THE REQUIREMENT OF PUTATIVE AUTOCHAPERONE MOTIFS FOR AUTOTRANSPORTER PASSENGER DOMAIN FOLDING

by

Elizabeth Gagnon

BSc., Dalhousie University, 2010

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

The Faculty of Graduate Studies

(Microbiology and Immunology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

October 2012

ABSTRACT

Protein secretion plays an essential role in the virulence of Gram-negative bacterial pathogens. Gram-negative bacteria have evolved multiple specialized secretion pathways in order to navigate proteins across the Gram-negative cell envelope. The simplest and most widespread secretion pathway is the type V secretion system and autotransporters (ATs) (Va) represent the largest class of secreted proteins in Gram-negative bacteria. ATs are structurally characterized by the presence of three distinct domains; an N-terminal signal sequence that targets the N-terminus of the polypeptide to the inner membrane, a passenger domain that often possesses β-helix structure and carries out the virulence function(s) and a conserved C-terminal translocation unit (TU) consisting of a β-domain that forms a porin-like structure in the outer membrane (OM) through which the passenger domain is thought to be extruded and an α-helical linker region that joins the β-domain to the C-terminus of the passenger domain. At the C-terminus of the majority of AT passenger domains is a conserved region termed the autochaperone (AC). The importance of conserved AC residues for passenger domain secretion and folding has been demonstrated for numerous ATs, yet the exact role of the AC region in OM translocation and passenger domain folding has yet to be clarified. In this study, the requirement of conserved C-terminal passenger domain motifs for the acquisition of passenger domain secondary structure and the interchangeability of these conserved AC motifs was investigated. A combination of far-UV CD spectroscopy and limited proteolysis with trypsin of full-length and AC-deleted Vag8, Ag43 and Smp passenger variants revealed that the requirement of conserved C-terminal AC motifs for the acquisition of passenger domain secondary structure varies among ATs. A cell wall fraction assay, in which the ability of BrkA, Ag43 and Smp AC regions to rescue BrkA AC-deleted passenger folding was tested, indicated that an AC region with similar structure to the cognate AC is necessary to rescue passenger domain folding. Altogether, the results of this study highlight the involvement of multiple factors in passenger domain folding and the likely variation that exists in the mechanism of passenger folding at the bacterial cell surface.

PREFACE

Biohazard Approval Certificate numbers under which this research was conducted are: H03-0255 and B06-0255 issued by the Biosafety Committee, Office of Research Services at the University of British Columbia.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AC	autochaperone
ΔAC	autochaperone-deleted
AT	autotransporter
CD	circular dichroism
FL	full-length
IM	inner membrane
IPTG	isopropyl β-D-1-thiogalactopyranoside
MCS	multiple cloning site
Ni-NTA	nickel-nitrilotriacetic acid
OD_{600}	optical density at 600 nm
OM	outer membrane
Pass	passenger
PCR	polymerase chain reaction
SAAT	self-associating autotransporter
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SS	signal sequence/signal peptide
TU	translocation-unit
UV	ultra-violet

ACKNOWLEDGEMENTS

Thanks to Dr. Rachel Fernandez for providing me the opportunity to work on the autotransporter project in her lab.

And many thanks to the ladies of the Fernandez lab and all my other lab friends (Jan, Manreet, Betty, Andrew, Mike, Kirsten and Mary) for their companionship.

I would also like to thank Fred for his CD spec wisdom and interesting conversation and Nomo for his cloning advice.

Thanks to CIHR for the Master Award/funding.

And finally, special thanks to Alex for his constant support, computer assistance and patience and without whom this thesis would never have been completed.

1. INTRODUCTION

1.1 Protein secretion in Gram-negative bacteria

Infections by Gram-negative bacteria are the cause of numerous life-threatening human diseases including bacterial dysentery, whooping cough, pneumonia, peptic ulcers, sepsis and meningitis (59). Protein secretion plays an essential role in the virulence of Gram-negative bacterial pathogens (8). In order for secreted protein virulence factors to reach the bacterial cell-surface or the extracellular environment, they must traverse at least one major barrier; the cell envelope (23). Gram-negative bacteria have evolved at least seven distinct, specialized secretion pathways in order to navigate proteins through the inner membrane (IM), periplasm and outer membrane (OM) of the Gram-negative cell envelope (12, 23, 24). These pathways, numbered I-VII, utilize specialized machineries that vary greatly between the different secretion systems (12, 22, 24). In type I and II secretion transports, effector molecules across the IM and OM to the extracellular environment in a continuous Sec-independent and two-step Sec-dependent manner respectively (22). Type III and IV secretion transports effectors across the IM , OM and often across the host cell membrane via formation of needle-like apparatuses (22). Type VI secretion, unlike the other secretion pathways, has been demonstrated to secrete effectors involved in interbacteria interactions as well as pathogenesis, and type VII secretion is important for mycobacterial host-pathogen interactions (6, 57, 67). Lastly, the simplest and most widespread secretion pathway is the type V secretion system, which is subdivided into three branches: (Va) classic monomeric autotransporter (AT), (Vb) two-partner secretion system, and (Vc) trimeric AT adhesin (TAA) (3, 21, 22, 34). ATs represent the largest class of secreted proteins in Gram-negative bacteria and numerous clinically important pathogens possess monomeric ATs including pathogenic *Salmonella*, *Neisseria*, *Bordetella*, *Chlamydia*, *Helicobacter*, and *Yersinia* species (3, 8, 22, 33). The majority of ATs characterized to-date are virulence factors and during infection they have a wide range of virulence functions including protease activity, cytotoxic activity, adherence, biofilm formation, intracellular mobility and invasion (3, 8, 20, 22).

1.2 Classic model of autotransporter secretion

ATs are structurally characterized by the presence of three distinct domains; an N-terminal signal sequence (SS), a passenger domain that carries out the virulence function(s) and a conserved C-terminal translocation unit (TU) consisting of a β-domain and an α-helical linker

region that joins the β-domain to the C-terminus of the passenger domain (Fig. 1) (3, 22, 23). In the classic model of AT secretion, proposed by Pohlner *et al*. for the secretion of *Neisseria gonorrhoeae* IgA protease (IgAP), the N-terminal signal sequence targets the N-terminus of the polypeptide to the inner membrane (Fig. 2i) and facilitates the export of the remainder of the polypeptide to the periplasmic space via the Sec pathway (Fig. 2ii) (11, 23, 55). The C-terminal β-domain then forms a porin-like structure in the outer membrane (OM) through which the passenger domain is thought to be extruded (Fig. 2iii) (23, 55). At the bacterial cell surface, the majority of ATs are cleaved from the TU and either remain associated with the bacterial cell surface or are released into the extracellular environment (Fig. 2iv) (3, 44, 75). Despite the fact that the first AT was described more than 20 years ago, numerous aspects of AT secretion, including the mechanisms of passenger domain translocation across the OM and passenger domain folding at the OM, have yet to be fully elucidated (8, 44, 75).

1.3 Autotransporter structure

1.3.1 Structure of the translocation unit

To-date, the crystal structures of the β-domains of five monomeric ATs have been solved. The first to be solved was the structure of the TU of NalP of *Neisseria meningitidis* by Oomen *et al*., followed by Hbp of *E. coli*, EspP of *E. coli*, EstA of *E. coli* and recently the β-domain of BrkA of *Bordetella pertussis* (Fig. 3) (2, 52, 53, 73, 82). Despite their diverse bacterial origins, the β-domains are all ~30 kDa and share a common structure of a 12 stranded β-barrel that are all almost perfectly superimposable (8, 44, 75). The 12 strands of the β-barrel are connected by extracellular loops and periplasmic turns of varying length (44, 52, 75). Similar to other β-barrel OM proteins, the β-strands of the AT β-barrel are amphipathic with hydrophilic residues that form a hydrophilic pore and hydrophobic residues that face the OM lipids (44, 52, 75). The pore of the β-barrel is narrow and has a measured diameter, determined from crystal structures, of ~10 Å (52). The AT β-domains also have an α-helix (α-helical linker) that spans the length of the pore of the β-barrel and is joined to the β-barrel by a periplasmic loop (52, 53, 73). After solving the crystal structure of the EspP β-domain, Barnard *et al*. described a mechanism where the short α-helix that remained following intra-barrel cleavage of the α-helical linker served to plug the β-barrel pore and stabilize its structure (2). Additionally, electrophysiological studies of SphB1of *B. pertussis*, as well as antibiotic sensitivity studies and dynamic simulations of NalP,

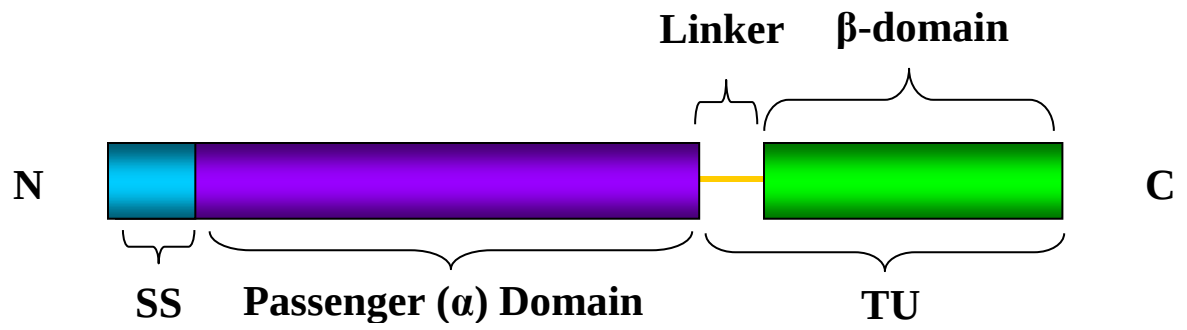


Figure 1. Organization of a typical autotransporter. A typical AT has a N-terminal signal sequence (SS), a passenger (α) domain and a C-terminal translocation unit (TU) consisting of an α-helical linker region and a β-domain that forms a β-barrel in the OM during secretion.

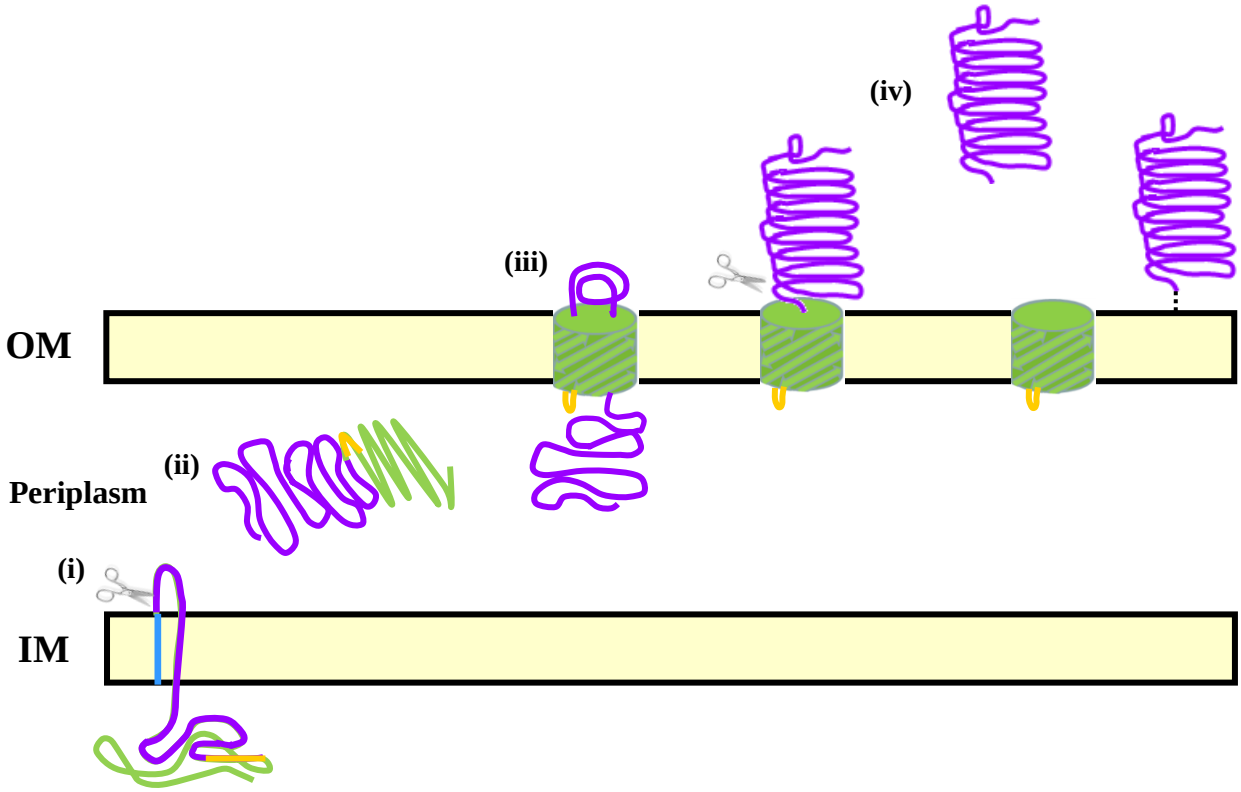


Figure 2. Classic model of AT secretion. (i) The signal sequence directs the N-terminus of the AT to the inner membrane (IM). (ii) The remainder of the polypeptide is exported to the periplasmic space via the Sec pathway. (iii) The C-terminal β-domain of the TU forms a β-barrel in the OM through which the passenger is extruded. (iv) The majority of ATs are cleaved from the TU and are released or remain associated with the bacterial cell surface. Modified from Pohlner *et al.*(55).

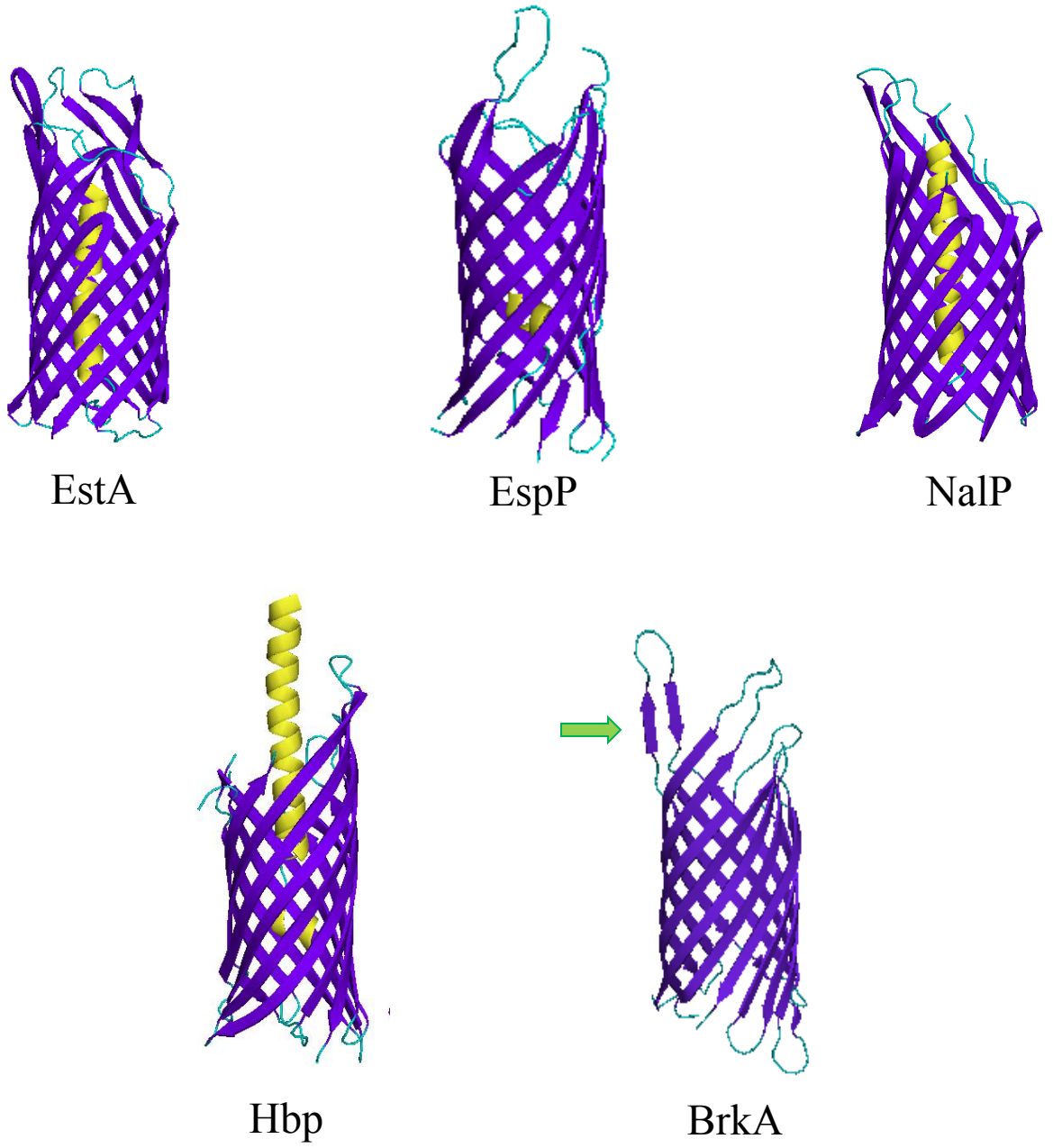


Figure 3. X-ray crystallography structures of all solved autotransporter β-domains. The crystal structures of five AT β-domains have been solved: EstA of *E. coli* (Protein Data Bank (PDB) ID 3KVN), EspP of *E. coli* (PDB ID 2QOM), NalP of *N. meningitides* (PDB ID 1WXR), Hbp of *E. coli* (PDB ID 3AEH), and of BrkA of *B. pertussis* (PDB ID 3QQ2). The β-domains are ~30 kDa and share a common structure of a 12 stranded β-barrel that are all almost perfectly superimposable. The 12 strands of the β-barrel are connected by extracellular loops and periplasmic turns of varying length. The α-helix of the TU passes through the pore formed by the β-domain β-barrel. The green arrow indicates L4 of the BrkA β-domain. β-sheets are colored purple, coil is colored aqua and α-helices are colored yellow. Images were generated by PyMOL v0.99.

suggest that the α-helical linker serves to both stabilize the β-domain structure and effectively close or plug the β-domain pore following passenger domain translocation across the OM (9, 34). Lastly, the BrkA β-domain crystal structure revealed an extracellular loop and hydrophobic pocket that is conserved among many AT β-barrels and is proposed to interact with the C-terminus of the passenger domain (Fig. 3) (39, 82).

1.3.2. Passenger domain structure

The crystal structures of multiple AT passenger domains have been solved including pertactin of *B. pertussis*, Hbp of *E. coli*, IgAP of *Haemophilus influenzae*, VacA of *Helicobacter pylori*, EspP of *E. coli* and EstA of *E. coli* (Fig. 4) (13, 16, 28, 35, 53, 73). Analysis of these structures has revealed that, with the exception of EstA, the passenger domain structures possess a common overall topology involving an extended right-handed parallel β-helical stem upon which other functional domains are attached (28, 35, 44, 53). Crystallographic analysis has also verified β-helix structure within the passenger domains of IcsA of *Shigella flexneri* and BrkA (40, 83). AT passenger domains are variable in size and sequence but the majority are predicted to possess extensive β-helical structure (32, 33, 44). The β-helix is a remarkably stable fold of β-strands and is not unique to prokaryotes and is also found in P22 tail-spike protein, members of the pectate lyase superfamily and a domain of the HET-s prion protein (70). However, the passenger domain crystal structures reveal that AT β-helices are less regular than the straight super helices of pectate lyases and other β-helix proteins found in nature (53).

The β-strands of the β-helix are joined by short loops and three β-sheets form a triangular rung of the helix (28, 35, 44, 53). N-terminal loops and domains of AT passengers make functional contributions to virulence (49). The β-strands of the β-helix are often comprised of six amino acid residues containing alternating non-polar and polar residue repeats. These repeat motifs are conserved and can be used to identify AT passenger domains in database searches (44, 49). The core of the helix is hydrophobic and is mostly filled with aliphatic and aromatic side chains (35, 53). The β-helix is stabilized by the presence of N and C-terminal caps. Crystal structures of pertactin, Hbp, EspP, VacA and IgAP passenger domains revealed that the β-helix is capped by a β-hairpin structure at the C-terminus of the passenger (49, 53, 58). The crystal structure of the VacA passenger also revealed the presence of several short α-helices at the C-

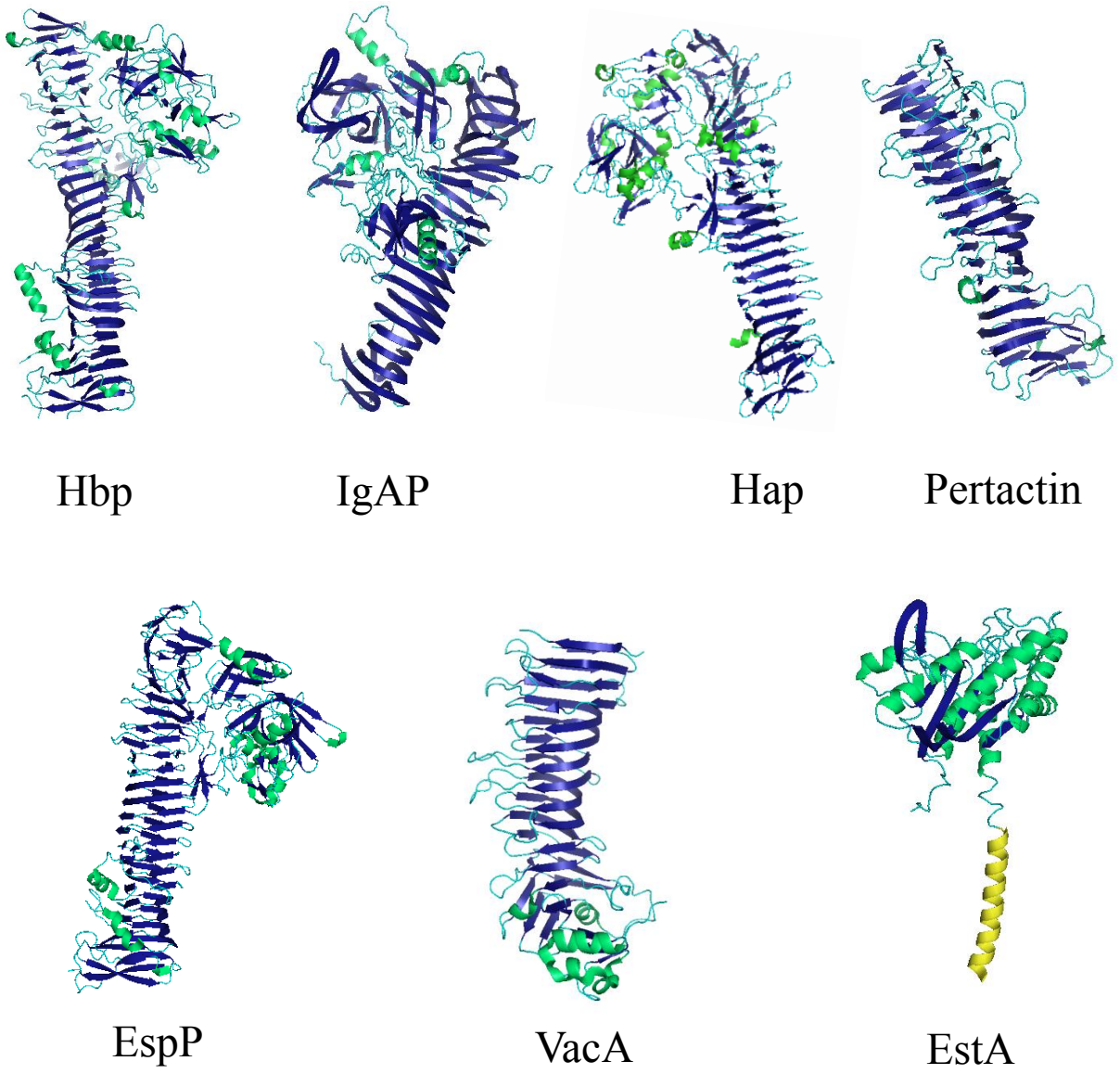


Figure 4. X-ray crystallography structures of all solved autotransporter passenger domains. The crystal structures of seven AT passenger domains have been solved: Hbp of *E. coli* (PDB ID 1WXR), IgAP of *H. influenza* (PDB ID 3H09), Hap of *H. influenza* (PDB ID 3SYJ), pertactin of *B. pertussis* (PDB ID 1DAB), EspP of *E. coli* (PDB 3SZE), VacA of *H. pylori* (PDB ID 2QV3), and EstA of *E. coli* (PDB ID 3KVN). With the exception of EstA, the passenger domain structures possess a common overall topology involving an extended right-handed parallel β-helical stem upon which other functional domains are attached. β-sheets are colored blue, coil is colored aqua and α-helices are green. The α-linker of the EstA TU is colored yellow. Images were generated by PyMOL v0.99.

terminus of the passenger that are absent from the other passenger domain structures (16). With the exception of pertactin and VacA, the passenger domains with solved crystal structures possess an N-terminal domain with α-helical structure (13, 16, 28, 35, 53, 73).

The majority of passenger domains are cleaved from the TU following OM translocation. However, cleavage of the passenger domain can occur in a variety of locations and mechanisms depending on the AT, including proteolytic cleavage of the passenger by a dedicated protease, intermolecular autocatalytic release, intramolecular cleavage within the channel of the β-domain and autocatalytic intramolecular cleavage (9, 44, 63). Therefore, the final secreted product of different ATs possess varying amounts of the passenger domain C-terminus and the α-helical linker of the TU depending on the site of cleavage. It should be noted that a few ATs, including Vag8 of *B. pertussis*, are not cleaved following OM translocation and remain directly associated with the β-domain of the TU via the α-helical linker (44, 71).

1.4 Autotransporter secretion

1.4.1 Translocation across the inner membrane and transit through the periplasm

The signal sequence at the N-terminus of the AT polypeptide targets the polypeptide to the Sec translocon in the IM, which catalyzes the energy-driven export of the polypeptide into the periplasm (66, 72). The majority of ATs possess a typical signal sequence consisting of a charged N domain, a hydrophobic H domain and a C domain which serves as the signal peptidase recognition site (8, 19, 22, 44). A small subset of ATs possess elongated signal sequences resulting from an amino terminal extension called the extended signal peptide region, and studies suggest that extended signal peptides may function to slow translocation of ATs across the IM in order to prevent accumulation of misfolded AT polypeptides within the periplasm (44, 72). Following export, the signal peptide is cleaved from the remainder of the AT polypeptide (8, 44, 63). Periplasmic chaperones likely mediate transit of the AT polypeptide through the periplasm and are important for the correct insertion of the β-domain β-barrel into the OM (69, 80). EspP was the first AT shown to directly interact with the periplasmic chaperones SurA, Skp and DegP (61). Other ATs, including Hbp and IcsA, have been demonstrated to at least transiently interact with periplasmic chaperones (64, 80). Periplasmic chaperones are also thought to be involved in maintaining the passenger part of the polypeptide in a secretion competent state and protecting it from degradation by DegP (44, 69).

1.4.2 Insertion of the β-domain β-barrel in the outer membrane

The name 'autotransporter' was derived from the supposition that an AT protein autonomously catalyzes the transport of its own passenger domain to bacterial cell surface (55). However, the Bam complex is required for the insertion of integral β-barrel OM proteins into the OM and has been demonstrated to be critical or required for the biogenesis of multiple ATs including AIDA-I, Pet, IgAP, BrkA and IcsA (27, 60, 63, 79). Multiple studies have also provided evidence for the direct interaction of ATs with BamA of the Bam complex (45, 61, 64). Photocrosslinking experiments by Ieva *et al*. have also identified potential interactions between EspP and BamD (25). How BamA and BamD mediate the incorporation of the β-domain β-barrel into the OM lipid bilayer is unknown and it is unclear whether ATs rely on the Bam complex solely for β-domain insertion in the OM or if components of the Bam complex also play a direct role in transport of the AT passenger domain across the OM (25, 44).

1.4.3 Translocation across the outer membrane

1.4.3.1 Requirement of the β-domain

It is generally accepted that the AT β-domain is required for passenger domain translocation across the OM and presentation of the passenger domain at the bacterial cell surface (44, 63, 75). Early studies demonstrated that β-domain deletion abrogates passenger domain transport across the OM (55). Recently, Sauri *et al.* replaced the Hbp β-domain with β-barrels of other OMPs and demonstrated that Hbp passenger processing and secretion were abolished (63). However, whether or not the passenger domain actually translocates across the OM through the β-domain β-barrel pore is under debate. Multiple potential mechanisms of passenger domain translocation across the OM have been proposed and the models that have accumulated a large amount of supporting evidence are the hairpin model and the BamA model(s) (44, 75).

1.4.3.2 The hairpin model

The hairpin model was originally proposed by Pohlner *et al*., with the description of IgAP (55). In this model, translocation of the passenger domain occurs directly through the β-domain β-barrel (55). The passenger domain is transported through the β-domain pore in an unfolded state in a C-to-N-terminal vectorial manner resulting in the formation of a hairpin structure

within the pore (55). One strand of the hairpin remains stationary within the pore as the other strand slides past it while moving through the pore. The C-to-N-terminal vectorial transport of the passenger domain to the bacterial cell surface has been confirmed by multiple studies and experimental approaches (31). However, the dimensions of the β-barrel pore, determined from β-domain crystal structures, indicate that the pore is too narrow to accommodate an α-helix as well as an unfolded sliding strand (34, 52). Additional studies have also demonstrated that passenger domain strands possessing small folded elements, including folded domains of truncated passengers, elements resulting from disulphide bond formation between endogenous passenger cysteines and disulphide bond formation between cysteines of heterologous passengers, are able to efficiently translocate (5, 68, 76, 77). Translocation of folded elements is not compatible with the hairpin model.

1.4.3.3. The Bam models

As mentioned above, the extent of the role of the Bam complex in AT biogenesis is currently a matter of great debate (37, 44, 45). In accordance with the hairpin model, the Bam complex is proposed to function only as an insertase and only be required for targeting and insertion of the AT β-domain into the OM (44). In recent years, with the increasing body of evidence to support the involvement of the Bam complex in AT biogenesis, additional models of OM translocation involving a direct role and requirement of the Bam complex have been proposed (27, 52, 60). In one such model, the β-domain of the AT TU functions to target the AT to the Bam complex which is proposed to actively transport the passenger domain across the OM through a channel inside the Bam complex (52, 68). Another model that has recently gained increasing support proposes a concerted function for the Bam complex and AT β-domain in passenger domain OM translocation (44, 45). In addition to inserting the β-domain β-barrel in the OM, Bam complex machinery would maintain the β-domain in an open, dilated conformation so that translocation of the passenger could occur (25, 45). A recent study by Ieva *et al*. demonstrated that a portion of the EspP passenger is incorporated in the β-domain prior to passenger secretion and that partial assembly of the β-domain of an EspP mutant occurred in the periplasm prior to β-domain integration into the OM (25).

1.5 Passenger domain folding

1.5.1. Tolerating folded elements during secretion

Given the small size of the β-domain β-barrel pore, it was originally hypothesized that the passenger domain polypeptide was maintained in an unfolded and linear state in the periplasm, or acquired an unfolded conformation, prior to translocation across the OM (34, 51). In accordance with this hypothesis, numerous studies, including a cysteine mutagenesis study in which disulphide bond formation between Hbp passenger subdomains blocked secretion, indicate limited tolerance toward folded elements during OM translocation (29-31, 36). However, a portion of the EspP passenger has been demonstrated to partially fold in the periplasm and a portion of the *E.coli* Ag43 passenger has been demonstrated to acquire partial resistance to degradation by proteases in the periplasm suggesting it partially folds prior to OM translocation (38, 78). Additionally, multiple studies, most involving disulphide-bond formation, have demonstrated that folded elements are tolerated during OM secretion (5, 68). Disulphide bond formation occurs in the periplasm and the presence of passenger domain disulphide bonds has been verified for several passenger domains (16, 43). Clarification was provided by a recent study by Leyton *et al.* in which a pair of endogenous and closely spaced cysteine residues of the *E. coli* AT Pet was used to determine the effect of disulphide-bond-induced folding on passenger domain OM translocation (45). It was revealed that passenger domain cysteine residues are intrinsically closely spaced to prevent congestion of the β-domain pore following disulphide-bond formation in the periplasm (45). Additionally, rigid structural elements are not formed between these closely spaced endogenous passenger cysteines which is why passenger domain disulphide bonds are tolerated during OM translocation (45).

1.5.2 Rapid folding of the passenger domain at the bacterial cell surface

Taking into account the predominantly unfolded nature of the passenger domain during OM translocation, rapid folding of the passenger domain at the bacterial cell surface must take place in order to protect the passenger domain against degradation by OM and extracellular proteases (51). Numerous studies have provided evidence for the C-terminal to N-terminal vectorial secretion of the passenger domain across the OM (31, 58). In one such study by Junker *et al*., a pair of cysteine residues was introduced into the passenger domain of pertactin resulting in a stalled translocation intermediate in which the C-terminus of the passenger was

demonstrated to be exposed at the bacterial cell surface and translocation of the N-terminal portion of the pertactin passenger was demonstrated to follow translocation of the C-terminal portion (31). Additionally, pertactin and Pet have been demonstrated to possess a stable folding core at the C-terminus of their passenger domain and an N-terminus with low autonomous folding propensity suggesting C to N-terminal sequential folding (32, 58). Sequential secretion has also been demonstrated for EspP, Hbp and Pet where secretion of the passenger N-terminus is dependent on secretion of the C-terminus (54, 59, 70). It has been suggested that folding of the C-terminus of the AT passenger domain at the bacterial cell surface, prior to secretion of the N-terminal portions, prevents backsliding of the passenger back into the periplasm during OM translocation (31, 32, 44). It has also been suggested that the free energy released upon folding of C-terminal passenger residues at the bacterial cell surface could provide a driving force for efficient OM secretion by providing a template structure that promotes fast and efficient folding of subsequently secreted N-terminal portions of the passenger domain (31, 42, 59, 70).

1.5.3 The Autochaperone

1.5.3.1 Discovery of the Autochaperone

Almost two decades ago, Ohnishi *et al*. identified a region at the C-terminus of the *Serratia marcescens* SSP passenger domain required for SSP passenger function and resistance to proteases (50). Interestingly, the C-terminus of the passenger was also able to rescue, *in trans*, the folding of a SSP variant lacking part of that region. It was suggested that the C-terminal region of the passenger played a role in guiding the folding of SSP into its native conformation (50). Oliver *et al*. identified a similar region of ~ 100 residues at the C-terminus of the BrkA passenger domain required for passenger domain folding and capable of *in trans* rescue of the folding of a BrkA variant lacking that region (51). This region has since been termed the autochaperone (AC) as a result of its intramolecular chaperone-like ability to promote the folding of the remainder of the passenger domain (10). Similar to intramolecular chaperones, the BrkA AC is translated as part of the protein it promotes the folding of (46). However, the BrkA AC remains a component of the passenger following passenger domain folding (51). Therefore, it is not a typical intramolecular chaperone as it does not function as an inhibitory pro-region that requires removal to allow folding to occur (46).

1.5.3.2 The role of the autochaperone

The BrkA AC region is conserved in a large group of unrelated ATs having diverse functions, and is always located at or in the vicinity of the C-terminus of the AT passenger domain (Fig. 5) (51). Given the location of the AC, and the C to N-terminal vectorial nature of passenger secretion, the AC would emerge early from the β-barrel (51). Therefore, the AC may act as a template that initiates or nucleates passenger domain folding as the remainder of the passenger is extruded across the OM onto the bacterial cell surface (Fig. 6) (24, 31). Pfam, a database of protein families, recognizes protein family (PF) 03212 from the primary sequence of the C-terminus of the BrkA passenger domain. The BrkA AC region is included within PF03212. The ATs IcsA and pertactin both possess the PF03212 motif and the crystal structure of this AC region of IcsA has been solved and characterized (40). The ACs of pertactin and IcsA, similar to the AC of BrkA, have been demonstrated to form stable β-helical structures (13, 40). The importance of the IscA AC for passenger domain secretion has also been demonstrated by linker mutagenesis studies (48).

While many ATs possess PF03212, many characterized ATs including AIDA-I, Pet, EspP, Hbp and SSP are not recognized by Pfam as possessing a protein family motif at the C-terminus of their passenger domains. However, the C-terminal regions or residues of these AT passenger domains lacking PF03212 have been demonstrated to play a role in AT secretion and or passenger domain folding (11, 50, 70). Linker mutagenesis of the Pet passenger resulted in passenger domain proteins that were secreted but degraded by proteases and, similar to BrkA and SSP, presentation of the C-terminal 200 residues of the EspP passenger at the bacterial cell surface was able to rescue the stability/secretion of some of the Pet mutants (11). Mutations in the EspP passenger C-terminus had drastic effects on passenger domain OM translocation (78). Similarly, in the Hbp passenger the last rung of the β-helix and a cap region were identified as the AC motif and the mutation of several conserved aromatic residues affected secretion and β-domain conformation (70). The results of these studies support a conserved role for the C-terminal passenger domain residues in AT biogenesis.

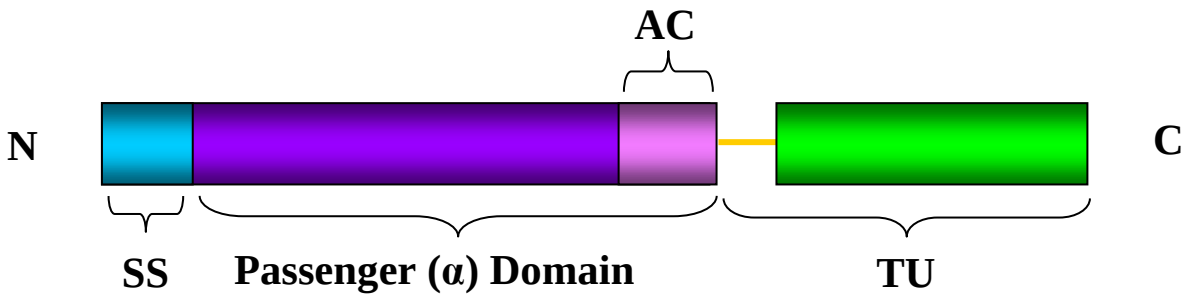


Figure 5. Location of the BrkA AC within the passenger domain. The autochaperone (AC) (pink) of BrkA consists of Glu^{601}-Ala^{692} and is located at the C-terminus of the passenger domain.

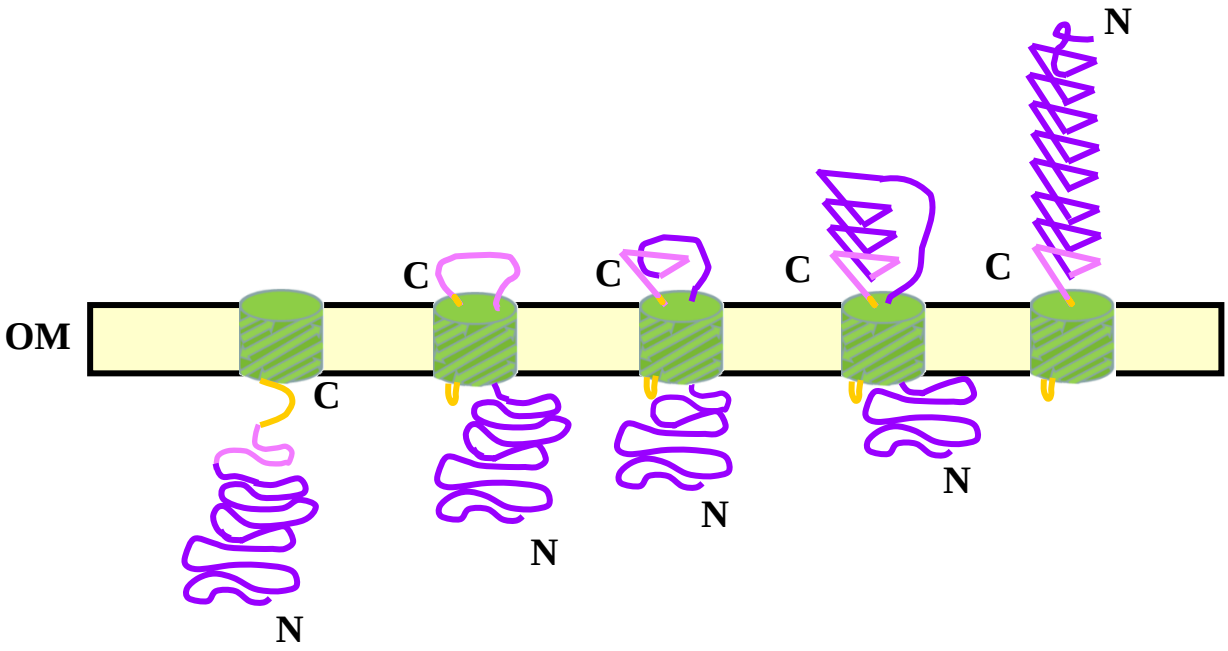


Figure 6. **Model of autochaperone mediated passenger domain folding during OM secretion.** The α-helical linker region initiates translocation of the passenger domain across the outer membrane in a C→N vectorial manner. After the AC emerges at the bacterial surface it assumes a conformation that acts as a template that initiates the folding of the remainder of the passenger domain as it is transported across the OM through the amphipathic β-barrel formed by the β-domain.

1.6 Project Overview

1.6.1 Purpose and Objectives

Despite extensive work to elucidate the role of conserved C-terminal passenger domain regions in AT biogenesis, the exact role of the AC regions in OM translocation and passenger domain folding has yet to be clarified and the structural folding mechanism has only been studied for a few ATs. Early experimentation with BrkA and Pet suggests a crucial role for the AC in passenger domain folding, whereas studies with EspP and recent work with Hbp and IcsA focus on the involvement of the C-terminus of the passenger domain in secretion across the OM rather than its role in passenger domain folding. No study to-date has investigated the effect of regions of α-helical structure present at the C-terminus of some β-helix passenger domains, N-terminal to the α-linker of the TU as seen in the VacA crystal structure and predicted in various other AT passengers, on the ability of the C-terminus of the passenger domain to act as a template to promote passenger domain folding. Additionally, the role of conserved C-terminal passenger domain regions in the ability of some AT passenger domains, such as Ag43, to partially fold in the periplasm has yet to be addressed.

The purpose of this study was to determine the requirement of conserved C-terminal regions of AT passenger domains for passenger domain folding *in vitro*. The specific objectives of the project were to (1) determine the requirement of these C-terminal regions for the acquisition of passenger domain secondary structure and (2) test the interchangeability of conserved C-terminal regions in passenger domain folding.

1.6.2 Hypotheses

Bioinformatic and phylogenetic analyses of the C-terminal sequences of over 200 AT passenger domains identified numerous putative AC motifs (S. Pleasance and R. Fernandez, unpublished). These putative AC motifs are categorized into three groups: (1) putative ACs of BrkA and related ATs, (2) putative ACs of Chlamydia ATs and (3) putative ACs of protease ATs (Fernandez lab, unpublished data). The BrkA and related AT group is further organized into three subgroups: BrkA/pertactin-like motifs (PF03212), AIDA-I-like AC motifs and EspP/Pet-like AC motifs. Based on these groupings of putative AC motifs and secondary structure prediction analysis of these motifs, two hypothesises were tested: (1) conserved C-terminal passenger domain motifs, similar to the BrkA AC, are required for the acquisition of secondary

structure by the passenger domain, and (2) an AC motif will only be able to rescue BrkA passenger domain folding if it is a member of the BrkA group of ACs.

To test the first hypothesis, ATs were chosen to represent two of the groups of putative ACs. Full-length and AC-deleted passenger constructs were cloned and the expressed proteins were purified under denaturing conditions and the ability of purified proteins to refold following dialysis was assessed by far-UV circular dichroism spectroscopy and limited proteolysis with trypsin.

Two *in vitro* experimental approaches were used to test the second hypothesis and assess the ability of different AC motifs to rescue the folding of AC-deleted BrkA passenger protein. The first approach tested the ability of AC motifs present in cell wall fractions to rescue AC-deleted BrkA passenger folding. Rescue of folding was detected as an increase in resistance to digestion with trypsin. For the second approach, passenger domain constructs encoding AC-deleted BrkA cloned in-frame with the ACs of representative ATs were expressed and the ability of purified fusion proteins to fold was assessed by far-UV circular dichroism spectroscopy and trypsinolysis.

2. MATERIALS AND METHODS

2.1 Bacterial strains and growth media

The *Escherichia coli* strains used in this study are BL21 (DE3) and DH5α and were purchased from Invitrogen (Burlington, ON). Expression strains were obtained by transforming pET30b and pET20b (Novagen, Billerica, MA) constructs into BL21 (DE3). The DH5α strain was used for cloning. *E. coli* strains were cultured at 37°C on Luria agar or Luria broth supplemented with the appropriate antibiotic. Ampicillin was used at a concentration of 100 µg/mL and kanamycin was used at a concentration of 50 µg/mL.

2.2 Molecular biology techniques

All DNA manipulations were performed using standard techniques (62). Restriction enzymes, DNA modification enzymes and DNA polymerases were purchased from New England BioLabs (Ipswich, MA). Primers were purchased from Alpha DNA (Montreal, QC). DNA sequencing was carried out at Genewiz (South Plainfield, NJ). All cloning PCR was performed using Phu polymerase with the following cycles: 1 cycle for 2.5 min at 98°C, 30 cycles of 15 sec at 98°C followed by 35 sec at annealing temperature followed by 30 sec/kb at 72°C, and finally 1 cycle of 72° C for 10 min. Amplified PCR products were separated on an agarose gel and a band of the expected size was gel extracted and column purified using the BioBasic (Markham, ON) Spin Column Gel Extraction Kit. The purified DNA was digested with the appropriate restriction enzymes and column purified using the BioBasic Spin Column PCR Purification Kit. Purified restriction digested PCR products were then ligated into the multiple cloning site (MCS) of the appropriate vector using T4 DNA Ligase. Primers used in this study are listed in Table 1 and all constructs made and used in this study are listed in Table 2.

2.2.1 Ag43 expression constructs

All Ag43 PCR was carried out using *E. coli* K12 TB1 genomic DNA as the template. *E. coli* K12 genomic DNA was purified using the DNeasy Blood & Tissue Kit (Qiagen, Valencia, CA). Ag43 Full-length (FL) and AC-deleted passengers (ΔAC-Pass1-3) were amplified using primer 11 as the forward primer and primers 13, 12, 21 and 22 as reverse primers respectively. All PCR products were digested with *Eco*RI/*Hind*III and ligated into *Eco*RI/*Hind*III digested pET30b. The AC and TU (AC+TU) construct insert was amplified using primer 23 as the

Table 1. **Primers used in this study**

Primer Number	Primer Name	Primer Sequence
1	BPvag8fw8	AAGGATCCGGTCACGGCAGCGCAGCG
2	Vag8pass-AC_R1	CGAAGCTTATCACCGGGAATCGCCCGTCACG
3	Vag8pass-AC_R2	CGAAGCTTATCAGGGCGAACCGTCCGGCTGC
4	BrkA 52-600_R HindIII	AATAAGCTTTTAGGGCGCGGGCGCCT
5	BrkA _R XhoI	AATCTCGAGTCAGAAGCTGTAGCGGTAG
6	BrkA Pass-SS (G1NCO)_F NcoI	TCAGTCCATGGCGCAGGAAGGAGAGTTCGAC
7	PrtS Pass-SS F_SacI	ATAGAGCTCCTTAGGTTCGCCCGAC
8	PrtS ΔAC645R_XhoI	ATTCTCGAGCGCCACGATGCTCGCCTC
9	PrtS ΔAC706R_XhoI	ATTCTCGAGGTCAAACGTGGCGTCGGT
10	PrtS AC+TU F4_NcoI	GCGCCATGGTCGTATCCGTTTGTTAAA
11	Ag43 Pass-SS53_F_EcoRI	CTAGAATTCGGCTGACATCGTTGTGCAC
12	Ag43 ΔAC634_R_HindIII	CTAAAGCTTTCATGCCCTGCCGCCGTCA
13	Ag43 FL Pass706_R_HindIII	CTAAAGCTTTCAATAAGCATTTTCACTGC
14	SPSM MP+JXN+linker_R_XhoI	TAATCTCGAGTCAGCGGCCATCGGCCT
15	SPSM Pass 28-447aa_R_XhoI	TAATCTCGAGTCAGCCCGACGTGCAAG
16	SPSM_R_XhoI	TCGACTCGAGTCAGAAGCGCCAGGT
17	SPSM 528+JXN+TU_F_NcoI	TAATCATGGGCGCGGGAAATGGCTAC
18	SPSM 641JXN+TU_F_NcoI	TCATCCATGGGCGCCACGTTTGACAGC
19	SPSM_F_SacI	TGCAGAGCTCATGATACTTAATA
20	Ag43_R_XhoI	TCGACTCGAGTCAGAAGGTCACATTC
21	Ag43 ΔAC522_R_HindIII	CTAAAGCTTTCAGTTCAGCGTCAGCGTG
22	Ag43 ΔAC383_R_HindIII	CTAAAGCTTTCAATTTCCCATGGATACGG
23	Ag43 523AC+TUp20_F_NcoI	TCACCATGGGCGACAGTACCGTCACCAC
24	BPvag8rev7	CCCAAGCTTACAACTCGTTGGTCGGC
25	BrkA 52-600_R_ XbaI	TAATTCTAGAGGGCGCGGGCGCCTGGA
26	Vag8 AC_F_XbaI	TAATTCTAGAGTGGCCGAGGTCAAG
27	Spm 528+JXN_F_XbaI	TAATTCTAGAGCGGGAAATGGCTAC
28	Vag8 p20b AC+TU_F_NcoI	ATATCCATGGGCGTGGCCGAGGTCAAG
29	Vag8 R with stop_HindIII	GGCAAGCTTTCACCAGCTGTAGCGATA
30	BrkA Np20b AC+TU_F_NcoI	AATACCATGGGCCCCTGACCCTGCAAA

Table 2. Strains and plasmids used in this study

*E. coli*Strain/ Plasmid	Relevant Characteristics/Sequence(s) cloned into MCS	Vector	Source/Reference
Strains			
DH5α	F- φ80lacZΔM15 Δ(lacZYA-argF) U169 recA1 endA1 hsdR17 (rk-, mk+) phoA supE44 λ- thi-1 gyrA96 relA1	N/A	Invitrogen
BL21 (DE3)	F- ompT hsdSB (rB-mB-) gal dcm (DE3)	N/A	Invitrogen
Plasmids			
pDO418	BrkA E^{61}-V^{699}	pET30b	(51)
pDO-JB5	Met^{1}–Ala^{42}+Glu^{693}–Phe^{1010}	pBS SKII	(51)
pBrkAΔACPass	BrkA E^{61}-P^{600}	pET30b	This study
pBrkAAC+TU	BrkA L^{607}-F^{1010}	pET20b	This study
pVag8-22	FL Vag8 AT	pBS SKII	(47)
pVag8FLPass	Vag8 V^{40}-R^{610}	pET30b	(47)
pVag8ΔACPass1	Vag8 V^{40}-R^{479}	pET30b	This study
pVag8ΔACPass2	Vag8 V^{40}-P^{438}	pET30b	This study
pVag8AC+TU	Vag8 V^{40}-W^{915}	pET20b	This study
pSmpFLPass	Smp Y^{29}-A^{706}	pET30b	This study
pSmpΔACPass1	Smp Y^{29}-D^{645}	pET30b	This study
pSmpΔACPass2	Smp Y^{29}-G^{447}	pET30b	This study
pSmpFLPass+linker	Smp Y^{29}-R^{763}	pET30b	This study
pSmpAC+TU1	Smp D^{641}-F^{1045}	pET20b	This study
pSmpAC+TU2	Smp A^{528}-F^{1045}	pET20b	This study
pSmppBSSK	FL Spm AT	pBS SKII	This study
pPrtSPass	Smp M^{1}-S^{731}	pUC57	This study
pPrtSAC+TU	Smp S^{618}-F^{1045}	pUC57	This study
pAg43FLPass	Ag43 A^{53}-L^{706}	pET30b	This study
pAg43ΔACPass1	Ag43 A^{53}-A^{634}	pET30b	This study
pAg43ΔACPass2	Ag43 A^{53}-N^{522}	pET30b	This study
pAg43ΔACPass3	Ag43 A^{53}-N^{383}	pET30b	This study
pAg43AC+TU	Ag43 D^{523}-F^{1039}	pET20b	This study
pBrkA-Vag8Pass	BrkA Pass E^{61}-P^{600}+ Vag8 R^{479}-R^{610}	pET30b	This study
pBrkA-SmpPass	BrkA Pass E^{61}-P^{600}+ Smp A^{528}-A^{706}	pET30b	This study

forward primer and primer 30 as the reverse primer. The PCR product was digested with *Nco*I/*Xho*I and ligated into *Nco*I/*Xho*I digested pET20b.

2.2.2 Vag8 expression constructs

All Vag8 PCR was carried out using pVag8-22 as the template. Vag8 ΔAC-Pass1 and 2 inserts were amplified using primer 1 as the forward primer and primers 2 and 3 as the reverse primers respectively. The PCR products were digested with *Hind*III/*BamH*I and ligated into *Hind*III/*BamH*I digested pET30b. The Vag8 AC+TU construct insert was amplified using primer 28 as the forward primer and primer 29 as the reverse primer. The PCR product was digested with *Nco*I/*Hind*III and ligated into *Nco*I/*Hind*III digested pET20b.

2.2.3 Smp expression constructs

Smp PCRs used pSmppBSSK as a template. pPrtSPass and pPrtSAC+TU encode the N-terminal 731 and C-terminal 427 residues respectively of Smp and were purchased from Genescript (Piscataway, NJ). Primer pairs 14/19 and 10/16 were used to amplify the pPrtSPass and pPrtSAC+TU inserts respectively. The resulting PCR products were digested with restriction enzyme pairs *Sac*I/*Pst*I and *Xho*I/*Pst*I respectively. These two products were ligated into *Sac*I/*Xho*I digested pBS SKII in a three way ligation reaction to yield pSmppBSSK. Smp FL Pass, FL Pass+Linker and ΔAC-Pass1 and 2 inserts were amplified from pSmppBSSK using primer 7 as the forward primer and primers 9, 14, 8 and 15 as the reverse primers respectively. PCR products were digested with *Sac*I/*Xho*I and ligated into *Sac*I/*Xho*I digested pET30b. Smp AC+TU1 and 2 inserts were amplified from pSmppBSSK using primer 16 as the reverse primer and primers 17 and 18 as the forward primers respectively. The PCR product was digested with *Nco*I/*Xho*I and ligated into *Nco*I/*Xho*I digested pET20b.

2.2.4 BrkA expression constructs

The BrkA AC-deleted passenger and AC+TU PCRs were carried out using pDO418 and pDO-JB5 respectively as templates. The BrkA ΔAC-Pass PCR used primer 6 as the forward primer and primer 4 as the reverse primer. The PCR product was digested with *Nco*I/*Hind*III and ligated into *Nco*I/*Hind*III digested pET30b. The BrkA AC+TU PCR used primer 30 as the

forward primer and primer 15as the reverse primer. The PCR product was digested with *Nco*I/*Xho*I and ligated into *Nco*I/*Xho*I digested pET20b.

BrkA fusion passenger constructs were generated by ligating the BrkA ΔAC-Pass insert with the Vag8 or Smp AC fragment into pET30b. The BrkA ΔAC-Pass insert was amplified from pDO418 using primer 6 as the forward primer and primer 25 as the reverse primer, the PCR product was digested with *Nco*I and ligated into *Nco*I/*EcoR*V digested pET30b. The Vag8 and Smp AC insert PCRs used pVag8FLPass and pSmpFLpass plasmids as templates and primer pairs 26/24 and 27/13 respectively. The Vag8 and Smp PCR products were digested with *Hind*III or *Xho*I and ligated into *EcoR*V/*Hind*III and *EcoR*V/*Xho*I digested pET30b. The BrkA ΔAC-Pass insert was digested out of pET30b with *Nco*I and *Xba*I and Vag8 and Smp AC inserts were digested out of pET30b with *Xba*I and *Blp*I followed by ligation of the BrkA ΔAC-Pass fragment with the Vag8 or Smp fragment via a three way ligation with pET30b digested with *Nco*I and *Blp*I.

2.3 Expression of BrkA, Vag8, Ag43 and Smp passenger variants

BrkA, Vag8, Ag43 and Smp full-length and AC-deleted passenger constructs were transformed into BL21 DE3 cells. Cultures of transformants were grown at 37°C to an optical density at 600 nm ($OD_{600)}$ of approximately 0.6 and expression of constructs was induced with 100 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) for 2.5 hrs.

2.4 Purification of BrkA, Vag8, Ag43 and Smp passenger variants under denaturing conditions

50 mL of induced culture was pelleted and prepared for nickel-affinity chromatography according to the Invitrogen Ni-NTA (nickel-nitrilotriacetic acid) Purification System denaturing purification protocol. In summary, culture pellets were resuspended in guanidine hydrochloride buffer followed by sonication on ice with a Sonicator Ultrasonic Processor XL (Mandel Scientific, Guelph, ON) at level 5 for 1 min. Sonicated samples were pelleted by centrifugation and supernatants were filtered using Millex-Hv 0.45 µm PVDF filters (EMD Millipore, Billerica, MA). Filtered samples were allowed to incubate with nickel-coated His Pur^{TM}Ni-NTA Resin (Thermo Fisher Scientific, Rockford, IL) at room temperature for 30 min followed by column draining by gravity and washing with denaturing binding buffer, low stringency wash buffer,

high stringency wash buffer and finally elution with 8 M urea pH4 elution buffer. 10 µL samples of each column elution were diluted in sample buffer and boiled for 5 minutes prior to SDS-PAGE and Coomassie Brilliant Blue staining. The best column elutions were chosen based on analysis of stained gels and dialyzed at 4^{o}C from 8M urea to 10 mM Tris pH8 via gradient dialysis for the step-wise replacement of urea with 10 mM Tris pH 8. For protracted dialysis, the number of buffer exchanges was doubled and the length of each exchange was increased from 2 to 4 hrs. The amino acid sequence of all His-tagged purified proteins can be found in Appendix A.

2.5 Purification of Vag8 passenger variants under hybrid conditions

50 mL of induced culture was pelleted and prepared according to a modified version of the Invitrogen Ni-NTA Purification System hybrid purification protocol. In summary, culture pellets were resuspended in guanidine hydrochloride buffer followed by sonication on ice at level 5 for 1 min. Sonicated samples were pelleted by centrifugation and supernatants were filtered using 0.45 µm filters. Filtered samples were allowed to incubate with nickel-coated resin (Thermo Fisher Scientific) at room temperature for 30 min followed by column draining by gravity and washing with denaturing binding buffer, low stringency denaturing wash buffer, and native wash buffer. Columns were incubated in native wash buffer for 0-4 days followed by elution in native elution buffer (Table 4). 10 µL samples of each column elution were diluted in sample buffer and boiled for 5 minutes prior to SDS-PAGE and Coomassie staining. The best column elutions were chosen based on analysis of stained gels and dialyzed to 10 mM Tris pH8 at 4^{o}C overnight.

Table 3. **Hybrid column protocol optimizations**

Sample	Number of Washes in Native Wash Buffer	Length of Incubation on Column
Hybrid 1	4	N/A
Hybrid 2	8	N/A
Hybrid 3	6	1 day in 6th Wash
Hybrid 4	10	4 days in 8th Wash

2.6 Far-UV circular dichroism spectroscopy of passenger variant proteins

Far-UV circular dichroism (CD) spectroscopic analysis was performed on dialyzed BrkA, Vag8, Ag43 and Smp passenger variants using a Jasco J-810 CD spectropolarimeter (Jasco, Easton, MD) at a temperature of 20° C using a cell path length of 1 mm. The CD spectrometer was blanked with 10 mM Tris pH 8. Purified passenger variants were analyzed at a concentration of 300 µM in 10 mM Tris buffer, pH 8. Three scans of each sample were taken and the spectra averaged. Sample analysis was carried out at the UBC laboratory of molecular biophysics.

2.7 *In vitro* limited proteolysis analysis of passenger variant proteins

100 µL of dialyzed purified passenger variant (300 µg/mL) was used in the limited proteolysis digestions with trypsin. 20 µL was removed and diluted in denaturing sample buffer as a 0 time point control prior to addition of 0.8 µL of 1 mg/mL trypsin. 20 µL samples were removed at 1, 5 and 15 min following addition of trypsin and were immediately diluted in denaturing sample buffer and boiled for 5 minutes prior to SDS-PAGE and Coomassie staining.

2.8 Zymogram

10 µL of dialyzed purified passenger proteins (300 µg/ml) were diluted in non-denaturing sample buffer and run on 1% skim milk 11% polyacrylamide gels for 2 hrs at 120 V. The gel was then rinsed in dH_2O prior to two 20 min washes in Triton-X buffer, a 10 min was in Tris buffer and overnight incubation at 37°C in Tris buffer pH8. Gels were then stained for 2 hrs in Coomassie followed by destaining for 30 min in 7% acetic acid. 1 µL of 0.1 mg/mL trypsin was included as a positive control.

2.9 Cell wall fraction preparations

Overnight cultures were used to inoculate 50 mL cultures. These cultures were grown to an OD_{600} of 0.6 prior to induction with 100 mM IPTG for 2.5 hrs. The induced culture was centrifuged at 6000xg for 5 min followed by resuspension of the pellet in 10 mM HEPES. The mixture was sonicated at level 5 for three 20 second intervals. The sonicated solution was centrifuged at 6000xg for 5 min after which the supernatant was centrifuged at 15000xg for 30 min. The pellet was resuspended in 500 µL of 10 mM HEPES and centrifuged at 14000 rpm for 30 min. The resultant pellet was resuspended in 75 µL of 10 mM HEPES.

2.10 Cell wall fraction folding assay

5 µL of cell wall fraction preparation was mixed with 20 µl of 100 µg/mL purified AC-deleted BrkA passenger domain (dialyzed in 10 mM Tris pH 8) and incubated together overnight at room temperature. The mixture was then diluted in 75 µL of 10 mM Tris resulting in a mixture with a total volume of 100 µL. A 20 µL 0 time point control was removed prior to addition of 0.8 µL of 0.5 mg/mL trypsin. 20 µL samples were removed at 1, 5 and 15 min time points and immediately boiled in denaturing SDS-PAGE disruption buffer. Samples were used for SDS-PAGE followed by Western blotting.

2.11 SDS-PAGE and western-blot analysis

SDS-PAGE was performed as previously described (15) and the separated proteins were visualized by staining with Coomassie. Unstained molecular weight markers were purchased from Thermo Fisher Scientific.

For Western blotting, samples were separated by SDS-PAGE prior to transfer to Immobilon-P membranes (Millipore, Bedford, MA) as previously described. Rabbit anti-His was used as the primary antibody (Santa Cruz Biotechnology, Santa Cruz, CA) and horseradish peroxidase-conjugated goat anti-rabbit (Santa Cruz Biotechnology) was used as the secondary antibody. The primary and secondary antibodies were used at 1:10 000 and 1:35 000 dilutions respectively. The blots were developed with the ECL Western Lightning reagent (PerkinElmer, Waltham, MA). Prestained molecular weight markers were purchased from Thermo Fisher Scientific.

2.12 Milk plate proteolysis test

Luria agar plates were prepared with 3% skim milk powder. 5 µL of ~ 200µg/mL of each protein sample was pipetted onto the plate. Trypsin was included as a positive control. The plate was incubated at 50° C overnight and serine protease activity was detected as a ‘halo’/zone of clearing.

3. THE REQUIREMENT OF PUTATIVE AC MOTIFS FOR THE ACQUISITION OF PASSENGER DOMAIN SECONDARY STRUCTURE

3.1 Vag8, Ag43 and Smp as representative autotransporters

The importance of conserved C-terminal passenger domain residues for passenger domain secretion and folding has been demonstrated for numerous ATs. However, the requirement of the passenger domain C-terminus for the acquisition of passenger domain secondary structure has only been demonstrated for BrkA (51). In the absence of their passenger domain C-terminus, SSP lacks functional protease activity and Pet is sensitive to digestion with proteases suggesting they at least lack native structure, but passenger domain secondary structure was not assessed (11, 50). Furthermore, the majority of studies on AT passenger domains found that manipulation of the passenger domain C-terminus resulted in severe defects in OM translocation. Therefore, the requirement and role of the C-terminus in passenger domain folding could not be determined.

To test the requirement of different putative ACs for the acquisition of passenger domain secondary structure and folding, ATs were chosen to represent two of the groups of putative ACs. Serratia marcescens protease (Smp) of *S. marsescens* was chosen to represent the Protease Group and virulence activated gene 8 (Vag8) of *B. pertussis* and antigen 43 (Ag43) of *E. coli* were chosen to represent the BrkA-like and AIDA-I-like subgroups of the BrkA Group, respectively. Vag8 has been demonstrated to bind C1-inhibitor of the human complement system and is predicted to be similar in structure to pertactin and BrkA (47). The folded Vag8 passenger is resistant to digestion with trypsin and is not cleaved from the TU following OM translocation (47). Ag43 of *E. coli* is a self-associating autotransporter (SAAT) that promotes autoaggregation and biofilm formation and is predicted to possess β-helical structure (74). The N-terminal portion of the Ag43 passenger, Ag43α, is cleaved from the remainder of the passenger domain following OM translocation, and is partly resistant to digestion with trypsin (38). Lastly, Smp is a metallo-protease that shares 88% sequence identity with SSP (41). The 32 kDa N-terminal portion of the Smp passenger is trypsin resistant and cleaved from the remainder of the passenger domain and TU following OM translocation (41). Analyses of BrkA passenger proteins were carried out as controls and for comparison to Vag8, Ag43 and Smp experimental results.

3.2 Identification of conserved C-terminal passenger domain motifs of Vag8, Ag43 and Smp that may function as ACs

Due to the lack of crystallographic analysis of the Vag8, Ag43 and Smp passenger domains, bioinformatic analyses of passenger domain amino acid sequences were carried-out using ProDom, Pfam, SUPERFAMILY and BetaWrap in order to provide insight into Vag8, Ag43 and Smp passenger domain structures (4, 7, 56, 81). Pfam and ProDom are databases of protein families and protein family domains. Pfam and ProDom analyses were used to identify conserved C-terminal regions and motifs of passenger domains that could potentially function as ACs (7, 56). BetaWrap is a program that predicts the right-handed parallel β-helix structural motif in protein amino acid sequences, and SUPERFAMILY is a database that provides protein domain/superfamily assignments where a superfamily is a grouping of domains from different families based on structure, function and sequence data (4, 81). BetaWrap and SUPERFAMILY analyses were used to identify and predict regions of β-helical structure. Lastly, SWISS-MODEL, a web based service that provides protein structure homology modelling, was used to generate protein structure alignment models based on passenger domain primary sequence (1).

3.2.1 BrkA Control

The BrkA AC region is recognized as part of the 'pertactin' family by Pfam (protein family 03212) (Fig. 7A). The BrkA passenger is predicted by BetaWrap analysis to have a high probability of containing β-helix structure (Table 4). The BrkA AC is included within the region of the BrkA passenger identified as belonging to the pectin lyase-like family by SUPERFAMILY and is predicted to be part of a β-helix structure (Fig. 8A). The AC region was also included in the SWISS-MODEL homology model of the C-terminus of the BrkA passenger domain which possessed β-helix structure (Fig. 9A). These predictions coincided with previous analyses and predictions of the BrkA passenger and the known β-helix structure of BrkA determined from crystallographic analysis of the BrkA passenger domain (51, 83).

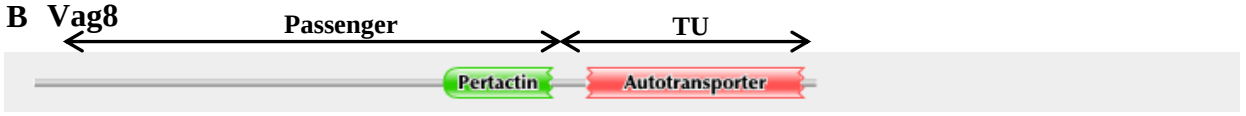
3.2.2 The Vag8 putative AC is the pertactin motif

Vag8 of *B. pertussis* was chosen to represent the BrkA group and subgroup of putative ACs. In addition to identifying the AT β-domain, Pfam also identified the 'pertactin' family from the Vag8 passenger amino acid sequence, indicating that the Vag8 putative AC is a member of the same AC subgroup as the BrkA and pertactin AC (Fig. 7B). The Pfam PF03212 spans

A BrkA

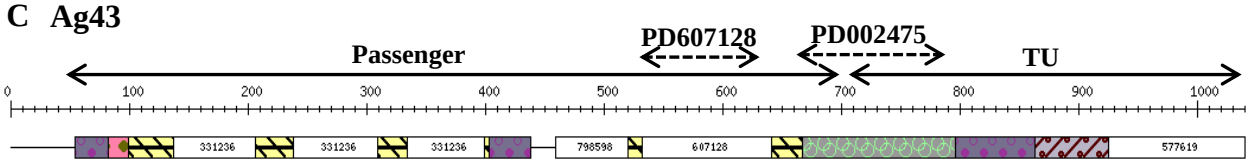
Family	Description	Entry type	Clan	Envelope		Alignment	
				Start	End	Start	End
Pertactin	Pertactin	Family	CL0268	583	704	583	703
Autotransporter	Autotransporter beta-domain	Domain	CL0193	743	996	746	994

B Vag8



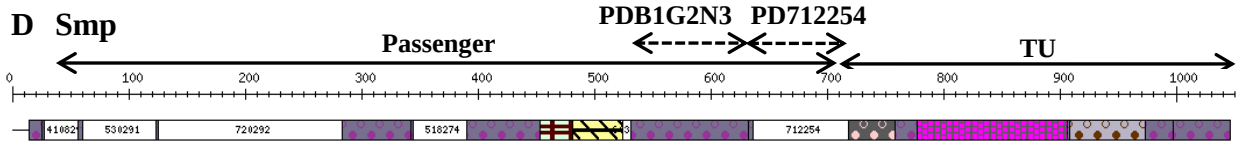
Family	Description	Entry type	Clan	Envelope		Alignment	
				Start	End	Start	End
Pertactin	Pertactin	Family	CL0268	480	606	480	605
Autotransporter	Autotransporter beta-domain	Domain	CL0193	648	901	650	898

C Ag43



Domain ID	Alignment
PD000317	99-402
PD607128	532-640
PD002475	667-795

D Smp



Domain ID	Alignment
PDB1G2N3	531-631
PD712254	636-717
PD011682	718-757

Figure 7. Pfam and ProDom analysis of passenger domain sequences to identify conserved C-terminal motifs. Pfam identified the Pertactin Family (PF03212) from the BrkA and Vag8 primary sequences. Pfam only identified the autotransporter beta-domain at the C-terminus of Ag43 and Smp sequences (not shown). Ag43 and Smp sequences were also analyzed by ProDom which identified multiple conserved domains in the vicinity of the C-terminus of the passenger domains. Locations of domains relative to passenger domain and TU boundaries are indicated with arrows. Analysis was carried out using Pfam 26.0 and ProDom release 2010.1.

Table 4. BetaWrap* analysis of representative autotransporter passenger domains.

Autotransporter	Best Wrap Raw Score	P-value
pertactin	-18.33	0.000018
BrkA	-21.92	0.0048
Vag8	-22.31	0.0075
Ag43	-21.50	0.0029
Smp	-23.13	0.017

*http://groups.csail.mit.edu/cb/betawrap/betawrap.html

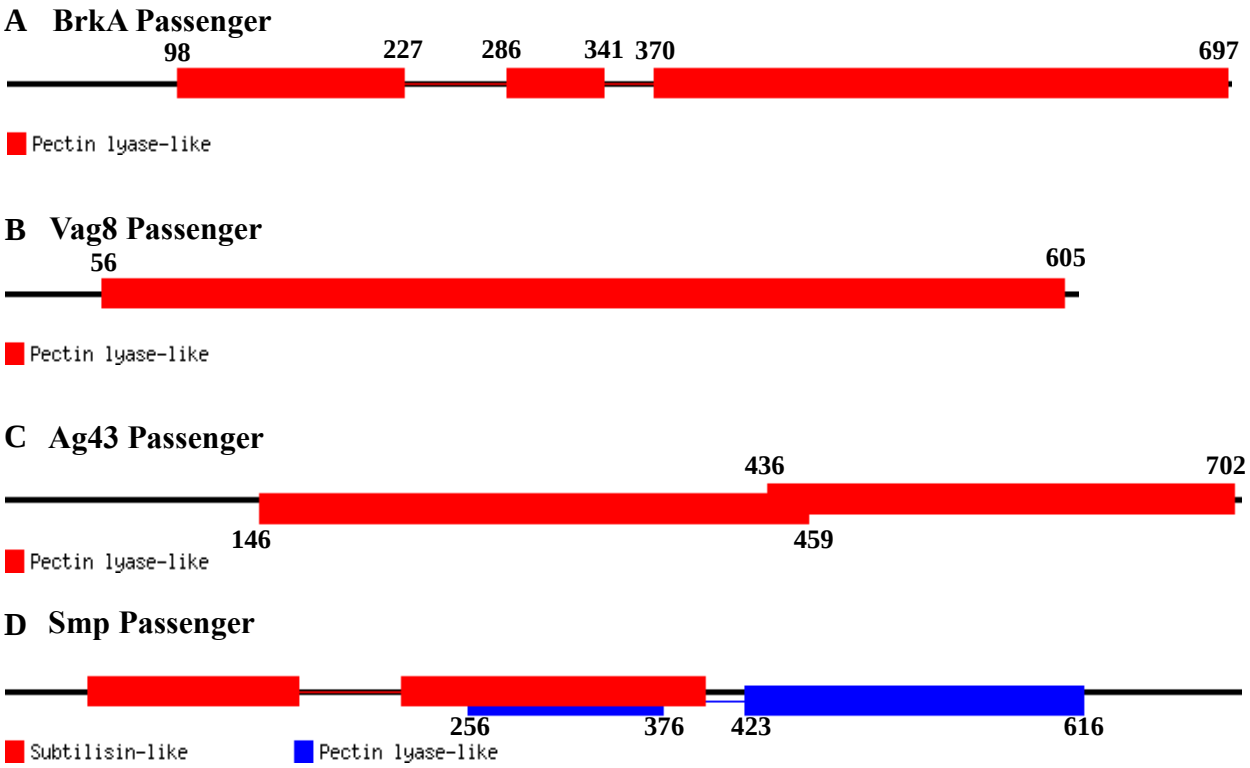


Figure 8. SUPERFAMILY analysis of passenger domain amino acid sequences. BrkA, Vag8, Ag43 and Smp signal sequence and full-length passenger sequences, extending to the alpha-helical linker of the translocation unit, were analyzed by the SUPERFAMILY database. Pectin lyase-like regions were identified in all of the passenger sequences and a subtilisin-like region was also identified within the N-terminal half of the Smp passenger. The boundaries of each Pectin lyase-like region are indicated above or below the SUPERFAMILY output. SUPERFAMILY 1.75 was used.

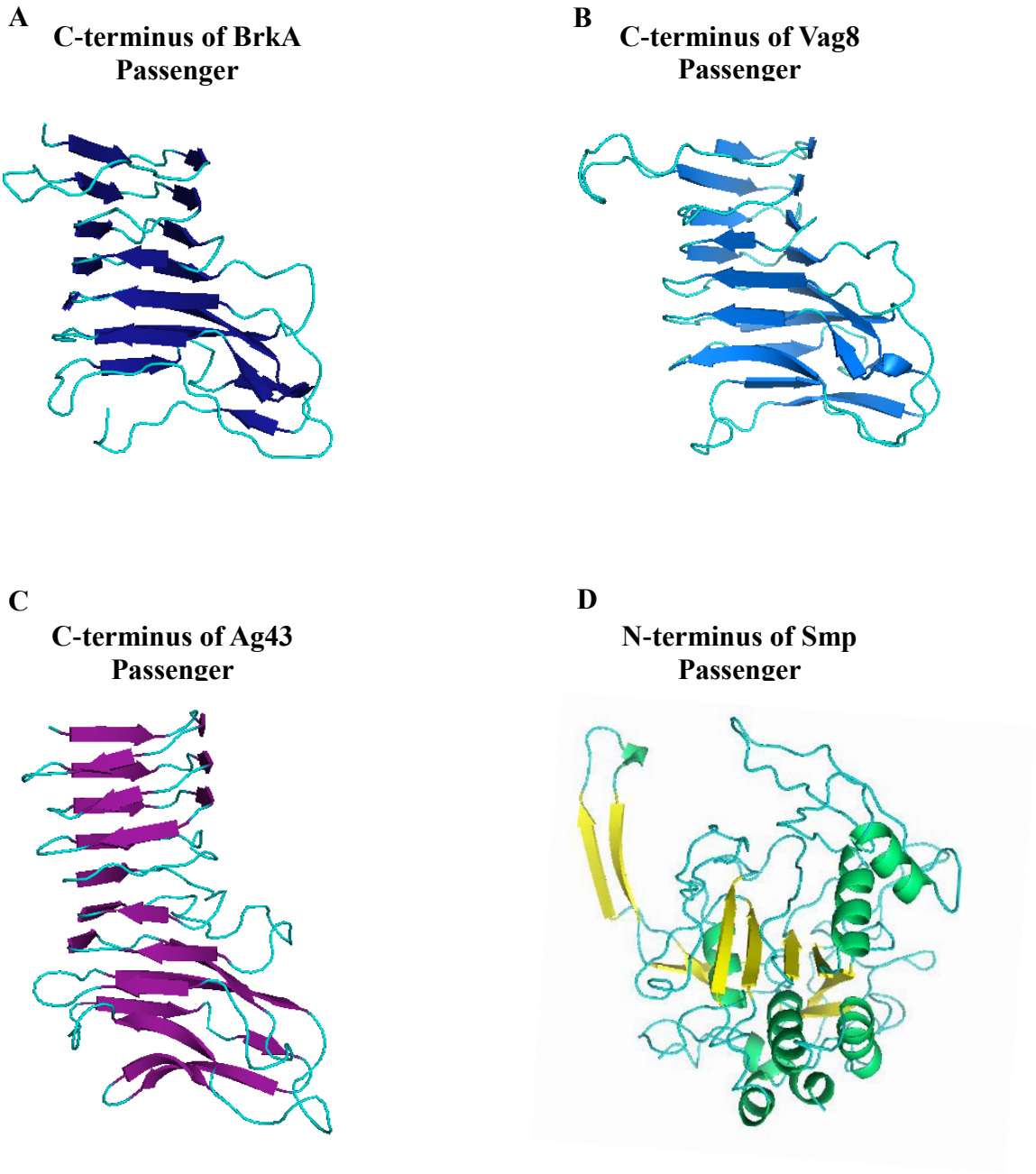


Figure 9. Passenger domain structure homology models generated by SWISS-MODEL. BrkA, Vag8, Ag43 and Smp passenger amino acid sequences were analyzed by SWISS-MODEL using the automated alignment mode. BrkA, Vag8 and Ag43 models shown were generated by alignment with PDB 2IOU, PDB 1DAB and PDB 3SYJ respectively. A model of the C-terminus of the Smp passenger was not generated. A model of the Smp N-terminus was generated by alignment with PDB 1THM.

residues V^{480}-T^{605} of the Vag8 passenger domain. Based on its BetaWrap raw score and P-value, the Vag8 passenger has high probability of possessing β-helical structure (Table 4). Furthermore, the Vag8 AC is included within the region of the Vag8 passenger domain that is recognized as the pectin-lyase family by SUPERFAMILY and is predicted to be a part of a β-helix (Fig. 8B). SWISS-MODEL also predicted that the C-terminal portion of the Vag8 passenger domain containing the putative AC would form a β-helix structure (Fig. 9B).

3.2.3 The Ag43 putative ACs are PD002475 and PD607128.

Ag43 of *E. coli* is a member of the BrkA group of putative ACs. However, it does not possess PF03212 and is assigned to a different putative AC subgroup than BrkA and Vag8. ProDom analysis of the Ag43 amino acid sequence identified multiple domains at the C-terminus of the Ag43 passenger domain that are conserved among ATs possessing AIDA-I subgroup putative ACs (Fig. 7C). These ProDom domains include PD002475 and PD607128. PD000317 spans the portion of the Ag43 passenger domain that corresponds to Ag43α, which is the region that functions in adhesion and aggregation. Therefore, PD000317 was determined to be a poor AC candidate. The BetaWrap raw score and P-value indicate that Ag43 likely possesses β-helical structure (Table 4). Based on SUPERFAMILY output, the Ag43 putative ACs span a region of the Ag43 passenger domain that is recognized as a member of the pectin-lyase family and is predicted to possess β-helix structure (Fig. 8C). SWISS-MODEL also included the C-terminal third of the Ag43 passenger domain containing the Ag43 putative ACs in its β-helix model of the C-terminus of the Ag43 passenger domain (Fig. 9C).

3.2.4 The Smp putative ACs are PD712254 and PDB1G2N3.

Smp of *S. marcescens* is a member of the protease group of putative ACs. ProDom analysis of the Smp amino acid sequence identified multiple domains at the C-terminus of the Smp passenger domain that are conserved among ATs of the protease group that could potentially function as ACs (Fig. 7D). These ProDom domains include PD712254, PD011682 and PDB1G2N3. PD011682 includes the very C-terminus of the passenger domain and the majority of the α-linker of the TU. As a result, PD011682 was determined to be a poor putative AC candidate. The Smp passenger domain has a lower probability of forming β-helical structure based on its P-value compared to that of BrkA, Vag8 and Ag43 (Table 4). However, SUPERFAMILY identified two pectin-lyase-like regions within the Smp passenger domain. The C-terminal region spans

residues 423-616 of the Smp passenger domain and includes PDB1G2N3 but not PD712254 (Fig. 8D). SWISS-MODEL was unable to generate a model of the C-terminus of the Smp passenger domain (Fig. 9D). Only a homology model of the Smp N-terminus with mixed α/β globular structure was generated. Therefore, the Smp passenger likely possesses a region of β-helical structure that includes PDB1G2N3 but not PD712254.

3.3 Secondary structure prediction and construct design

To test the requirement of conserved C-terminal passenger domain regions in the acquisition of passenger domain secondary structure, a set of pET30b expression constructs encoding full-length and AC-deleted passenger variants were designed for Vag8, Ag43 and Smp (Fig. 10). In order to avoid interrupting an important secondary structure element, PSIPRED secondary structure prediction analysis was used to identify predicted regions of coil near the N-terminal boundaries of the putative AC regions for the AC-deleted constructs and near the N-terminal boundary of the α-helical linker of the TU for the full-length passenger constructs.

3.3.1 BrkA Control.

PSIPRED secondary structure prediction analysis predicted only β-sheet and coil structure for the BrkA passenger domain (Appendix B). The PF03212 spans residues V^{583}-Q^{703} of the BrkA passenger domain. An AC-deleted passenger domain pET30b construct encoding BrkA passenger residues E^{61}-P^{600} was cloned and a full-length BrkA passenger construct encoding residues E^{61}-V^{699} was previously cloned (Fig. 11A) (51).

3.3.2 The Vag8 passenger is predicted to possess β-sheet and coil structure.

PSIPRED secondary structure prediction analysis revealed that the Vag8 passenger domain is predicted to possess β-sheet and coil structure (Appendix B). The secondary structure prediction was used to identify predicted regions of coil near the N-terminal boundary of PF03212 (Fig. 10). Two pET30b AC-deleted passenger domain constructs encoding Vag8 residuesV^{40}-R^{479} and V^{40}-P^{438} were cloned (Fig. 11B). A full-length Vag8 passenger construct encoding residues V^{40}-R^{610} was previously cloned (47).

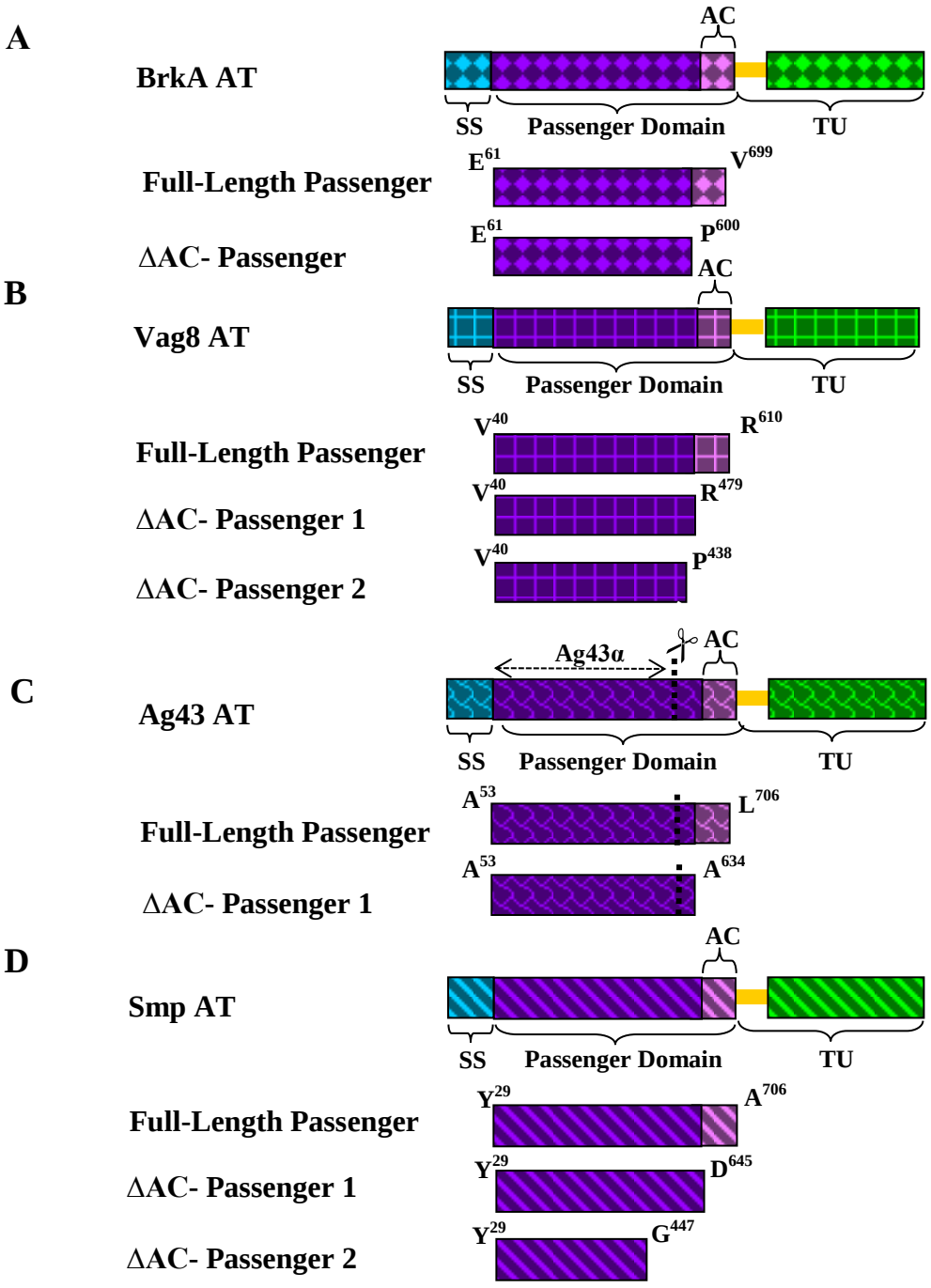


Figure 10. Structural representations of full-length and ΔAC-deleted passenger variants. The full-length autotransporter is characterized by the presence of three distinct domains; an N-terminal signal sequence (SS), a passenger and a C-terminal translocation unit (TU). A construct encoding a full length passenger variant extending from the SS cleavage site to the N-terminus of the alpha-linker (yellow) of the TU was cloned for each autotransporter (AT). At least one autochaperone-deleted variant (ΔAC) was cloned for each autotransporter.

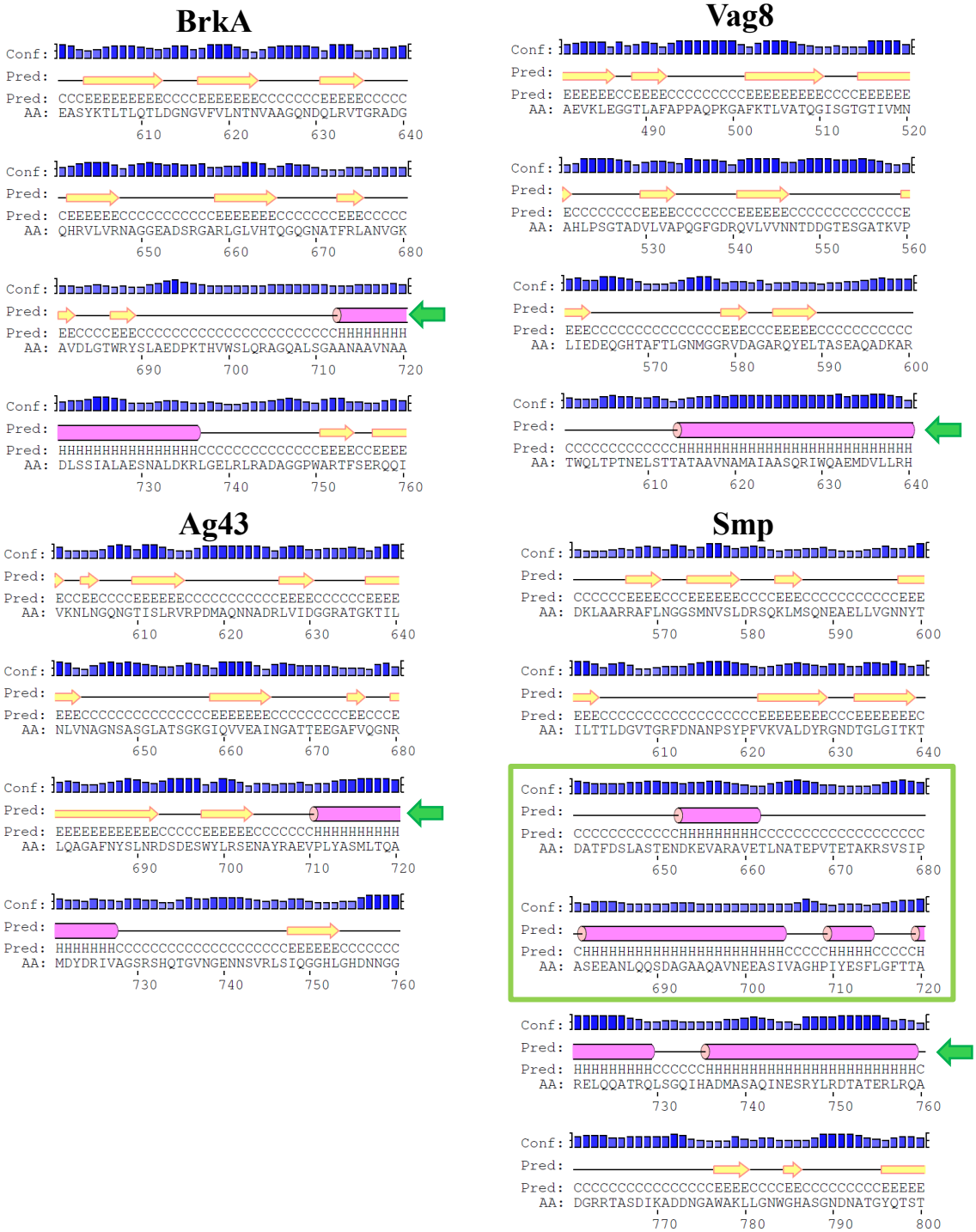


Figure 11. PSIPRED secondary prediction analysis of the C-terminus of passenger domains. Yellow arrows and pink cylinders represent predicted β-sheet and α-helical structure respectively. The confidence of the prediction is indicated by the blue bars above the predicted structure, where darker blue and a higher bar indicate a higher level of confidence. The green arrows indicate the α-helix of the α-helical linker of the TU. The green box in the Smp prediction indicates a region of coil located at the C-terminus of the Smp passenger domain that is absent from the BrkA, Vag8 and Ag43 passenger domain predictions shown. PSIPRED v3.0 was used.

3.3.3 The Ag43 passenger is predicted to possess β-sheet and coil structure.

PD002475 was previously identified as an important region in the secretion of other β-helix ATs (40, 51). Therefore, it was chosen as the best candidate for an AC motif of Ag43. PSIPRED secondary structure prediction analysis predicted only β-sheet and coil structure for the Ag43 passenger domain (Appendix B). The secondary structure prediction was used to identify predicted regions of coil near the N-terminal boundary of PD002475 (Fig. 10). pET30b constructs encoding the full-length passenger domain as well as an AC-deleted passenger variant, in which PD002475 was the predicted AC, were cloned from *E. coli* K12 genomic DNA (Fig. 11C).

3.3.4 The Smp passenger is predicted to possess mixed α/β-structure with a region of α-helical structure at the C-terminus of the passenger domain.

PSIPRED secondary structure prediction analysis predicted regions of coil, β-sheet and α-helical structure within the Smp passenger domain (Appendix B). Similar to SSP, Smp is predicted to have a region of α-helical structure at the C-terminus of the passenger N-terminal to the α-helical linker of the TU. The secondary structure prediction was used to identify predicted regions of coil near the N-terminal boundaries of PD712254 and PDB1G2N3 (Fig. 10D). pET30b constructs encoding the full-length passenger domain as well as two AC-deleted passenger variants, in which either the majority of PD712254 (ΔAC-Pass1) or PD712254 and PDB1G2N3 (ΔAC-Pass2) were deleted, were cloned from pSmppBSSK (Fig. 12D). Smp ΔAC-Pass1 lacks the predicted region of coil at the Smp passenger C-terminus and Smp ΔAC-Pass2 lacks the C-terminal region of coil as well as a significant region of predicted β-sheet structure N-terminal to it.

3.4 Assessing the folded-state of purified AC-deleted passenger variants

Expression of AC-deleted and full-length passenger constructs was induced in BL21 (DE3) and culture lysates were prepared for Ni-NTA under denaturing conditions (Figs. 12 & 13). Purified proteins were dialyzed from 8 M urea to 10 mM Tris pH 8 via step-wise dialysis to replace urea with folding buffer, and incubated at 4^{o} C for 7 days to allow folding to occur. Dialyzed proteins were then prepared for and analyzed by far-UV CD spectroscopy and limited-proteolysis with trypsin. Pure samples of each type of protein secondary structure have characteristic CD spectra in the far-UV range; an elipticity minimum of ~218 nm indicates β-sheet structure, elipticity minima at ~208 and ~222 nm indicates α-helical structure and an elipticity minimum of ~200 nm indicates random coil or disordered structure (Fig.14A) (18). Bacterial proteins that get secreted into the

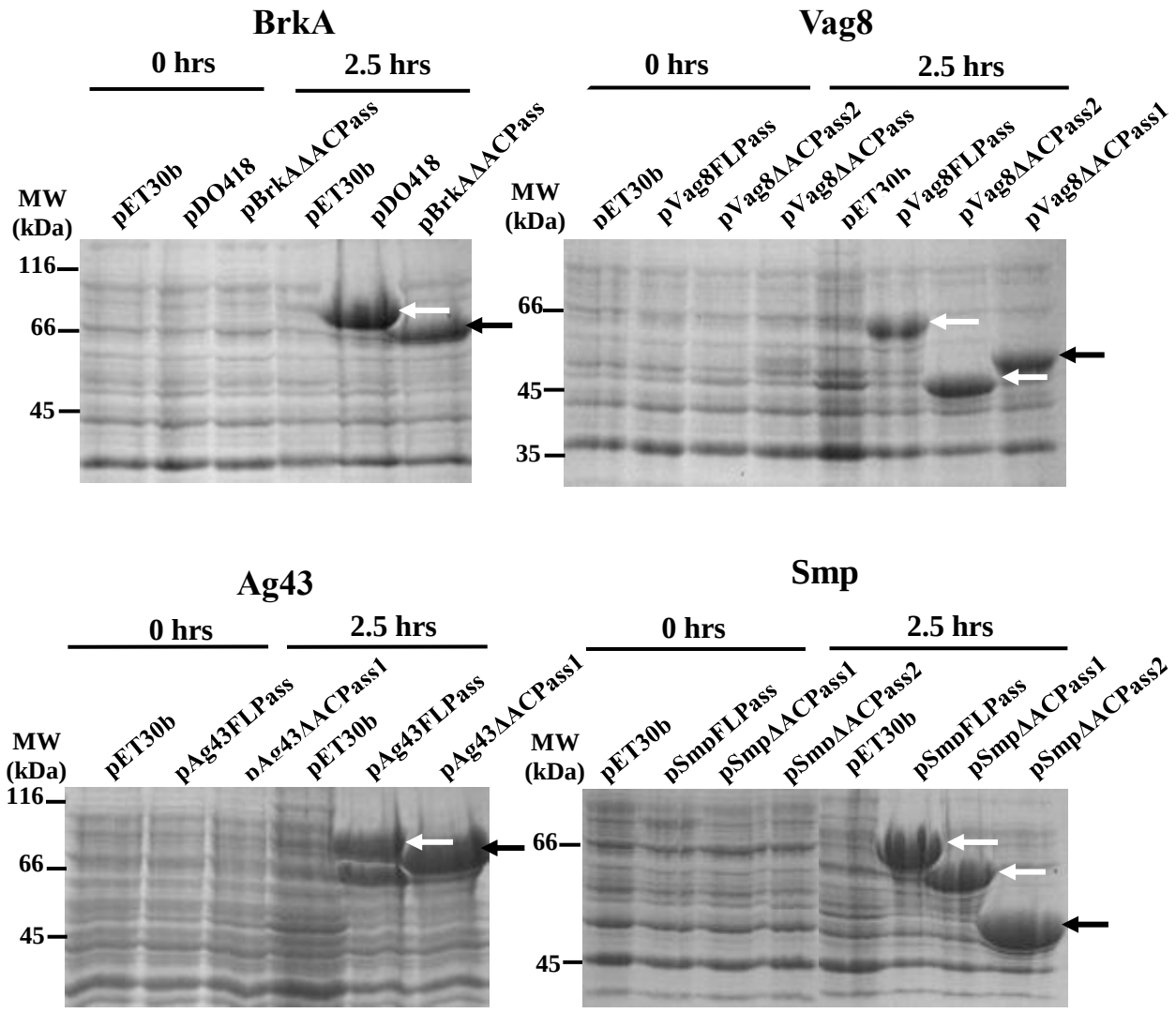


Figure 12. Induction of full-length and AC-deleted passenger domain expression constructs. Full-length (FL) and AC-deleted (ΔAC) passenger (pass) constructs were transformed into BL21 (DE3) and their expression was induced for 2.5 hrs with 100 mM IPTG. Arrows indicate the bands corresponding to induced passenger variants. Samples were visualized by SDS-PAGE followed by staining with Coomassie Brilliant Blue. The larger Ag43 FL Pass band corresponds to the full-length Ag43 passenger domain and the small band corresponds to the cleaved Ag43α. The numbers on the left correspond to molecular weight (MW) in kDa.

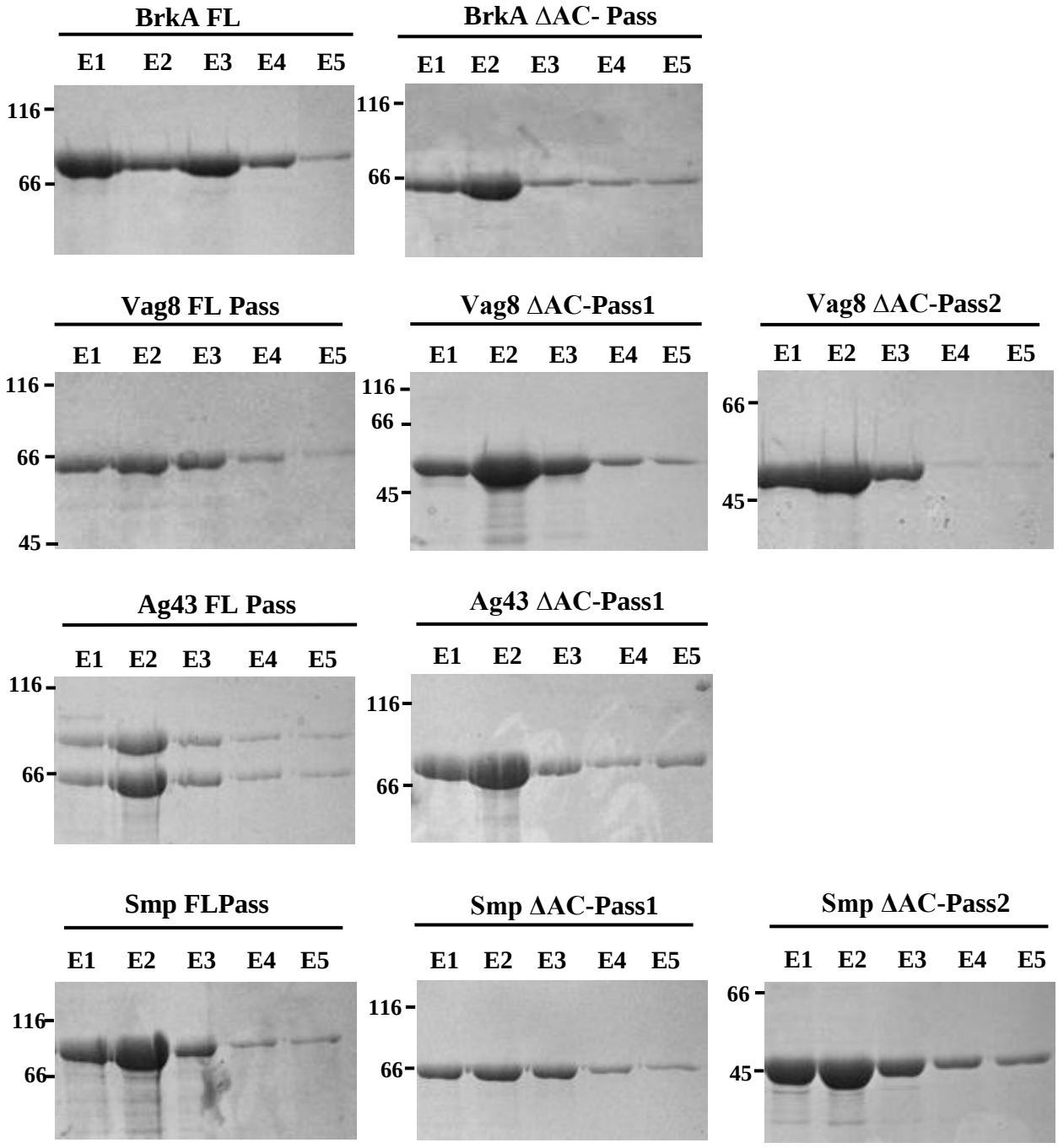


Figure 13. Purification of full-length and AC-deleted passenger variants via nickel affinity chromatography under denaturing conditions. Induced cell lysates were resuspended in guanidine hydrochloride, sonicated and centrifuged. The supernatants were loaded onto 2 mL columns with nickel-coated beads. The columns were washed with two rounds of denaturing wash buffers followed by elution in 5 mL of 8M urea elution buffer. 10 µL of each elution was run on SDS-PAGE followed by Coomassie staining. The cleanest elutions with the highest amounts of expressed protein were chosen for dialysis and folding experiments. The numbers on the left correspond to molecular weight in kDa.

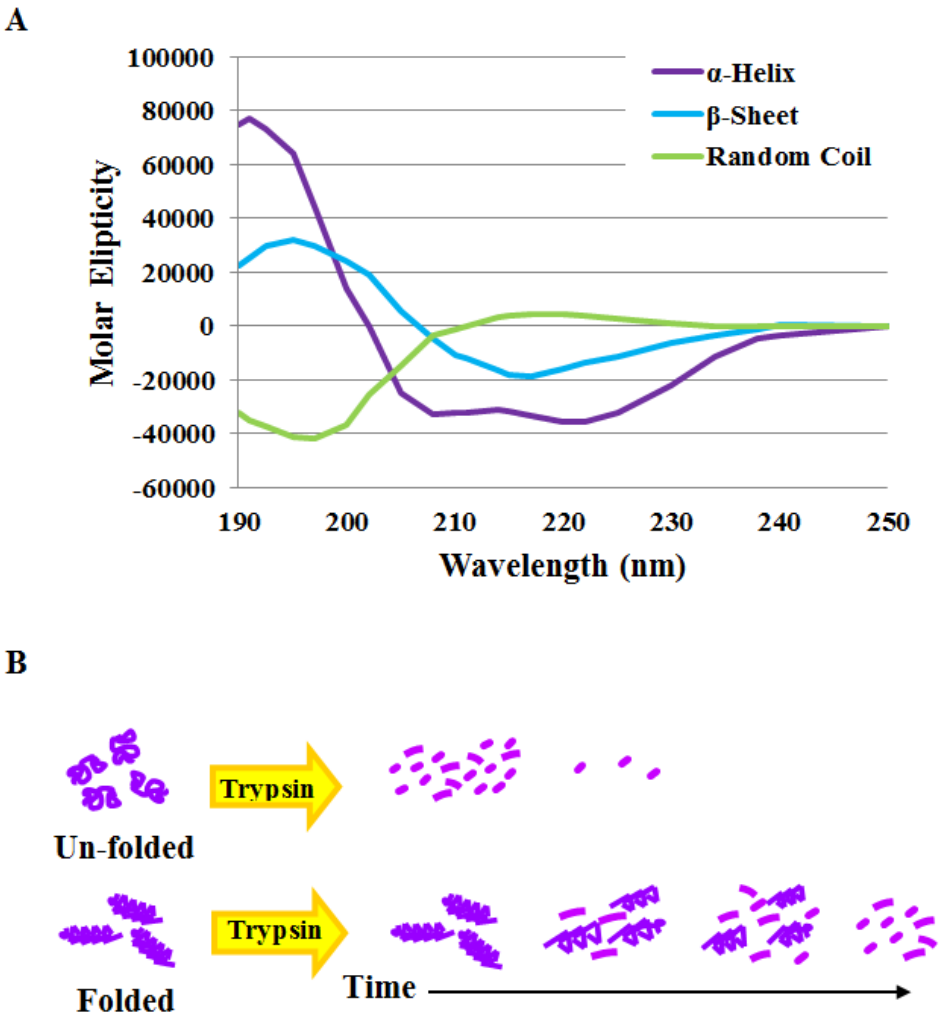


Figure 14. Far-UV circular dichroism spectroscopy standard spectra and limited proteolysis with trypsin diagram. (A) Poly-lysine standard spectra for 100% coil, 100% β-sheet and 100% α-helical structure. Data points were derived by Greenfield *et al*. (17). (B) Diagramatic representation of the high resistance to trypsin of folded proteins over time relative to un-folded or denatured proteins.

extracellular environment generally have a compactly folded native state that is significantly more resistant to digestion by proteases relative to the unfolded state (38). Therefore, unfolded proteins are more sensitive to digestion with trypsin compared to natively folded proteins and trypsin sensitivity of AC-deleted passenger variants relative to the folded full-length passenger domain can be used to assess the folded state of purified AC-deleted passenger proteins (Fig 14B).

3.4.1 BrkA Control

Prior CD spectroscopic analysis, in addition to crystallographic analysis of the BrkA passenger, indicated that the passenger domain of BrkA is comprised of β-sheet structure (51, 83). Far-UV CD spectroscopy of purified full-length and AC-deleted passenger proteins produced spectra with a minimum of approximately 220 nm and 200 nm, respectively (Fig. 15A). The wavelength minimum of 220 nm is consistent with the predicted β-sheet structure of the native BrkA passenger domain, and the minimum of 200 nm of the AC-deleted passenger variant indicated that the purified protein lacked secondary structure, which is consistent with the known requirement of the BrkA AC motif for BrkA passenger domain folding (51). The results of the limited proteolysis of purified passenger proteins with trypsin were consistent with the CD spectroscopy results (Fig. 15B). Native BrkA passenger domain has been previously demonstrated to be highly trypsin resistant relative to the unfolded BrkA passenger domain. Dialyzed full-length passenger proteins were detectable 15 min following addition of trypsin, whereas the purified AC-deleted variant was rapidly degraded by trypsin and was no longer detectable 5 min following addition of trypsin, suggesting that AC-deleted BrkA passenger variant lacked native structure.

3.4.2 The Vag8 AC motif is required for passenger domain folding in vitro.

Secondary structure prediction analysis of the Vag8 passenger domain predicted that, similar to the BrkA passenger domain, the Vag8 passenger was comprised of β-sheet structure. Therefore, the far-UV CD spectrum of natively folded full-length Vag8 passenger domain was expected to be similar to that of the BrkA spectrum and, in the absence of the region encoding the putative AC, the spectra of the AC-deleted passenger variants were expected to be similar the AC-deleted BrkA spectrum. The spectra of the AC-deleted Vag8 passenger variants had a minimum at ~200 nm that indicated they lacked secondary structure (Fig. 16A). However, the spectrum of the purified full-length passenger protein suggested that purified full-length Vag8 passenger protein

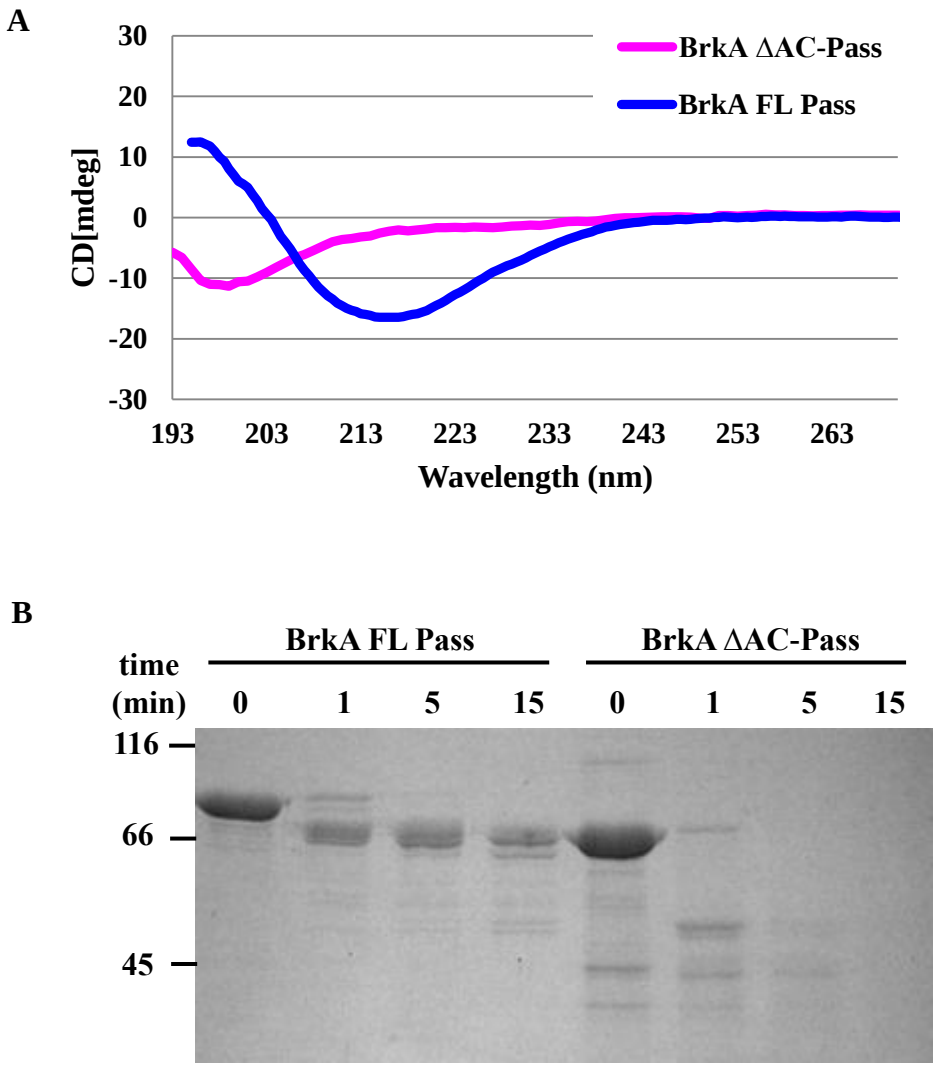


Figure 15. Assessing the folded state of purified BrkA passenger domain variants. Following a 7 day incubation in 10 mM Tris pH8 folding buffer, the folded state of purified BrkA passenger variants was assessed with far-UV CD spectroscopy and limited proteolysis with trypsin. (A) Dialyzed protein samples were diluted to 3 µM and submitted to three CD spectroscopy scans and measurements were taken in the 190-265 nm range at 20^{o}C. (B) 20 µg of dialyzed samples were subjected to proteolysis with 0.5 µg of trypsin and samples were removed at 0, 1, 5 and 15 min following the addition of trypsin, and samples were resolved by SDS-PAGE followed staining with Coomassie Brilliant Blue. The numbers on the left correspond to molecular weight in kDa.

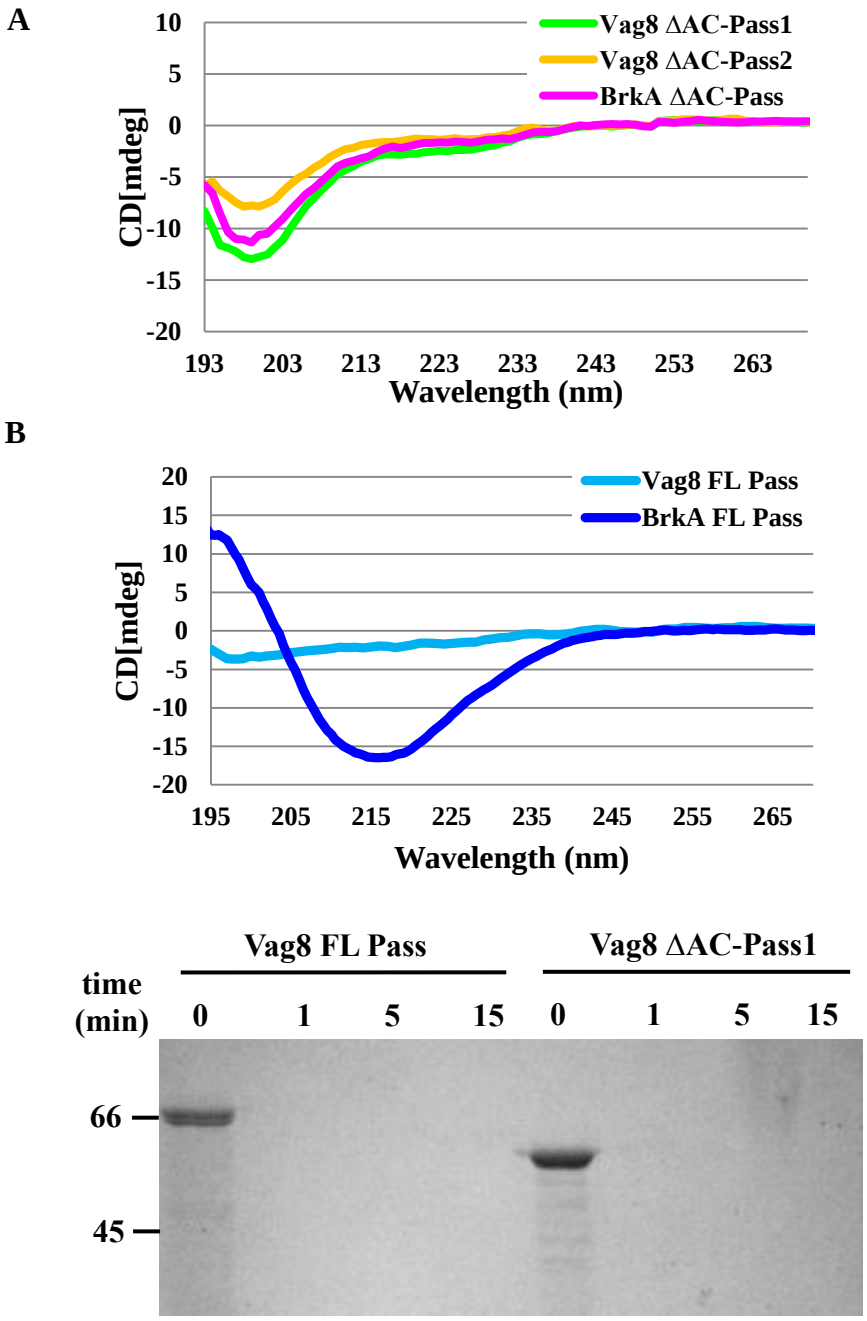


Figure 16. Assessing the folded state of Vag8 passenger variants purified under denaturing conditions. Following a 7 day incubation in 10 mM Tris pH 8 folding buffer, the folded state of purified Vag8 full-length (Vag8 FL Pass) and AC-deleted (ΔAC-Pass) passenger variants was assessed by far-UV CD spectroscopy or limited proteolysis with trypsin. (A) Overlay of the CD spectra of the two ΔAC-Pass Vag8 passenger variants and the BrkA ΔAC-Pass protein. (B) Overlay of Vag8 FL Pass CD spectrum and BrkA FL Pass spectrum. (C) Trypsin was added to purified and dialyzed FL and ΔAC-Pass Vag8 proteins and samples were removed at 1, 5 and 15 min following addition of trypsin. Samples were visualized by SDS-PAGE followed by staining with Coomassie Brilliant Blue. The numbers on the left correspond to molecular weight in kDa.

also lacked native structure (Fig. 16B). Limited proteolysis with trypsin also suggested that dialyzed full-length Vag8 passenger lacked secondary structure, as natively folded full-length Vag8 passenger was expected to be resistant to digestion with trypsin relative the AC-deleted variants (Fig. 16C). The purification and dialysis of the full-length Vag8 passenger was repeated and new samples were analyzed by far-UV CD spectroscopy. The poor quality of the spectra and inspection of dialyzed protein samples suggested that full-length passenger proteins either failed to fold or precipitated during dialysis. TANGO analysis of the Vag8 passenger domain amino acid sequence for propensity for hydrophobic region aggregation revealed that, unlike the BrkA and Ag43 passenger domains, the Vag8 passenger possessed multiple regions, particularly in the N-terminal half of the passenger, that were prone to aggregation (Fig. 17) (14).

To acquire folded full-length passenger, a hybrid-condition Ni-NTA purification protocol involving preparation of induced culture lysates under denaturing conditions, an initial denaturing pH wash, native washes and elution under native conditions, was used. The sample eluted from the column (Hybrid 1) and dialyzed to 10 mM Tris pH8 did not aggregate or precipitate, yet far-UV CD spectroscopy indicated that the dialyzed protein lacked secondary structure following a 2 week incubation at 4^{o} C (Fig. 18A). A series of hybrid-purification protocol optimizations, involving increasing numbers of native washes and addition of room-temperature incubation steps between washes, were attempted. An optimized protocol (Hybrid 4) including six additional native washes and a 4-day room temperature incubation of the column in the 8^{th} native wash resulted in a sample with a CD spectrum with a minimum of ~220 nm consistent with the expected β-sheet secondary structure of folded Vag8 passenger domain (Fig. 18B). Vag8 ΔAC-Pass1 purified under optimized hybrid conditions had a CD spectrum minimum of ~200 nm which was consistent with unfolded protein (Fig. 19A). The spectral peaks located in 190-215 nm range of Hybrid 4 protocol purified proteins were caused by imidazole contamination resulting from prolonged incubation of the sample on the column in native wash buffer. The folded state of the full-length Vag8 Hybrid 4 sample was verified by limited proteolysis with trypsin. The optimized hybrid sample was relatively more resistant to digestion with trypsin than were the AC-deleted variants, as full-length hybrid purified proteins were present 15 min following addition of trypsin whereas the AC-deleted variants were not detectable 1 min following addition of trypsin (Fig. 19B). The optimized protocol was repeated to verify reproducibility of the results and that the protocol could reliably purify folded full-length Vag8 passenger following a two week incubation at 4° C (Fig. 18C).

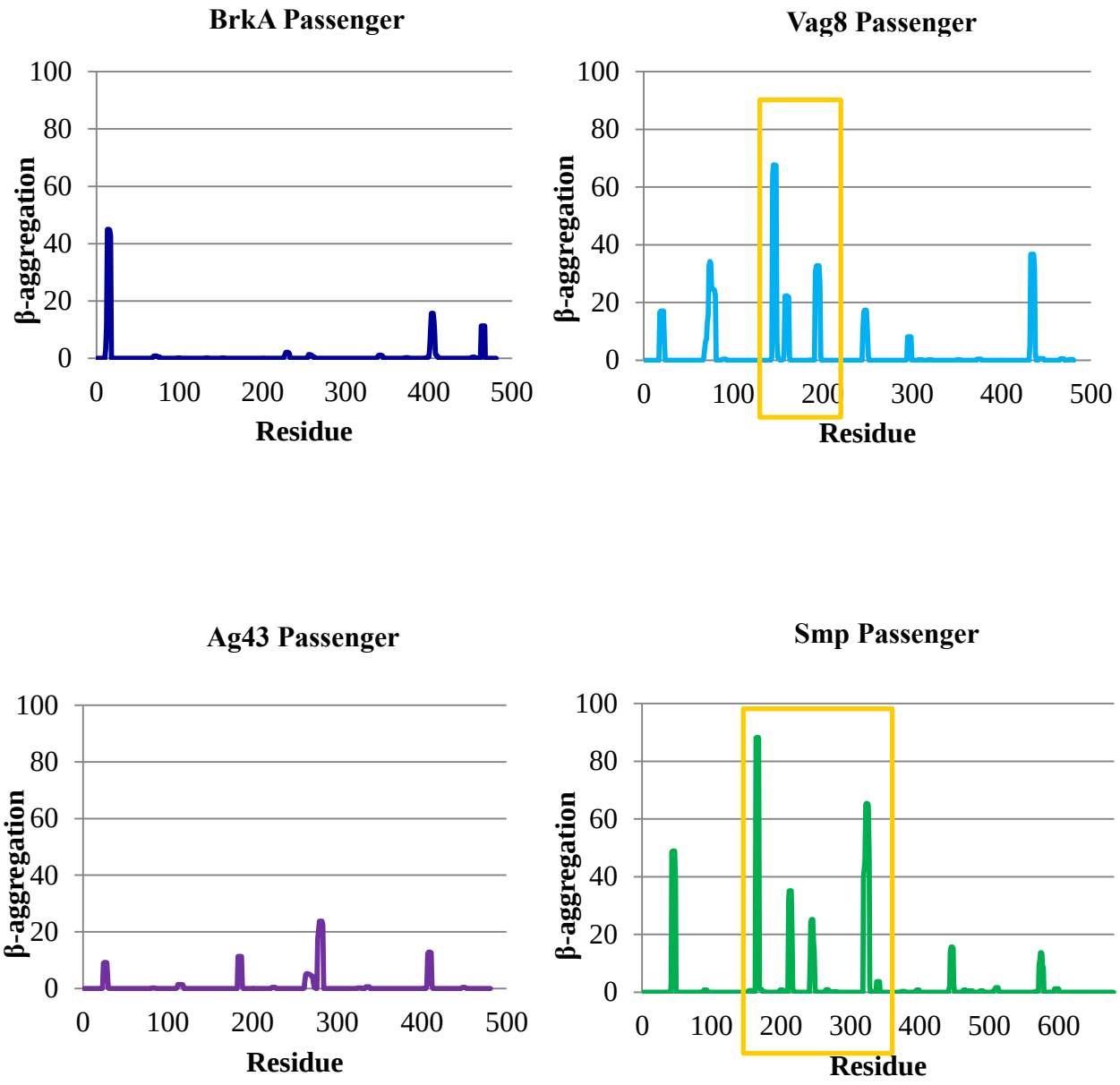
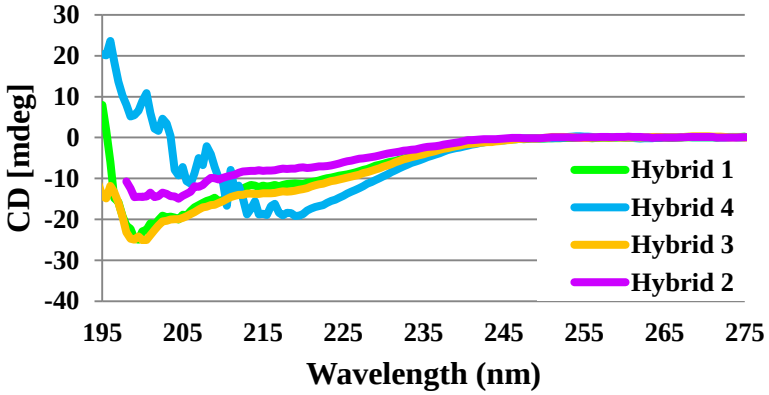
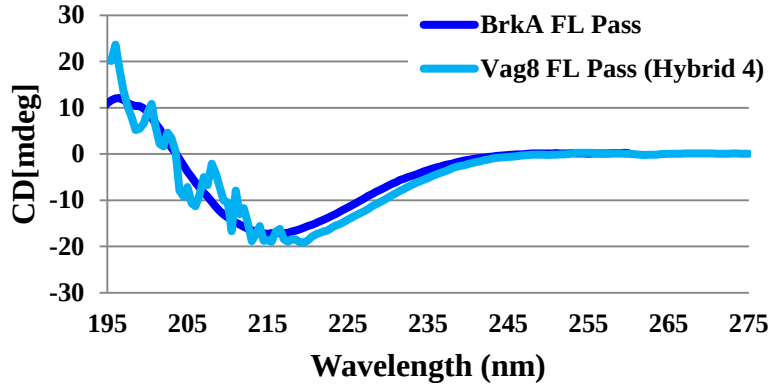


Figure 17. TANGO analysis of passenger domain sequences. TANGO software was used to analyse the amino acid sequences of the BrkA, Vag8, Ag43 and Smp passenger domains (14). Each residue was assigned an aggregation score from 0-100 where a higher score coincides with a higher propensity to aggregate. Yellow boxes indicate passenger domain regions with residues that have an aggregation score of greater than 50 and have high propensity for aggregation upon folding.

A



B



C

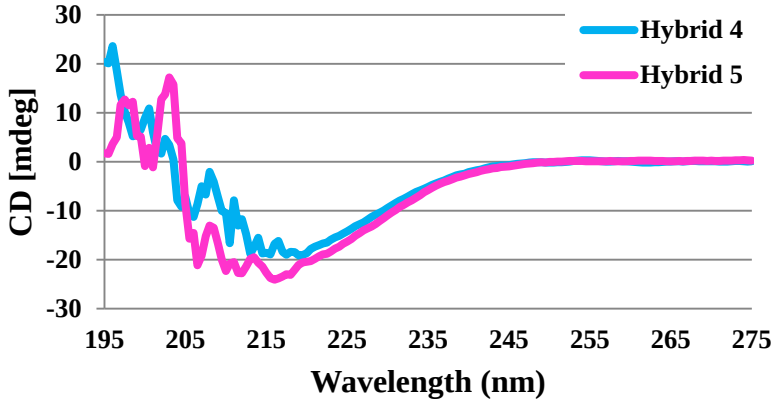
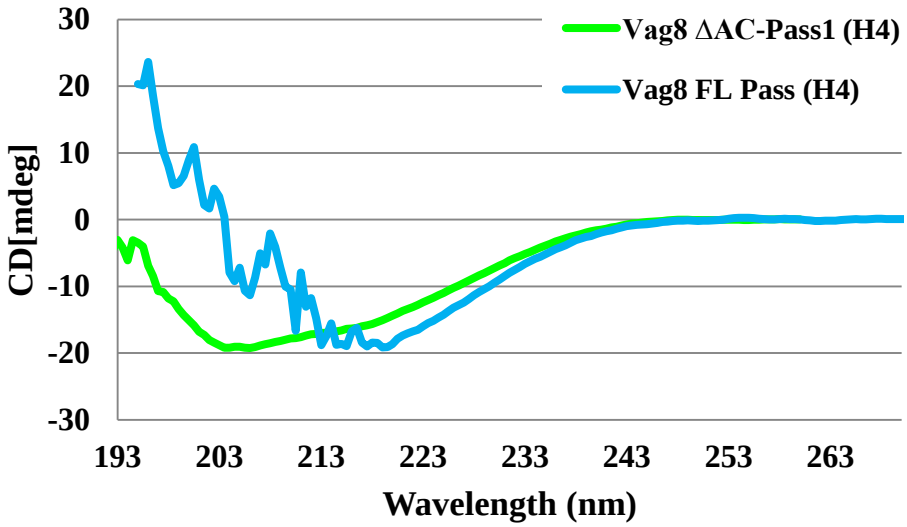


Figure 18. Hybrid-condition purification of full-length Vag8 passenger proteins. Four hybrid protocol optimizations (Hybrid 1-4) were tested for the purification of the full-length (FL) Vag8 passenger under hybrid conditions. (A) Overlay of the CD spectra of the Vag8 FL Pass samples purified under the four optimizations of the hybrid condition Ni-NTA protocol. (B) Overlay of the CD spectra of the BrkA FL Pass sample purified under denaturing conditions and the Vag8 FL Pass sample purified under Hybrid 4 conditions. (C) Overlay of the CD spectra of two Vag8 FL Pass samples purified under Hybrid 4 conditions.

A



B

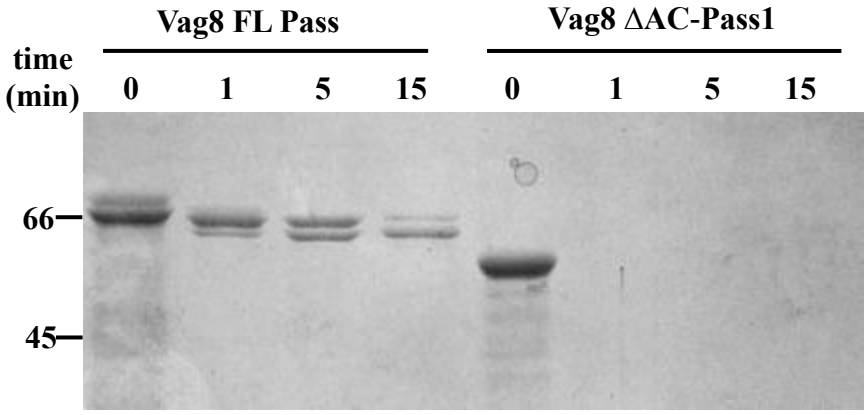


Figure 19. Assessing the folded state of hybrid-condition purified full-length and AC-deleted Vag8 passenger variants. Following dialysis to replace the native elution buffer with 10 mM Tris pH 8, the folded state of Hybrid 4 (H4) purified Vag8 full-length passenger (FL Pass) and AC-deleted passenger (ΔAC-Pass) was assessed with far-UV CD spectroscopy and limited proteolysis with trypsin. (A) Dialyzed protein samples were diluted to 3 µM and submitted to three CD spectroscopy scans and measurements were taken in the 190-265 nm range at 20^{o} C. (B) 20 µg of dialyzed samples were subjected to proteolysis with 0.5 µg of trypsin and samples were removed at 0, 1, 5 and 15 min following the addition of trypsin, and samples were resolved by SDS-PAGE followed staining with Coomassie Brilliant Blue. The numbers on the left correspond to molecular weight in kDa.

Extended refolding periods of 1-3 months were necessary for Hybrid1-3 samples to acquire the same level of structure and resistance to digestion with trypsin. The CD spectra combined with the results of limited proteolysis with trypsin of the optimized hybrid-protocol purified full-length and AC-deleted proteins indicated that, similar to BrkA, the Vag8 AC is required for acquisition of passenger domain secondary structure *in vitro* (Fig. 19B).

3.4.3 The conserved regions at the C-terminus of the Ag43 passenger domain are not required for the acquisition of passenger domain secondary structure.

Secondary structure prediction analysis of the Ag43 passenger domain and previous CD spectroscopic analysis indicated that, similar to the BrkA and Vag8 passenger domains, the Ag43 passenger domain was comprised of β-sheet structure. Therefore, the far-UV CD spectrum of natively folded full-length Ag43 passenger domain was expected to be similar to that of the BrkA and Vag8 passenger spectra. The spectrum of the purified full-length Ag43 passenger protein had a minimum of ~220 nm consistent with the expected β-sheet structure (Fig. 20A). The spectrum of the AC-deleted variant (ΔAC-Pass1) however, also possessed a minimum of ~220 nm which was consistent with β-sheet structure (Fig. 20A). Limited proteolysis with trypsin revealed that the AC-deleted variant had relatively the same level of resistance to digestion by trypsin as the full-length passenger protein (Fig. 20B). The ability of Ag43 ΔAC-Pass1 to acquire near native structure suggested that PD002475 was not required for the acquisition of Ag43 passenger domain secondary structure *in vitro*. However, Ag43 ΔAC-Pass1 is not cleaved to release Ag43α following dialysis indicating PD002475 is required for the acquisition of native structure.

To determine if PD607128 or the C-terminus of PD000317 was required for passenger domain folding, two additional AC-deleted passenger domain pET30b constructs were cloned, expressed, purified and dialyzed (Fig. 21A-D). One of the constructs (ΔAC-Pass2) encoded Ag43α (residues A^{53}-N^{522}) and the second (ΔAC-Pass3) encoded residues A^{53}-N^{383} of the Ag43 passenger domain. Far-UV CD spectroscopy of Ag43 ΔAC-Pass2 generated a spectrum with a wide trough that extends from ~223 nm to ~200 nm with a minimum of ~220 nm which, indicates that the purified sample had partially folded with regions of coil and β-sheet structure (Fig. 22A). Limited proteolysis with trypsin indicates that Ag43 ΔAC-Pass2 also retained similar resistance to degradation with trypsin as the full-length Ag43 passenger protein (Fig. 22C). Far-UV CD spectroscopy of Ag43 ΔAC-Pass3 generated a spectrum with a minimum of ~200 nm, consistent

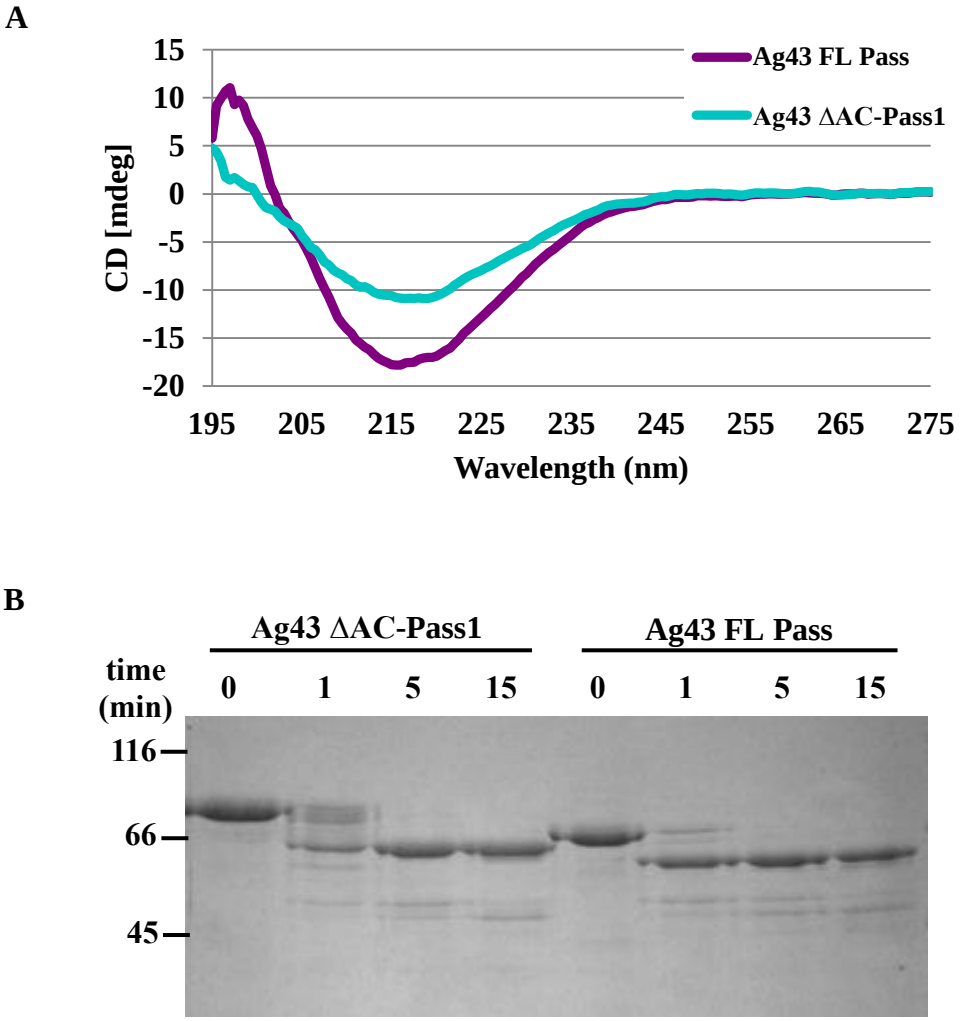


Figure 20. Assessing the folded state of full-length and ΔAC-Passenger 1 Ag43 proteins. Eluted proteins were dialyzed to 10 mM Tris pH 8 and analyzed by far-UV CD spectroscopy and limited proteolysis with trypsin. (A) Dialyzed protein samples were diluted to 3 µM and submitted to three CD spectroscopy scans and measurements were taken in the 190-265 nm range at 20° C. (B) 20 µg of dialyzed samples were subjected to proteolysis with 0.5 µg of trypsin and samples were removed at 0, 1, 5 and 15 min following the addition of trypsin, and samples were resolved by SDS-PAGE followed staining with Coomassie Brilliant Blue. The single Ag43 FL Pass band present at the 0 time point is the cleaved Ag43 passenger domain (Ag43α). The numbers on the left correspond to molecular weight in kDa.

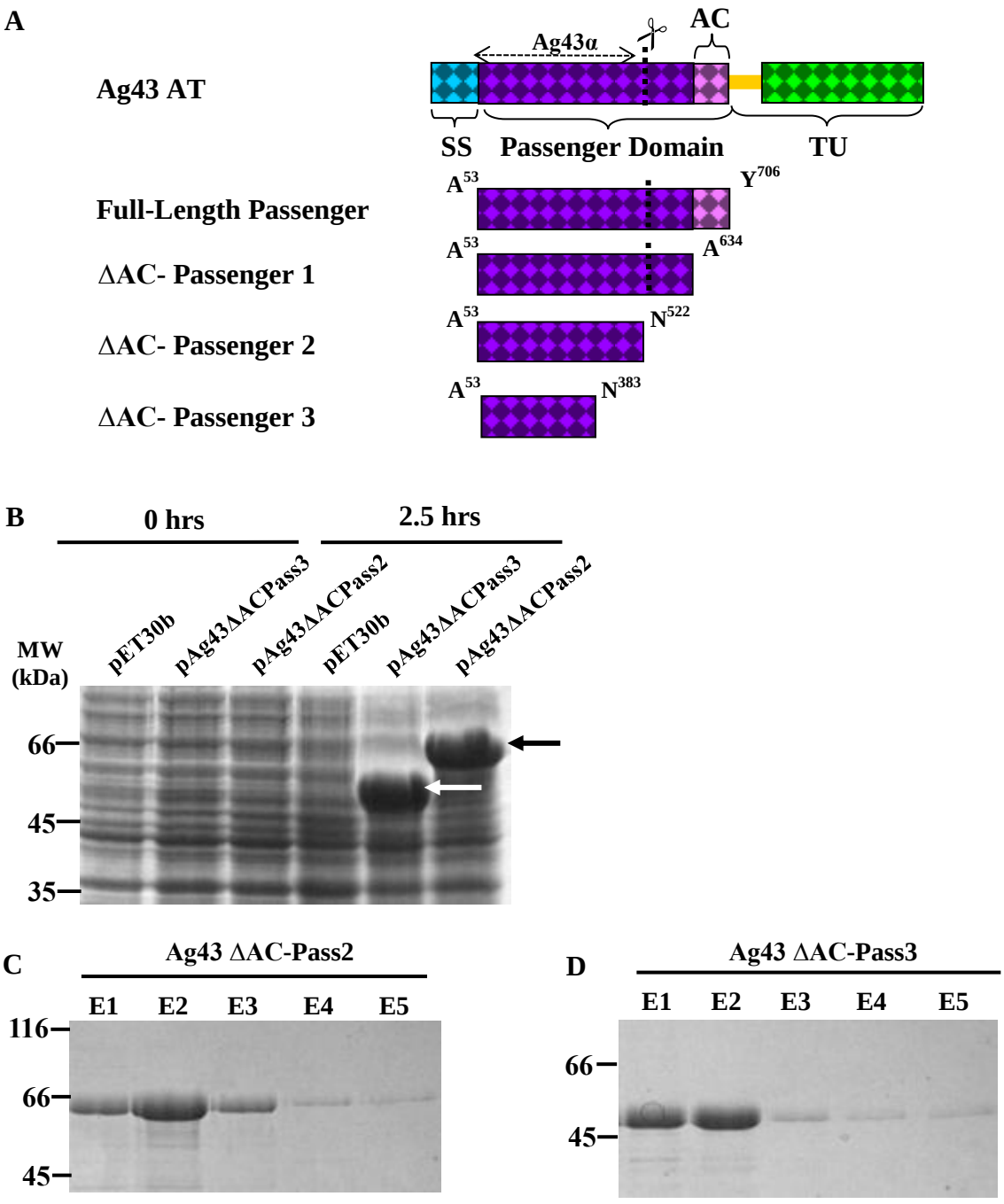


Figure 21. Expression and purification of Ag43 AC-deleted passenger variants. (A) Two additional Ag43 AC-deleted passenger (ΔAC-Pass) constructs were designed encoding passenger variants with C-terminal truncations of increasing size. (B) Expression of constructs was inducted for 2.5 hrs with 100 mM IPTG. (C&D) Induced proteins were harvested from cell lysates and purified by Ni-NTA under denaturing conditions. Samples were visualized by SDS-PAGE followed by Coomassie Brilliant Blue staining. The numbers on the left correspond to molecular weight in kDa.

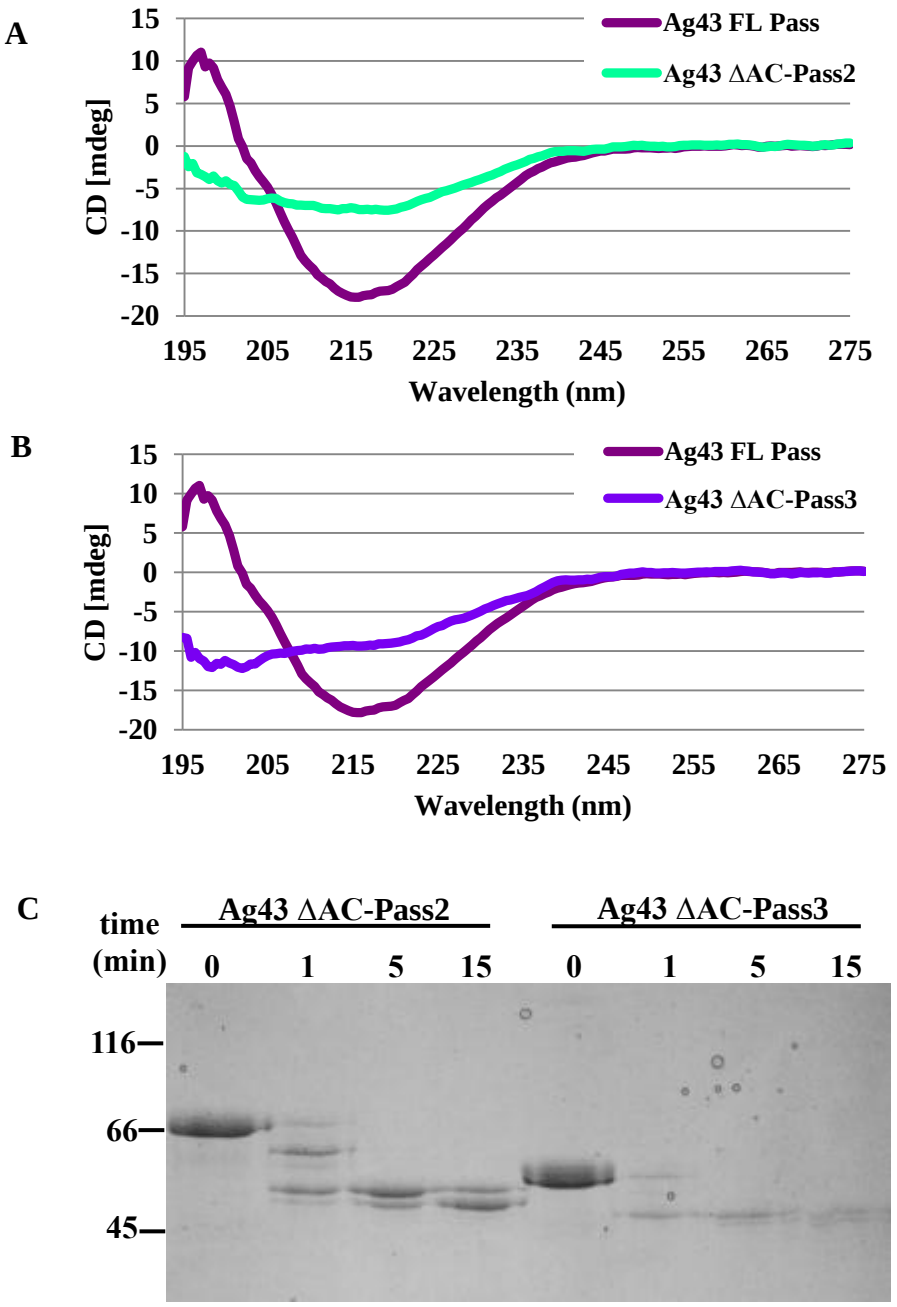


Figure 22. Analysis of the folded state of additional Ag43 AC-deleted passenger variants. Following a 7 day incubation in 10 mM Tris pH 8 folding buffer, the folded state of purified Ag43 ΔAC-Pass2 and Ag43 ΔAC-Pass3 samples were assessed by far-UV CD spectroscopy or limited proteolysis with trypsin. (A) Overlay of the CD spectra of the Ag43 ΔAC-Pass2 and FL Ag43 passenger proteins. (B) Overlay of the CD spectra of the Ag43 ΔAC-Pass3 and FL Ag43 passenger proteins. (C) Trypsin was added to purified and dialyzed Ag43 ΔAC-Pass2 and ΔAC-Pass3 proteins and samples were removed at 1, 5 and 15 min following addition of trypsin. Samples were visualized with SDS-PAGE followed by Coomassie Brilliant Blue staining. The numbers on the left correspond to molecular weight in kDa.

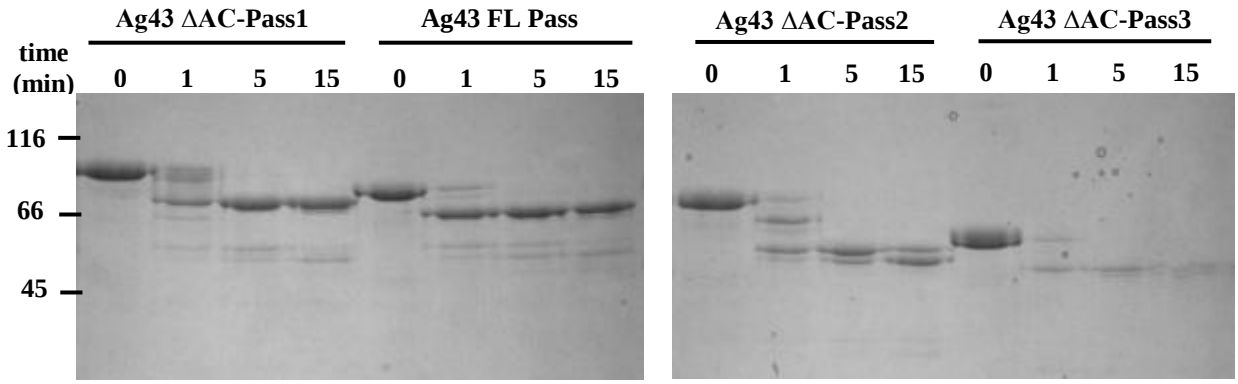
with unfolded/denatured protein (Fig. 22B). However, the spectrum had an additional minimum at ~220 nm indicating that, although the majority of the sample was unfolded, a portion of the sample had acquired β-sheet structure (Fig. 22B). Limited proteolysis with trypsin revealed that Ag43 ΔAC-Pass3 was sensitive to digestion with trypsin relative to the full-length Ag43 passenger protein. However, a portion of the sample retained its resistance to digestion with trypsin (Fig. 22C). Samples were re-analyzed by far-UV CD 5 months later to determine if they had acquired additional secondary structure. Unlike the full-length Vag8 passenger samples purified under optimized hybrid conditions, the partially folded Ag43 passenger variants did not acquire additional secondary structure following a prolonged incubation period at 4° C (data not shown).

Increasingly larger C-terminal truncations of the Ag43 passenger domain resulted in an increasing loss of β-sheet structure and resistance to digestion with trypsin relative to the Ag43 full-length passenger protein (Fig. 23 A,B). The largest C-terminal truncation passenger variant (Ag43 ΔAC-Pass3) lacked over 320 of C-terminal residues of the passenger which is approximately half of the Ag43 passenger domain, including more than 130 residues of Ag43α, suggesting that the conserved motifs in the C-terminal half of the passenger domain are not required to achieve secondary structure and a low level of resistance to digestion with trypsin, but are necessary to achieve a native folded state.

3.4.4 The Smp passenger does not require conserved C-terminal residues to acquire α-helical secondary structure.

Secondary structure prediction analysis of the Smp passenger domain predicted mixed α/β-structure. Therefore, the far-UV CD spectrum of natively folded full-length Smp passenger domain was expected to reflect the predicted mixed α/β-structure. The requirement of C-terminal passenger residues for passenger domain function of SSP, a homologue of Smp, has been previously demonstrated (50). Therefore, deletion of the C-terminal putative AC motifs was expected to affect the ability of the Smp passenger domain to fold *in vitro*. The Smp ΔAC-Pass2 sample repeatedly precipitated upon dialysis and could not be analyzed by CD spectroscopy. The spectra of the purified full-length and ΔAC-Pass1 Smp passenger proteins had wide troughs spanning ~205-223 nm suggesting α-helical structure (Fig. 24A), and both the full-length Smp passenger and the ΔAC-Pass1 samples were highly trypsin sensitive and rapidly degraded following the addition of trypsin in a limited proteolysis with trypsin assay (Fig. 24B). To determine if the dialyzed Smp full-length passenger protein had acquired functional activity, a zymogram testing for

A



B

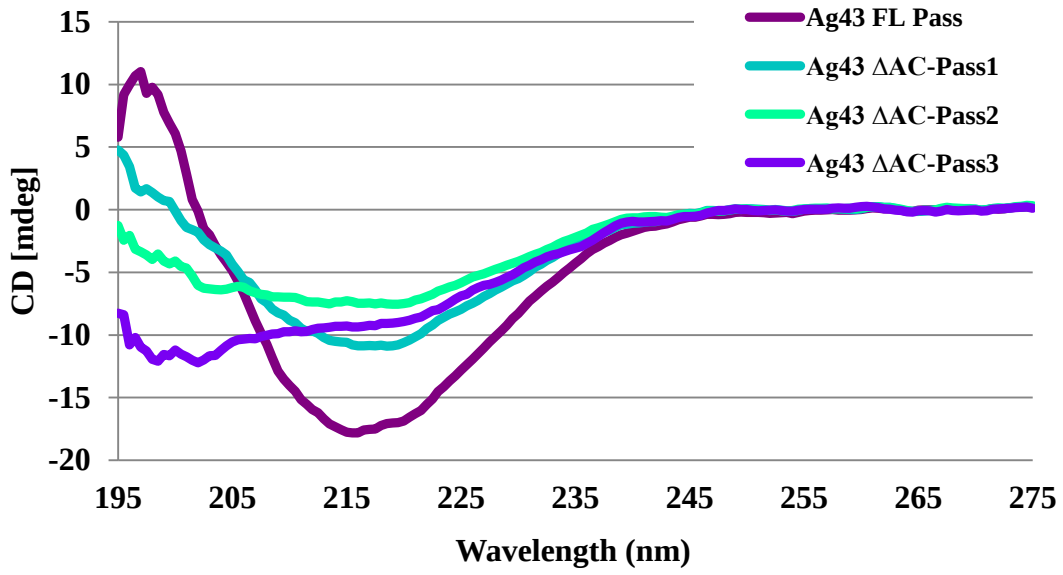


Figure 23. Comparison and overlay of Ag43 passenger variant trypsin sensitivity and far-UV CD spectra. (A) Limited proteolysis of Ag43 passenger variants with trypsin. (B) Overlay of CD spectra of full-length Ag43 passenger and AC-deleted Ag43 passenger variants. All samples were diluted to 3 µM and submitted to three CD spectroscopy scans and measurements were taken in the 190-275 nm range at 20° C.

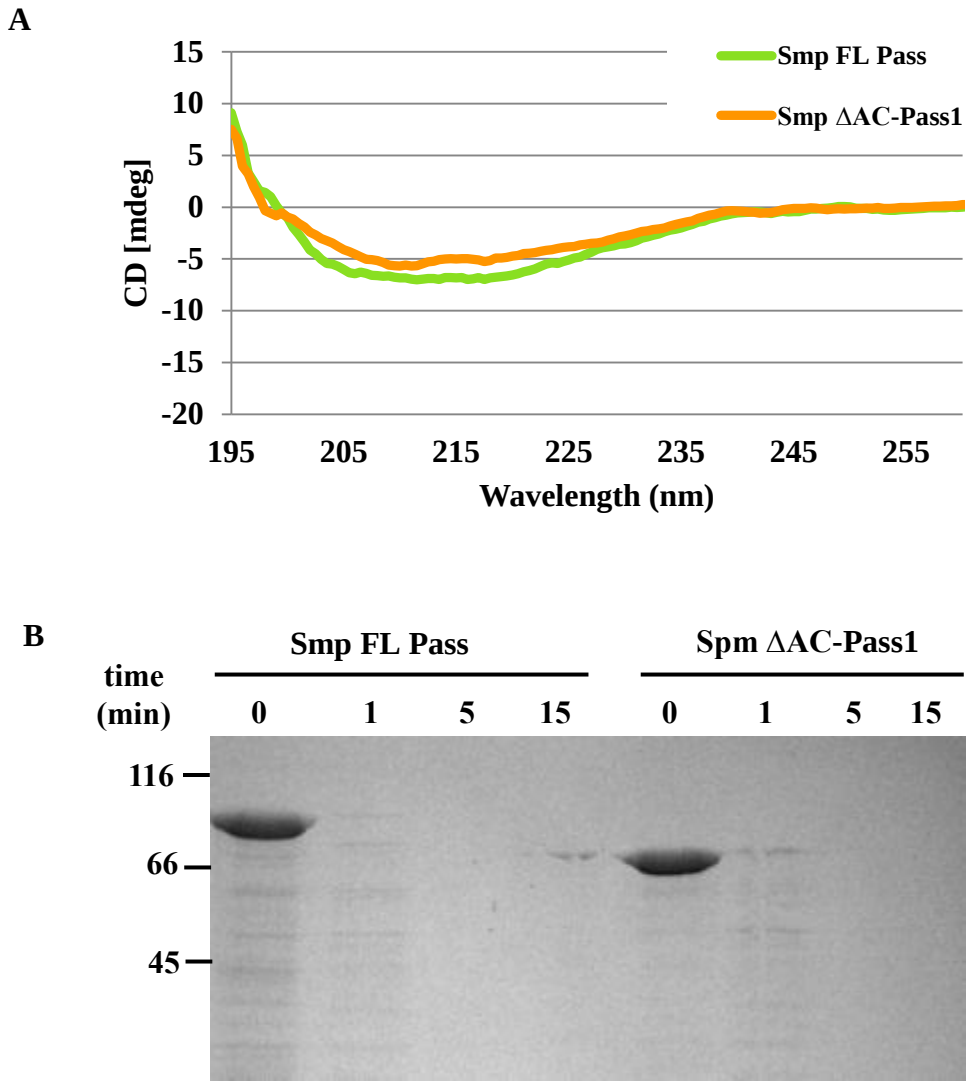


Figure 24. Assessing the folded state of full-length and ΔAC-Passenger 1 Smp proteins. Eluted proteins were dialyzed to 10 mM Tris pH8 and analyzed by far-UV CD spectroscopy and limited proteolysis with trypsin. (A) Dialyzed Smp full length (FL) and AC-deleted (ΔAC-Pass1) protein samples were diluted to 3 µM and submitted to three CD spectroscopy scans and measurements were taken in the 190-265 nm range at 20° C. (B) 20 µg of dialyzed samples were subjected to proteolysis with 0.5 µg of trypsin and samples were removed at 0, 1, 5 and 15 min following the addition of trypsin. Samples were resolved by SDS-PAGE followed staining with Coomassie Brilliant Blue. The numbers on the left correspond to molecular weight in kDa.

serine protease activity was performed using trypsin as a positive control. Both the full-length and ΔAC-deleted passenger proteins lacked serine protease activity under the tested conditions suggesting the Smp variants had failed to acquire native structure (Fig. 26C).

Due to the helical nature of the C-terminus of the Smp passenger domain, it was difficult to determine where the α-helical linker of the TU ended and the passenger began. Therefore, an extended Smp full-length passenger pET30b construct, encoding the α-helical linker in addition to the passenger domain of Smp, was cloned, expressed, purified and dialyzed to 10 mM Tris pH 8 (Fig. 25 A-C). The spectrum of the purified extended full-length Smp passenger protein had two well-defined minima at ~218 and ~223 nm consistent with primarily α-helical structure (Fig. 26A). The extended passenger protein was also readily degraded by trypsin and lacked protease activity when tested via zymogram suggesting it had failed to obtain native structure (Fig. 26B,C).

The spectrum of the purified extended full-length Smp passenger protein exhibited noticeably higher levels of circular dichroism than full-length and AC-deleted passenger samples despite preparation of all Smp passenger samples to a similar molarity. Close examination of full-length and ΔAC-Pass1 samples by Western blot and centrifugation revealed the presence of small precipitates. TANGO analysis of the Smp passenger domain amino acid sequence revealed that, similar to the Vag8 passenger domain, the Smp passenger possessed multiple regions, particularly in the N-terminal half of the passenger, that were prone to aggregation (Fig. 18).

To generate similar quality purified Smp passenger protein samples, new samples were column-purified under denaturing conditions followed by a protracted and gradual step-wise dialysis to remove urea and to allow proper folding to occur. Samples were then filtered prior to determination of sample molarity and preparation for far-UV CD. The filtered full-length passenger sample had similar levels of circular dichroism as the extended full-length passenger sample and extensive α-helical secondary structure (Fig. 27A). The filtered ΔAC-Pass1 and ΔAC-Pass2 samples also had similar levels of dichroism as the full-length passenger samples. The double minima of the spectra suggested extensive α-helical content and the trough extending from ~200-230 nm is characteristic of a partially folded protein (Fig. 27B,C). To verify that the filtered samples also lacked protease activity filtered Smp passenger samples were aliquoted on a milk-Luria agar plate and incubated at 50° C overnight. Trypsin, included as a positive control, resulted in a zone of clearing whereas the Smp samples did not, indicating that purified proteins lacked protease activity (Fig. 28).

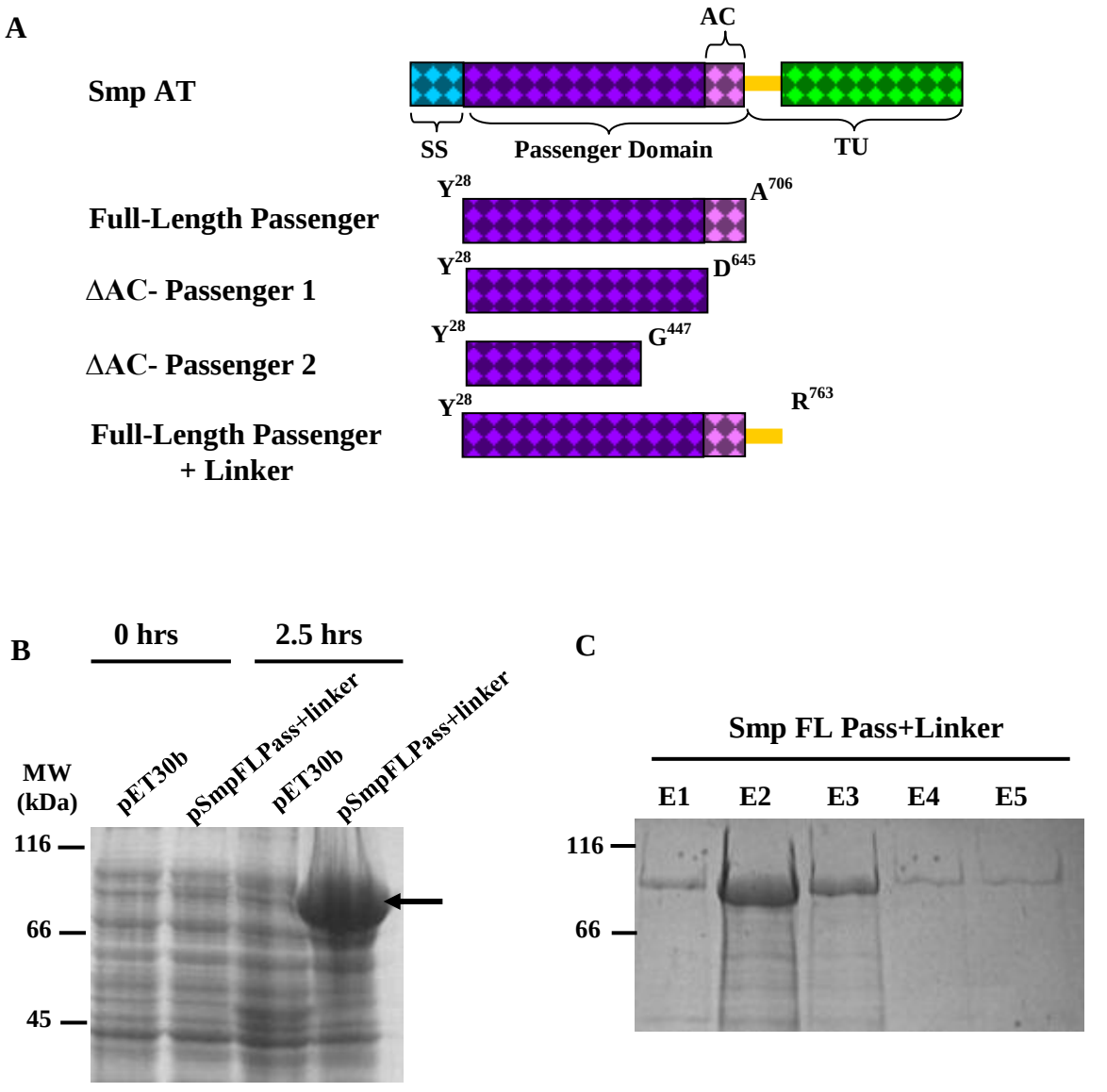


Figure 25. Expression and purification of the extended full-length Smp passenger variant. A pET30b construct was designed to encode an extended full-length Smp passenger variant. (A) Smp FL Pass+Linker encodes the full-length Smp passenger and the α-linker of the TU. (B) Expression of the construct was induced in BL21 (DE3) with 100 mM IPTG for 2.5 hrs. (C) Induced proteins were harvested from cell lysates and purified by Ni-NTA under denaturing conditions. Samples were visualized by SDS-PAGE followed by staining with Coomassie Brilliant Blue. The numbers on the left correspond to molecular weight in kDa.

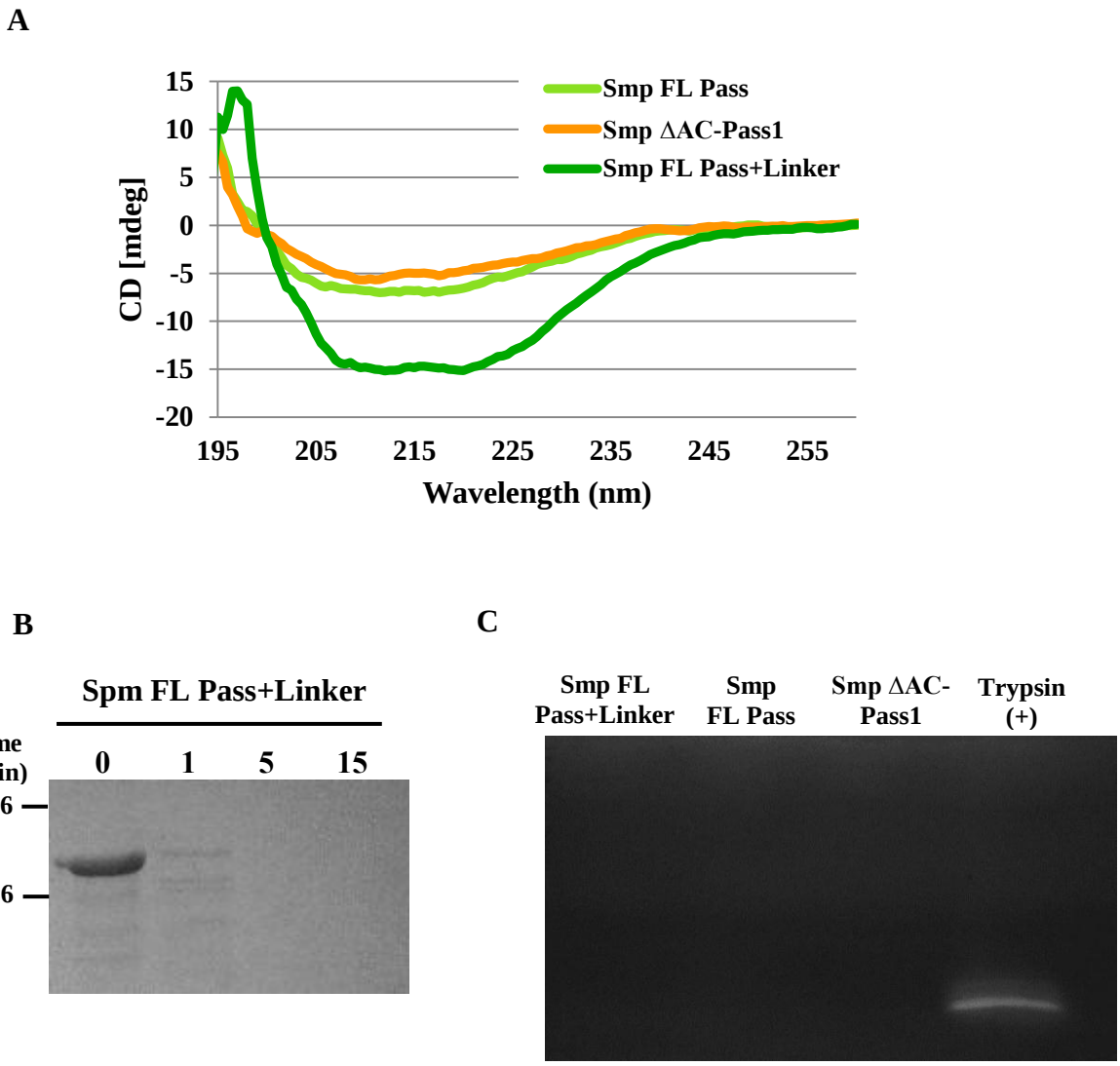


Figure 26. Assessing the folded state of the extended full-length Smp passenger variant. Following a 7 day incubation in 10 mM Tris pH8 folding buffer, the folded state of purified extended full-length Smp passenger (Smp FL Pass+Linker) was assessed by far-UV CD spectroscopy or limited proteolysis with trypsin. (A) Overlays of the CD spectra of FL, ΔAC-Pass1 and FL Pass+Linker Smp passenger variants. (B) Limited proteolysis analysis of Smp Pass+Linker with trypsin. Following addition of trypsin, samples were removed at 0, 1, 5, and 15 min. Samples were visualized by SDS-PAGE followed by Coomassie staining. (C) Zymogram of Smp passenger variants for serine protease activity in a 11% SDS-PAGE gel with 1% casein. Following washes to remove SDS from the PAGE gel, the gel was incubated at 37^{o} C overnight followed staining with Coomassie Brilliant Blue. The numbers on the left correspond to molecular weight in kDa.

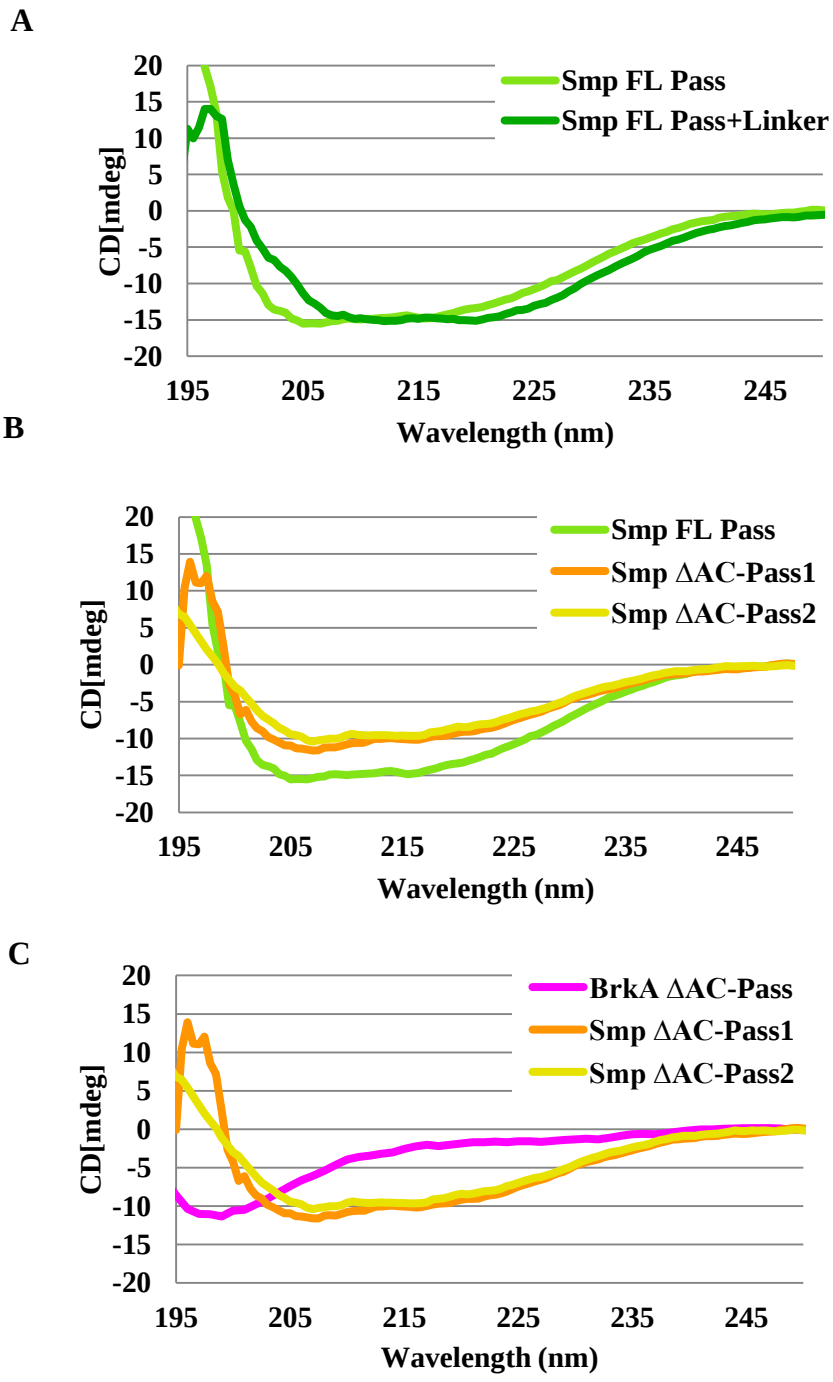


Figure 27. Far-UV CD spectroscopic analysis of filtered Smp passenger variants. Smp FL and ΔAC-passenger variants were re-purified under denaturing conditions followed by gentle gradient dialysis to replace urea with Tris folding buffer. (A) Overlay of the CD spectra of the Smp FL Pass+Linker and filtered Smp FL Pass proteins. (B) Overlay of the CD spectra of filtered Smp FL Pass and AC-deleted passenger proteins. (C) Overlay of the CD spectra of filtered AC-deleted Smp passenger proteins and unfolded BrkA ΔAC-Pass protein.

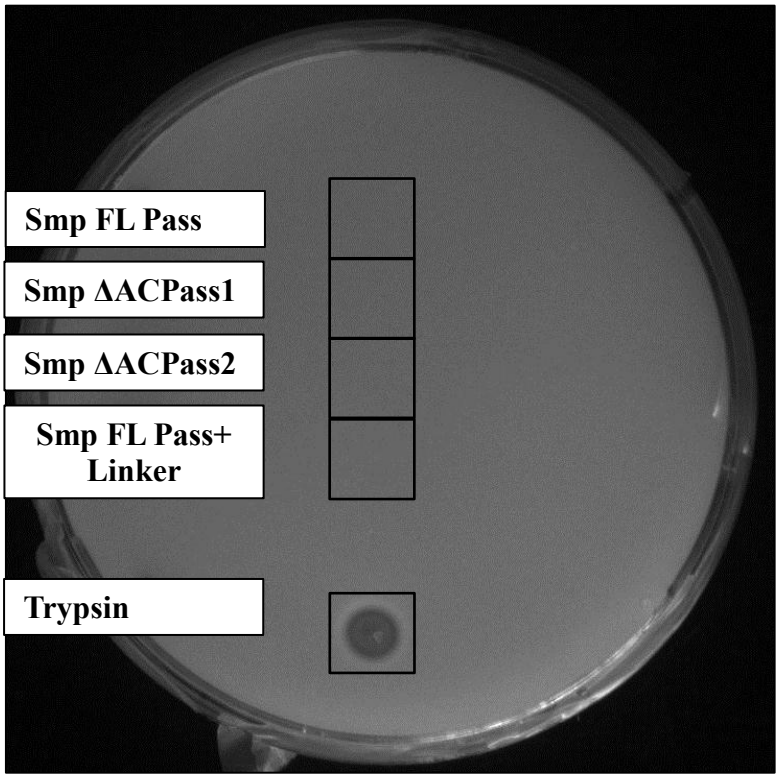


Figure 28. Milk plate test for protease activity of purified Smp passenger variants. Equal amounts of purified, filtered Smp passenger variants were aliquoted on a 3% skim milk Luria agar plate and were incubated at 50° C overnight. Trypsin was included as a positive control. Protease activity was detected as a presence of a halo/clearing of casein within the agar medium. Boxes indicate location of aliquoted samples.

Deletion of a C-terminal region of the passenger predicted to form β-sheet structure (ΔAC-Pass2) resulted in a CD spectrum nearly superimposable to the spectrum of the sample possessing that region (ΔAC-Pass1), indicating that Smp variants purified *in vitro* were unable to acquire at least a portion of the predicted β-helical structure (Fig. 27B,C). Altogether, extensive α-helical structure of the full-length and AC-deleted variants and the lack of protease activity of the full-length passenger domain proteins suggests that the conserved motifs at the C-terminus of the Smp passenger were not necessary for the acquisition of α-helical secondary structure, and that the presence of the C-terminal motifs alone, or in addition to the α-helical linker of the TU, were not sufficient to promote native folding of the Smp passenger domain *in vitro* (Fig. 29A,B).

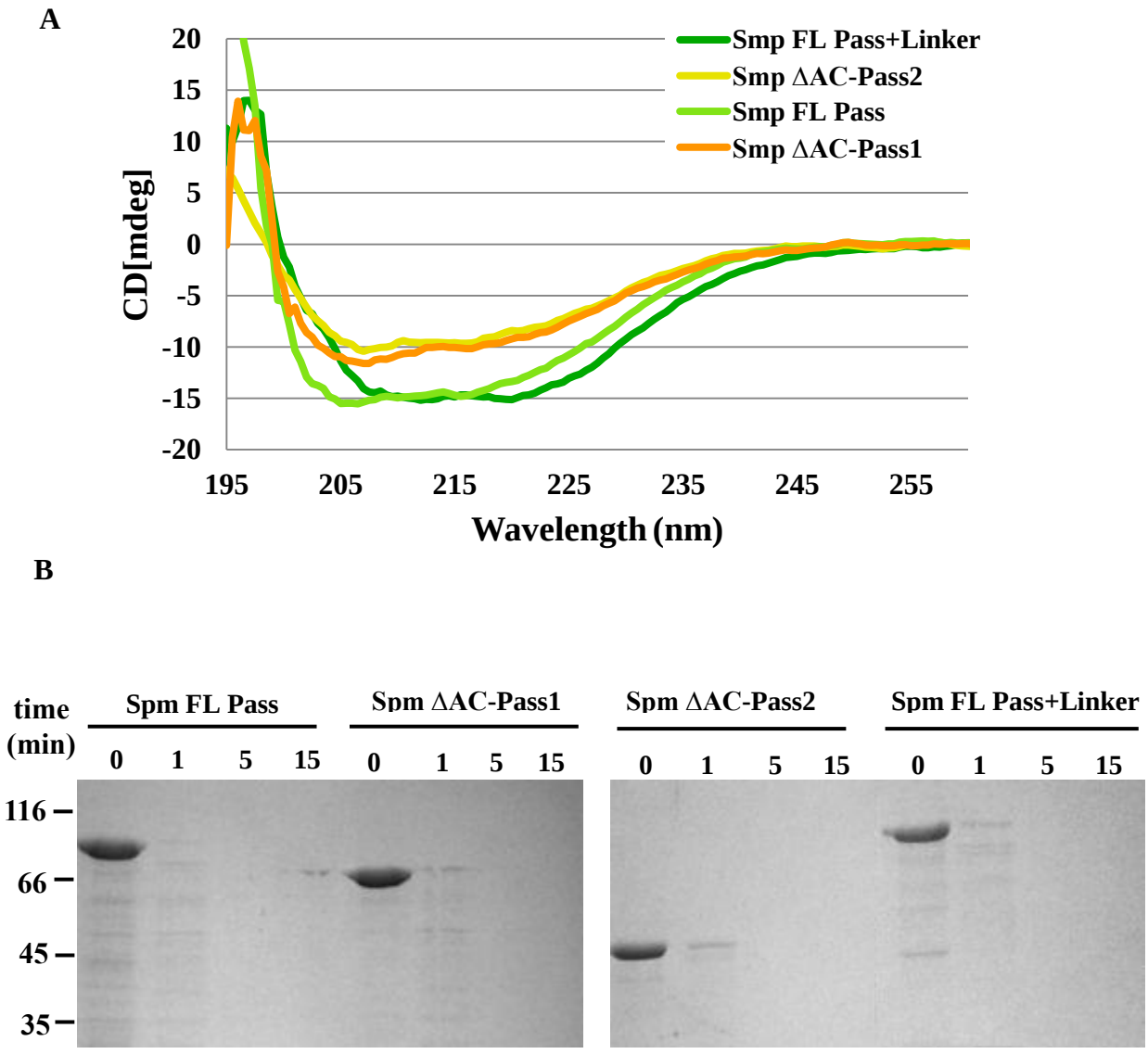


Figure 29. CD spectra overlay and comparison of trypsin sensitivity of Smp passenger variants. (A) Overlay of CD spectra of purified, filtered Smp passenger variants. All samples were diluted to 3 µM and submitted to three CD spectroscopy scans and measurements were taken in the 190-275 nm range at 20° C. (B) Limited proteolysis of Ag43 passenger variants with trypsin. Samples were resolved by SDS-PAGE followed staining with Coomassie Brilliant Blue. The numbers on the left correspond to molecular weight in kDa.

4. INTERCHANGEABILITY OF ACS

4.1 Design and expression of AC+TU constructs

The ability of a conserved C-terminal AC region to rescue passenger domain resistance to degradation by proteases has been demonstrated for BrkA and SSP, and the C-terminal AC region of EspP has been demonstrated to rescue trypsin resistance of the Pet passenger domain (11, 50, 51). Presumably, the AC region at the C-terminus of the EspP passenger is able to act as a template to promote the proper folding of the unfolded Pet passenger as both EspP and Pet possess EspP-like ACs that form templates with similar structure. All putative ACs of the BrkA putative AC group are predicted to form similar β-helix structures, therefore; I hypothesized that any BrkA group AC would be able to act as a template to promote the proper folding of the AC-deleted BrkA passenger domain, whereas the conserved C-terminal passenger domain motifs belonging to a different putative AC group would form a template with different structure and would be unable to rescue folding of the BrkA passenger domain.

To determine the ability of different AC motifs to rescue purified AC-deleted BrkA passenger folding, the ability of Spm, Ag43 and Vag8 AC regions to rescue BrkA passenger domain folding *in trans* was tested. A set of pET20b constructs encoding the BrkA, Vag8, Ag43 or Smp AC/conserved C-terminal passenger domain regions and TU were cloned (Fig. 30). The pET20b vector encodes the PelB SS N-terminal to its MCS. As a result, the AC+TU polypeptides possess the PelB SS which targets the AC+TU polypeptides for translocation across the IM. The TU targets the polypeptide to the OM. Two Smp AC+TU constructs were cloned. The first construct (pSmpAC+TU1) encodes the conserved α-helical region at the C-terminus of Smp passenger domain in addition to the Smp TU, and the second construct (pSmpAC+TU2) encodes a portion of the predicted β-sheet region of the Smp passenger in addition to the C-terminal α-helical region and the TU.

Expression of the AC+TU pET20b constructs was induced in BL21 (DE3) with IPTG. Cell wall fractions of induced lysates were harvested and run on SDS-PAGE followed by Coomassie staining to verify the presence of AC+TU proteins in the cell wall fraction (Fig. 31). Bands corresponding to the Ag43 and Smp AC+TU proteins were detected. The Ag43 and Smp conserved C-terminal motifs are located C-terminal to their passenger domain cleavage sites. Therefore, the Ag43 and Smp AC+TU proteins are present as single intact proteins ~55, 44 and 56 kDa in size. BrkA however, undergoes an intra-barrel cleavage event following OM translocation. As a result,

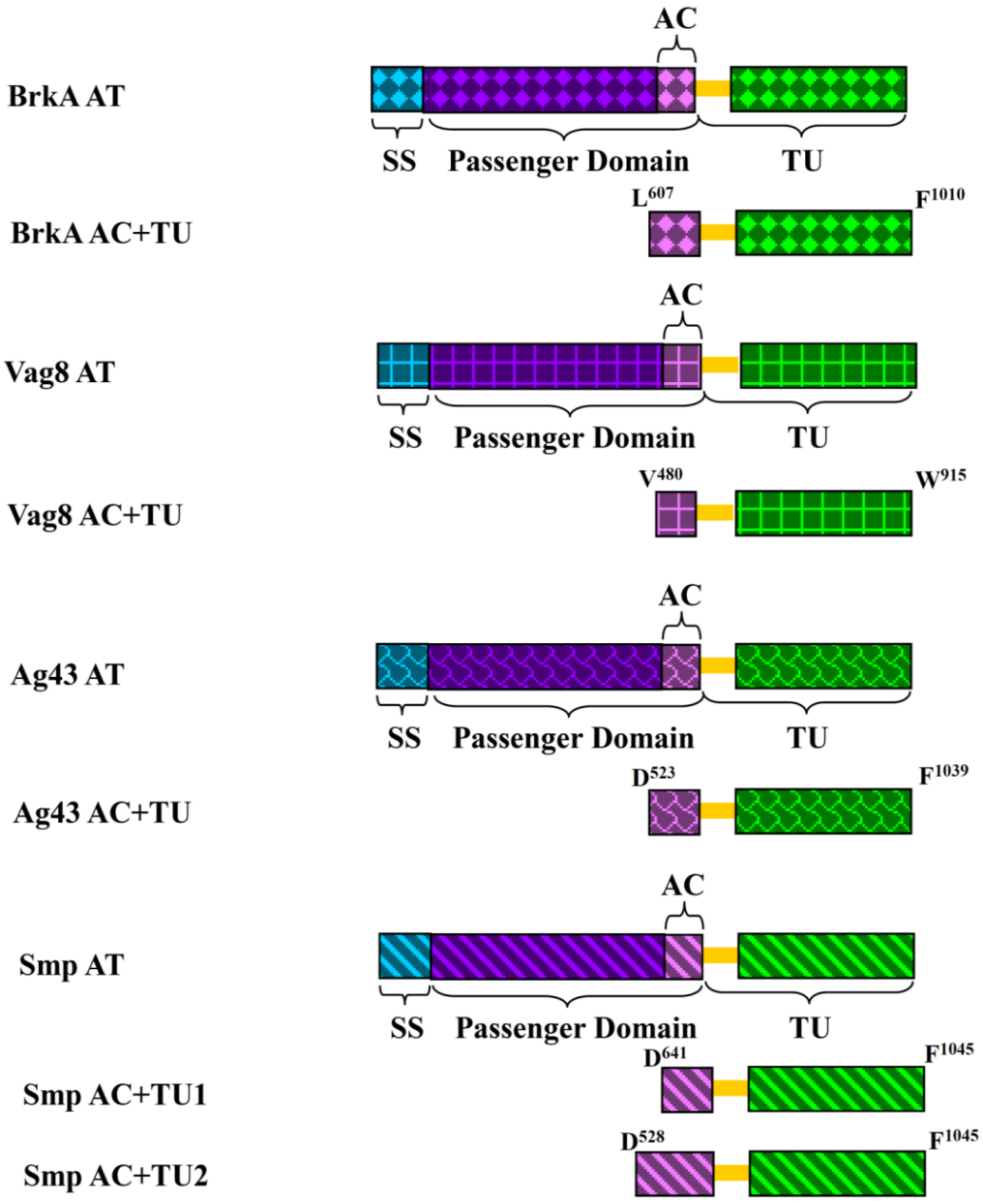


Figure 30. Structural representations of AC+TU proteins. The full-length autotransporter is characterized by the presence of three distinct domains; an N-terminal signal sequence (SS), a passenger and a C-terminal translocation unit (TU). Conserved regions at the C-terminus of the passenger domain that are important for secretion and folding are autochaperones (AC). A construct encoding the AC and TU (AC+TU) was cloned for BrkA, Vag8 and Ag43. Two AC+TU constructs were designed and cloned for Smp. Smp AC+TU1 encodes the predicted α-helical region at the C-terminus of the Smp passenger and Smp AC+TU2 encodes the C-terminal 178 amino acids of the passenger domain which includes a region of predicted β-sheet structure in addition to the predicted α-helical region.

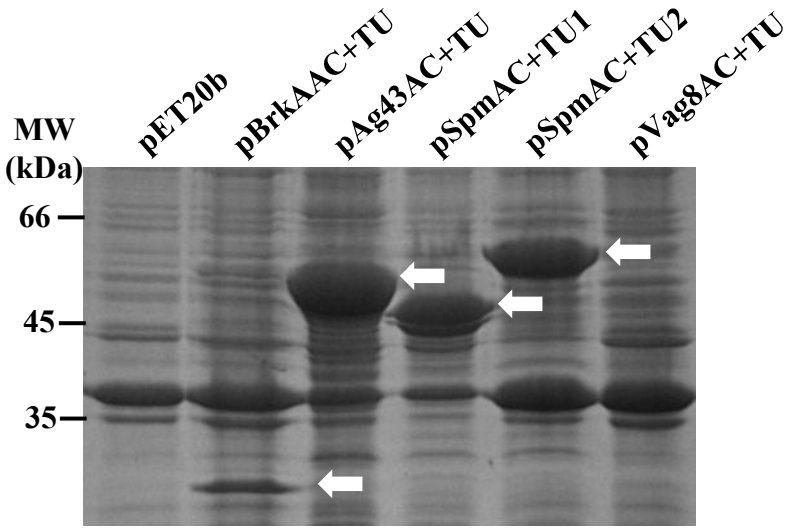


Figure 31. Cell wall fraction of whole cell lysates to verify the presence of AC+TU variants in the cell envelope. BL21 (DE3) were transformed with empty vector (pET20b), BrkA AC+TU, Vag8 AC+TU, Ag43 AC+TU and Spm AC+TU constructs. Transformed bacteria were induced with IPTG and cell wall fractions were obtained. Samples were run on SDS-PAGE followed by Coomassie blue staining of the gels. Arrows indicate induced AC+TU proteins. The numbers on the left correspond to molecular weight (MW) in kDa.

the passenger component of the BrkA AC+TU protein is cleaved from the TU. A ~30 kDa band corresponding to the BrkA TU was detected on the Coomassie stained gel and the uncleaved AC+TU was detected by Western blotting with BrkA passenger antiserum (data not shown). A band corresponding to the Vag8 AC+TU was not detected indicating low or no expression from the Vag8 AC+TU construct.

4.2 C-terminal regions of the Smp passenger are not able to rescue folding of the BrkA passenger domain.

Purified AC-deleted BrkA passenger with an N-terminal 6xHis-tag, verified to be unfolded by CD spectroscopy and trypsinolysis, was incubated with a cell wall fraction containing one of the AC+TU proteins. Following an overnight incubation, samples were subjected to limited proteolysis with trypsin followed by SDS-PAGE and Western blotting with anti-His primary antibody to determine if the different AC+TUs had rescued AC-deleted BrkA passenger domain folding. Rescue of folding was detected as an increased resistance of the BrkA passenger protein to digestion with trypsin relative to BrkA AC-deleted passenger incubated with the pET20b control cell wall fraction. AC-deleted BrkA passenger incubated with the control fraction was almost completely degraded 5 min following addition of trypsin and undetectable 15 min following addition of trypsin (Fig. 32). An increase in resistance to digestion with trypsin was detected for BrkA passenger incubated with BrkA and Ag43 AC+TU cell wall fractions, as AC-deleted BrkA was detectable 15 min following addition of trypsin (Fig. 32). Similar to the pET20b control fraction, BrkA passenger proteins incubated with the Smp AC+TU cell wall fractions and the cell wall fraction of cells transformed with pVag8 AC+TU, which had no detectable expression following induction with IPTG, were undetectable 15 min following the addition of trypsin. As hypothesized, the conserved C-terminal Smp passenger domain regions were unable to rescue folding of the BrkA passenger domain *in trans*.

4.3 Construct design of BrkA fusion passenger proteins

In order to verify the results of the cell wall fraction folding assay, and in order to test the ability of the Vag8 AC to rescue BrkA passenger domain folding, a set of BrkA fusion passenger

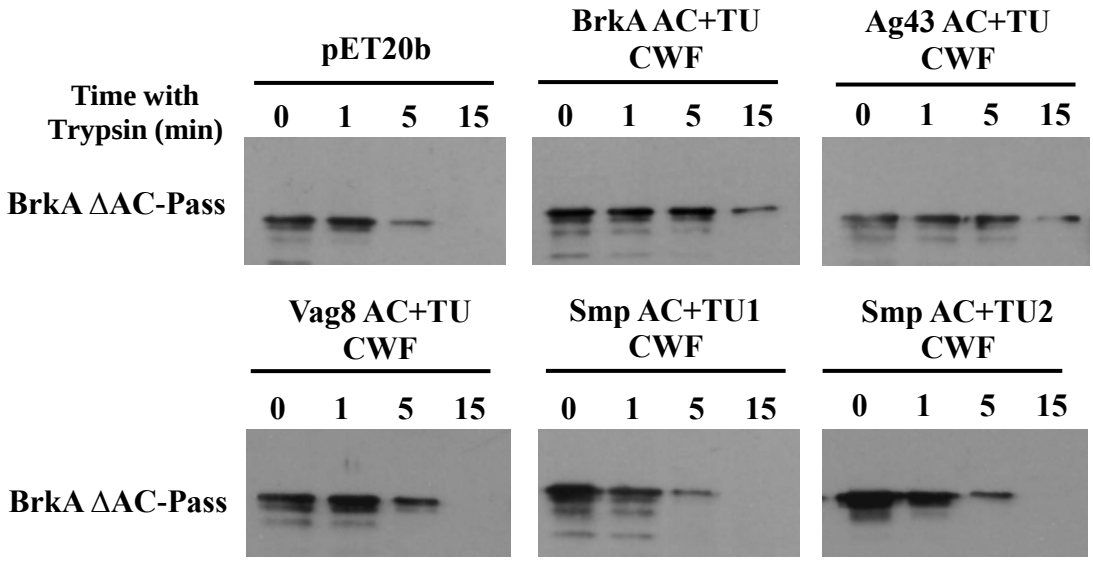


Figure 32. Cell wall fraction assay to test the ability of different ACs to rescue BrkA passenger folding. The cell wall fraction containing the AC+TU proteins were mixed with purified AC-deleted BrkA passenger protein and incubated together overnight at room temperature. The mixture was submitted to limited proteolysis with trypsin and samples were removed 0, 1, 5 and 15 min following the addition of trypsin. Samples were visualized by SDS-PAGE followed by western blotting with α-His primary antibody.

pET30b expression constructs were cloned. These constructs encoded the AC-deleted BrkA passenger cloned in frame with either the Vag8 or Smp passenger domain C-terminus (Fig. 33). In order to avoid interrupting an important secondary structure element that could impede folding, PSIPRED secondary structure prediction analysis was used to locate regions of coil at the N-terminal boundaries of the Vag8 and Smp conserved C-terminal domains. The Smp passenger C-terminal region, coinciding with the portion of the Smp passenger encoded by pSmpAC+TU2, was encoded by the BrkA-Smp fusion passenger construct.

4.4 Expression, purification and analysis of the folded state of BrkA passenger fusion proteins

Expression of BrkA fusion constructs was induced in BL21 (DE3) with IPTG (Fig. 34A). Proteins were harvested from induced cell lysates and purified via Ni-NTA under denaturing conditions (Fig. 34B). Urea in eluted protein samples was replaced with 10 mM Tris pH 8 folding buffer by step-wise dialysis and incubated at 4^{o} C for a week to allow folding to occur. Full-length BrkA passenger was also expressed, purified and dialyzed as a folding control.

To assess the folded state of BrkA fusion proteins and the full-length passenger control, samples were prepared for and analyzed by far-UV CD spectroscopy and limited proteolysis with trypsin. The full-length BrkA control had a spectrum minimum of ~218 nm and was detected 15 min following addition of trypsin, indicating that the full-length BrkA passenger protein had folded (Fig. 35&36). The BrkA fusion proteins however, were rapidly degraded with trypsin compared to the full-length control (Fig. 35A, B) and both had a spectrum minimum of ~200 nm which was consistent with unfolded protein (Fig. 36 A, B). It was hypothesized that the BrkA-Vag8 passenger would fold, therefore, the BrkA-Vag8 Pass protein was reanalyzed by CD spectroscopy and limited proteolysis with trypsin following an additional week of incubation in folding buffer. The BrkA-Vag8 passenger protein did not possess β-sheet structure according to CD spectroscopy and had not acquired resistance to digestion with trypsin. These results indicate that with the construct design and experimental conditions used, the Vag8 AC and Smp passenger C-terminus were unable to promote the folding of the BrkA passenger domain *in cis*.

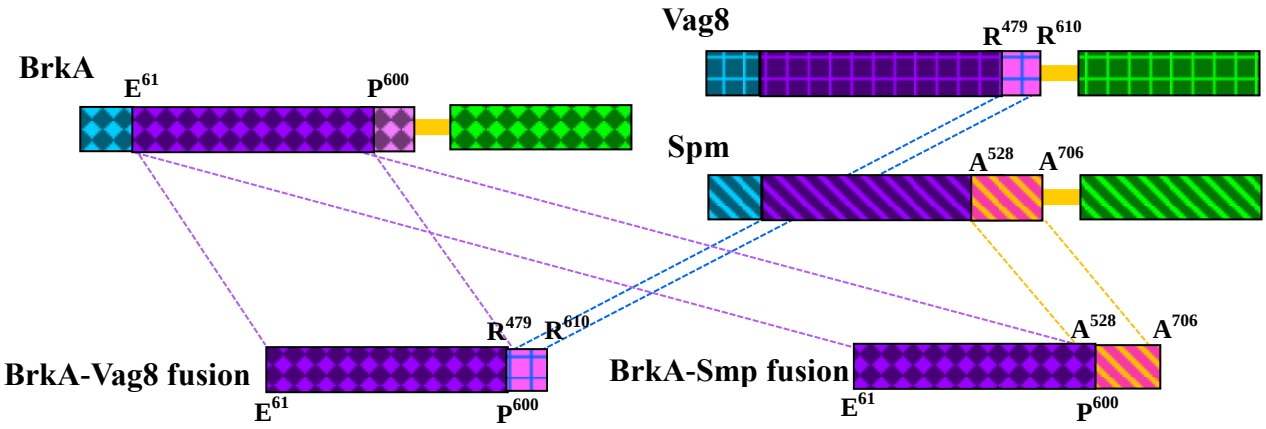


Figure 33. Structural representations of the BrkA passenger domain fusion proteins. Two BrkA full-length passenger fusion constructs were designed to test the ability of different ACs to rescue the folding of the BrkA passenger domain *in cis*. The BrkA-Vag8 fusion construct encodes the AC-deleted BrkA passenger cloned in frame with the C-terminal AC region of the Vag8 passenger domain. The BrkA-Smp fusion construct encodes the AC-deleted BrkA passenger cloned in frame with the C-terminal 178 residues of the Smp passenger domain which correspond to the passenger fragment encoded by the Smp AC+TU2 construct.

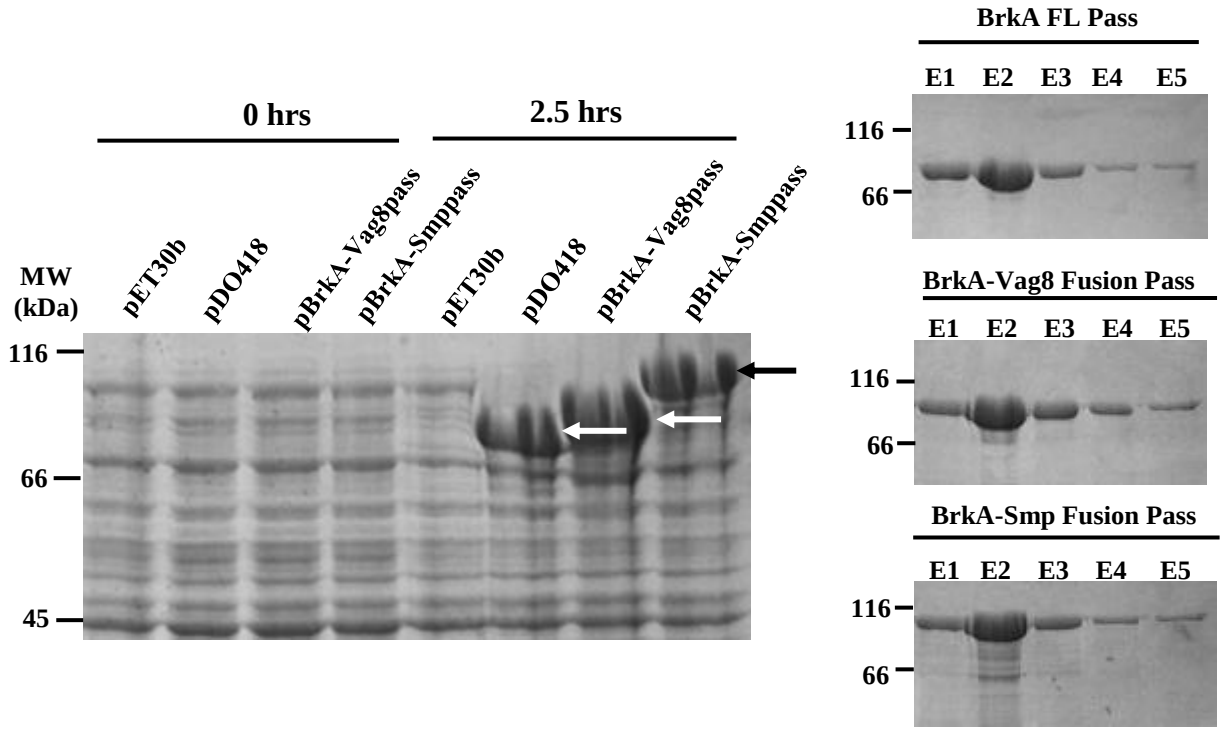


Figure 34. Expression and purification of BrkA fusion passenger proteins. (A) Expression of BrkA full-length passenger and fusion constructs was induced for 2.5 hrs with 100 mM IPTG. (B) Induced proteins were harvested from cell lysates and purified by Ni-NTA under denaturing conditions. Samples were visualized by SDS-PAGE followed by staining with Coomassie Brilliant Blue. The numbers on the left correspond to molecular weight in kDa.

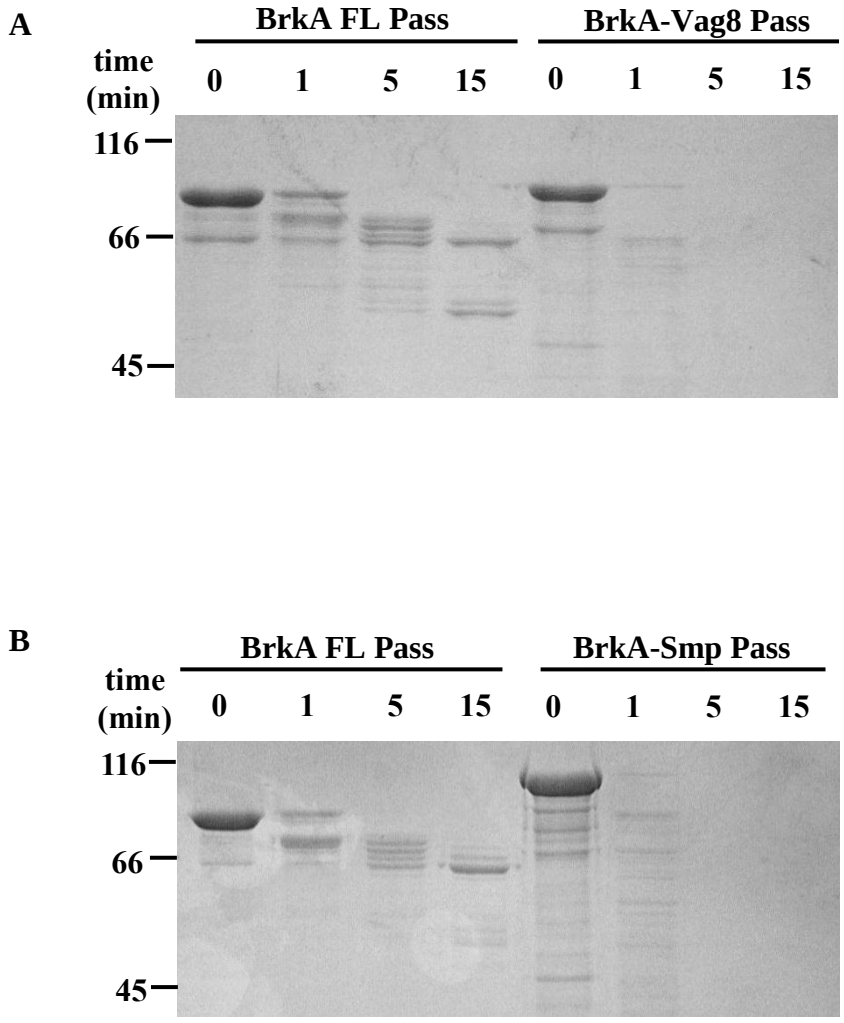
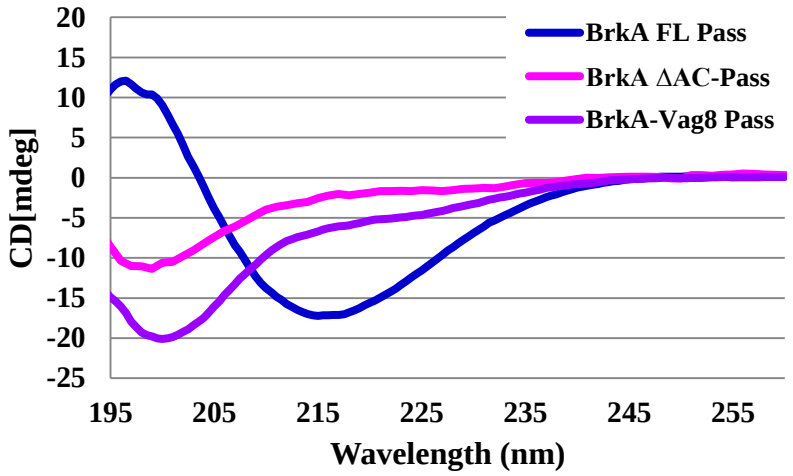


Figure 35. Limited proteolysis of BrkA passenger fusion proteins. Following a seven day incubation in 10 mM Tris pH 8 folding buffer, the folded state of purified BrkA fusion passenger proteins was assessed with limited proteolysis with trypsin. 20 µg of dialyzed BrkA-Vag8 passenger (A) and BrkA-Smp passenger (B) samples were subjected to proteolysis with 0.5 µg of trypsin and samples were removed at 0, 1, 5 and 15 min following the addition of trypsin. BrkA full-length passenger was refolded and digested with trypsin as a folding control. Samples were resolved by SDS-PAGE followed staining with Coomassie Brilliant Blue.

A



B

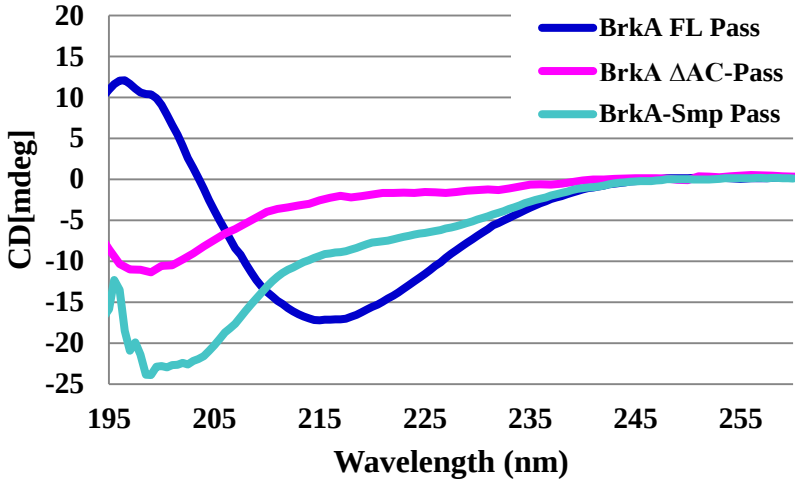


Figure 36. Far-UV circular dichroism spectroscopic analysis of BrkA passenger fusion proteins. Following a seven day incubation in 10 mM Tris pH8 folding buffer, the folded state of purified BrkA passenger fusion proteins was assessed with far-UV CD spectroscopy. Dialyzed protein samples were diluted to 3 µM and submitted to three CD spectroscopy scans and measurements were taken in the 190-265 nm range at 20°C. (A) CD spectra overlay of BrkA-Vag8 fusion passenger protein and BrkA full-length and AC-deleted passenger proteins. (B) CD spectra overlay of BrkA-Smp fusion passenger protein and BrkA full-length and AC-deleted passenger proteins.

5. DISCUSSION AND FUTURE DIRECTIONS

5.1 Project summary and conclusions

The purpose of this study was to determine the requirement of conserved C-terminal regions of AT passenger domains for passenger domain folding *in vitro*. The requirement of conserved C-terminal passenger domain motifs for the acquisition of passenger domain secondary structure and the interchangeability of these conserved AC motifs was investigated. A combination of far-UV CD spectroscopy and limited proteolysis with trypsin of full-length and AC-deleted Vag8, Ag43 and Smp passenger variants revealed that the requirement of conserved C-terminal AC motifs for the acquisition of passenger domain secondary structure varies among ATs. A cell wall fraction folding assay, which tested the ability of BrkA, Ag43 and Smp AC regions to rescue BrkA AC-deleted passenger folding, indicated that an AC region with similar structure to the cognate AC is necessary to rescue passenger domain folding.

5.2 Vag8 passenger domain folding and aggregation.

In accordance with the hypothesis that conserved C-terminal AC motifs are required for the acquisition of passenger domain secondary structure, the PF03212 motif, which is required for BrkA passenger domain folding, is required for the acquisition of Vag8 passenger domain structure *in vitro*. Also similar to BrkA, the Vag8 passenger domain is predicted to be comprised of entirely β-sheet and coil structure. However, unlike the BrkA passenger, the Vag8 passenger domain was prone to aggregation during dialysis. TANGO analysis of the Vag8 passenger domain amino acid sequence revealed aggregation-prone regions at the N-terminus of the Vag8 passenger, where the functional region is presumably located, that are absent from the easily purified and refolded passenger domains of BrkA, pertactin and Ag43. A hybrid column approach was ultimately used to prevent aggregation of the Vag8 passenger domain upon removal of urea. During hybrid column purification, the N-terminus of the Vag8 passenger is anchored via a 6xHis-tag to nickel-coated beads. This anchoring likely prevents or decreases non-productive/off-pathway interactions within and among Vag8 passenger domains that could result in aggregation upon the gradual removal of urea from the column with native washes. Vag8 passenger domain folding, akin to BrkA and pertactin passenger domain folding, occurs slowly *in vitro* and does not occur on a biologically feasible time scale (32, 51). Tethering of the Vag8 passenger during the 4 day incubation of the column in native wash buffer likely increased

the rate of passenger domain folding, compared to Hybrid1-3 purifications, by decreasing the number of off-pathway species thereby increasing the amount of on-pathway species capable of obtaining native conformation (26). Given that the Vag8 passenger domain *in vivo* folding environment differs greatly from the *in vitro* folding environment, aggregation is likely not an issue for secretion across the OM or folding at the bacterial cell surface. The vectorial transport of the Vag8 passenger across the OM, and the resultant sequential folding, likely increases the rate of folding and prevents aggregation by smoothing the Vag8 passenger folding/energy landscape in favour of the natively folded protein. This likely occurs by limiting possible off-pathway interactions including aggregation promoting inter- and intramolecular interactions among Vag8 passenger domains and other OM moieties (26).

5.3 Acquisition of Ag43 passenger domain secondary structure in the absence of the AC

Unlike Vag8 and BrkA, the C-terminal half of the Ag43 passenger domain is not necessary for the acquisition of secondary structure. However, the C-terminal PD002475 appears to be necessary for the acquisition of native structure. In its absence there were noticeable differences in the CD spectrum compared to that of the full-length passenger protein. Additionally, the truncated variant was not cleaved to release Ag43α suggesting it lacked native structure. Increasingly larger C-terminal passenger domain truncations also resulted in decreasing amounts of β-sheet structure, as detected by CD spectroscopy, and increasing sensitivity to digestion with trypsin. Nonetheless, the Ag43 passenger variant corresponding to Ag43α (Ag43 ΔAC-Pass2) was able to obtain a significant degree of secondary structure and resistance to trypsin *in vitro*. These results coincide with the results of a study by Knudsen *et al*. who demonstrated that Ag43α acquired a partially protease-resistant state in the periplasm (38). The authors proposed that Ag43α acquired partial structure but could not rule out the involvement of protective periplasmic chaperones in the apparent resistance of Ag43α to periplasmic proteases (38).

If the Ag43 passenger does not require an AC-template to acquire extensive β-sheet structure and can acquire this structure in the periplasm, how is the Ag43 passenger translocated across the OM to the bacterial cell surface? There is no evidence for chaperone-mediated unfolding of passenger domains in the periplasm (32). Therefore, it is possible that the pore of the β-domain is maintained in a sufficiently open and flexible conformation to allow

translocation of a passenger domain with partial structure onto the bacterial cell surface, where vectorial secretion would promote vectorial assembly of the passenger domain into its native structure. Translocation of a passenger domain with partial structure, and the necessity of a β-domain with a pore of sufficient size to accommodate that structure, is consistent with the model of OM translocation in which the Bam complex machinery interacts with the AT β-domain to maintain it in a transiently dilated conformation (25, 78, 79). However, it is more likely that the Ag43 passenger domain lacks the opportunity to acquire extensive secondary structure during its short duration within the periplasm given the slow rate of passenger domain folding *in vitro* (32, 51). Additionally, Ag43 possesses an extended signal peptide region (38, 74). The extended signal sequence of EspP, another AT whose passenger domain has been demonstrated to acquire partial structure in the periplasm, is required to prevent the EspP passenger from acquiring a translocation-incompatible conformation within the periplasm (68, 72). The extended signal sequence of EspP is thought to transiently anchor the N-terminus of the EspP polypeptide to the IM thereby preventing compact folding and facilitating the binding of periplasmic chaperones before folding takes place (72). It is possible that the extended signal sequence of Ag43α is effective in preventing passenger domain folding to the same extent as the EspP extended signal sequence (38).

5.4 Inability of the full-length Smp passenger to acquire native structure *in vitro*

Given the inability of the full-length Smp passenger to acquire native structure *in vitro*, it was not possible to determine the requirement of conserved C-terminal putative AC regions of the Smp passenger for acquiring native structure. However, the undeniable presence of α-helical structure detected in both the full-length and AC-deleted variant CD spectra revealed that the Smp passenger C-terminus was not required for the acquisition of passenger domain α-helical structure. The ability of the predicted α-helical regions to acquire secondary structure is not surprising given that the α-helices of the Smp passenger subdomains would not require a template to promote rapid folding (65). In addition to the predicted α-helical region at the C-terminus of the passenger domain, there is an extensive region of predicted α-helical structure at the N-terminus of the passenger within the protease domain. The sensitivity of these regions of α-helical structure to digestion with trypsin, and the lack of protease activity in the full-length and AC-deleted Smp passenger variants, indicated that regions of the passenger proteins that had

acquired α-helical structure lacked a native conformation in the absence of the predicted β-helix structure of the passenger. Additionally, the full-length Smp passenger was not cleaved to release the mature protease, further suggesting that the Smp passenger domain failed to acquire native structure *in vitro.*

Interestingly, denatured full-length SSP passenger was also unable to acquire functional protease activity following removal of urea, indicating that, similar to Smp, the SSP passenger was unable to refold to native structure following treatment with urea (50). Rescue of unfolded SSP passenger function was successful upon addition of an OM fraction containing the C-terminus of the SSP passenger as well as the TU (50). Therefore, native folding of the Smp passenger may require additional factors present in the OM to acquire native structure. It is more likely that the C-terminus of the Smp passenger requires interaction with the β-domain of the TU in order to promote native folding. The recently solved crystal structure and characterization of the BrkA β-domain identified an extracellular loop (L4) and hydrophobic patch that is conserved among many AT β-domains and is proposed to interact with the C-terminus of the passenger domain (Fig. 3) (82). Studies of other ATs, including AIDA-I, have demonstrated interaction of the β-domain with the C-terminus of the passenger domain (39). It is possible that the conserved loop and hydrophobic patch of the β-domain serves as a platform or anchor for the AC so that it swiftly acquires the correct structure to act as a template and nucleate the folding of the remainder of the passenger domain as it is extruded onto the bacterial cell surface.

To determine which regions at the C-terminus of the Smp passenger domain are required for Smp passenger native structure, a series of expression constructs encoding the complete Smp AT with a series of internal deletions in the putative AC region could be employed. If OM protease deficient *E. coli* exhibited additional protease activity when expressing one of the constructs, it would suggest that the deleted region of the Smp passenger is not required for the acquisition of Smp native structure *in vivo*. Additionally, purified Smp passenger could be incubated on a column with nickel-bound AC+TU proteins, followed by elution of the Smp passenger and assessment of its folded state with limited proteolysis to determine if incubation with the Smp AC+TU could rescue its resistance to digestion with trypsin. Similarly, far-UV CD spectroscopy of the eluted passenger could be used to detect the presence of additional β-sheet structure, which would indicate folding of the predicted β-helix component of the Smp passenger domain.

If the C-terminus of the passenger domain of some ATs require interaction with the β-domain in order to promote native passenger domain folding, why can the BrkA, Vag8 and Ag43 passengers acquire native structure *in vitro* in the absence of their β-domains? The answer to this question may be structure of the passenger domains themselves. ATs with protease function possess globular subdomains absent from the pure or predominantly β-sheet and coil passengers of Vag8, pertactin, BrkA and Ag43. In comparison to a pure β-sheet protein, a protein with globular subdomains would have a rougher and more complex energy landscape for folding (26). A complex folding mechanism and the abundance of aggregation prone regions in the Smp passenger could also explain the aggregation prone-nature of the Smp passenger during renaturation *in vitro* and once again highlights the importance of the vectorial OM translocation of the passenger domain for passenger domain folding at the bacterial cell surface.

5.5 What's in a name? Clarification of an inconsistency

Oliver *et al*. initially reported that the location of the BrkA AC motif coincides with PD002475 (51). However, recent analysis of the BrkA AT sequence with ProDom failed to identify PD002475 from the BrkA primary sequence. ProDom also failed to identify PD002475 from other ATs such as pertactin and Hap, which were previously reported by Oliver *et al*. to possess PD002475. On the other hand, ATs such as IcsA, Ag43 and AIDA-I still possess PD002475 but the region of the IcsA passenger corresponding to the Pfam pertactin motif does not coincide with the current location of PD002475, which is identified to span a region at the extreme C-terminus of the passenger domain and includes the C-terminal ~40 residues of the passenger domain, the α-helical linker and the N-terminal region of the β-domain. ProDom reports that this domain corresponds to the Pfam autotransoporter (β) domain. However, ProDom does identify domains in the regions of the BrkA, pertactin and IcsA passengers that are conserved and correspond to the PF03212 motif and the previously described region of the BrkA AC. PD607128 is identified from all three and is described by ProDom as corresponding to the Pfam pertactin family. For ATs such as BrkA and pertactin, the ProDom domain located C-terminal to PD607128 also corresponds to the Pfam pertactin family. PD607128 is identified in all AT passenger domains possessing BrkA Group ACs. However, only passengers with the full Pfam pertactin motif are identified by Pfam as possessing PF03212. Passengers with EspP and AIDA-I-like ACs possess different conserved motifs C-terminal to PD607128 that can be used to

distinguish them. The change in domain annotation likely reflects the evolution and growth of the database as a result of substantial sequence addition in recent years.

5.6 The requirement of a similar template to rescue BrkA passenger domain folding

It was expected that the C-terminus of the Ag43 passenger domain would rescue folding of the AC-deleted BrkA passenger. CD spectroscopy results revealed that full-length Ag43 passenger had similar secondary structure as the full-length BrkA passenger and, based on primary sequence alignments, BrkA and Ag43 had similar C-terminal passenger domain homology models with β-helix and β-hairpin structure. It was also expected that the C-terminus of the Smp passenger would fail to promote the folding of the AC-deleted BrkA passenger. The groupings of putative ACs were based on bioinformatic and phylogenetic analyses. As a result, it was not surprising that the Smp passenger C-terminus was predicted to possess different secondary structure than the ATs with BrkA Group ACs. Similarly, based on the Smp passenger primary sequence, SWISS-MODEL was unable to generate a homology model of the Smp passenger C-terminus. The inability to generate a C-terminal passenger domain model with β-helix structure, and the presence of predicted α-helical structure at the C-terminus of the Smp passenger, suggests that the Smp passenger C-terminus provides a template with different structure from that of the BrkA Group ACs.

EspP, Pet and other members of the EspP-like AC subgroup are predicted to possess ACs with similar structure as the BrkA-like and AIDA-I like ACs. To verify that an AC template with similar structure is sufficient to rescue BrkA passenger domain folding, the cell wall fraction folding assay could be carried-out using the Pet or EspP AC+TU cell wall fraction. Additionally, the ability of the BrkA AC+TU to promote folding of the Pet AC-deleted passenger domain and the ability of a Chlamydia Group AC+TU cell wall fractions to promote AC-deleted BrkA passenger domain folding could also be tested. *Chlamydia* AT passenger domains and putative ACs were not investigated in this study as a result of the cloning difficulties associated with the sequence similarity of *Chlamydia* ATs and the highly adenine- and thymine-rich repetitive nature of *Chlamydia* genomes.

5.7 The inability of the BrkA-Vag8 fusion passenger to fold *in vitro*

The Vag8 AC was expected to rescue folding of the AC-deleted BrkA passenger given that the EspP AC was previously demonstrated to rescue the folding of the Pet passenger at the bacterial cell surface (11). The Vag8 and BrkA ACs belong to the same BrkA AC subgroup, similar to the EspP and Pet ACs which also share the same AC subgroup. Additionally, CD spectroscopy revealed that BrkA and Vag8 share similar secondary structure and SWISS-MODEL generated nearly superimposable homology structure models of the C-terminus of their passenger domains. The inability of the Vag8 AC to promote folding of the AC-deleted BrkA passenger *in cis* (in the BrkA-Vag8 fusion passenger protein) does not necessarily indicate that the Vag8 AC forms a template that differs in structure from that of the BrkA AC, or that the Vag8 AC is unable to rescue folding of the BrkA passenger. In the absence of a BrkA fusion passenger protein that was able to successfully refold, it is possible that the fusion proteins were unable to fold as a result of the conditions or cloning strategy and construct design used. The crystal structures of Vag8 and Smp passengers have yet to be solved and as a result, construct and fusion boundaries were based on secondary structure prediction. Therefore, it is possible that an important structural element was interrupted, or the addition of two residues (a serine and arginine) as part of the cloning strategy implemented had an impact on the ability of the BrkA fusion passengers to fold *in vitro*. Future AC interchangeability experiments should take advantage of the solved crystal structure of model ATs such as pertactin, IgAP or EspP.

5.8 The autochaperone is likely just as ‘auto’ as the autotransporter

The term ‘autotransporter’ was initially coined as a result of the supposition that an AT protein autonomously catalyzes the transport of its own passenger domain to the bacterial cell surface (55). However, in recent years multiple factors including periplasmic chaperones and Bam complex machinery have been shown to play important roles in AT biogenesis. Nonetheless, the term autotransporter has endured. Similarly, the term ‘autochaperone’ was coined as a result of the supposition that the AC autonomously promotes passenger domain folding *in vitro* and at the bacterial cell surface (10). However, the AC may not act autonomously given that: a) full-length passenger domains fold slowly *in vitro*, b) the Vag8 and Smp passengers have a tendency to aggregate in the presence of their AC regions, c) the full-length Smp and SSP passenger domains are unable to acquire native structure *in vitro* and d) the Ag43

passenger can acquire extensive structure without its AC. The slow rate of passenger domain folding *in vitro*, in combination with the propensity of AT passenger domains to aggregate during folding *in vitro* (in the presence of the AC regions), emphasizes the importance of the *in vivo* energy/folding landscape for fast and native folding of the passenger domain at the bacterial cell surface. The inability of some AT passenger domains to acquire native structure *in vitro*, even in the presence of their AC regions, suggests that additional OM factors or contributions from the TU are necessary to achieve native passenger domain structure *in vivo*. Finally, the ability of the Ag43 passenger to fold in the absence of its AC region, and the proposed role of the extended signal sequence and periplasmic chaperones in the maintenance of passenger polypeptides in a secretion-competent state within the periplasm, suggests that in some ATs other regions within the passenger domain (in addition to the AC) that contribute to passenger domain folding. In summary, the propensity of only some AT passenger domains to aggregate during refolding *in vitro*, and the varied requirement of conserved C-terminal AC regions of different passenger domains for the acquisition of secondary structure and folding *in vitro*, highlights the involvement of multiple factors in passenger domain folding and the likely variation that exists in the mechanism of passenger domain secretion across the OM and folding at the bacterial cell surface.

5.9 Implications and applications

The results of this study demonstrate that the role of the AC in passenger domain folding varies among ATs, and highlights the roles of additional factors in passenger domain folding. In so doing, it provides additional insight into the mechanism of passenger domain folding at the bacterial cell surface and AT biogenesis. Additionally, this study supports the idea that inhibitors of the AT passenger folding mechanism that target the AC may have therapeutic potential as antimicrobials. AT passenger domains must fold into native structure following OM translocation in order to carry out their virulence function (44). The similar structure of ACs belonging to the same AC group is predicted by bioinformatic analyses and is supported by the ability of an AC of the AIDA-I-like subgroup to rescue the folding of a passenger domain with a BrkA-like subgroup AC. The similar structure and function of ACs within an AC group suggests that an AC-targeted folding inhibitor would be effective in inhibiting the pathogenesis of a wide range of clinically important Gram-negative bacterial pathogens.

BIBLIOGRAPHY

1. **Arnold, K., L. Bordoli, J. Kopp, and T. Schwede.** 2006. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics **22:**195-201.
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:**1214-20.
3. **Benz, I., and M. A. Schmidt.** 2011. Structures and functions of autotransporter proteins in microbial pathogens. Int J Med Microbiol **301:**461-8.
4. **Bradley, P., L. Cowen, M. Menke, J. King, and B. Berger.** 2001. BETAWRAP: successful prediction of parallel beta -helices from primary sequence reveals an association with many microbial pathogens. Proc Natl Acad Sci U S A **98:**14819-24.
5. **Brandon, L. D., and M. B. Goldberg.** 2001. Periplasmic transit and disulfide bond formation of the autotransported Shigella protein IcsA. J Bacteriol **183:**951-8.
6. **Cascales, E., and C. Cambillau.** 2012. Structural biology of type VI secretion systems. Philos Trans R Soc Lond B Biol Sci **367:**1102-11.
7. **Corpet, F., F. Servant, J. Gouzy, and D. Kahn.** 2000. ProDom and ProDom-CG: tools for protein domain analysis and whole genome comparisons. Nucleic Acids Res **28:**267-9.
8. **Dautin, N., and H. D. Bernstein.** 2007. Protein secretion in gram-negative bacteria via the autotransporter pathway. Annu Rev Microbiol **61:**89-112.
9. **De, E., N. Saint, K. Glinel, A. C. Meli, D. Levy, and F. Jacob-Dubuisson.** 2008. Influence of the passenger domain of a model autotransporter on the properties of its translocator domain. Mol Membr Biol **25:**192-202.
10. **Desvaux, M., N. J. Parham, and I. R. Henderson.** 2004. Type V protein secretion: simplicity gone awry? Curr Issues Mol Biol **6:**111-24.
11. **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:**39912-20.
12. **Economou, A., P. J. Christie, R. C. Fernandez, T. Palmer, G. V. Plano, and A. P. Pugsley.** 2006. Secretion by numbers: Protein traffic in prokaryotes. Mol Microbiol **62:**308-19.
13. **Emsley, P., I. G. Charles, N. F. Fairweather, and N. W. Isaacs.** 1996. Structure of Bordetella pertussis virulence factor P.69 pertactin. Nature **381:**90-2.
14. **Fernandez-Escamilla, A. M., F. Rousseau, J. Schymkowitz, and L. Serrano.** 2004. Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. Nat Biotechnol **22:**1302-6.
15. **Fernandez, R. C., and A. A. Weiss.** 1994. Cloning and sequencing of a Bordetella pertussis serum resistance locus. Infect Immun **62:**4727-38.
16. **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:**16293-8.
17. **Greenfield, N., and G. D. Fasman.** 1969. Computed circular dichroism spectra for the evaluation of protein conformation. Biochemistry **8:**4108-16.

18. **Greenfield, N. J.** 2006. Using circular dichroism spectra to estimate protein secondary structure. Nat Protoc **1:**2876-90.
19. **Hegde, R. S., and H. D. Bernstein.** 2006. The surprising complexity of signal sequences. Trends Biochem Sci **31:**563-71.
20. **Henderson, I. R., and J. P. Nataro.** 2001. Virulence functions of autotransporter proteins. Infect Immun **69:**1231-43.
21. **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.
22. **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.
23. **Henderson, I. R., F. Navarro-Garcia, and J. P. Nataro.** 1998. The great escape: structure and function of the autotransporter proteins. Trends Microbiol **6:**370-8.
24. **Holland, I. B.** 2010. The extraordinary diversity of bacterial protein secretion mechanisms. Methods Mol Biol **619:**1-20.
25. **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.
26. **Jahn, T. R., and S. E. Radford.** 2008. Folding versus aggregation: polypeptide conformations on competing pathways. Arch Biochem Biophys **469:**100-17.
27. **Jain, S., and M. B. Goldberg.** 2007. Requirement for YaeT in the outer membrane assembly of autotransporter proteins. J Bacteriol **189:**5393-8.
28. **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:**559-74.
29. **Jong, W. S., C. M. ten Hagen-Jongman, T. den Blaauwen, D. J. Slotboom, J. R. Tame, D. Wickstrom, J. W. de Gier, B. R. Otto, and J. Luirink.** 2007. Limited tolerance towards folded elements during secretion of the autotransporter Hbp. Mol Microbiol **63:**1524-36.
30. **Jose, J., J. Kramer, T. Klauser, J. Pohlner, and T. F. Meyer.** 1996. Absence of periplasmic DsbA oxidoreductase facilitates export of cysteine-containing passenger proteins to the Escherichia coli cell surface via the Iga beta autotransporter pathway. Gene **178:**107-10.
31. **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:**1323-32.
32. **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:**4918-23.
33. **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:**306-15.
34. **Khalid, S., and M. S. Sansom.** 2006. Molecular dynamics simulations of a bacterial autotransporter: NalP from Neisseria meningitidis. Mol Membr Biol **23:**499-508.

35. **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.
36. **Klauser, T., J. Pohlner, and T. F. Meyer.** 1992. Selective extracellular release of cholera toxin B subunit by Escherichia coli: dissection of Neisseria Iga beta-mediated outer membrane transport. EMBO J **11:**2327-35.
37. **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:**206-14.
38. **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:**563-77.
39. **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.
40. **Kuhnel, K., and D. Diezmann.** 2011. Crystal structure of the autochaperone region from the Shigella flexneri autotransporter IcsA. J Bacteriol **193:**2042-5.
41. **Kwon, Y. T., H. H. Lee, and H. M. Rho.** 1993. Cloning, sequencing, and expression of a minor protease-encoding gene from Serratia marcescens ATCC21074. Gene **125:**75-80.
42. **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:**1088-101.
43. **Letley, D. P., J. L. Rhead, K. Bishop, and J. C. Atherton.** 2006. Paired cysteine residues are required for high levels of the Helicobacter pylori autotransporter VacA. Microbiology **152:**1319-25.
44. **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:**213-25.
45. **Leyton, D. L., Y. R. Sevastsyanovich, 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:**42283-91.
46. **Ma, B., C. J. Tsai, and R. Nussinov.** 2000. Binding and folding: in search of intramolecular chaperone-like building block fragments. Protein Eng **13:**617-27.
47. **Marr, N., N. R. Shah, R. Lee, E. J. Kim, and R. C. Fernandez.** 2011. Bordetella pertussis autotransporter Vag8 binds human C1 esterase inhibitor and confers serum resistance. PLoS One **6:**e20585.
48. **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:**4666-76.
49. **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 (Berl) **88:**451-8.

50. **Ohnishi, Y., M. Nishiyama, S. Horinouchi, and T. Beppu.** 1994. Involvement of the COOH-terminal pro-sequence of Serratia marcescens serine protease in the folding of the mature enzyme. J Biol Chem **269:**32800-6.
51. **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:**1367-83.
52. **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:**1257-66.
53. **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:**17339-45.
54. **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 **107:**17739-44.
55. **Pohlner, J., R. Halter, K. Beyreuther, and T. F. Meyer.** 1987. Gene structure and extracellular secretion of Neisseria gonorrhoeae IgA protease. Nature **325:**458-62.
56. **Punta, M., P. C. Coggill, R. Y. Eberhardt, J. Mistry, J. Tate, C. Boursnell, N. Pang, K. Forslund, G. Ceric, J. Clements, A. Heger, L. Holm, E. L. Sonnhammer, S. R. Eddy, A. Bateman, and R. D. Finn.** 2012. The Pfam protein families database. Nucleic Acids Res **40:**D290-301.
57. **Records, A. R.** 2011. The type VI secretion system: a multipurpose delivery system with a phage-like machinery. Mol Plant Microbe Interact **24:**751-7.
58. **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:**420-7.
59. **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:**287-96.
60. **Rossiter, A. E., D. L. Leyton, K. Tveen-Jensen, D. F. Browning, Y. Sevastsyanovich, T. J. Knowles, K. B. Nichols, A. F. Cunningham, M. Overduin, M. A. Schembri, and I. R. Henderson.** 2011. The essential beta-barrel assembly machinery complex components BamD and BamA are required for autotransporter biogenesis. J Bacteriol **193:**4250-3.
61. **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:**6571-83.
62. **Sambrook, J., Fritsch, E. F. & Maniatis, T. .** 1989. Molecular cloning : a laboratory manual. 2nd edn., Cold Spring Harbor Laboratory.
63. **Sauri, A., N. Oreshkova, Z. Soprova, W. S. Jong, M. Sani, P. J. Peters, J. Luirink, and P. van Ulsen.** 2011. Autotransporter beta-domains have a specific function in protein secretion beyond outer-membrane targeting. J Mol Biol **412:**553-67.
64. **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:**3982-91.

65. **Shih, C. T., Z. Y. Su, J. F. Gwan, B. L. Hao, C. H. Hsieh, and H. C. Lee.** 2000. Mean-field HP model, designability and alpha-helices in protein structures. Phys Rev Lett **84:**386-9.
66. **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:**4654-9.
67. **Simeone, R., D. Bottai, and R. Brosch.** 2009. ESX/type VII secretion systems and their role in host-pathogen interaction. Curr Opin Microbiol **12:**4-10.
68. **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:**945-58.
69. **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:**2473-84.
70. **Soprova, Z., A. Sauri, P. van Ulsen, J. R. Tame, T. den Blaauwen, W. S. Jong, and J. Luirink.** 2010. A conserved aromatic residue in the autochaperone domain of the autotransporter Hbp is critical for initiation of outer membrane translocation. J Biol Chem **285:**38224-33.
71. **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:**6005-13.
72. **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:**221-6.
73. **van den Berg, B.** 2010. Crystal structure of a full-length autotransporter. J Mol Biol **396:**627-33.
74. **van der Woude, M. W., and I. R. Henderson.** 2008. Regulation and function of Ag43 (flu). Annu Rev Microbiol **62:**153-69.
75. **van Ulsen, P.** 2011. Protein folding in bacterial adhesion: secretion and folding of classical monomeric autotransporters. Adv Exp Med Biol **715:**125-42.
76. **Veiga, E., V. de Lorenzo, and L. A. Fernandez.** 1999. Probing secretion and translocation of a beta-autotransporter using a reporter single-chain Fv as a cognate passenger domain. Mol Microbiol **33:**1232-43.
77. **Veiga, E., V. de Lorenzo, and L. A. Fernandez.** 2004. Structural tolerance of bacterial autotransporters for folded passenger protein domains. Mol Microbiol **52:**1069-80.
78. **Velarde, J. J., and J. P. Nataro.** 2004. Hydrophobic residues of the autotransporter EspP linker domain are important for outer membrane translocation of its passenger. J Biol Chem **279:**31495-504.
79. **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:**262-5.
80. **Wagner, J. K., J. E. Heindl, A. N. Gray, S. Jain, and M. B. Goldberg.** 2009. Contribution of the periplasmic chaperone Skp to efficient presentation of the autotransporter IcsA on the surface of Shigella flexneri. J Bacteriol **191:**815-21.

81. **Wilson, D., M. Madera, C. Vogel, C. Chothia, and J. Gough.** 2007. The SUPERFAMILY database in 2007: families and functions. Nucleic Acids Res **35:**D308-13.
82. **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:**577-87.
83. **Zhao, L., N. T. Nguyen, R. C. Fernandez, and M. E. Murphy.** 2009. Crystallographic characterization of the passenger domain of the Bordetella autotransporter BrkA. Acta Crystallogr Sect F Struct Biol Cryst Commun **65:**608-11.

APPENDICES

Appendix A: Primary sequences of purified proteins

BrkA ΔAC-Pass

MHHHHHHSSGLVPRGSGMKQTAAAKFQREHMDSPDLGTDDDDKAMAQEGEFDHRDNTLIAVFDD
GVGINLDDDPDELGETAPPTLKDIHISVEHKNPMSKPAIGVRVSGAGRALTLAGSTIDATEGGI
PAVVRRGGTLELDGVTVAGGEGMEPMTVSDAGSRLSVRGGVLGGEAPGVGLVRAAQGGQASIID
ATLQSILGPALIADGGSISVAGGSIDMDMGPGFPPPPPPLPGAPLAAHPPLDRVAAVHAGQDGK
VTLREVALRAHGPQATGVYAYMPGSEITLQGGTVSVQGDDGAGVVAGAGLLDALPPGGTVRLDG
TTVSTDGANTDAVLVRGDAARAEVVNTVLRTAKSLAAGVSAQHGGRVTLRQTRIETAGAGAEGI
SVLGFEPQSGSGPASVDMQGGSITTTGNRAAGIALTHGSARLEGVAVRAEGSGSSAAQLANGTL
VVSAGSLASAQSGAISVTDTPLKLMPGALASSTVSVRLTDGATAQGGNGVFLQQHSTIPVAVAL
ESGALARGDIVADGNKPLDAGISLSVASGAAWHGATQVLQSATLGKGGTWVVNADSRVQDMSMR
GGRVEFQAPAP

Vag8 FL Pass

MHHHHHHSSGLVPRGSGMKQTAAAKFQREHMDSPDLGTDDDDKAMAISDPVTAAQRIDGGAAFL
GDVAIATTKASEHGINVTGRTAEVRVTGGTIRTSGNQAQGLRVGTENAPDNTALGASVFLQNLI
IETSGTGALGVSVHEPQGGGGTRLSMSGTTVRTRGDDSFALQLSGPASATLNDVALETAGQQAP
AVVLWQGAQLNAQGLVVQVNGAGVSAIHAQDAGSFTLSGSDITARGLEVVGIYVQEGMQGTLTG
TRVTTQGDTAPALQVEDAGTHVSMNGGALSTSGANSPAAWLLAGGSAQFRDTVLRTVGEASHGV
DVAAHSEVELAHAQVRADGQGAHGLVVTRSSAMVRAGSLVEST∆GDGAAALLESGHLTVDGSVV
HGHGAAGLEVDGESNVSLLNGARLSSDQPTAIRLIDPRSVLNLDIKDRAQLLGDIAPEAQQPDG
SPEQARVRVALADGGTWAGRTDGAVHTVRLLDRGVWTVTGDSRVAEVKLEGGTLAFAPPAQPKG
AFKTLVATQGISGTGTIVMNAHLPSGTADVLVAPQGFGDRQVLVVNNTDDGTESGATKVPLIED
EQGHTAFTLGNMGGRVDAGARQYELTASEAQADKARTWQLTPTNEL

Vag8 ΔAC-Pass1

MHHHHHHSSGLVPRGSGMKQTAAAKFQREHMDSPDLGTDDDDKAMAISDPVTAAQRIDGGAAFL
GDVAIATTKASEHGINVTGRTAEVRVTGGTIRTSGNQAQGLRVGTENAPDNTALGASVFLQNLI
IETSGTGALGVSVHEPQGGGGTRLSMSGTTVRTRGDDSFALQLSGPASATLNDVALETAGQQAP
AVVLWQGAQLNAQGLVVQVNGAGVSAIHAQDAGSFTLSGSDITARGLEVVGIYVQEGMQGTLTG
TRVTTQGDTAPALQVEDAGTHVSMNGGALSTSGANSPAAWLLAGGSAQFRDTVLRTVGEASHGV
DVAAHSEVELAHAQVRADGQGAHGLVVTRSSAMVRAGSLVESTGDGAAALLESGHLTVDGSVVH
GHGAAGLEVDGESNVSLLNGARLSSDQPTAIRLIDPRSVLNLDIKDRAQLLGDIAPEAQQPDGS
PEQARVRVALADGGTWAGRTDGAVHTVRLLDRGVWTVTGDSR

Vag8 ΔAC-Pass2

MHHHHHHSSGLVPRGSGMKQTAAAKFQREHMDSPDLGTDDDDKAMA**ISDP**VTAAQRIDGGAAFL
GDVAIATTKASEHGINVTGRTAEVRVTGGTIRTSGNQAQGLRVGTENAPDNTALGASVFLQNLI
IETSGTGALGVSVHEPQGGGGTRLSMSGTTVRTRGDDSFALQLSGPASATLNDVALETAGQQAP
AVVLWQGAQLNAQGLVVQVNGAGVSAIHAQDAGSFTLSGSDITARGLEVVGIYVQEGMQGTLTG
TRVTTQGDTAPALQVEDAGTHVSMNGGALSTSGANSPAAWLLAGGSAQFRDTVLRTVGEASHGV
DVAAHSEVELAHAQVRADGQGAHGLVVTRSSAMVRAGSLVESTGDGAAALLESGHLTVDGSVVH
GHGAAGLEVDGESNVSLLNGARLSSDQPTAIRLIDPRSVLNLDIKDRAQLLGDIAPEAQQPDGS
P

Smp FL Pass

MHHHHHHSSGLVPRGSGMKQTAAAKFQREHMDSPDLGTDDDDKAMAISDPNSSSYREPGQLGSP
DSWKNAEFNRQWGLEAISAEFAYARAYTGKGVTIGVIDDAILSHPEFAGKLTRLDNGSYNFSYD
KQDNMSFGTHGTHVAGIAAAKRDGSGMHGVAYDADIIGTKLNDYGNRNGREELIQSAARVINNS
WGIRPDIRRDAKGDIIWLPNGRPDYVAWVKTDVINEVMRNKSNLEWGSEQPVPTGGHSAMATLL
RAAKHGKLIVFSAGNYNNYNIPEAQKSLPYAFPEVLNNYLIVTNLSNNDKLSVSSTSCGHTASF
LACQPGSSIYSSVGELVSNTGGAVNREAYNKGELTVKPDYGNMSGTSMAPDVTGFAAVLMQRFP
YMSAAQISAVIKTTATDLGEVGIDHLFGWGRVNLRDAINGPKMFITQEDIPQEFYVPGSYSEKQ
FVVNIPGLGNIVEAGTPVERRCTSGECDFDSWSNDIRGHGGLTKTGAGTLAVLGNNTYSGDTWV
KQGVLAYNGSVASNVYIENSGTVAGDRTVGAFRAVRGCEHGDAGNGYGTLHVLLDAVFDRGSQY
NVELADKGRSDKLAARRAFLNGGSMNVSLDRSQKLMSQNEAELLVGNNYTILTTLDGVTGRFDN
ANPSYPFVKVALDYRGNDTGLGITKTDATFDSLASTENDKEVARAVETLNATEPVTETAKRSVS
IPASEEANLQQSDAGAAQAVNEEASIVA

Smp ΔAC-Pass1

MHHHHHHSSGLVPRGSGMKQTAAAKFQREHMDSPDLGTDDDDKAMAISDPNSSSYREPGQLGSP
DSWKNAEFNRQWGLEAISAEFAYARAYTGKGVTIGVIDDAILSHPEFAGKLTRLDNGSYNFSYD
KQDNMSFGTHGTHVAGIAAAKRDGSGMHGVAYDADIIGTKLNDYGNRNGREELIQSAARVINNS
WGIRPDIRRDAKGDIIWLPNGRPDYVAWVKTDVINEVMRNKSNLEWGSEQPVPTGGHSAMATLL
RAAKHGKLIVFSAGNYNNYNIPEAQKSLPYAFPEVLNNYLIVTNLSNNDKLSVSSTSCGHTASF
LACQPGSSIYSSVGELVSNTGGAVNREAYNKGELTVKPDYGNMSGTSMAPDVTGFAAVLMQRFP
YMSAAQISAVIKTTATDLGEVGIDHLFGWGRVNLRDAINGPKMFITQEDIPQEFYVPGSYSEKQ
FVVNIPGLGNIVEAGTPVERRCTSGECDFDSWSNDIRGHGGLTKTGAGTLAVLGNNTYSGDTWV
KQGVLAYNGSVASNVYIENSGTVAGDRTVGAFRAVRGCEHGDAGNGYGTLHVLLDAVFDRGSQY
NVELADKGRSDKLAARRAFLNGGSMNVSLDRSQKLMSQNEAELLVGNNYTILTTLDGVTGRFDN
ANPSYPFVKVALDYRGNDTGLGITKTDATFD

Smp ΔAC-Pass2

MHHHHHHSSGLVPRGSGMKQTAAAKFQREHMDSPDLGTDDDDKAMAISDPNSSSYREPGQLGSP
DSWKNAEFNRQWGLEAISAEFAYARAYTGKGVTIGVIDDAILSHPEFAGKLTRLDNGSYNFSYD
KQDNMSFGTHGTHVAGIAAAKRDGSGMHGVAYDADIIGTKLNDYGNRNGREELIQSAARVINNS
WGIRPDIRRDAKGDIIWLPNGRPDYVAWVKTDVINEVMRNKSNLEWGSEQPVPTGGHSAMATLL
RAAKHGKLIVFSAGNYNNYNIPEAQKSLPYAFPEVLNNYLIVTNLSNNDKLSVSSTSCGHTASF
LACQPGSSIYSSVGELVSNTGGAVNREAYNKGELTVKPDYGNMSGTSMAPDVTGFAAVLMQRFP
YMSAAQISAVIKTTATDLGEVGIDHLFGWGRVNLRDAINGPKMFITQEDIPQEFYVPGSYSEKQ
FVVNIPGLGNIVEAGTPVERRCTSG

Smp FL Pass+Linker

MHHHHHHSSGLVPRGSGMKQTAAAKFQREHMDSPDLGTDDDDKAMAISDPNSSSYREPGQLGSP
DSWKNAEFNRQWGLEAISAEFAYARAYTGKGVTIGVIDDAILSHPEFAGKLTRLDNGSYNFSYD
KQDNMSFGTHGTHVAGIAAAKRDGSGMHGVAYDADIIGTKLNDYGNRNGREELIQSAARVINNS
WGIRPDIRRDAKGDIIWLPNGRPDYVAWVKTDVINEVMRNKSNLEWGSEQPVPTGGHSAMATLL
RAAKHGKLIVFSAGNYNNYNIPEAQKSLPYAFPEVLNNYLIVTNLSNNDKLSVSSTSCGHTASF
LACQPGSSIYSSVGELVSNTGGAVNREAYNKGELTVKPDYGNMSGTSMAPDVTGFAAVLMQRFP
YMSAAQISAVIKTTATDLGEVGIDHLFGWGRVNLRDAINGPKMFITQEDIPQEFYVPGSYSEKQ
FVVNIPGLGNIVEAGTPVERRCTSGECDFDSWSNDIRGHGGLTKTGAGTLAVLGNNTYSGDTWV
KQGVLAYNGSVASNVYIENSGTVAGDRTVGAFRAVRGCEHGDAGNGYGTLHVLLDAVFDRGSQY
NVELADKGRSDKLAARRAFLNGGSMNVSLDRSQKLMSQNEAELLVGNNYTILTTLDGVTGRFDN
ANPSYPFVKVALDYRGNDTGLGITKTDATFDSLASTENDKEVARAVETLNATEPVTETAKRSVS
IPASEEANLQQSDAGAAQAVNEEASIVAGHPIYESFLGFTTARELQQATRQLSGQIHADMASAQ
INESRYLRDTATERLRQADGR

Ag43 FL Pass

MHHHHHHSSGLVPRGSGMKQTAAAKFQREHMDSPDLGTDDDDKAMAISDPNSADIVVHPGETVN
GGTLANHDNQIVFGTTNGMTISTGLEYGPDNEANTGGQWVQDGGTANKTTVTSGGLQRVNPGGS
VSDTVISAGGGQSLQGRAVNTTLNGGEQWMHEGAIATGTVINDKGWQVVKPGTVATDTVVNTGA
EGGPDAENGDTGQFVRGDAVRTTINKNGRQIVRAEGTANTTVVYAGGDQTVHGHALDTTLNGGY
QYVHNGGTASDTVVNSDGWQIVKNGGVAGNTTVNQKGRLQVDAGGTATNVTLKQGGALVTSTAA
TVTGINRLGAFSVVEGKADNVVLENGGRLDVLTGHTATNTRVDDGGTLDVRNGGTATTVSMGNG
GVLLADSGAAVSGTRSDGKAFSIGGGQADALMLEKGSSFTLNAGDTATDTTVNGGLFTARGGTL
AGTTTLNNGAILTLSGKTVNNDTLTIREGDALLQGGSLTGNGSVEKSGSGTLTVSNTTLTQKAV
NLNEGTLTLNDSTVTTDVIAQRGTALKLTGSTVLNGAIDPTNVTLASGATWNIPDNATVQSVVD
DLSHAGQIHFTSTRTGKFVPATLKVKNLNGQNGTISLRVRPDMAQNNADRLVIDGGRATGKTIL
NLVNAGNSASGLATSGKGIQVVEAINGATTEEGAFVQGNRLQAGAFNYSLNRDSDESWYLRSEN
AY

Ag43 ΔAC-Pass1

MHHHHHHSSGLVPRGSGMKQTAAAKFQREHMDSPDLGTDDDDKAMAISDPNSADIVVHPGETVN
GGTLANHDNQIVFGTTNGMTISTGLEYGPDNEANTGGQWVQDGGTANKTTVTSGGLQRVNPGGS
VSDTVISAGGGQSLQGRAVNTTLNGGEQWMHEGAIATGTVINDKGWQVVKPGTVATDTVVNTGA
EGGPDAENGDTGQFVRGDAVRTTINKNGRQIVRAEGTANTTVVYAGGDQTVHGHALDTTLNGGY
QYVHNGGTASDTVVNSDGWQIVKNGGVAGNTTVNQKGRLQVDAGGTATNVTLKQGGALVTSTAA
TVTGINRLGAFSVVEGKADNVVLENGGRLDVLTGHTATNTRVDDGGTLDVRNGGTATTVSMGNG
GVLLADSGAAVSGTRSDGKAFSIGGGQADALMLEKGSSFTLNAGDTATDTTVNGGLFTARGGTL
AGTTTLNNGAILTLSGKTVNNDTLTIREGDALLQGGSLTGNGSVEKSGSGTLTVSNTTLTQKAV
NLNEGTLTLNDSTVTTDVIAQRGTALKLTGSTVLNGAIDPTNVTLASGATWNIPDNATVQSVVD
DLSHAGQIHFTSTRTGKFVPATLKVKNLNGQNGTISLRVRPDMAQNNADRLVIDGGRA

Ag43 ΔAC-Pass2

MHHHHHHSSGLVPRGSGMKQTAAAKFQREHMDSPDLGTDDDDKAMAISDPNSADIVVHPGETVN
GGTLANHDNQIVFGTTNGMTISTGLEYGPDNEANTGGQWVQDGGTANKTTVTSGGLQRVNPGGS
VSDTVISAGGGQSLQGRAVNTTLNGGEQWMHEGAIATGTVINDKGWQVVKPGTVATDTVVNTGA
EGGPDAENGDTGQFVRGDAVRTTINKNGRQIVRAEGTANTTVVYAGGDQTVHGHALDTTLNGGY
QYVHNGGTASDTVVNSDGWQIVKNGGVAGNTTVNQKGRLQVDAGGTATNVTLKQGGALVTSTAA
TVTGINRLGAFSVVEGKADNVVLENGGRLDVLTGHTATNTRVDDGGTLDVRNGGTATTVSMGNG
GVLLADSGAAVSGTRSDGKAFSIGGGQADALMLEKGSSFTLNAGDTATDTTVNGGLFTARGGTL
AGTTTLNNGAILTLSGKTVNNDTLTIREGDALLQGGSLTGNGSVEKSGSGTLTVSNTTLTQKAV
NLNEGTLTLN

Ag43 ΔAC-Pass3

MHHHHHHSSGLVPRGSGMKQTAAAKFQREHMDSPDLGTDDDDKAMAISDPNSADIVVHPGETVN
GGTLANHDNQIVFGTTNGMTISTGLEYGPDNEANTGGQWVQDGGTANKTTVTSGGLQRVNPGGS
VSDTVISAGGGQSLQGRAVNTTLNGGEQWMHEGAIATGTVINDKGWQVVKPGTVATDTVVNTGA
EGGPDAENGDTGQFVRGDAVRTTINKNGRQIVRAEGTANTTVVYAGGDQTVHGHALDTTLNGGY
QYVHNGGTASDTVVNSDGWQIVKNGGVAGNTTVNQKGRLQVDAGGTATNVTLKQGGALVTSTAA
TVTGINRLGAFSVVEGKADNVVLENGGRLDVLTGHTATNTRVDDGGTLDVRNGGTATTVSMGN

BrkA-Vag8 Pass

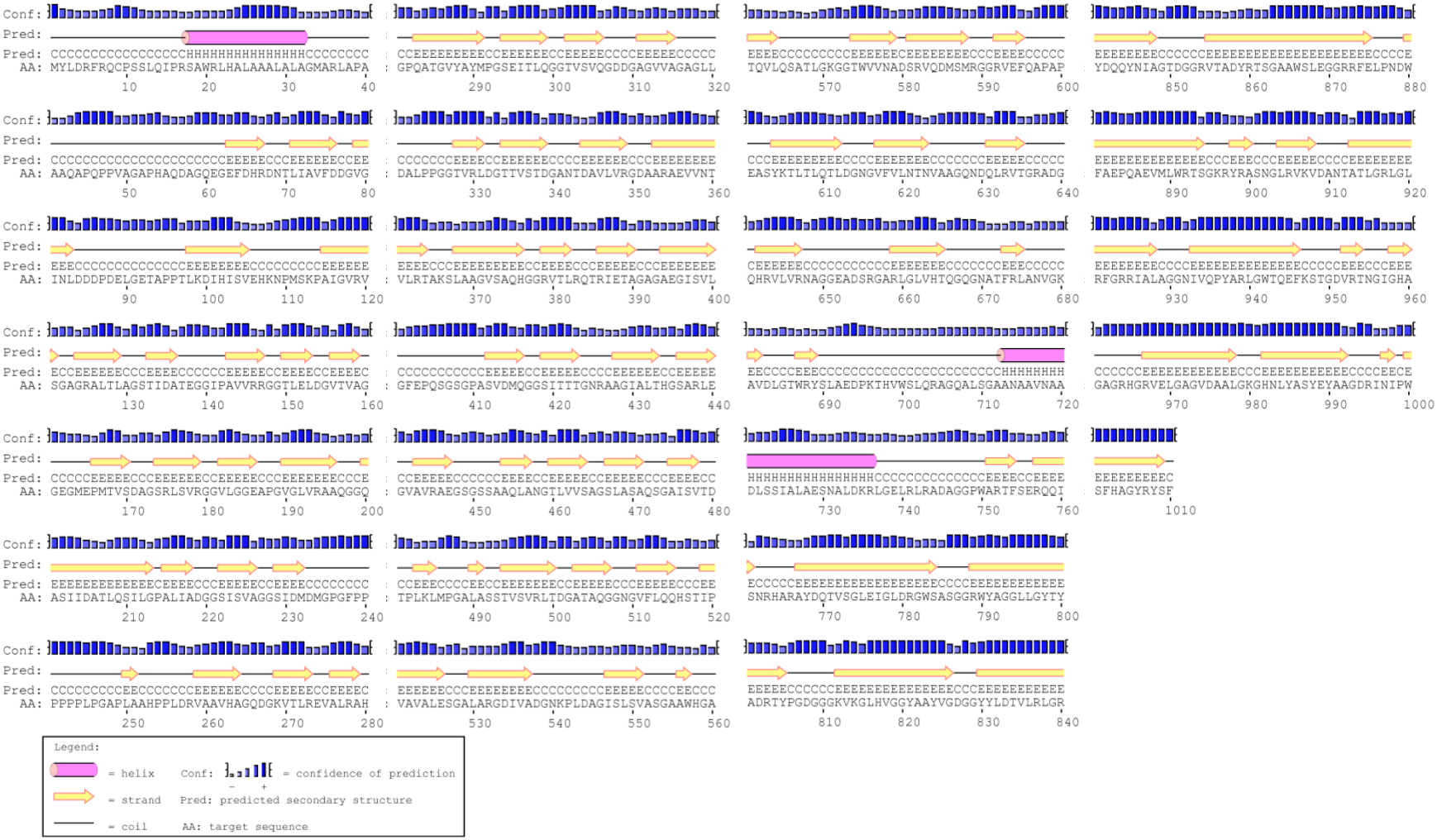
MHHHHHHSSGLVPRGSGMKQTAAAKFQREHMDSPDLGTDDDDKAMAQEGEFDHRDNTLIAVFDD
GVGINLDDDPDELGETAPPTLKDIHISVEHKNPMSKPAIGVRVSGAGRALTLAGSTIDATEGGI
PAVVRRGGTLELDGVTVAGGEGMEPMTVSDAGSRLSVRGGVLGGEAPGVGLVRAAQGGQASIID
ATLQSILGPALIADGGSISVAGGSIDMDMGPGFPPPPPPLPGAPLAAHPPLDRVAAVHAGQDGK
VTLREVALRAHGPQATGVYAYMPGSEITLQGGTVSVQGDDGAGVVAGAGLLDALPPGGTVRLDG
TTVSTDGANTDAVLVRGDAARAEVVNTVLRTAKSLAAGVSAQHGGRVTLRQTRIETAGAGAEGI
SVLGFEPQSGSGPASVDMQGGSITTTGNRAAGIALTHGSARLEGVAVRAEGSGSSAAQLANGTL
VVSAGSLASAQSGAISVTDTPLKLMPGALASSTVSVRLTDGATAQGGNGVFLQQHSTIPVAVAL
ESGALARGDIVADGNKPLDAGISLSVASGAAWHGATQVLQSATLGKGGTWVVNADSRVQDMSMR
GGRVEFQAPAPSRVAEVKLEGGTLAFAPPAQPKGAFKTLVATQGISGTGTIVMNAHLPSGTADV
LVAPQGFGDRQVLVVNNTDDGTESGATKVPLIEDEQGHTAFTLGNMGGRVDAGARQYELTASEA
QADKARTWQLTPTNEL

BrkA-Smp Pass

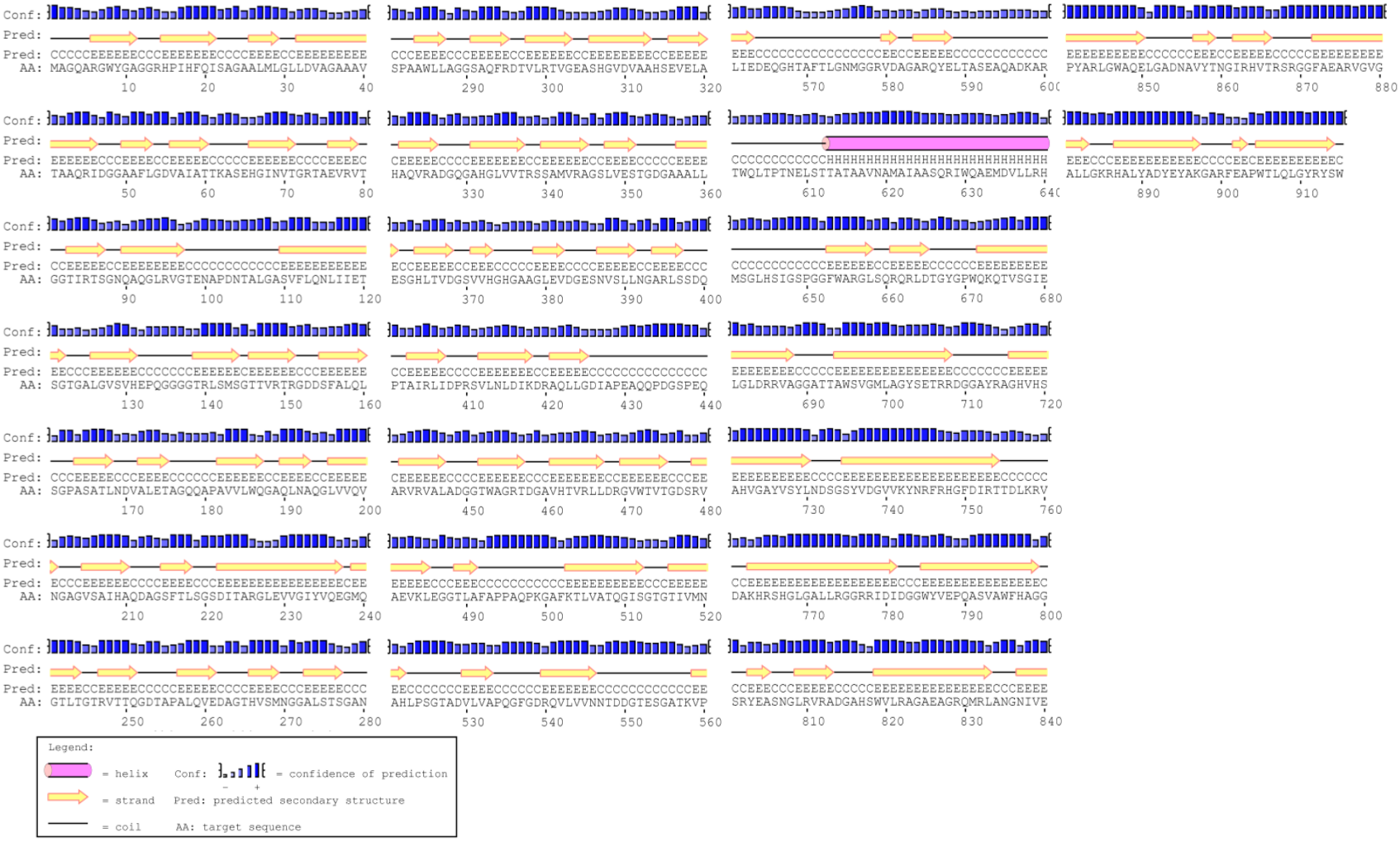
MHHHHHHSSGLVPRGSGMKQTAAAKFQREHMDSPDLGTDDDDKAMAQEGEFDHRDNTLIAVFDD
GVGINLDDDPDELGETAPPTLKDIHISVEHKNPMSKPAIGVRVSGAGRALTLAGSTIDATEGGI
PAVVRRGGTLELDGVTVAGGEGMEPMTVSDAGSRLSVRGGVLGGEAPGVGLVRAAQGGQASIID
ATLQSILGPALIADGGSISVAGGSIDMDMGPGFPPPPPPLPGAPLAAHPPLDRVAAVHAGQDGK
VTLREVALRAHGPQATGVYAYMPGSEITLQGGTVSVQGDDGAGVVAGAGLLDALPPGGTVRLDG
TTVSTDGANTDAVLVRGDAARAEVVNTVLRTAKSLAAGVSAQHGGRVTLRQTRIETAGAGAEGI
SVLGFEPQSGSGPASVDMQGGSITTTGNRAAGIALTHGSARLEGVAVRAEGSGSSAAQLANGTL
VVSAGSLASAQSGAISVTDTPLKLMPGALASSTVSVRLTDGATAQGGNGVFLQQHSTIPVAVAL
ESGALARGDIVADGNKPLDAGISLSVASGAAWHGATQVLQSATLGKGGTWVVNADSRVQDMSMR
GGRVEFQAPAPSRAGNGYGTLHVLLDAVFDRGSQYNVELADKGRSDKLAARRAFLNGGSMNVSL
DRSQKLMSQNEAELLVGNNYTILTTLDGVTGRFDNANPSYPFVKVALDYRGNDTGLGITKTDAT
FDSLASTENDKEVARAVETLNATEPVTETAKRSVSIPASEEANLQQSDAGAAQAVNEEASIVA

Appendix B: Autotransporter secondary structure predictions

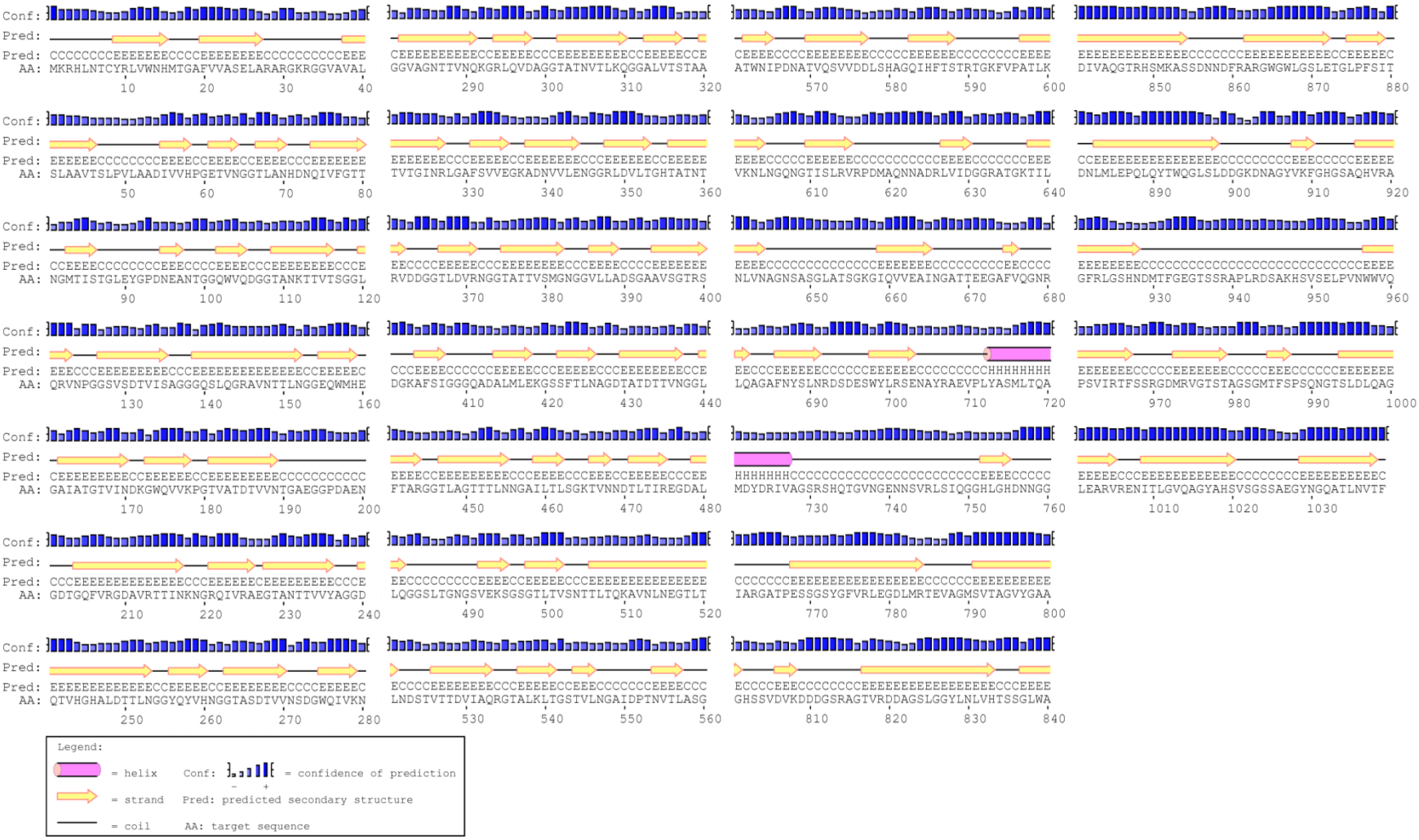
BrkA PSIPRED



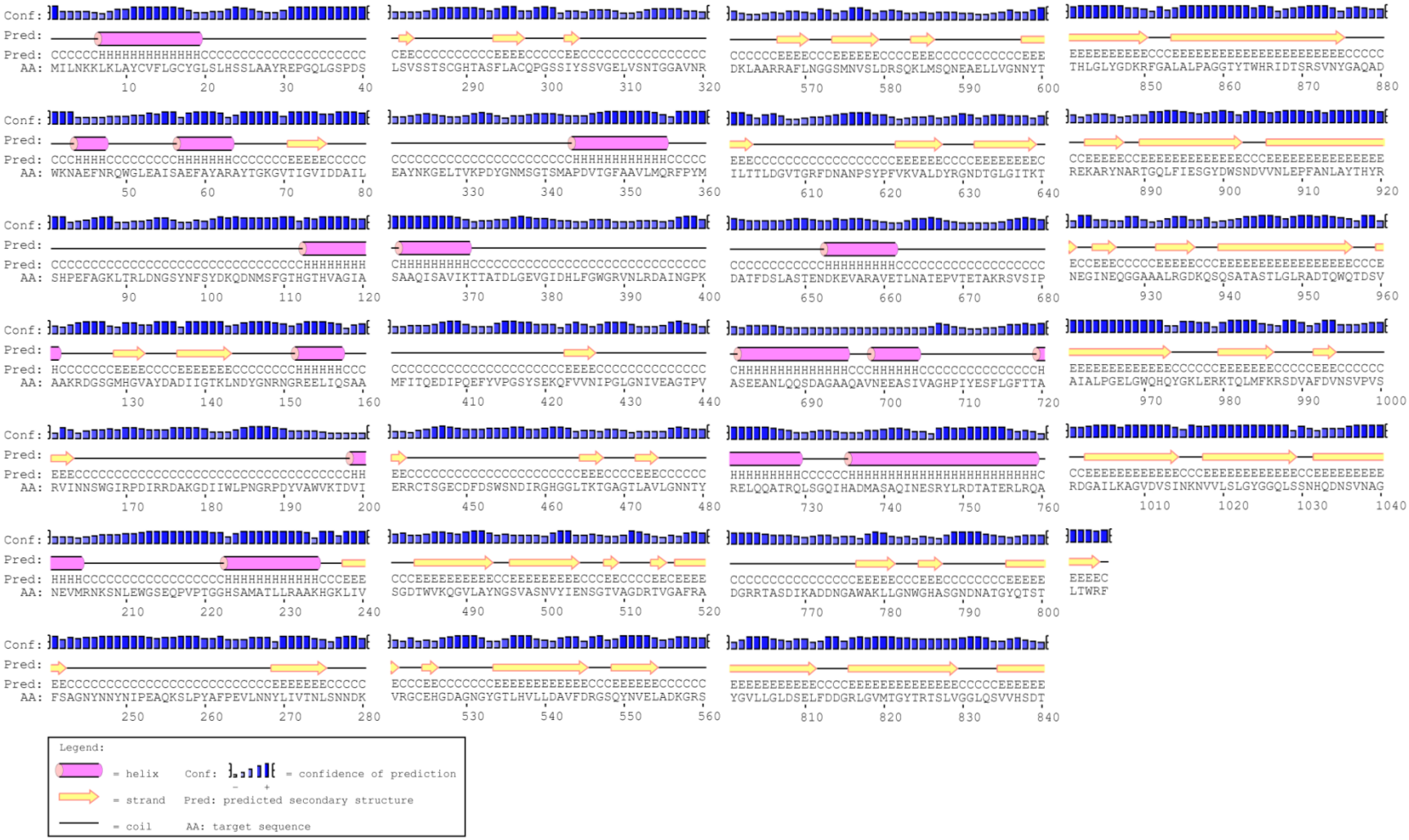
Vag8 PSIPRED



Ag43 PSIPRED



Smp PSIPRED


Conf:
Pred:
AA:
Legend:
= helix
= strand
= coil
Conf: = confidence of prediction
Pred: predicted secondary structure
AA: target sequence