, G. G. Hardy, S. Grass, T. Kehl-Fie, and J. W. St Geme, 3rd.** 2006. Translocator proteins in the two-partner secretion family have multiple domains. *J Biol Chem* **281**:18051-18058.
114. **Li, H., S. Grass, T. Wang, T. Liu, and J. W. St Geme, 3rd.** 2007. Structure of the *Haemophilus influenzae* HMW1B translocator protein: evidence for a twin pore. *J Bacteriol* **189**:7497-7502.
115. **Fan, E., S. Fiedler, F. Jacob-Dubuisson, and M. Muller.** 2012. Two-partner secretion of gram-negative bacteria: a single beta-barrel protein enables transport across the outer membrane. *J Biol Chem* **287**:2591-2599.

116. **Baud, C., H. Hodak, E. Willery, H. Drobecq, C. Loch, M. Jamin, and F. Jacob-Dubuisson.** 2009. Role of DegP for two-partner secretion in *Bordetella*. *Mol Microbiol* **74**:315-329.
117. **Linke, D., T. Riess, I. B. Autenrieth, A. Lupas, and V. A. Kempf.** 2006. Trimeric autotransporter adhesins: variable structure, common function. *Trends Microbiol* **14**:264-270.
118. **Lyskowski, A., J. C. Leo, and A. Goldman.** 2011. Structure and biology of trimeric autotransporter adhesins. *Adv Exp Med Biol* **715**:143-158.
119. **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**:2179-2189.
120. **Nummelin, H., M. C. Merckel, J. C. Leo, H. Lankinen, M. Skurnik, and A. Goldman.** 2004. The *Yersinia* adhesin YadA collagen-binding domain structure is a novel left-handed parallel beta-roll. *EMBO J* **23**:701-711.
121. **Riess, T., S. G. Andersson, A. Lupas, M. Schaller, A. Schafer, P. Kyme, J. Martin, J. H. Walzlein, U. Eehalt, H. Lindroos, M. Schirle, A. Nordheim, I. B. Autenrieth, and V. A. Kempf.** 2004. *Bartonella* adhesin a mediates a proangiogenic host cell response. *J Exp Med* **200**:1267-1278.
122. **St Geme, J. W., 3rd, and D. Cutter.** 2000. The *Haemophilus influenzae* Hia adhesin is an autotransporter protein that remains uncleaved at the C terminus and fully cell associated. *J Bacteriol* **182**:6005-6013.
123. **Cotter, S. E., N. K. Surana, S. Grass, and J. W. St Geme, 3rd.** 2006. Trimeric autotransporters require trimerization of the passenger domain for stability and adhesive activity. *J Bacteriol* **188**:5400-5407.
124. **Meng, G., N. K. Surana, J. W. St Geme, 3rd, and G. Waksman.** 2006. Structure of the outer membrane translocator domain of the *Haemophilus influenzae* Hia trimeric autotransporter. *EMBO J* **25**:2297-2304.
125. **Szczesny, P., and A. Lupas.** 2008. Domain annotation of trimeric autotransporter adhesins--daTAA. *Bioinformatics* **24**:1251-1256.
126. **Koretke, K. K., P. Szczesny, M. Gruber, and A. N. Lupas.** 2006. Model structure of the prototypical non-fimbrial adhesin YadA of *Yersinia enterocolitica*. *J Struct Biol* **155**:154-161.
127. **Salacha, R., F. Kovacic, C. Brochier-Armanet, S. Wilhelm, J. Tommassen, A. Filloux, R. Voulhoux, and S. Bleves.** 2010. The *Pseudomonas aeruginosa* patatin-like protein PlpD is the archetype of a novel Type V secretion system. *Environ Microbiol* **12**:1498-1512.
128. **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**:e14403.
129. **Holland, I. B., L. Schmitt, and J. Young.** 2005. Type 1 protein secretion in bacteria, the ABC-transporter dependent pathway (review). *Mol Membr Biol* **22**:29-39.
130. **Young, J., and I. B. Holland.** 1999. ABC transporters: bacterial exporters-revisited five years on. *Biochim Biophys Acta* **1461**:177-200.
131. **Hinsa, S. M., M. Espinosa-Urgel, J. L. Ramos, and G. A. O'Toole.** 2003. Transition from reversible to irreversible attachment during biofilm formation by

- Pseudomonas fluorescens* WCS365 requires an ABC transporter and a large secreted protein. *Mol Microbiol* **49**:905-918.
132. **Holland, I. B., B. Kenny, B. Steipe, and A. Pluckthun.** 1990. Secretion of heterologous proteins in *Escherichia coli*. *Methods Enzymol* **182**:132-143.
 133. **Goebel, W., and J. Hedgpeth.** 1982. Cloning and functional characterization of the plasmid-encoded hemolysin determinant of *Escherichia coli*. *J Bacteriol* **151**:1290-1298.
 134. **Ladant, D., and A. Ullmann.** 1999. Bordetella pertussis adenylate cyclase: a toxin with multiple talents. *Trends Microbiol* **7**:172-176.
 135. **Kleanthous, C.** 2010. Swimming against the tide: progress and challenges in our understanding of colicin translocation. *Nat Rev Microbiol* **8**:843-848.
 136. **Koronakis, V., J. Eswaran, and C. Hughes.** 2004. Structure and function of TolC: the bacterial exit duct for proteins and drugs. *Annu Rev Biochem* **73**:467-489.
 137. **Chervaux, C., and I. B. Holland.** 1996. Random and directed mutagenesis to elucidate the functional importance of helix II and F-989 in the C-terminal secretion signal of *Escherichia coli* hemolysin. *J Bacteriol* **178**:1232-1236.
 138. **Wolff, N., G. Sapriel, C. Bodenreider, A. Chaffotte, and P. Delepelaire.** 2003. Antifolding activity of the SecB chaperone is essential for secretion of HasA, a quickly folding ABC pathway substrate. *J Biol Chem* **278**:38247-38253.
 139. **Koronakis, V., A. Sharff, E. Koronakis, B. Luisi, and C. Hughes.** 2000. Crystal structure of the bacterial membrane protein TolC central to multidrug efflux and protein export. *Nature* **405**:914-919.
 140. **Koronakis, V., C. Hughes, and E. Koronakis.** 1991. Energetically distinct early and late stages of HlyB/HlyD-dependent secretion across both *Escherichia coli* membranes. *EMBO J* **10**:3263-3272.
 141. **Cornelis, G. R.** 2006. The type III secretion injectisome. *Nat Rev Microbiol* **4**:811-825.
 142. **Gophna, U., E. Z. Ron, and D. Graur.** 2003. Bacterial type III secretion systems are ancient and evolved by multiple horizontal-transfer events. *Gene* **312**:151-163.
 143. **Galan, J. E.** 2009. Common themes in the design and function of bacterial effectors. *Cell Host Microbe* **5**:571-579.
 144. **Dean, P.** 2011. Functional domains and motifs of bacterial type III effector proteins and their roles in infection. *FEMS Microbiol Rev* **35**:1100-1125.
 145. **Enninga, J., and I. Rosenshine.** 2009. Imaging the assembly, structure and activity of type III secretion systems. *Cell Microbiol* **11**:1462-1470.
 146. **Blocker, A., N. Jouihri, E. Larquet, P. Gounon, F. Ebel, C. Parsot, P. Sansonetti, and A. Allaoui.** 2001. Structure and composition of the *Shigella flexneri* "needle complex", a part of its type III secreton. *Mol Microbiol* **39**:652-663.
 147. **Feldman, M. F., S. Muller, E. Wuest, and G. R. Cornelis.** 2002. SycE allows secretion of YopE-DHFR hybrids by the *Yersinia enterocolitica* type III Ysc system. *Mol Microbiol* **46**:1183-1197.
 148. **Blocker, A., P. Gounon, E. Larquet, K. Niebuhr, V. Cabiliaux, C. Parsot, and P. Sansonetti.** 1999. The tripartite type III secreton of *Shigella flexneri* inserts IpaB and IpaC into host membranes. *J Cell Biol* **147**:683-693.

149. **Veenendaal, A. K., J. L. Hodgkinson, L. Schwarzer, D. Stabat, S. F. Zenk, and A. J. Blocker.** 2007. The type III secretion system needle tip complex mediates host cell sensing and translocon insertion. *Mol Microbiol* **63**:1719-1730.
150. **Deane, J. E., P. Abrusci, S. Johnson, and S. M. Lea.** 2010. Timing is everything: the regulation of type III secretion. *Cell Mol Life Sci* **67**:1065-1075.
151. **Journet, L., C. Agrain, P. Broz, and G. R. Cornelis.** 2003. The needle length of bacterial injectisomes is determined by a molecular ruler. *Science* **302**:1757-1760.
152. **Kubori, T., A. Sukhan, S. I. Aizawa, and J. E. Galan.** 2000. Molecular characterization and assembly of the needle complex of the *Salmonella typhimurium* type III protein secretion system. *Proc Natl Acad Sci U S A* **97**:10225-10230.
153. **Mota, L. J., L. Journet, I. Sorg, C. Agrain, and G. R. Cornelis.** 2005. Bacterial injectisomes: needle length does matter. *Science* **307**:1278.
154. **Parsot, C., C. Hamiaux, and A. L. Page.** 2003. The various and varying roles of specific chaperones in type III secretion systems. *Curr Opin Microbiol* **6**:7-14.
155. **Galan, J. E., and H. Wolf-Watz.** 2006. Protein delivery into eukaryotic cells by type III secretion machines. *Nature* **444**:567-573.
156. **Anderson, D. M., and O. Schneewind.** 1997. A mRNA signal for the type III secretion of Yop proteins by *Yersinia enterocolitica*. *Science* **278**:1140-1143.
157. **Niemann, G. S., R. N. Brown, I. T. Mushamiri, N. T. Nguyen, R. Taiwo, A. Stufkens, R. D. Smith, J. N. Adkins, J. E. McDermott, and F. Heffron.** 2013. RNA Type III Secretion Signals That Require Hfq. *J Bacteriol* **195**:2119-2125.
158. **Alvarez-Martinez, C. E., and P. J. Christie.** 2009. Biological diversity of prokaryotic type IV secretion systems. *Microbiol Mol Biol Rev* **73**:775-808.
159. **Hofreuter, D., S. Odenbreit, and R. Haas.** 2001. Natural transformation competence in *Helicobacter pylori* is mediated by the basic components of a type IV secretion system. *Mol Microbiol* **41**:379-391.
160. **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**:1704-1721.
161. **Locht, C., L. Coutte, and N. Mielcarek.** 2011. The ins and outs of pertussis toxin. *FEBS J* **278**:4668-4682.
162. **Zechner, E. L., S. Lang, and J. F. Schildbach.** 2012. Assembly and mechanisms of bacterial type IV secretion machines. *Philos Trans R Soc Lond B Biol Sci* **367**:1073-1087.
163. **Fronzes, R., E. Schafer, L. Wang, H. R. Saibil, E. V. Orlova, and G. Waksman.** 2009. Structure of a type IV secretion system core complex. *Science* **323**:266-268.
164. **Chandran, V., R. Fronzes, S. Duquerroy, N. Cronin, J. Navaza, and G. Waksman.** 2009. Structure of the outer membrane complex of a type IV secretion system. *Nature* **462**:1011-1015.
165. **Lai, E. M., and C. I. Kado.** 1998. Processed VirB2 is the major subunit of the promiscuous pilus of *Agrobacterium tumefaciens*. *J Bacteriol* **180**:2711-2717.
166. **Schmidt-Eisenlohr, H., N. Domke, C. Angerer, G. Wanner, P. C. Zambryski, and C. Baron.** 1999. Vir proteins stabilize VirB5 and mediate its association with the T pilus of *Agrobacterium tumefaciens*. *J Bacteriol* **181**:7485-7492.
167. **Aly, K. A., and C. Baron.** 2007. The VirB5 protein localizes to the T-pilus tips in *Agrobacterium tumefaciens*. *Microbiology* **153**:3766-3775.

168. **de Jong, M. F., Y. H. Sun, A. B. den Hartigh, J. M. van Dijl, and R. M. Tsolis.** 2008. Identification of VceA and VceC, two members of the VjbR regulon that are translocated into macrophages by the Brucella type IV secretion system. *Mol Microbiol* **70**:1378-1396.
169. **Nagai, H., E. D. Cambronne, J. C. Kagan, J. C. Amor, R. A. Kahn, and C. R. Roy.** 2005. A C-terminal translocation signal required for Dot/Icm-dependent delivery of the Legionella RalF protein to host cells. *Proc Natl Acad Sci U S A* **102**:826-831.
170. **Schulein, R., P. Guye, T. A. Rhomberg, M. C. Schmid, G. Schroder, A. C. Vergunst, I. Carena, and C. Dehio.** 2005. A bipartite signal mediates the transfer of type IV secretion substrates of Bartonella henselae into human cells. *Proc Natl Acad Sci U S A* **102**:856-861.
171. **Hohlfeld, S., I. Pattis, J. Puls, G. V. Plano, R. Haas, and W. Fischer.** 2006. A C-terminal translocation signal is necessary, but not sufficient for type IV secretion of the Helicobacter pylori CagA protein. *Mol Microbiol* **59**:1624-1637.
172. **Ding, Z., K. Atmakuri, and P. J. Christie.** 2003. The outs and ins of bacterial type IV secretion substrates. *Trends Microbiol* **11**:527-535.
173. **Vincent, C. D., and J. P. Vogel.** 2006. The Legionella pneumophila IcmS-LvgA protein complex is important for Dot/Icm-dependent intracellular growth. *Mol Microbiol* **61**:596-613.
174. **Ninio, S., and C. R. Roy.** 2007. Effector proteins translocated by Legionella pneumophila: strength in numbers. *Trends Microbiol* **15**:372-380.
175. **Pukatzki, S., A. T. Ma, D. Sturtevant, B. Krastins, D. Sarracino, W. C. Nelson, J. F. Heidelberg, and J. J. Mekalanos.** 2006. Identification of a conserved bacterial protein secretion system in Vibrio cholerae using the Dictyostelium host model system. *Proc Natl Acad Sci U S A* **103**:1528-1533.
176. **Silverman, J. M., Y. R. Brunet, E. Cascales, and J. D. Mougous.** 2012. Structure and regulation of the type VI secretion system. *Annu Rev Microbiol* **66**:453-472.
177. **Cascales, E.** 2008. The type VI secretion toolkit. *EMBO Rep* **9**:735-741.
178. **Boyer, F., G. Fichant, J. Berthod, Y. Vandenbrouck, and I. Attree.** 2009. Dissecting the bacterial type VI secretion system by a genome wide in silico analysis: what can be learned from available microbial genomic resources? *BMC Genomics* **10**:104.
179. **Kanamaru, S.** 2009. Structural similarity of tailed phages and pathogenic bacterial secretion systems. *Proc Natl Acad Sci U S A* **106**:4067-4068.
180. **Mougous, J. D., M. E. Cuff, S. Raunser, A. Shen, M. Zhou, C. A. Gifford, A. L. Goodman, G. Joachimiak, C. L. Ordonez, S. Lory, T. Walz, A. Joachimiak, and J. J. Mekalanos.** 2006. A virulence locus of Pseudomonas aeruginosa encodes a protein secretion apparatus. *Science* **312**:1526-1530.
181. **Ballister, E. R., A. H. Lai, R. N. Zuckermann, Y. Cheng, and J. D. Mougous.** 2008. In vitro self-assembly of tailorable nanotubes from a simple protein building block. *Proc Natl Acad Sci U S A* **105**:3733-3738.
182. **Pukatzki, S., A. T. Ma, A. T. Revel, D. Sturtevant, and J. J. Mekalanos.** 2007. Type VI secretion system translocates a phage tail spike-like protein into target cells where it cross-links actin. *Proc Natl Acad Sci U S A* **104**:15508-15513.

183. **Sexton, J. A., J. L. Miller, A. Yoneda, T. E. Kehl-Fie, and J. P. Vogel.** 2004. *Legionella pneumophila* DotU and IcmF are required for stability of the Dot/Icm complex. *Infect Immun* **72**:5983-5992.
184. **Bonemann, G., A. Pietrosiuk, A. Diemand, H. Zentgraf, and A. Mogk.** 2009. Remodelling of VipA/VipB tubules by ClpV-mediated threading is crucial for type VI protein secretion. *EMBO J* **28**:315-325.
185. **Basler, M., M. Pilhofer, G. P. Henderson, G. J. Jensen, and J. J. Mekalanos.** 2012. Type VI secretion requires a dynamic contractile phage tail-like structure. *Nature* **483**:182-186.
186. **Ma, A. T., S. McAuley, S. Pukatzki, and J. J. Mekalanos.** 2009. Translocation of a *Vibrio cholerae* type VI secretion effector requires bacterial endocytosis by host cells. *Cell Host Microbe* **5**:234-243.
187. **Russell, A. B., R. D. Hood, N. K. Bui, M. LeRoux, W. Vollmer, and J. D. Mougous.** 2011. Type VI secretion delivers bacteriolytic effectors to target cells. *Nature* **475**:343-347.
188. **Russell, A. B., P. Singh, M. Brittnacher, N. K. Bui, R. D. Hood, M. A. Carl, D. M. Agnello, S. Schwarz, D. R. Goodlett, W. Vollmer, and J. D. Mougous.** 2012. A widespread bacterial type VI secretion effector superfamily identified using a heuristic approach. *Cell Host Microbe* **11**:538-549.
189. **Bonemann, G., A. Pietrosiuk, and A. Mogk.** 2010. Tubules and donuts: a type VI secretion story. *Mol Microbiol* **76**:815-821.
190. **Henderson, I. R., F. Navarro-Garcia, and J. P. Nataro.** 1998. The great escape: structure and function of the autotransporter proteins. *Trends Microbiol* **6**:370-378.
191. **Bernstein, H. D.** 2007. Are bacterial 'autotransporters' really transporters? *Trends Microbiol* **15**:441-447.
192. **Kuhnel, K., and D. Diezmann.** 2011. Crystal structure of the autochaperone region from the *Shigella flexneri* autotransporter IcsA. *J Bacteriol* **193**:2042-2045.
193. **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.
194. **Peterson, J. H., C. A. Woolhead, and H. D. Bernstein.** 2003. Basic amino acids in a distinct subset of signal peptides promote interaction with the signal recognition particle. *J Biol Chem* **278**:46155-46162.
195. **Sijbrandi, R., M. L. Urbanus, C. M. ten Hagen-Jongman, H. D. Bernstein, B. Oudega, B. R. Otto, and J. Luirink.** 2003. Signal recognition particle (SRP)-mediated targeting and Sec-dependent translocation of an extracellular *Escherichia coli* protein. *J Biol Chem* **278**:4654-4659.
196. **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**:9038-9048.
197. **Lambert-Buisine, C., E. Willery, C. Locht, and F. Jacob-Dubuisson.** 1998. N-terminal characterization of the *Bordetella pertussis* filamentous haemagglutinin. *Mol Microbiol* **28**:1283-1293.
198. **Brandon, L. D., N. Goehring, A. Janakiraman, A. W. Yan, T. Wu, J. Beckwith, and M. B. Goldberg.** 2003. IcsA, a polarly localized autotransporter with an

- atypical signal peptide, uses the Sec apparatus for secretion, although the Sec apparatus is circumferentially distributed. *Mol Microbiol* **50**:45-60.
199. **Leyton, D. L., M. D. de Luna, Y. R. Sevastyanovich, K. Tveen Jensen, D. F. Browning, A. Scott-Tucker, and I. R. Henderson.** 2010. The unusual extended signal peptide region is not required for secretion and function of an *Escherichia coli* autotransporter. *FEMS Microbiol Lett.*
 200. **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**:221-226.
 201. **Jong, W. S., and J. Luirink.** 2008. The conserved extension of the Hbp autotransporter signal peptide does not determine targeting pathway specificity. *Biochem Biophys Res Commun* **368**:522-527.
 202. **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**:39682-39690.
 203. **van Ulsen, P., L. van Alphen, J. ten Hove, F. Franssen, P. van der Ley, and J. Tommassen.** 2003. A Neisserial autotransporter NalP modulating the processing of other autotransporters. *Mol Microbiol* **50**:1017-1030.
 204. **Coutte, L., R. Antoine, H. Drobecq, C. Locht, and F. Jacob-Dubuisson.** 2001. Subtilisin-like autotransporter serves as maturation protease in a bacterial secretion pathway. *EMBO J* **20**:5040-5048.
 205. **Ashgar, S. S., N. J. Oldfield, K. G. Wooldridge, M. A. Jones, G. J. Irving, D. P. Turner, and D. A. Ala'Aldeen.** 2007. CapA, an autotransporter protein of *Campylobacter jejuni*, mediates association with human epithelial cells and colonization of the chicken gut. *J Bacteriol* **189**:1856-1865.
 206. **Coutte, L., E. Willery, R. Antoine, H. Drobecq, C. Locht, and F. Jacob-Dubuisson.** 2003. Surface anchoring of bacterial subtilisin important for maturation function. *Mol Microbiol* **49**:529-539.
 207. **Henderson, I. R., and J. P. Nataro.** 2001. Virulence functions of autotransporter proteins. *Infect Immun* **69**:1231-1243.
 208. **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**:e43245.
 209. **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**:4918-4923.
 210. **Emsley, P., I. G. Charles, N. F. Fairweather, and N. W. Isaacs.** 1996. Structure of *Bordetella pertussis* virulence factor P.69 pertactin. *Nature* **381**:90-92.
 211. **Otto, B. R., R. Sijbrandi, J. Luirink, B. Oudega, J. G. Heddle, K. Mizutani, S. Y. Park, and J. R. Tame.** 2005. Crystal structure of hemoglobin protease, a heme binding autotransporter protein from pathogenic *Escherichia coli*. *J Biol Chem* **280**:17339-17345.
 212. **Khan, S., H. S. Mian, L. E. Sandercock, N. Y. Chirgadze, and E. F. Pai.** 2011. Crystal structure of the passenger domain of the *Escherichia coli* autotransporter EspP. *J Mol Biol* **413**:985-1000.

213. **Johnson, T. A., J. Qiu, A. G. Plaut, and T. Holyoak.** 2009. Active-site gating regulates substrate selectivity in a chymotrypsin-like serine protease the structure of haemophilus influenzae immunoglobulin A1 protease. *J Mol Biol* **389**:559-574.
214. **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**:3864-3874.
215. **Gangwer, K. A., D. J. Mushrush, D. L. Stauff, B. Spiller, M. S. McClain, T. L. Cover, and D. B. Lacy.** 2007. Crystal structure of the Helicobacter pylori vacuolating toxin p55 domain. *Proc Natl Acad Sci U S A* **104**:16293-16298.
216. **van den Berg, B.** 2010. Crystal structure of a full-length autotransporter. *J Mol Biol* **396**:627-633.
217. **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**:1367-1383.
218. **Peterson, J. H., P. Tian, R. Ieva, N. Dautin, and H. D. Bernstein.** 2010. Secretion of a bacterial virulence factor is driven by the folding of a C-terminal segment. *Proc Natl Acad Sci U S A*.
219. **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**:19-34.
220. **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**:39912-39920.
221. **May, K. L., and R. Morona.** 2008. Mutagenesis of the Shigella flexneri autotransporter IcsA reveals novel functional regions involved in IcsA biogenesis and recruitment of host neural Wiscott-Aldrich syndrome protein. *J Bacteriol* **190**:4666-4676.
222. **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**:1257-1266.
223. **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**:1214-1220.
224. **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**:645-56.
225. **Zhai, Y., K. Zhang, Y. Huo, Y. Zhu, Q. Zhou, J. Lu, I. Black, X. Pang, A. W. Roszak, X. Zhang, N. W. Isaacs, and F. Sun.** 2011. Autotransporter passenger domain secretion requires a hydrophobic cavity at the extracellular entrance of the beta-domain pore. *Biochem J* **435**:577-587.
226. **Khalid, S., and M. S. Sansom.** 2006. Molecular dynamics simulations of a bacterial autotransporter: NalP from Neisseria meningitidis. *Mol Membr Biol* **23**:499-508.
227. **Purdy, G. E., C. R. Fisher, and S. M. Payne.** 2007. IcsA surface presentation in Shigella flexneri requires the periplasmic chaperones DegP, Skp, and SurA. *J Bacteriol* **189**:5566-5573.

228. **Ruiz-Perez, F., I. R. Henderson, D. L. Leyton, A. E. Rossiter, Y. Zhang, and J. P. Nataro.** 2009. Roles of periplasmic chaperone proteins in the biogenesis of serine protease autotransporters of Enterobacteriaceae. *J Bacteriol* **191**:6571-6583.
229. **Ieva, R., and H. D. Bernstein.** 2009. Interaction of an autotransporter passenger domain with BamA during its translocation across the bacterial outer membrane. *Proc Natl Acad Sci U S A* **106**:19120-19125.
230. **Sauri, A., Z. Soprova, D. Wickstrom, J. W. de Gier, R. C. Van der Schors, A. B. Smit, W. S. Jong, and J. Luirink.** 2009. The Bam (Omp85) complex is involved in secretion of the autotransporter haemoglobin protease. *Microbiology* **155**:3982-3991.
231. **Ieva, R., P. Tian, J. H. Peterson, and H. D. Bernstein.** 2011. Sequential and spatially restricted interactions of assembly factors with an autotransporter beta domain. *Proc Natl Acad Sci U S A* **108**:E383-391.
232. **Mogensen, J. E., J. H. Kleinschmidt, M. A. Schmidt, and D. E. Otzen.** 2005. Misfolding of a bacterial autotransporter. *Protein Sci* **14**:2814-2827.
233. **Volokhina, E. B., J. Grijpstra, M. Stork, I. Schilders, J. Tommassen, and M. P. Bos.** 2011. Role of the periplasmic chaperones Skp, SurA, and DegQ in outer membrane protein biogenesis in *Neisseria meningitidis*. *J Bacteriol* **193**:1612-1621.
234. **Jain, S., and M. B. Goldberg.** 2007. Requirement for YaeT in the outer membrane assembly of autotransporter proteins. *J Bacteriol* **189**:5393-5398.
235. **Rossiter, A. E., D. L. Leyton, K. Tveen-Jensen, D. F. Browning, Y. Sevastyanovich, 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**:4250-4253.
236. **Renn, J. P., and P. L. Clark.** 2008. A conserved stable core structure in the passenger domain beta-helix of autotransporter virulence proteins. *Biopolymers* **89**:420-427.
237. **Junker, M., R. N. Besingi, and P. L. Clark.** 2009. Vectorial transport and folding of an autotransporter virulence protein during outer membrane secretion. *Mol Microbiol* **71**:1323-1332.
238. **Marin, E., G. Bodelon, and L. A. Fernandez.** 2010. Comparative analysis of the biochemical and functional properties of C-terminal domains of autotransporters. *J Bacteriol* **192**:5588-5602.
239. **Roussel-Jazede, V., P. Van Gelder, R. Sijbrandi, L. Rutten, B. R. Otto, J. Luirink, P. Gros, J. Tommassen, and P. Van Ulsen.** 2011. Channel properties of the translocator domain of the autotransporter Hbp of *Escherichia coli*. *Mol Membr Biol* **28**:158-170.
240. **Ieva, R., K. M. Skillman, and H. D. Bernstein.** 2008. Incorporation of a polypeptide segment into the beta-domain pore during the assembly of a bacterial autotransporter. *Mol Microbiol* **67**:188-201.
241. **Sauri, A., C. M. Ten Hagen-Jongman, P. van Ulsen, and J. Luirink.** 2012. Estimating the size of the active translocation pore of an autotransporter. *J Mol Biol* **416**:335-345.
242. **Leyton, D. L., Y. R. Sevastyanovich, D. F. Browning, A. E. Rossiter, T. J. Wells, R. E. Fitzpatrick, M. Overduin, A. F. Cunningham, and I. R.**

- Henderson.** 2011. Size and conformation limits to secretion of disulfide-bonded loops in autotransporter proteins. *J Biol Chem* **286**:42283-42291.
243. **Skillman, K. M., T. J. Barnard, J. H. Peterson, R. Ghirlando, and H. D. Bernstein.** 2005. Efficient secretion of a folded protein domain by a monomeric bacterial autotransporter. *Mol Microbiol* **58**:945-958.
244. **Pavlova, O., J. H. Peterson, R. Ieva, and H. D. Bernstein.** 2013. Mechanistic link between beta barrel assembly and the initiation of autotransporter secretion. *Proc Natl Acad Sci U S A* **110**:E938-947.
245. **Phan, G., H. Remaut, T. Wang, W. J. Allen, K. F. Pirker, A. Lebedev, N. S. Henderson, S. Geibel, E. Volkan, J. Yan, M. B. Kunze, J. S. Pinkner, B. Ford, C. W. Kay, H. Li, S. J. Hultgren, D. G. Thanassi, and G. Waksman.** 2011. Crystal structure of the FimD usher bound to its cognate FimC-FimH substrate. *Nature* **474**:49-53.
246. **Renn, J. P., and P. L. Clark.** 2011. Disulfide bond-mediated passenger domain stalling as a structural probe of autotransporter outer membrane secretion in vivo. *Methods Enzymol* **492**:233-251.
247. **Renn, J. P., M. Junker, R. N. Besingi, E. Braselmann, and P. L. Clark.** 2012. ATP-independent control of autotransporter virulence protein transport via the folding properties of the secreted protein. *Chem Biol* **19**:287-296.
248. **Braselmann, E., and P. L. Clark.** 2012. Autotransporters: The Cellular Environment Reshapes a Folding Mechanism to Promote Protein Transport. *J Phys Chem Lett* **3**:1063-1071.
249. **Selkrig, J., K. Mosbahi, C. T. Webb, M. J. Belousoff, A. J. Perry, T. J. Wells, F. Morris, D. L. Leyton, M. Totsika, M. D. Phan, N. Celik, M. Kelly, C. Oates, E. L. Hartland, R. M. Robins-Browne, S. H. Ramarathinam, A. W. Purcell, M. A. Schembri, R. A. Strugnell, I. R. Henderson, D. Walker, and T. Lithgow.** 2012. Discovery of an archetypal protein transport system in bacterial outer membranes. *Nat Struct Mol Biol* **19**:506-510, S501.
250. **Cover, T. L., and S. R. Blanke.** 2005. *Helicobacter pylori* VacA, a paradigm for toxin multifunctionality. *Nat Rev Microbiol* **3**:320-332.
251. **Navarro-Garcia, F.** 2010. Enteroaggregative *Escherichia coli* plasmid-encoded toxin. *Future Microbiol* **5**:1005-1013.
252. **Charbonneau, M. E., F. Berthiaume, and M. Mourez.** 2006. Proteolytic processing is not essential for multiple functions of the *Escherichia coli* autotransporter adhesin involved in diffuse adherence (AIDA-I). *J Bacteriol* **188**:8504-8512.
253. **Shere, K. D., S. Sallustio, A. Manassis, T. G. D'Aversa, and M. B. Goldberg.** 1997. Disruption of IcsP, the major *Shigella* protease that cleaves IcsA, accelerates actin-based motility. *Mol Microbiol* **25**:451-462.
254. **Turner, D. P., A. G. Marietou, L. Johnston, K. K. Ho, A. J. Rogers, K. G. Wooldridge, and D. A. Ala'Aldeen.** 2006. Characterization of MspA, an immunogenic autotransporter protein that mediates adhesion to epithelial and endothelial cells in *Neisseria meningitidis*. *Infect Immun* **74**:2957-2964.
255. **van Ulsen, P., B. Adler, P. Fassler, M. Gilbert, M. van Schilfhaarde, P. van der Ley, L. van Alphen, and J. Tommassen.** 2006. A novel phase-variable autotransporter serine protease, AusI, of *Neisseria meningitidis*. *Microbes Infect* **8**:2088-2097.

256. **Fink, D. L., L. D. Cope, E. J. Hansen, and J. W. Geme, 3rd.** 2001. The Hemophilus influenzae Hap autotransporter is a chymotrypsin clan serine protease and undergoes autoproteolysis via an intermolecular mechanism. *J Biol Chem* **276**:39492-39500.
257. **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**:1942-1952.
258. **Charbonneau, M. E., J. Janvore, and M. Mourez.** 2009. Autoprocessing of the Escherichia coli AIDA-I autotransporter: a new mechanism involving acidic residues in the junction region. *J Biol Chem* **284**:17340-17351.
259. **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**:2132-2141.
260. **Nguyen, V. Q., R. M. Caprioli, and T. L. Cover.** 2001. Carboxy-terminal proteolytic processing of Helicobacter pylori vacuolating toxin. *Infect Immun* **69**:543-546.
261. **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**:1403-1413.
262. **Lindenthal, C., and E. A. Elsinghorst.** 1999. Identification of a glycoprotein produced by enterotoxigenic Escherichia coli. *Infect Immun* **67**:4084-4091.
263. **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**:389-401.
264. **Easton, D. M., M. Totsika, L. P. Allsopp, M. D. Phan, A. Idris, D. J. Wurple, 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.
265. **Raetz, C. R., and C. Whitfield.** 2002. Lipopolysaccharide endotoxins. *Annu Rev Biochem* **71**:635-700.
266. **Whitfield, C.** 2006. Biosynthesis and assembly of capsular polysaccharides in Escherichia coli. *Annu Rev Biochem* **75**:39-68.
267. **Benz, I., and M. A. Schmidt.** 2002. Never say never again: protein glycosylation in pathogenic bacteria. *Mol Microbiol* **45**:267-276.
268. **Drickamer, K., and M. E. Taylor.** 1998. Evolving views of protein glycosylation. *Trends Biochem Sci* **23**:321-324.
269. **Calo, D., L. Kaminski, and J. Eichler.** 2010. Protein glycosylation in Archaea: sweet and extreme. *Glycobiology* **20**:1065-1076.
270. **Szymanski, C. M., and B. W. Wren.** 2005. Protein glycosylation in bacterial mucosal pathogens. *Nat Rev Microbiol* **3**:225-237.
271. **Nothhaft, H., and C. M. Szymanski.** 2010. Protein glycosylation in bacteria: sweeter than ever. *Nat Rev Microbiol* **8**:765-778.
272. **Szymanski, C. M., S. M. Logan, D. Linton, and B. W. Wren.** 2003. Campylobacter--a tale of two protein glycosylation systems. *Trends Microbiol* **11**:233-238.

273. **Scott, N. E., B. L. Parker, A. M. Connolly, J. Paulech, A. V. Edwards, B. Crossett, L. Falconer, D. Kolarich, S. P. Djordjevic, P. Hojrup, N. H. Packer, M. R. Larsen, and S. J. Cordwell.** 2011. Simultaneous glycan-peptide characterization using hydrophilic interaction chromatography and parallel fragmentation by CID, higher energy collisional dissociation, and electron transfer dissociation MS applied to the N-linked glycoproteome of *Campylobacter jejuni*. *Mol Cell Proteomics* **10**:M000031-MCP000201.
274. **Logan, S. M.** 2006. Flagellar glycosylation - a new component of the motility repertoire? *Microbiology* **152**:1249-1262.
275. **Grass, S., C. F. Lichti, R. R. Townsend, J. Gross, and J. W. St Geme, 3rd.** 2010. The *Haemophilus influenzae* HMW1C protein is a glycosyltransferase that transfers hexose residues to asparagine sites in the HMW1 adhesin. *PLoS Pathog* **6**:e1000919.
276. **Charbonneau, M. E., J. P. Cote, M. F. Haurat, B. Reiz, S. Crepin, F. Berthiaume, C. M. Dozois, M. F. Feldman, and M. Mourez.** 2012. A structural motif is the recognition site for a new family of bacterial protein O-glycosyltransferases. *Mol Microbiol* **83**:894-907.
277. **Moormann, C., I. Benz, and M. A. Schmidt.** 2002. Functional substitution of the TibC protein of enterotoxigenic *Escherichia coli* strains for the autotransporter adhesin heptosyltransferase of the AIDA system. *Infect Immun* **70**:2264-2270.
278. **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**:1798-1807.
279. **Charbonneau, M. E., V. Girard, A. Nikolakakis, M. Campos, F. Berthiaume, F. Dumas, F. Lepine, and M. Mourez.** 2007. O-linked glycosylation ensures the normal conformation of the autotransporter adhesin involved in diffuse adherence. *J Bacteriol* **189**:8880-8889.
280. **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**:563-577.
281. **Otzen, D.** 2012. N for AsN - O for StrOcture? A strand-loop-strand motif for prokaryotic O-glycosylation. *Mol Microbiol* **83**:879-883.
282. **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**:1954-1963.
283. **Tang, G., T. Ruiz, and K. P. Mintz.** 2012. O-polysaccharide glycosylation is required for stability and function of the collagen adhesin EmaA of *Aggregatibacter actinomycetemcomitans*. *Infect Immun* **80**:2868-2877.
284. **Goldberg, M. B., and J. A. Theriot.** 1995. *Shigella flexneri* surface protein IcsA is sufficient to direct actin-based motility. *Proc Natl Acad Sci U S A* **92**:6572-6576.
285. **Goldberg, M. B., O. Barzu, C. Parsot, and P. J. Sansonetti.** 1993. Unipolar localization and ATPase activity of IcsA, a *Shigella flexneri* protein involved in intracellular movement. *J Bacteriol* **175**:2189-2196.
286. **Charles, M., M. Perez, J. H. Kobil, and M. B. Goldberg.** 2001. Polar targeting of *Shigella* virulence factor IcsA in *Enterobacteriaceae* and *Vibrio*. *Proc Natl Acad Sci U S A* **98**:9871-9876.

287. **Janakiraman, A., K. R. Fixen, A. N. Gray, H. Niki, and M. B. Goldberg.** 2009. A genome-scale proteomic screen identifies a role for DnaK in chaperoning of polar autotransporters in *Shigella*. *J Bacteriol* **191**:6300-6311.
288. **Fixen, K. R., A. Janakiraman, S. Garrity, D. J. Slade, A. N. Gray, N. Karahan, A. Hochschild, and M. B. Goldberg.** 2012. Genetic reporter system for positioning of proteins at the bacterial pole. *MBio* **3**.
289. **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**:4841-4850.
290. **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**:1506-1511.
291. **Sherlock, O., M. A. Schembri, A. Reisner, and P. Klemm.** 2004. Novel roles for the AIDA adhesin from diarrheagenic *Escherichia coli*: cell aggregation and biofilm formation. *J Bacteriol* **186**:8058-8065.
292. **Barnes, M. G., and A. A. Weiss.** 2001. BrkA protein of *Bordetella pertussis* inhibits the classical pathway of complement after C1 deposition. *Infect Immun* **69**:3067-3072.
293. **Lomholt, H., K. Poulsen, and M. Kilian.** 1995. Comparative characterization of the iga gene encoding IgA1 protease in *Neisseria meningitidis*, *Neisseria gonorrhoeae* and *Haemophilus influenzae*. *Mol Microbiol* **15**:495-506.
294. **Yen, Y. T., M. Kostakioti, I. R. Henderson, and C. Stathopoulos.** 2008. Common themes and variations in serine protease autotransporters. *Trends Microbiol* **16**:370-379.
295. **Dautin, N.** 2010. Serine protease autotransporters of enterobacteriaceae (SPATEs): biogenesis and function. *Toxins (Basel)* **2**:1179-1206.
296. **Ruiz-Perez, F., and J. P. Nataro.** 2013. Bacterial serine proteases secreted by the autotransporter pathway: classification, specificity, and role in virulence. *Cell Mol Life Sci*.
297. **Weiss, A., and J. Brockmeyer.** 2013. Prevalence, biogenesis, and functionality of the serine protease autotransporter EspP. *Toxins (Basel)* **5**:25-48.
298. **Hoopman, T. C., W. Wang, C. A. Brautigam, J. L. Sedillo, T. J. Reilly, and E. J. Hansen.** 2008. *Moraxella catarrhalis* synthesizes an autotransporter that is an acid phosphatase. *J Bacteriol* **190**:1459-1472.
299. **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**:951-966.
300. **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**:589-604.
301. **Leininger, E., M. Roberts, J. G. Kenimer, I. G. Charles, N. Fairweather, P. Novotny, and M. J. Brennan.** 1991. Pertactin, an Arg-Gly-Asp-containing *Bordetella pertussis* surface protein that promotes adherence of mammalian cells. *Proc Natl Acad Sci U S A* **88**:345-349.

302. **Bokhari, H., I. Bilal, and S. Zafar.** 2012. BapC autotransporter protein of *Bordetella pertussis* is an adhesion factor. *Journal of basic microbiology* **52**:390-396.
303. **Wells, T. J., M. Totsika, and M. A. Schembri.** 2010. Autotransporters of *Escherichia coli*: a sequence-based characterization. *Microbiology* **156**:2459-2469.
304. **Girard, V., and M. Mourez.** 2006. Adhesion mediated by autotransporters of Gram-negative bacteria: structural and functional features. *Res Microbiol* **157**:407-416.
305. **Pizarro-Cerda, J., and P. Cossart.** 2006. Bacterial adhesion and entry into host cells. *Cell* **124**:715-727.
306. **Barnhart, M. M., and M. R. Chapman.** 2006. Curli biogenesis and function. *Annu Rev Microbiol* **60**:131-147.
307. **Cotter, S. E., N. K. Surana, and J. W. St Geme, 3rd.** 2005. Trimeric autotransporters: a distinct subfamily of autotransporter proteins. *Trends Microbiol* **13**:199-205.
308. **Poolman, J. T., and H. O. Hallander.** 2007. Acellular pertussis vaccines and the role of pertactin and fimbriae. *Expert Rev Vaccines* **6**:47-56.
309. **Hannan, T. J., M. Totsika, K. J. Mansfield, K. H. Moore, M. A. Schembri, and S. J. Hultgren.** 2012. Host-pathogen checkpoints and population bottlenecks in persistent and intracellular uropathogenic *Escherichia coli* bladder infection. *FEMS Microbiol Rev* **36**:616-648.
310. **Campellone, K. G., and J. M. Leong.** 2003. Tails of two Tirs: actin pedestal formation by enteropathogenic *E. coli* and enterohemorrhagic *E. coli* O157:H7. *Curr Opin Microbiol* **6**:82-90.
311. **Sokurenko, E. V., V. Chesnokova, D. E. Dykhuizen, I. Ofek, X. R. Wu, K. A. Krogfelt, C. Struve, M. A. Schembri, and D. L. Hasty.** 1998. Pathogenic adaptation of *Escherichia coli* by natural variation of the FimH adhesin. *Proc Natl Acad Sci U S A* **95**:8922-8926.
312. **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**:105-107.
313. **Charbonneau, M. E., and M. Mourez.** 2007. Functional organization of the autotransporter adhesin involved in diffuse adherence. *J Bacteriol* **189**:9020-9029.
314. **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**:2114-2118.
315. **Czczulin, 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**:4135-4145.
316. **Smukalla, S., M. Caldara, N. Pochet, A. Beauvais, S. Guadagnini, C. Yan, M. D. Vinces, A. Jansen, M. C. Prevost, J. P. Latge, G. R. Fink, K. R. Foster, and K. J. Verstrepen.** 2008. FLO1 is a variable green beard gene that drives biofilm-like cooperation in budding yeast. *Cell* **135**:726-737.
317. **Menozi, F. D., P. E. Boucher, G. Riveau, C. Gantiez, and C. Locht.** 1994. Surface-associated filamentous hemagglutinin induces autoagglutination of *Bordetella pertussis*. *Infect Immun* **62**:4261-4269.

318. **Flemming, H. C., and J. Wingender.** 2010. The biofilm matrix. *Nat Rev Microbiol* **8**:623-633.
319. **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**:285-293.
320. **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**:822-880.
321. **Kaper, J. B., J. P. Nataro, and H. L. Mobley.** 2004. Pathogenic *Escherichia coli*. *Nat Rev Microbiol* **2**:123-140.
322. **Zhao, L., X. Chen, X. Xu, G. Song, and X. Liu.** 2009. Analysis of the AIDA-I gene sequence and prevalence in *Escherichia coli* isolates from pigs with post-weaning diarrhoea and oedema disease. *Vet J* **180**:124-129.
323. **Jallat, C., V. Livrelli, A. Darfeuille-Michaud, C. Rich, and B. Joly.** 1993. *Escherichia coli* strains involved in diarrhea in France: high prevalence and heterogeneity of diffusely adhering strains. *J Clin Microbiol* **31**:2031-2037.
324. **Jallat, C., A. Darfeuille-Michaud, C. Rich, and B. Joly.** 1994. Survey of clinical isolates of diarrhoeogenic *Escherichia coli*: diffusely adhering *E. coli* strains with multiple adhesive factors. *Res Microbiol* **145**:621-632.
325. **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**:143-149.
326. **Mainil, J. G., E. Jacquemin, P. Pohl, A. Kaeckenbeeck, and I. Benz.** 2002. DNA sequences coding for the F18 fimbriae and AIDA adhesin are localised on the same plasmid in *Escherichia coli* isolates from piglets. *Vet Microbiol* **86**:303-311.
327. **Ngeleka, M., J. Pritchard, G. Appleyard, D. M. Middleton, and J. M. Fairbrother.** 2003. Isolation and association of *Escherichia coli* AIDA-I/STb, rather than EAST1 pathotype, with diarrhea in piglets and antibiotic sensitivity of isolates. *J Vet Diagn Invest* **15**:242-252.
328. **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**:378-381.
329. **Zhang, W., M. Zhao, L. Ruesch, A. Omot, and D. Francis.** 2007. Prevalence of virulence genes in *Escherichia coli* strains recently isolated from young pigs with diarrhea in the US. *Vet Microbiol* **123**:145-152.
330. **Dubreuil, J. D.** 2010. STb and AIDA-I: the missing link? *Crit Rev Microbiol* **36**:212-220.
331. **Ravi, M., M. Ngeleka, S. H. Kim, C. Gyles, F. Berthiaume, M. Mourez, D. Middleton, and E. Simko.** 2007. Contribution of AIDA-I to the pathogenicity of a porcine diarrheagenic *Escherichia coli* and to intestinal colonization through biofilm formation in pigs. *Vet Microbiol* **120**:308-319.
332. **Laarmann, S., and M. A. Schmidt.** 2003. The *Escherichia coli* AIDA autotransporter adhesin recognizes an integral membrane glycoprotein as receptor. *Microbiology* **149**:1871-1882.

333. **Fang, Y., M. Ngeleka, D. M. Middleton, and E. Simko.** 2008. Isolation and identification of AIDA-I receptors in porcine intestinal mucus. *Vet Microbiol* **126**:345-355.
334. **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**:1539-1546.
335. **Suhr, M., I. Benz, and M. A. Schmidt.** 1996. Processing of the AIDA-I precursor: removal of AIDAc and evidence for the outer membrane anchoring as a beta-barrel structure. *Mol Microbiol* **22**:31-42.
336. **Berthiaume, F., N. Rutherford, and M. Mourez.** 2007. Mutations affecting the biogenesis of the AIDA-I autotransporter. *Res Microbiol* **158**:348-354.
337. **Berthiaume, F., M. F. Leblond, J. Harel, and M. Mourez.** 2010. Growth-phase-dependent expression of the operon coding for the glycosylated autotransporter adhesin AIDA-I of pathogenic *Escherichia coli*. *FEMS Microbiol Lett* **311**:176-184.
338. **Benz, I., T. van Alen, J. Bolte, M. E. Wormann, and M. A. Schmidt.** 2010. Modulation of transcription and characterization of the promoter organization of the autotransporter adhesin heptosyltransferase and the autotransporter adhesin AIDA-I. *Microbiology* **156**:1155-1166.
339. **Charbonneau, M. E., and M. Mourez.** 2008. The *Escherichia coli* AIDA-I autotransporter undergoes cytoplasmic glycosylation independently of export. *Res Microbiol* **159**:537-544.
340. **Diderichsen, B.** 1980. flu, a metastable gene controlling surface properties of *Escherichia coli*. *J Bacteriol* **141**:858-867.
341. **van der Woude, M. W., and A. J. Baumler.** 2004. Phase and antigenic variation in bacteria. *Clin Microbiol Rev* **17**:581-611, table of contents.
342. **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.
343. **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**:877-887.
344. **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**:2203-2209.
345. **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**:115-120.
346. **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**:283-296.
347. **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**:63-76.
348. **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**:1553-1562.

349. **van der Woude, M. W., and I. R. Henderson.** 2008. Regulation and function of Ag43 (flu). *Annu Rev Microbiol* **62**:153-169.
350. **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**:3233-3244.
351. **Elsinghorst, E. A., and J. A. Weitz.** 1994. Epithelial cell invasion and adherence directed by the enterotoxigenic *Escherichia coli* tib locus is associated with a 104-kilodalton outer membrane protein. *Infect Immun* **62**:3463-3471.
352. **Evans, D. J., Jr., and D. G. Evans.** 1973. Three characteristics associated with enterotoxigenic *Escherichia coli* isolated from man. *Infect Immun* **8**:322-328.
353. **Crossman, L. C., R. R. Chaudhuri, S. A. Beatson, T. J. Wells, M. Desvaux, A. F. Cunningham, N. K. Petty, V. Mahon, C. Brinkley, J. L. Hobman, S. J. Savarino, S. M. Turner, M. J. Pallen, C. W. Penn, J. Parkhill, A. K. Turner, T. J. Johnson, N. R. Thomson, S. G. Smith, and I. R. Henderson.** 2010. A commensal gone bad: complete genome sequence of the prototypical enterotoxigenic *Escherichia coli* strain H10407. *J Bacteriol* **192**:5822-5831.
354. **Elsinghorst, E. A., and D. J. Kopecko.** 1992. Molecular cloning of epithelial cell invasion determinants from enterotoxigenic *Escherichia coli*. *Infect Immun* **60**:2409-2417.
355. **Yamashita, S., P. Lukacik, T. J. Barnard, N. Noinaj, S. Felek, T. M. Tsang, E. S. Krukonis, B. J. Hinnebusch, and S. K. Buchanan.** 2011. Structural insights into Ail-mediated adhesion in *Yersinia pestis*. *Structure* **19**:1672-1682.
356. **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**:2256-2265.
357. **Espert, S. M., E. A. Elsinghorst, and G. P. Munson.** 2011. The tib Adherence Locus of Enterotoxigenic *Escherichia coli* Is Regulated by Cyclic AMP Receptor Protein. *J Bacteriol* **193**:1369-1376.
358. **Kolb, A., S. Busby, H. Buc, S. Garges, and S. Adhya.** 1993. Transcriptional regulation by cAMP and its receptor protein. *Annu Rev Biochem* **62**:749-795.
359. **Lindenthal, C., and E. A. Elsinghorst.** 2001. Enterotoxigenic *Escherichia coli* TibA glycoprotein adheres to human intestine epithelial cells. *Infect Immun* **69**:52-57.
360. **Dehio, C., M. Meyer, J. Berger, H. Schwarz, and C. Lanz.** 1997. Interaction of *Bartonella henselae* with endothelial cells results in bacterial aggregation on the cell surface and the subsequent engulfment and internalisation of the bacterial aggregate by a unique structure, the invasome. *J Cell Sci* **110 (Pt 18)**:2141-2154.
361. **Kline, K. A., S. Falker, S. Dahlberg, S. Normark, and B. Henriques-Normark.** 2009. Bacterial adhesins in host-microbe interactions. *Cell Host Microbe* **5**:580-592.
362. **Kay, B. K., M. P. Williamson, and M. Sudol.** 2000. The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. *FASEB J* **14**:231-241.
363. **Kohler, S. D., A. Weber, S. P. Howard, W. Welte, and M. Drescher.** 2010. The proline-rich domain of TonB possesses an extended polyproline II-like conformation of sufficient length to span the periplasm of Gram-negative bacteria. *Protein Sci* **19**:625-630.

364. **Williamson, M. P.** 1994. The structure and function of proline-rich regions in proteins. *Biochem J* **297** (Pt 2):249-260.
365. **Loveless, B. J., and M. H. Saier, Jr.** 1997. A novel family of channel-forming, autotransporting, bacterial virulence factors. *Mol Membr Biol* **14**:113-123.
366. **Henderson, I. R., R. Cappello, and J. P. Nataro.** 2000. Autotransporter proteins, evolution and redefining protein secretion. *Trends Microbiol* **8**:529-532.
367. **Brockmeyer, J., M. Bielaszewska, A. Fruth, M. L. Bonn, A. Mellmann, H. U. Humpf, and H. Karch.** 2007. Subtypes of the plasmid-encoded serine protease EspP in Shiga toxin-producing *Escherichia coli*: distribution, secretion, and proteolytic activity. *Appl Environ Microbiol* **73**:6351-6359.
368. **Davis, J., A. L. Smith, W. R. Hughes, and M. Golomb.** 2001. Evolution of an autotransporter: domain shuffling and lateral transfer from pathogenic *Haemophilus* to *Neisseria*. *J Bacteriol* **183**:4626-4635.
369. **Soding, J., and A. N. Lupas.** 2003. More than the sum of their parts: on the evolution of proteins from peptides. *Bioessays* **25**:837-846.
370. **Knight, J., R. Deora, D. G. Assimos, and R. P. Holmes.** 2013. The genetic composition of *Oxalobacter formigenes* and its relationship to colonization and calcium oxalate stone disease. *Urolithiasis* **41**:187-196.
371. **De Muynck, C., C. S. Pereira, M. Naessens, S. Parmentier, W. Soetaert, and E. J. Vandamme.** 2007. The genus *Gluconobacter oxydans*: comprehensive overview of biochemistry and biotechnological applications. *Crit Rev Biotechnol* **27**:147-171.
372. **Frick, I. M., M. Morgelin, and L. Bjorck.** 2000. Virulent aggregates of *Streptococcus pyogenes* are generated by homophilic protein-protein interactions. *Mol Microbiol* **37**:1232-1247.
373. **Fink, D. L., B. A. Green, and J. W. St Geme, 3rd.** 2002. The *Haemophilus influenzae* Hap autotransporter binds to fibronectin, laminin, and collagen IV. *Infect Immun* **70**:4902-4907.
374. **Hendrixson, D. R., and J. W. St Geme, 3rd.** 1998. The *Haemophilus influenzae* Hap serine protease promotes adherence and microcolony formation, potentiated by a soluble host protein. *Mol Cell* **2**:841-850.
375. **May, K. L., M. Grabowicz, S. W. Polyak, and R. Morona.** 2012. Self-association of the *Shigella flexneri* IcsA autotransporter protein. *Microbiology* **158**:1874-1883.
376. **Bernardini, M. L., J. Mounier, H. d'Hauteville, M. Coquis-Rondon, and P. J. Sansonetti.** 1989. Identification of *icsA*, a plasmid locus of *Shigella flexneri* that governs bacterial intra- and intercellular spread through interaction with F-actin. *Proc Natl Acad Sci U S A* **86**:3867-3871.
377. **Teh, M. Y., and R. Morona.** 2013. Identification of *Shigella flexneri* IcsA residues affecting interaction with N-WASP, and evidence for IcsA-IcsA co-operative interaction. *PLoS One* **8**:e55152.
378. **Padrick, S. B., and M. K. Rosen.** 2010. Physical mechanisms of signal integration by WASP family proteins. *Annu Rev Biochem* **79**:707-735.
379. **Takeichi, M.** 1990. Cadherins: a molecular family important in selective cell-cell adhesion. *Annu Rev Biochem* **59**:237-252.
380. **Boggon, T. J., J. Murray, S. Chappuis-Flament, E. Wong, B. M. Gumbiner, and L. Shapiro.** 2002. C-cadherin ectodomain structure and implications for cell adhesion mechanisms. *Science* **296**:1308-1313.

381. **Pokutta, S., K. Herrenknecht, R. Kemler, and J. Engel.** 1994. Conformational changes of the recombinant extracellular domain of E-cadherin upon calcium binding. *Eur J Biochem* **223**:1019-1026.
382. **Leckband, D., and S. Sivasankar.** 2012. Cadherin recognition and adhesion. *Curr Opin Cell Biol* **24**:620-627.
383. **Wu, Y., X. Jin, O. Harrison, L. Shapiro, B. H. Honig, and A. Ben-Shaul.** 2010. Cooperativity between trans and cis interactions in cadherin-mediated junction formation. *Proc Natl Acad Sci U S A* **107**:17592-17597.
384. **Leckband, D., and A. Prakasam.** 2006. Mechanism and dynamics of cadherin adhesion. *Annu Rev Biomed Eng* **8**:259-287.
385. **Brieher, W. M., A. S. Yap, and B. M. Gumbiner.** 1996. Lateral dimerization is required for the homophilic binding activity of C-cadherin. *J Cell Biol* **135**:487-496.
386. **Shan, W. S., H. Tanaka, G. R. Phillips, K. Arndt, M. Yoshida, D. R. Colman, and L. Shapiro.** 2000. Functional cis-heterodimers of N- and R-cadherins. *J Cell Biol* **148**:579-590.
387. **Ozawa, M.** 2002. Lateral dimerization of the E-cadherin extracellular domain is necessary but not sufficient for adhesive activity. *J Biol Chem* **277**:19600-19608.
388. **Klingelhofer, J., O. Y. Laur, R. B. Troyanovsky, and S. M. Troyanovsky.** 2002. Dynamic interplay between adhesive and lateral E-cadherin dimers. *Mol Cell Biol* **22**:7449-7458.
389. **Troyanovsky, R. B., E. Sokolov, and S. M. Troyanovsky.** 2003. Adhesive and lateral E-cadherin dimers are mediated by the same interface. *Mol Cell Biol* **23**:7965-7972.
390. **Chappuis-Flament, S., E. Wong, L. D. Hicks, C. M. Kay, and B. M. Gumbiner.** 2001. Multiple cadherin extracellular repeats mediate homophilic binding and adhesion. *J Cell Biol* **154**:231-243.
391. **Sivasankar, S., W. Brieher, N. Lavrik, B. Gumbiner, and D. Leckband.** 1999. Direct molecular force measurements of multiple adhesive interactions between cadherin ectodomains. *Proc Natl Acad Sci U S A* **96**:11820-11824.
392. **Bayas, M. V., A. Leung, E. Evans, and D. Leckband.** 2006. Lifetime measurements reveal kinetic differences between homophilic cadherin bonds. *Biophys J* **90**:1385-1395.
393. **Ahrens, T., O. Pertz, D. Haussinger, C. Fauser, T. Schulthess, and J. Engel.** 2002. Analysis of heterophilic and homophilic interactions of cadherins using the c-Jun/c-Fos dimerization domains. *J Biol Chem* **277**:19455-19460.
394. **Harrison, O. J., X. Jin, S. Hong, F. Bahna, G. Ahlsen, J. Brasch, Y. Wu, J. Vendome, K. Felsovalyi, C. M. Hampton, R. B. Troyanovsky, A. Ben-Shaul, J. Frank, S. M. Troyanovsky, L. Shapiro, and B. Honig.** 2011. The extracellular architecture of adherens junctions revealed by crystal structures of type I cadherins. *Structure* **19**:244-256.
395. **Katsamba, P., K. Carroll, G. Ahlsen, F. Bahna, J. Vendome, S. Posy, M. Rajebhosale, S. Price, T. M. Jessell, A. Ben-Shaul, L. Shapiro, and B. H. Honig.** 2009. Linking molecular affinity and cellular specificity in cadherin-mediated adhesion. *Proc Natl Acad Sci U S A* **106**:11594-11599.
396. **Gibbs, K. A., M. L. Urbanowski, and E. P. Greenberg.** 2008. Genetic determinants of self identity and social recognition in bacteria. *Science* **321**:256-259.

397. **Pathak, D. T., X. Wei, A. Bucuvalas, D. H. Haft, D. L. Gerloff, and D. Wall.** 2012. Cell contact-dependent outer membrane exchange in myxobacteria: genetic determinants and mechanism. *PLoS Genet* **8**:e1002626.
398. **Gibbs, K. A., L. M. Wenren, and E. P. Greenberg.** 2011. Identity gene expression in *Proteus mirabilis*. *J Bacteriol* **193**:3286-3292.
399. **Wei, X., D. T. Pathak, and D. Wall.** 2011. Heterologous protein transfer within structured myxobacteria biofilms. *Mol Microbiol* **81**:315-326.
400. **Frantz, C., K. M. Stewart, and V. M. Weaver.** 2010. The extracellular matrix at a glance. *J Cell Sci* **123**:4195-4200.
401. **Schaefer, L., and R. V. Iozzo.** 2008. Biological functions of the small leucine-rich proteoglycans: from genetics to signal transduction. *J Biol Chem* **283**:21305-21309.
402. **Lappann, M., H. Claus, T. van Alen, M. Harmsen, J. Elias, S. Molin, and U. Vogel.** 2010. A dual role of extracellular DNA during biofilm formation of *Neisseria meningitidis*. *Mol Microbiol* **75**:1355-1371.
403. **Hamilton, W. D.** 1964. The genetical evolution of social behaviour. I. *J Theor Biol* **7**:1-16.
404. **Keller, L., and K. G. Ross.** 1998. Selfish genes: a green beard in the red fire ant. *Nature* **394**:573-575.
405. **Gardner, A., and S. A. West.** 2010. Greenbeards. *Evolution* **64**:25-38.
406. **Queller, D. C., E. Ponte, S. Bozzaro, and J. E. Strassmann.** 2003. Single-gene greenbeard effects in the social amoeba *Dictyostelium discoideum*. *Science* **299**:105-106.

Annexe

Résultats supplémentaires - Structure et oligomérisation d'AIDA-I

Les résultats précédents montrent qu'AIDA-I permet l'auto-agrégation des bactéries en interagissant directement avec lui-même sans avoir recours à un facteur accessoire. Ils montrent aussi qu'AIDA-I subit un changement de conformation lors de l'interaction, passant d'une conformation monomérique à une conformation oligomérique compétente pour l'auto-agrégation (article 4). Il y a toutefois encore plusieurs questions qui restent sans réponse, notamment quant à la région d'AIDA-I qui est responsable de l'oligomérisation de la protéine et l'auto-agrégation des bactéries. Avec les résultats précédents, nous avons pu déterminer que c'est le domaine passager qui est responsable de l'oligomérisation. Toutefois, le domaine passager est composé de différents modules qui sont associés aux diverses fonctionnalités de la protéine. Par exemple, dans la partie N-terminale du domaine passager, on retrouve un module d'adhésion (1). Il est donc possible qu'il existe une région précise qui soit responsable de l'oligomérisation et des interactions AIDA-I/AIDA-I. Il est aussi possible que tout le domaine passager participe à l'oligomérisation. De plus, selon la région d'AIDA-I impliquée dans l'oligomérisation, différents types d'interactions sont possibles.

Nous avons donc poursuivi notre étude du mécanisme d'auto-agrégation plus en détail. Les résultats qui suivent ne sont toujours pas publiés et portent sur la structure et l'oligomérisation de AIDA-I. Nous avons d'abord collaboré avec le groupe du Dr Isabelle Rouiller, une experte en microscopie électronique de l'Université McGill, afin de visualiser AIDA-I et ainsi d'améliorer notre compréhension du mécanisme d'auto-agrégation. Nous avons ensuite caractérisé les oligomères d'AIDA-I à l'aide d'une approche de FRET (Fluorescence ou Förster Resonance Energy Transfer).

1) Structure de AIDA-I

Nous avons d'abord voulu obtenir une structure d'AIDA-I. Pour ce faire, nous avons utilisé la microscopie électronique, en collaboration avec Brian Beckett et Dre Isabelle Rouiller de l'Université McGill. Pour visualiser la protéine, nous avons donc

purifié AIDA-I possédant un « His-tag » par chromatographie d'affinité (comme dans l'article 3). La protéine purifiée a ensuite été solubilisée dans un tampon TBS 1% b-OG (n-octyl- β -D-glucoside). Dans ces conditions, AIDA-I est monomérique (Fig. 1).

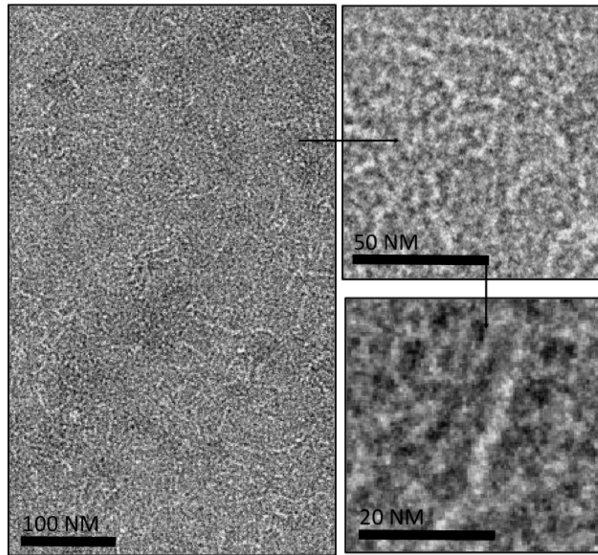


Fig. 1 : AIDA-I est monomérique en présence de détergent

AIDA-I a été purifié par chromatographie d'affinité à partir de *E. coli* C600, solubilisé en TBS 1% b-OG à une concentration de 150 $\mu\text{g}/\text{mL}$. Un échantillon de 5 μL a été déposé sur une grille de cuivre couverte de carbone pendant 2 min pour laisser les protéines adhérentes à la grille et a été coloré à l'acétate d'uranyle 2% pendant 1 min. La grille a ensuite été visualisée à l'aide d'un microscope à transmission FEI Tecnai G2 F20 200kV équipé d'une caméra CCD Gatan Ultrascan 4000.

Afin d'obtenir plus de détails quant à la structure d'AIDA-I, les molécules, à partir de micrographes comme dans la figure 1, ont été normalisées et alignées et dix classes de molécules (ou dix axes de symétrie différents) ont été identifiées en deux dimensions (Fig. 2A). Alternativement, plusieurs séries d'inclinaison des échantillons en microscopie électronique ont été effectuées afin d'obtenir différents tomogrammes. Les tomogrammes ont ensuite été utilisés pour effectuer une reconstitution en trois dimensions d'AIDA-I (Fig. 2B).

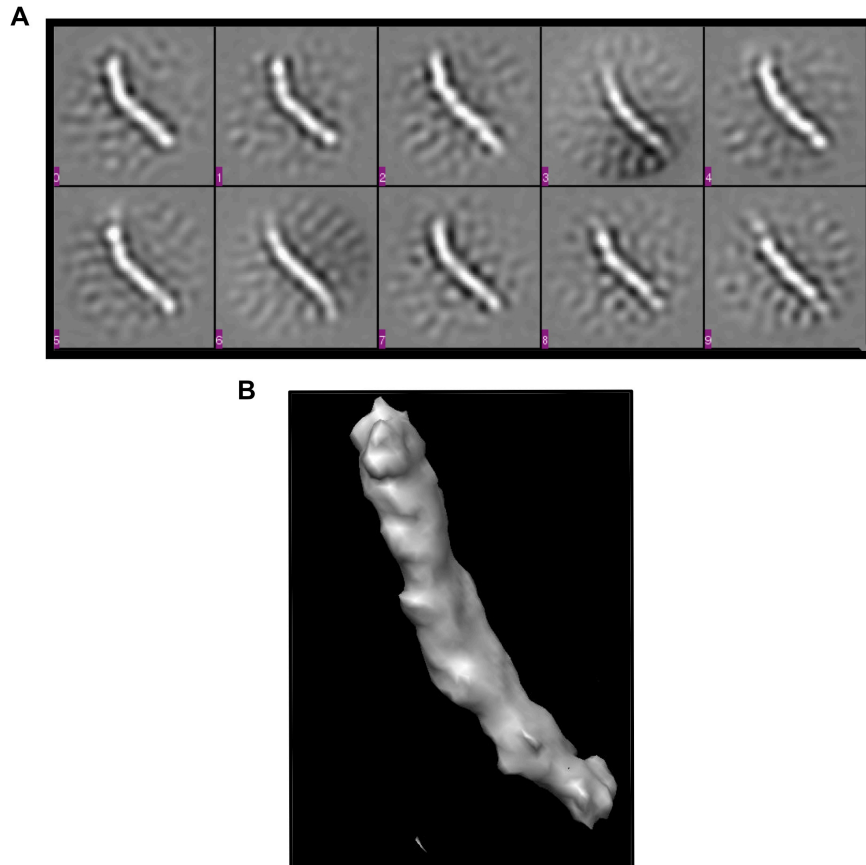


Fig. 2 : Structure d'AIDA-I monomérique

(A) Avec l'aide du logiciel SPARX (6), 27 000 molécules à partir de micrographes (comme en Fig. 1) ont été normalisées, centrées, regroupées par similarités grâce à un « k-means clustering » et finalement alignées. (B) Structure d'AIDA-I en 3D générée à partir de l'alignement de tomogrammes. Plusieurs séries d'inclinaison ont été effectuées avec les échantillons à 50 000X à l'aide d'un microscope à transmission FEI Tecnai G2 F20 200kV équipé d'une caméra CCD Gatan Ultrascan 4000 et du module automatique d'acquisition de tomogrammes de FEI. Les images ont été récupérées avec des sauts de 2° de -70° à +50°. Les séries d'inclinaison ont été normalisées et alignées à l'aide du logiciel IMOD afin de générer les tomogrammes (7).

Les images montrent des monomères d'AIDA-I qui font environ 320 Å en longueur. En général, les tonneaux β d'autotransporteur sont très similaires; dans le cas de l'autotransporteur EspP, celui-ci a une longueur de 70 Å (jusqu'à la boucle périplasmique la plus éloignée) (2). Ceci veut donc dire que la partie extracellulaire de AIDA-I, comprenant le domaine passager et la région de jonction, a une longueur d'environ 250 Å. De plus, il est intéressant de constater que les monomères d'AIDA-I ne sont pas des bâtons

droits, mais adoptent plutôt une structure courbée. Ceci contraste avec l'image habituelle de la structure du domaine passager en hélice β plutôt rigide que l'on voit selon les prédictions de structure. Finalement, AIDA-I ressemble à un enfillement de billes sur une ficelle (structure de « beads on a string »). Ceci représente bien le fait qu'AIDA-I soit une protéine modulaire. Ces résultats suggèrent donc que l'hélice β droite formée par les répétitions de la séquence consensus de 19 acides aminés n'est pas rigide et uniforme, comme le montrent les modélisations de structure se basant sur de la pertactine (voir Fig. 9E de la section «Recension de la littérature» et (3)). En revanche, la structure d'AIDA-I pourrait ressembler plutôt à celle de VacA où plusieurs bris sont présents dans l'hélice β , ce qui a pour conséquence de séparer le domaine passager en plusieurs sous-domaines (4). Nous n'avons toutefois pas pu identifier clairement les parties N- et C-terminales. Il est possible que, dans la structure 3D de AIDA-I, le N-terminal de AIDA-I soit situé en haut et le C-terminal soit en bas. Il semble que le tonneau β soit la structure la plus étroite d'un autotransporteur, avec un diamètre de 20 Å en son point le plus étroit (5). L'ajout d'un anticorps dirigé contre le « His-tag », présent en N-terminal, permettrait de confirmer l'orientation de la protéine.

2) Oligomérisation de AIDA-I

Pour déclencher l'oligomérisation d'AIDA-I, il suffit de diluer la solution de protéines pour que la concentration de détergent se retrouve sous la concentration critique micellaire (CCM). Donc en diluant la concentration de b-OG de 1% à 0,3%, l'oligomérisation de AIDA-I est déclenchée et, en microscopie électronique, on remarque l'apparition d'« araignées » qui contiennent de 4 à 12 molécules d'AIDA-I (Fig.3). Toutefois, comme ces « araignées » sont plutôt hétérogènes, de grosseurs et de formes différentes, il n'a pas été possible d'aligner et de normaliser les complexes ni de faire un reconstitution en trois dimensions, comme pour la protéine monomérique.

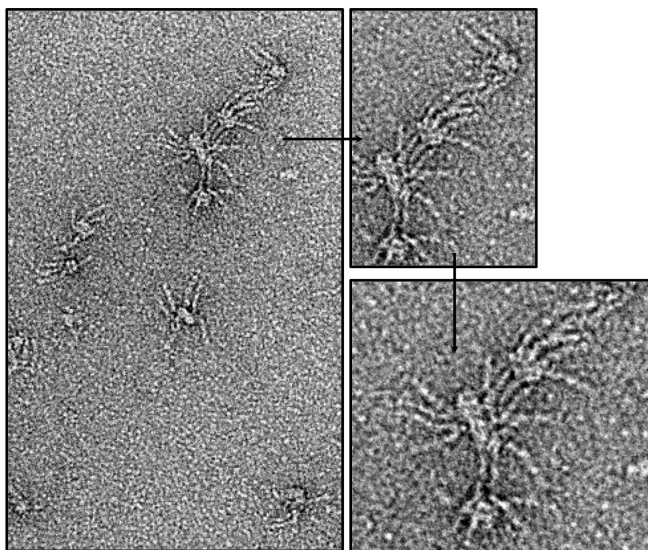


Fig. 3: Oligomérisation d'AIDA-I en absence de détergent

Pour déclencher l'oligomérisation d'AIDA-I, la solution de protéines purifiées a été diluée à une concentration de 150 $\mu\text{g}/\text{mL}$ en TBS, pour que la concentration de b-OG soit de 0,1%, c'est-à-dire sous la concentration critique micellaire. L'échantillon a ensuite été déposée sur une grille et colorée comme à la Fig. 1 et a été observée à l'aide d'un microscope à transmission FEI Tecnai G2 F20 200kV équipé d'une caméra CCD Gatan Ultrascan 4000.

Pour caractériser les oligomères formés par les molécules d'AIDA-I en solution, nous avons utilisé une approche de FRET. Le principe de la FRET repose sur le transfert d'énergie, sous forme non radiative (donc sans lumière), entre un fluorophore donneur et un fluorophore accepteur pour lesquels la longueur d'onde d'excitation de l'accepteur correspond à la longueur d'onde d'émission du donneur. En bref, la FRET se produit lorsqu'en excitant le donneur, celui-ci transfère une partie de l'énergie qu'il a acquise à l'accepteur, ce qui fait que l'on détecte alors la fluorescence de l'accepteur. Pratiquement, la réaction de FRET se voit donc par l'augmentation de la fluorescence de l'accepteur ainsi que par la diminution de la fluorescence du donneur. Le phénomène de FRET est dépendant de la distance entre les fluorophores et se produit uniquement lorsque les deux fluorophores sont à proximité l'un de l'autre, ce qui en fait une technique intéressante pour évaluer les interactions protéine-protéine. La distance maximale pour qu'il y ait transfert d'énergie varie pour chaque paire de fluorophores et chaque paire de fluorophores est donc

caractérisée par un rayon de Förster propre, qui représente la distance entre les fluorophores pour avoir une efficacité de transfert de 50%.

Pour pouvoir utiliser notre approche de FRET avec AIDA-I, nous avons utilisé des mutants contenant une insertion de 5 acides aminés et ajoutant une cystéine. Nous avons choisi plusieurs mutants qui ne montrent pas de défauts de fonctionnalité et qui sont distribués dans tout le domaine passager (1). Les différents mutants ont d'abord été purifiés par chromatographie d'affinité grâce au « His-tag » et les fluorophores ont ensuite été couplés sur la cystéine présente dans l'insertion à l'aide d'une molécule de liaison contenant un groupement maléimide. Nous avons utilisé le couple de fluorophores Alexa fluor 488 et Alexa fluor 546, déjà utilisé en FRET auparavant et ayant un rayon de Förster théorique de 64 Å. (8). Il est aussi possible de suivre la FRET en calculant une efficacité de transfert. Pour le couple de fluorophores AlexaFluor 488-AlexaFluor 546, l'efficacité se calcule comme suit :

$$\text{efficacité} = 1 - \frac{D + A}{A}$$

où D et A représente le ratio de fluorescence du donneur et de l'accepteur respectivement après excitation à la longueur d'onde du donneur : (Dans notre cas, la longueur d'onde d'excitation du donneur est de 488 nm et les maximums d'intensité (I) d'émission du donneur et de l'accepteur sont respectivement aux longueurs d'onde de 516 nm et 571 nm.)

$$D = \frac{I_{516}(\text{Donneur}) - I_{516}(\text{Tampon})}{I_{571}(\text{Donneur}) - I_{571}(\text{Tampon})}$$

$$A = \frac{I_{516}(\text{donneur} + \text{accepteur}) - I_{516}(\text{Donneur})}{I_{571}(\text{donneur} + \text{accepteur}) - I_{571}(\text{Donneur})}$$

Nous avons d'abord utilisé un mutant qui se retrouve au milieu du domaine passager de AIDA-I, le mutant i10. Le mutant i10 a été marqué par l'AlexaFluor 488 (i10-488) et par l'AlexaFluor 546 (i10-546) et les deux protéines ont été mélangées ensemble en TBS. La

solution a ensuite été excitée à une longueur d'onde de 488 nm et l'émission de fluorescence a été mesurée entre les longueurs d'ondes de 500 et 600 nm. Le spectre d'émission montre une augmentation de la fluorescence de l'accepteur à 571 nm ainsi qu'une diminution de la fluorescence du donneur à 516 nm, ce qui montre un transfert d'énergie entre le donneur et l'accepteur (Fig. 4A). Lorsque i10-488 et i10-546 sont mélangées en présence de détergent, donc en empêchant l'oligomérisation, il n'y a plus de transfert d'énergie, ce qui confirme que le transfert d'énergie résulte bien de l'interaction entre les protéines. (Fig. 4B). En présence de détergent, l'efficacité de FRET devient presque nulle en comparaison avec celle du mélange de protéines fait en TBS (Fig. 4C). Alternativement, l'inhibition de l'oligomérisation par le sel empêche aussi le transfert d'énergie entre les fluorophores (Fig. 4C). Ces résultats confirment donc que la FRET est un bon moyen pour étudier l'oligomérisation d'AIDA-I.

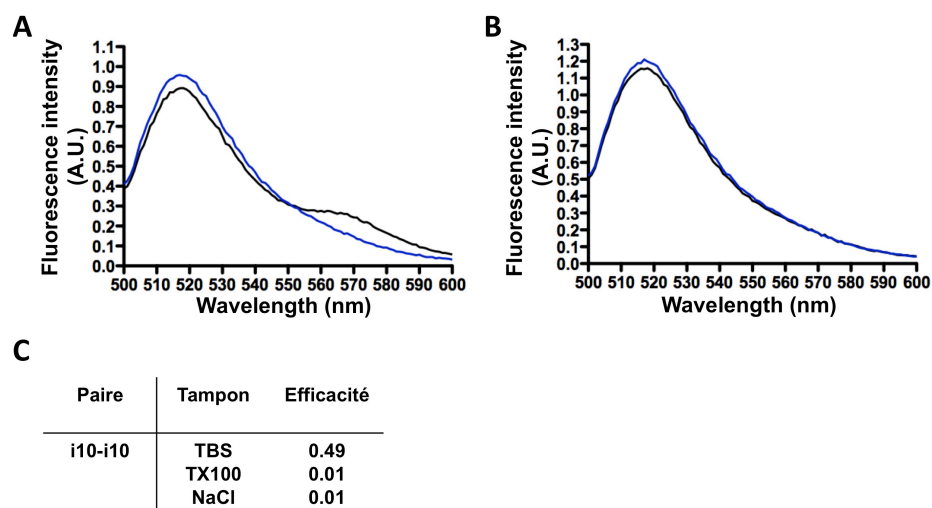


Fig. 4: Interaction du mutant i10 observé par FRET

Le mutant i10 a été purifié par chromatographie d'affinité, solubilisé en TBS 1% b-OG et a été marqué sur la cystéine (présente dans l'insertion) avec le fluorophore donneur AlexaFluor 488 ou bien avec le fluorophore accepteur AlexaFluor 546. Les deux protéines ont ensuite été mélangées ensemble et diluées en TBS (A) ou bien en TBS 1% b-OG (B) pour obtenir une concentration finale de chaque protéine de 23 µg/mL. L'émission de fluorescence a été mesurée entre 500 et 600 nm suite à une excitation à 488 nm pour le donneur seul (i10-488) (ligne bleue) ou pour le mélange donneur et accepteur (i10-488 + i10-546) (ligne noire). L'efficacité de fluorescence a été calculée pour le mélange donneur et accepteur (i10-488 + i10-546) en TBS, en présence de 1% b-OG et en présence de 1M NaCl (C).

Nous avons donc ensuite choisi des positions situées tout le long du domaine passager d'AIDA-I et nous avons regardé systématiquement les interactions entre les différentes paires donneur/accepteur possibles (Fig. 5A). De façon intéressante, nous avons observé un transfert d'énergie entre le donneur et l'accepteur pour toutes les paires homologues et, à différents degrés, pour plusieurs paires hétérologues (Fig. 5B). Il est à noter que, pour les paires hétérologues, l'efficacité de FRET peut varier selon que les insertions se retrouvent sur une surface de l'hélice β ou bien sur une autre, ce qui peut faire varier la distance entre les deux fluorophores. L'observation de FRET pour toutes les paires homologues suggère que tout le domaine passager participe aux interactions AIDA-I/AIDA-I. Ceci est en accord avec le modèle d'oligomérisation de l'autotransporteur Hap pour lequel l'interaction entre deux molécules d'Hap se fait par le « domaine SAAT » en entier (9). Dans le cas de AIDA-I, le domaine passager au complet représente un « domaine SAAT ».

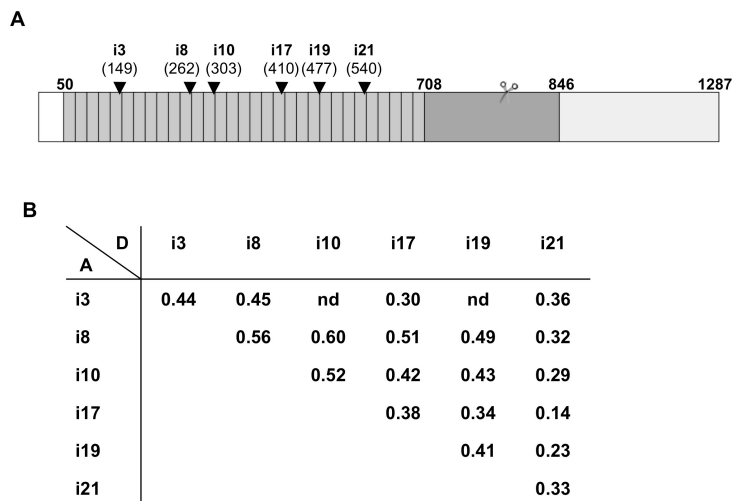


Fig. 5: Caractérisation des interactions AIDA-I/AIDA-I par FRET

(A) Position des mutants d'insertion d'AIDA-I choisis pour la FRET. Ces mutants n'ont pas de défaut de fonctionnalité, mais l'insertion ajoute une cystéine. Les fluorophores ont été couplés sur les cystéines présentes dans ces insertions. (B) Les différentes paires donneurs/accepteurs ont été mélangées ensemble et l'efficacité de fluorescence a été calculée pour chaque mélange. nd signifie que ce mélange n'a pas été effectué (no data). D, position du donneur; A, position de l'accepteur.

Deux types d'interaction sont possibles entre les protéines purifiées. Les protéines peuvent interagir en *cis*, « côte-à-côte », ou bien en *trans*, « face-à-face » (Fig. 6A-B). Nos résultats semblent montrer les deux types d'interactions. D'un côté, les paires i3-488/i3-546 et i21-488/i21-546 permettent un transfert d'énergie seulement lorsque les protéines interagissent en *cis*, de façon « côte-à-côte ». Ces deux insertions sont retrouvées à chaque extrémité du domaine passager, ce qui implique que, dans le cas d'une interaction en *trans*, les fluorophores seraient trop éloignés pour qu'il y ait transfert d'énergie. D'un autre côté, la paire i21-488/i3-546 permet donc un transfert d'énergie seulement si les protéines interagissent en *trans*. Effectivement, nous observons un transfert d'énergie pour les trois paires: i3-i3, i21-i21 et i3-i21 (Fig. 5B et Fig. 6C-E). Ces résultats montrent donc que AIDA-I peut interagir autant en *cis* et qu'en *trans*. Par ailleurs, il est intéressant de constater que l'efficacité du transfert d'énergie est plus important pour les mutants retrouvés au centre du domaine passager, notamment les mutants i8 et i10 (Fig. 5B). Dans ces cas-là, il est donc probable que la plus grande efficacité de FRET résulte de la combinaison des deux types d'interactions. Nos résultats suggèrent donc que les oligomères d'AIDA-I sont formés par des protéines qui interagissent en *cis* et des protéines qui interagissent en *trans* et que tout le domaine passager participe à l'interaction.

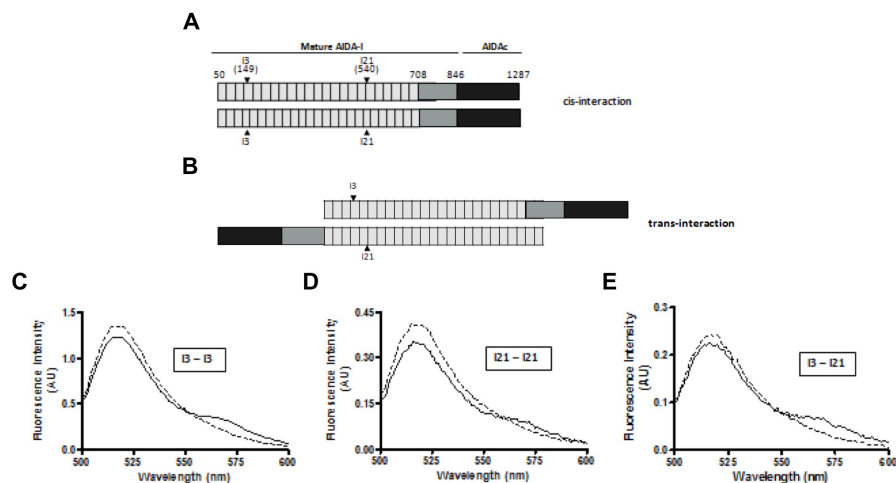


Fig. 6: AIDA-I interagit en *cis* et en *trans*

(A-B) Schématisation de l'interaction en *cis* et en *trans*. Les paires i3-488/i3-546 et i21-488/i21-546 représentent des paires donneur/accepteur pour lesquelles la FRET n'est possible que s'il y a des interactions en *cis* alors que la FRET n'est possible que s'il y a des

interactions en *trans* pour la paire i21-488/i3-546. (C-E) Spectres d'émission entre 500 et 600 nm pour le donneur seul (lignes pointillées) et pour le couple donneur/accepteur (lignes pleines) pour les couples i3-488/i3-546 (C), i21-488/i21-546 (D) et i21-488/i3-546 (E).

3) Interactions en *cis* vs en *trans*

Pour qu'il y ait auto-agrégation entre les bactéries, les molécules d'AIDA-I doivent obligatoirement interagir en *trans*. Les interactions en *cis* représentent, quant à elles, des interactions à la surface d'une même bactérie et représentent possiblement un artefact provenant de l'oligomérisation d'AIDA-I *in vitro*. D'un autre côté, il est aussi possible que ces interactions soient réelles et servent, par exemple, à réguler le processus d'auto-agrégation ou bien à augmenter la force de l'interaction entre deux bactéries. Nous avons donc voulu vérifier si ces interactions en *cis* étaient bien réelles. Pour ce faire, nous avons repris l'expérience d'inhibition de l'auto-agrégation des bactéries par des protéines purifiées (voir l'article 3). Par contre, cette fois-ci, nous avons mis les protéines purifiées en contact des bactéries en présence de détergent. Il est à noter que l'ajout de détergent seul n'affecte pas la capacité d'auto-agrégation de bactéries exprimant AIDA-I (Fig. 7; colonne 2). Comme il a été vu dans l'article 3, l'ajout de AIDA-I purifié sans détergent inhibe l'auto-agrégation de bactéries exprimant AIDA-I (Fig 7; colonne 3). Toutefois, l'ajout d'AIDA-I purifié en présence de détergent n'a aucun effet sur l'auto-agrégation des bactéries exprimant AIDA-I (Fig. 7; colonne 4). Le détergent empêche l'oligomérisation d'AIDA-I, comme cela a été montré en microscopie électronique. La différence entre les deux suspensions de protéines purifiées est donc que, en présence de détergent, AIDA-I est monomérique tandis que, en TBS, AIDA-I forme des oligomères (Fig.1 et 3). La présence des dimères d'AIDA-I en *cis* à la surface des bactéries a aussi été observée lors d'une expérience de pontage chimique fait avec du BS³, un agent pontant incapable de pénétrer dans les bactéries (Victoria Girard, résultats non publiés). Encore une fois, les dimères en *cis* semblaient importants pour l'oligomérisation en *trans*.

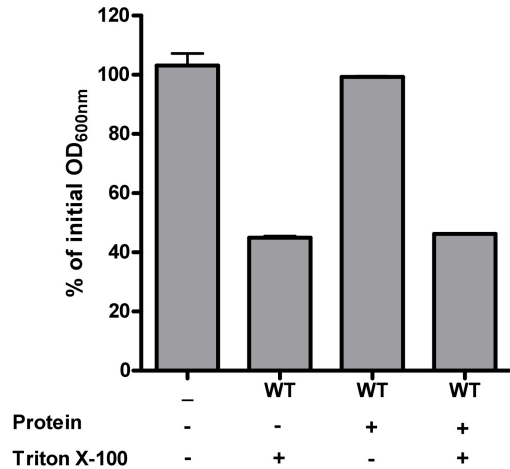


Fig. 7: L'interaction préalable en *cis* est nécessaire pour l'interaction en *trans*

Des cultures d'E. coli C600 exprimant AIDA-I (WT) ou non (-) ont été normalisées à une DO₆₀₀ de 1,5 et incubées à 4°C pendant 3h sans agitation. La DO₆₀₀ après 3h a été mesurée et a été comparée à la DO₆₀₀ initiale. Les bactéries ont aussi été incubées en présence ou en absence d'AIDA-I purifié et en présence ou en absence de détergent.

Le modèle pour l'auto-agrégation qui en découle est donc qu'AIDA-I interagit d'abord en *cis*, sur la même bactérie, et ce sont donc ces dimères d'AIDA-I qui interagissent ensuite en *trans* pour permettre l'auto-agrégation des bactéries (Fig. 8). AIDA-I monomérique n'est donc pas capable d'inhiber l'auto-agrégation des bactéries puisque AIDA-I se retrouve sous la forme de dimères à la surface des bactéries.

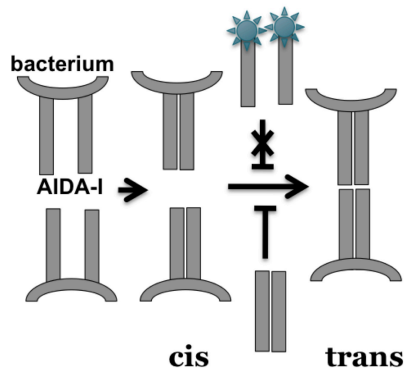


Fig. 8: Modèle pour l'interaction d'AIDA-I à la surface des bactéries

AIDA-I interagit d'abord avec une deuxième molécule d'AIDA-I à la surface des bactéries pour former des dimères en *cis* et ce sont ces dimères qui interagissent entre eux pour promouvoir l'auto-agrégation des bactéries. Lorsque l'on ajoute des molécules d'AIDA-I

purifiées en TBS, il y a inhibition de l'auto-agrégation des bactéries car les molécules d'AIDA-I sont capables d'oligomériser et de former des dimères en solution. Toutefois, l'ajout d'AIDA-I purifié en présence de détergent, donc l'ajout d'AIDA-I monomérique, n'est pas capable d'inhiber l'auto-agrégation.

La raison de l'interaction préalable en *cis* reste à déterminer. Toutefois, il est possible que l'interaction entre des dimères d'AIDA-I soit plus forte que l'interaction entre deux monomères. La présence de dimères en *cis* contraste toutefois avec le mécanisme d'interaction utilisé par l'autotransporteur Hap (9). En effet, dans le cas de Hap, un monomère de Hap sur une bactérie interagit avec un autre monomère sur une deuxième bactérie. Il est à noter que Hap est une SPATE et possède un gros domaine globulaire permettant l'activité sérine protéase. Ce domaine empêche donc stériquement l'interaction en *cis* entre deux molécules de Hap. Il est cependant intéressant de constater que les interactions Hap-Hap en *trans* permettent le recrutement d'autres molécules de Hap, ce qui permet la formation d'oligomères plus important. Donc, encore une fois, pour permettre l'auto-agrégation, il y a recrutement de plusieurs molécules de Hap à un même site sur une même bactérie.

Le mécanisme d'auto-agrégation médié par AIDA-I implique donc des interactions en *cis* et des interactions en *trans*, ainsi qu'un changement de conformation dans la protéine (article 3). Un mécanisme similaire est aussi utilisé par les cadhérines des cellules eucaryotes (10). Les cadhérines sont les molécules responsables des interactions cellules-cellules chez les eucaryotes et permettent la séparation entre les différentes couches de cellules dans les tissus. Comme pour AIDA-I, l'interaction entre les cadhérines induit un changement de conformation (11) et les cadhérines interagissent autant en *cis* qu'en *trans* (12). Dans le cas des cadhérines, l'interaction en *cis* permet de stabiliser les jonctions serrées et d'augmenter la force d'interaction entre les deux cellules (13). Une autre caractéristique des cadhérines est leur préférence vis-à-vis des interactions homophiliques (14). Il est donc possible qu'AIDA-I et les autres SAATs soient des molécules « cadhérine-like » chez les bactéries et qu'ils contribuent à l'organisation des communautés microbienne possiblement en montrant une spécificité d'interaction similaire aux cadhérines.

Référence

1. Charbonneau, M. E., and M. Mourez. 2007. Functional organization of the autotransporter adhesin involved in diffuse adherence. *J Bacteriol* 189:9020-9029.
2. 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* 14:1214-1220.
3. Emsley, P., I. G. Charles, N. F. Fairweather, and N. W. Isaacs. 1996. Structure of *Bordetella pertussis* virulence factor P.69 pertactin. *Nature* 381:90-92.
4. Gangwer, K. A., D. J. Mushrush, D. L. Stauff, B. Spiller, M. S. McClain, T. L. Cover, and D. B. Lacy. 2007. Crystal structure of the *Helicobacter pylori* vacuolating toxin p55 domain. *Proc Natl Acad Sci U S A* 104:16293-16298.
5. 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):645-56.
6. Hohn, M., G. Tang, G. Goodyear, P. R. Baldwin, Z. Huang, P. A. Penczek, C. Yang, R. M. Glaeser, P. D. Adams, and S. J. Ludtke. 2007. SPARX, a new environment for Cryo-EM image processing. *J Struct Biol* 157:47-55.
7. Mastrorade, D. N. 1997. Dual-axis tomography: an approach with alignment methods that preserve resolution. *J Struct Biol* 120:343-352.
8. Kong, H. J., T. R. Polte, E. Alsberg, and D. J. Mooney. 2005. FRET measurements of cell-traction forces and nano-scale clustering of adhesion ligands varied by substrate stiffness. *Proc Natl Acad Sci U S A* 102:4300-4305.
9. 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* 30:3864-3874.
10. Takeichi, M. 1990. Cadherins: a molecular family important in selective cell-cell adhesion. *Annu Rev Biochem* 59:237-252.
11. Petrova, Y. I., M. M. Spano, and B. M. Gumbiner. 2012. Conformational epitopes at cadherin calcium-binding sites and p120-catenin phosphorylation regulate cell adhesion. *Mol Biol Cell* 23:2092-2108.
12. Shan, W. S., H. Tanaka, G. R. Phillips, K. Arndt, M. Yoshida, D. R. Colman, and L. Shapiro. 2000. Functional cis-heterodimers of N- and R-cadherins. *J Cell Biol* 148:579-590.
13. Wu, Y., X. Jin, O. Harrison, L. Shapiro, B. H. Honig, and A. Ben-Shaul. 2010. Cooperativity between trans and cis interactions in cadherin-mediated junction formation. *Proc Natl Acad Sci U S A* 107:17592-17597.
14. Katsamba, P., K. Carroll, G. Ahlsen, F. Bahna, J. Vendome, S. Posy, M. Rajebhosale, S. Price, T. M. Jessell, A. Ben-Shaul, L. Shapiro, and B. H. Honig. 2009. Linking molecular affinity and cellular specificity in cadherin-mediated adhesion. *Proc Natl Acad Sci U S A* 106:11594-11599.