

ARTICLE

Received 18 Dec 2013 | Accepted 28 May 2014 | Published 26 Jun 2014

DOI: 10.1038/ncomms5239

A mortise–tenon joint in the transmembrane domain modulates autotransporter assembly into bacterial outer membranes

Denisse L. Leyton^{1,2}, Matthew D. Johnson², Rajiv Thapa^{2,3}, Gerard H.M. Huysmans^{4,5}, Rhys A. Dunstan¹, Nermin Celik¹, Hsin-Hui Shen^{2,3}, Dorothy Loo², Matthew J. Belousoff¹, Anthony W. Purcell², Ian R. Henderson⁶, Travis Beddoe^{2,7}, Jamie Rossjohn², Lisandra L. Martin³, Richard A. Strugnell⁸ & Trevor Lithgow¹

Bacterial autotransporters comprise a 12-stranded membrane-embedded β -barrel domain, which must be folded in a process that entraps segments of an N-terminal passenger domain. This first stage of autotransporter folding determines whether subsequent translocation can deliver the N-terminal domain to its functional form on the bacterial cell surface. Here, paired glycine-aromatic ‘mortise and tenon’ motifs are shown to join neighbouring β -strands in the C-terminal barrel domain, and mutations within these motifs slow the rate and extent of passenger domain translocation to the surface of bacterial cells. In line with this, biophysical studies of the autotransporter Pet show that the conserved residues significantly quicken completion of the folding reaction and promote stability of the autotransporter barrel domain. Comparative genomics demonstrate conservation of glycine-aromatic residue pairings through evolution as a previously unrecognized feature of all autotransporter proteins.

¹ Department of Microbiology, Monash University, Melbourne, Victoria 3800, Australia. ² Department of Biochemistry and Molecular Biology, Monash University, Melbourne, Victoria 3800, Australia. ³ School of Chemistry, Monash University, Melbourne, Victoria 3800, Australia. ⁴ Institut Pasteur, Molecular Genetics Unit, Departments of Microbiology and of Structural Biology and Chemistry, rue du Dr Roux, 75724 Paris Cedex 15, France. ⁵ CNRS ERL3526, rue du Dr Roux, 75724 Paris Cedex 15, France. ⁶ School of Immunity and Infection, University of Birmingham, Birmingham B15 2TT, UK. ⁷ Department of Agriculture Sciences, La Trobe University, Melbourne, Victoria 3083, Australia. ⁸ Department of Microbiology and Immunology, The University of Melbourne, Melbourne, Victoria 3010, Australia. Correspondence and requests for materials should be addressed to T.L. (email: trevor.lithgow@monash.edu).

Autotransporters are integral membrane proteins that play key roles in bacterial pathogenesis. Most autotransporters have large, β -helical passenger domains that are presented on the extracellular surface of bacterial cells; some of these passenger domains remain attached to the surface to function as adhesins, while others are proteolytically processed and released into the environment as effector proteins^{1–4}. The folding of autotransporters is complex, requiring coordination of three reactions, namely (i) the completion of a 12-stranded carboxy-terminal (C-terminal) β -sheet that will wrap to form a β -barrel, but ensuring that (ii) segments of the amino-terminal (N-terminal) domain are entrapped in the luminal cavity of the C-terminal barrel domain and (iii) the folding of the N-terminal passenger domain into its functional form⁵.

How these folding reactions are coordinated is unclear. A range of methods have been used to address the folding state of autotransporters in their transit through the periplasm and initiation of assembly into the outer membrane^{5–8}. Several lines of evidence indicate that a threading reaction enabling the passenger domain to emerge from the β -barrel can be completed after insertion into the outer membrane^{8–12}. For at least some autotransporters, folding of the passenger domain, from its membrane proximal to distal end, assists the process of translocation through the barrel domain¹² and could be driven in part by the free energy change associated with this sort of vectorial protein folding⁹. But what features of the barrel domain enforce control over these processes to ensure the outcome of the folding and assembly process delivers a functional autotransporter onto the bacterial cell surface?

In a cellular context, the β -barrel assembly machinery (the BAM complex) catalyses steps in the assembly of autotransporters as it does for other types of β -barrel proteins^{12–15}. Periplasmic chaperones such as SurA also participate in the folding and insertion stages for autotransporters, as does the translocation and assembly module that spans the outer and inner membrane^{14,16–18}. At least some of these factors might facilitate or retard aspects of protein folding, to assist an overall coordination of the process. Intrinsic features too must play a role in the regulation of the assembly process. Both functional and sequence-based analysis support this supposition: (i) chimeric autotransporters carrying β -barrel domains from other outer membrane proteins showed that efficient passenger domain translocation depends on the barrel domain of an autotransporter¹⁹ and (ii) genome-wide analysis found sequence profile PF03797 (ref. 20) and discrete motifs²¹ in the sequences of autotransporters that have been conserved through evolution. One conserved motif referred to as ‘the β -motif’²¹, appears to act as targeting sequence recognized by the BAM complex^{22–26}. A second conserved ‘ α -linker motif’ was found to be situated within the lumen of several autotransporter barrel domains for which crystal structures are available²¹.

Here, we report on structural analysis demonstrating that two other highly conserved motifs, originally referred to simply as motif-3 and motif-4 (ref. 21), occur in neighbouring strands of the barrel domain. The sequence motifs reflect a structural arrangement of highly conserved residues facing into the lumen of the barrel domain, which we here refer to as a mortise–tenon joint that aligns and locks the two neighbouring β -strands, by analogy to the joinery used in woodworking. We sought to understand why these residues have been conserved through evolution, and what they signify in the mechanism of autotransporter assembly. Using a combination of circular dichroism (CD) and fluorescence spectroscopy, limited proteolysis and thermal titrations we find that these conserved residues assist in the controlled folding of barrel domains to accommodate the luminal segment of polypeptide that is so

important for correct assembly of autotransporters in bacterial outer membranes.

Results

Two sequence motifs converge as a mortise–tenon structure.

Through structure-based analysis of the sequence motifs conserved in the barrel domains of autotransporters, we identified a highly conserved aromatic residue adopting a notable rotamer conformation: this residue is usually a tyrosine, always faces the β -barrel lumen, and the rotamer has χ -angles^{27,28} of ‘+60,90’ leaving the residue positioned over the neighbouring β -strand. In this +60,90 conformation, a tyrosine side-chain is compact, extending only 3 Å out from the luminal surface of the β -barrel.

Figure 1a shows the conserved tyrosine residue at position 13 in the sequence motif previously referred to as motif-4 (ref. 21). We now refer to this as the tenon motif (Fig. 1a). This tyrosine side-chain locks into a void created by the presence of a highly conserved glycine in position 3 of the neighbouring strand; previously referred to as motif-3 (ref. 21), we now refer to this as the mortise motif. The nomenclature makes analogy with the mortise–tenon joint used in construction to stabilize adjoining structural elements: Fig. 1b depicts how in the structure of an autotransporter barrel domain, the aromatic residue (dark green) slots into a ‘hole’ created between the backbone of the neighbouring β -strand and the α -helical plug in the barrel lumen. This is made possible only because it is a glycine residue (light blue) at the corresponding position in the neighbouring β -strand (Fig. 1b). Crystal structures of the barrel domains of several autotransporters illustrate how the mortise–tenon structure minimizes potential conflicts between the luminal surface of the barrel domain and the ensconced α -linker segment of the autotransporter (Fig. 1c,d).

In Esp^{29,30} [pdb 2QOM] and Hbp³¹ [pdb 3AEH] the mortise motif maps to β -strand number 3 of the barrel domain, while a tenon motif maps to β -strand number 4 (Fig. 1c). In each case, the aromatic residue in the tenon motif locks in a space created by a conserved glycine in the neighbouring strand (Fig. 1d). While the co-conservation of glycine residues in the neighbouring β -strand explains how aromatic residues can be accommodated in the luminal space, the question becomes what functional value might exert the evolutionary pressure to maintain the aromatic residues, given their size and potential conflict with other residues?

We used the autotransporter Pet as an experimental model to address this question. Pet was the first identified autotransporter with enterotoxic activity and is a critical virulence factor of enteroaggregative *Escherichia coli*³². The homologous protein from shiga toxin-producing *Escherichia coli*, called Esp^{29,30}, has been crystallized thereby providing a means to model the topology of Pet (Supplementary Fig. 1a) and three-dimensional structure of Pet (Supplementary Fig. 1b). The conserved glycine residue in the mortise motif on β -strand 3 of the Pet β -barrel domain corresponds to residue G¹⁰⁷⁶ (Supplementary Fig. 1c), and this residue was substituted with alanine to create PetG^{1076A}.

Substitution in the mortise motif perturbs protein stability.

To directly monitor the folding of the barrel domain, plasmids were engineered for the expression of proteins Pet ^{Δ 1–902} and Pet ^{Δ 1–902}G^{1076A}, encompassing the barrel domain and the 116 residues of the passenger domain immediately N-terminal to the barrel domain (Supplementary Fig. 1d). Pet ^{Δ 1–902} and the Pet ^{Δ 1–902}G^{1076A} variant protein were purified in urea and refolded *in vitro* by rapid dilution of the denaturant in detergent micelles³³. Pet belongs to the SPATE (serine protease autotransporters of the *Enterobacteriaceae*) subfamily of autotransporters, which are characterized by an autocatalytic

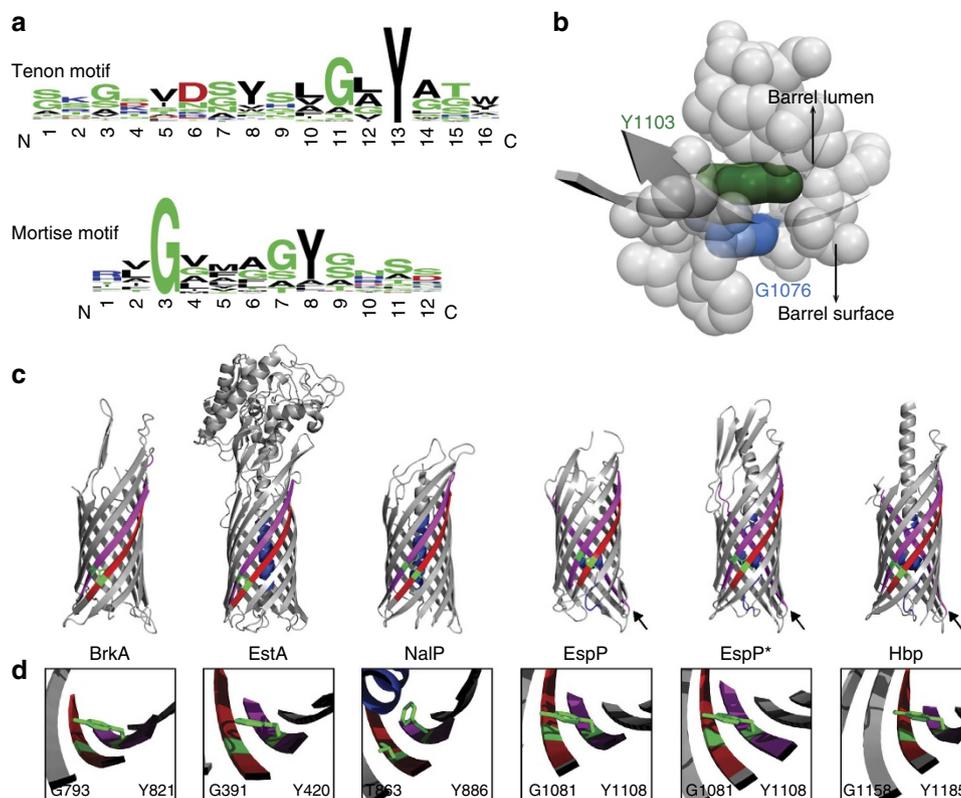


Figure 1 | A mortise-tenon structure is a highly conserved feature in autotransporters. (a) Sequence Logos describing the tenon motif and mortise motif. Amino acids are coloured according to chemical properties; acidic, red; basic, blue; hydrophilic, green and hydrophobic, black. The height of each letter is proportional to sequence conservation. (b) Representation of the mortise-tenon join in the crystal structure of EspP [pdb code 2QOM]²⁹ with Y1103 highlighted (dark green). The depicted view through the axis of the barrel, has translucent grey spheres representing the residues surrounding Y1103 with G1076 shaded light blue, demonstrates that the tyrosine is slotted into the neighbouring residues. Graphic produced in Pymol(The PyMOL Molecular Graphics System, V.r.p., Schrödinger, LLC.). (c) Tenon motif (magenta), mortise motif (red) and, where evident in the crystal structure, the α -linker motif (blue)²¹ were mapped onto the β -barrel domain region of Brk [pdb code 3QO2]³³, EstA [pdb code 3KVN]⁵³, NalP [pdb code 1UYN]⁵⁴, EspP [pdb code 2QOM]²⁹, ‘precleaved’ EspP* [pdb code 3SLJ]³⁰ and Hbp [pdb code 3AEH]³¹. Arrows indicate position of other strands with statistically significant motifs. The conserved glycine residue and aromatic residue are shown in green. (d) In all β -barrel domain structures other than that of NalP, the conserved aromatic residue in the tenon motif is positioned over the conserved glycine residue in the mortise motif. In the NalP structure, glycine is substituted for threonine. The conserved phenylalanine in the neighbouring strand adopts a $-60,90$ rotamer, minimizing steric clash with the α -linker segment.

cleavage of the passenger domain, a reaction that depends on precise positioning of the α -linker segment within the lumen of the correctly folded barrel domain^{30,31,34}. In the cleavage of Pet³², the passenger domain is cleaved between asparagine residues N¹⁰¹⁸ and N¹⁰¹⁹, and autocatalytic processing of the Pet passenger domain can be used diagnostically as a readout of correct β -barrel folding.

The efficiency of the refolding reaction was monitored by the appearance of an ~ 30 kDa cleavage product from autocatalytic processing of the full-length protein (Supplementary Fig. 1e). Cleavage kinetics of Pet ^{$\Delta 1-902$} and Pet ^{$\Delta 1-902$ G¹⁰⁷⁶A} barrel domains were obtained by mixing aliquots of the reaction mixture at increasing incubation times with SDS to stop folding (Fig. 2a, left panel). The unfolded and folded forms were readily separated by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and quantified by densitometry (Fig. 2a, right panel), which showed that both Pet ^{$\Delta 1-902$} and Pet ^{$\Delta 1-902$ G¹⁰⁷⁶A} refolded with similar efficiencies ($46 \pm 3\%$ and $35 \pm 1\%$). Cleavage of Pet ^{$\Delta 1-902$} followed complex kinetics. A burst amplitude indicated that cleavage commenced within the dead time of the experiment ($13 \pm 1\%$ of the total amplitude). The burst phase was followed by exponential kinetics that resolved into two phases: the majority of the cleaved protein ($22 \pm 2\%$) was cleaved with a faster rate constant of $0.13 \pm 0.02 \text{ min}^{-1}$, whilst a smaller population

($11 \pm 1\%$ of the amplitude) was cleaved with a significantly slower rate constant of $0.002 \pm 0.001 \text{ min}^{-1}$. Pet ^{$\Delta 1-902$ G¹⁰⁷⁶A} cleavage was slower compared with Pet ^{$\Delta 1-902$} with kinetics that were well described by a single exponential phase with a rate constant of $0.004 \pm 0.001 \text{ min}^{-1}$. The rate constant of Pet ^{$\Delta 1-902$ G¹⁰⁷⁶A} cleavage agreed well with the slower rate constant of Pet ^{$\Delta 1-902$} cleavage. Two likely interpretations are that folding to the native structure is rate limiting over the slow path, or that the smallest of the Pet ^{$\Delta 1-902$} populations and Pet ^{$\Delta 1-902$ G¹⁰⁷⁶A} are folded in an alternative conformation impaired in cleavage compared with the native structure. The latter would suggest that the G¹⁰⁷⁶A substitution not merely destabilizes the proteins, but increases the frequency of accessing a non-native conformation. To discriminate whether the delayed cleavage of the Pet ^{$\Delta 1-902$ G¹⁰⁷⁶A} protein was direct or simply concomitant with a delay or change in folding of the β -barrel structure, we monitored protein folding by CD and tryptophan fluorescence spectroscopy and by limited proteolysis.

Far-UV CD spectra of Pet ^{$\Delta 1-902$} and Pet ^{$\Delta 1-902$ G¹⁰⁷⁶A} both showed the same characteristic β -signature with minima at 218 nm (Fig. 2b), whereas the proteins were completely unfolded in urea. For both variants, the acquisition of secondary structure occurred immediately after rapid dilution of the denaturant with only minor changes observed between the spectra taken at initial

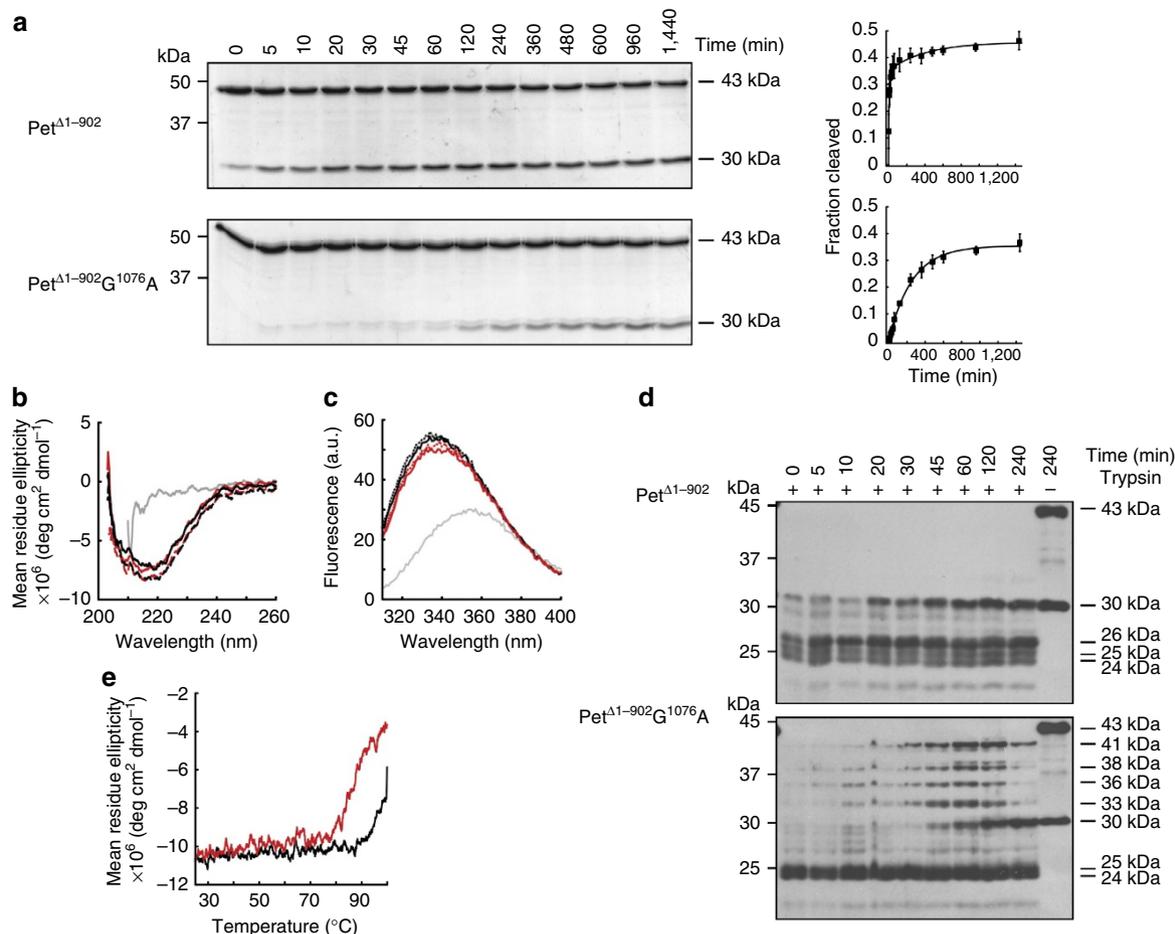


Figure 2 | Substitution of G¹⁰⁷⁶ reduces the cleavage rate, stability and structure of the barrel domain. (a) Folding kinetics of $\text{Pet}^{\Delta 1-902}$ and $\text{Pet}^{\Delta 1-902}\text{G}^{1076}\text{A}$ into detergent micelles was monitored by the appearance of the autocatalytically cleaved form (~ 30 kDa) at the expense of the uncleaved form (~ 43 kDa) of Pet, analysed by SDS-PAGE and Coomassie blue staining. Densitometry analysis was plotted as a function of time and fitted to exponential functions (right panel). The stated values and error bars correspond to the average and s.d., respectively, of three independent experiments. Far-UV CD (b) and tryptophan fluorescence (c) spectra of unfolded protein (grey), and *in vitro* folded and cleaved $\text{Pet}^{\Delta 1-902}$ (black) and $\text{Pet}^{\Delta 1-902}\text{G}^{1076}\text{A}$ mutant (red) at $t=0$ min (full lines) and $t=8$ h (dashed lines). (d) Folding of $\text{Pet}^{\Delta 1-902}$ and $\text{Pet}^{\Delta 1-902}\text{G}^{1076}\text{A}$ monitored by the accumulation of trypsin ($20 \mu\text{g ml}^{-1}$)-resistant fragments; samples were analysed by SDS-PAGE and immunoblot analysis with anti-barrel domain antibodies. (e) Thermal denaturation of *in vitro* folded, cleaved and purified $\text{Pet}^{\Delta 1-902}$ (black) and $\text{Pet}^{\Delta 1-902}\text{G}^{1076}\text{A}$ (red) monitored at 218 nm.

and final time points (Fig. 2b). Tryptophan fluorescence spectra also indicated a rapid folding of the barrel domain (Fig. 2c). The unfolded proteins were characterized by spectra of relatively low intensity and a maximum emission wavelength (λ_{max}) around 355 nm, indicating the solvent exposure of the tryptophan residues in this conformation. However, on dilution into detergent micelles, spectra immediately increased in intensity and showed blue-shifted maxima (Fig. 2c). The tryptophan residues in $\text{Pet}^{\Delta 1-902}$ adopted their native state determinants within the dead time of the experiment giving indistinguishable spectra from initial and final samples (Fig. 2c). Whilst $\text{Pet}^{\Delta 1-902}\text{G}^{1076}\text{A}$ also adopted a rapid shift within the experimental dead time, it reproducibly had a diminished fluorescent intensity relative to that in $\text{Pet}^{\Delta 1-902}$, even after prolonged incubation (Fig. 2c). Importantly, the λ_{max} of the mutant remained red-shifted 5 nm towards higher wavelengths relative to that of the wild-type spectrum (339 nm and 334 nm, respectively), indicating that the tryptophan residues in $\text{Pet}^{\Delta 1-902}\text{G}^{1076}\text{A}$ do not completely adopt the orientation found in $\text{Pet}^{\Delta 1-902}$. Together, these data indicate that both $\text{Pet}^{\Delta 1-902}$ and $\text{Pet}^{\Delta 1-902}\text{G}^{1076}\text{A}$ fold rapidly; while perhaps near-native, the conformation of $\text{Pet}^{\Delta 1-902}\text{G}^{1076}\text{A}$ is at least somewhat distorted.

The distorted conformation of $\text{Pet}^{\Delta 1-902}\text{G}^{1076}\text{A}$ was examined further using limited treatment with trypsin. Titrations of trypsin concentration and digestion times showed that limited digest with $20 \mu\text{g ml}^{-1}$ trypsin for 20 min generated discrete sets of proteolytic products. Immediately after the initiation of folding by rapid dilution out of urea, trypsinolysis of $\text{Pet}^{\Delta 1-902}$ generated two dominant fragments of ~ 30 kDa and 26 kDa and two other fragments of 25 kDa and 24 kDa (Fig. 2d). These fragments provided a kinetic fingerprint by the transient exposure of specific arginine and lysine residues into trypsin-accessible positions during the folding of $\text{Pet}^{\Delta 1-902}$. Digestion of $\text{Pet}^{\Delta 1-902}\text{G}^{1076}\text{A}$ generated a dominant set of 25/24 kDa fragments and a ladder of larger fragments (Fig. 2d). The delay in the appearance of a 30-kDa fragment, present at $t=0$ min for $\text{Pet}^{\Delta 1-902}$ (Fig. 2d), is consistent with the delayed self-cleavage observed for $\text{Pet}^{\Delta 1-902}\text{G}^{1076}\text{A}$ (Fig. 2a). We suggest that the high-molecular weight species between 33 and 41 kDa generated through trypsin digestion represent N-terminal truncated forms of folded $\text{Pet}^{\Delta 1-902}\text{G}^{1076}\text{A}$ passenger domains that are still attached to the barrel domain (Fig. 2d), and that the tryptic fingerprint of the mutant protein is as a result of the folded passenger domain exposing only limited access to trypsin sites.

N-terminal sequencing revealed that the 30, 26, 25 and 24 kDa forms of Pet^{Δ1-902} all start with the cleavage site sequence N¹⁰¹⁹LNKRM (Supplementary Fig. 1a). Therefore, the size differences of these fragments result from trypsin cleavages towards the C-terminal end of the barrel domain; trypsin cleavage at arginine and lysine residues in loop L5 and turn T5 correlate well with the size of the fragments (Supplementary Fig. 1a). Consistent with this observation, the N-terminal sequence of the 30 kDa form of Pet^{Δ1-902}G¹⁰⁷⁶A is also initiated at N¹⁰¹⁹, evidencing an intact and autocatalytically cleaved barrel domain. However, sequencing of the 25/24 kDa forms of Pet^{Δ1-902}G¹⁰⁷⁶A showed that they initiate from I¹⁰⁴⁰ within the first transmembrane β -strand: if the barrel domain was folded in a native conformation, this residue would not be accessible to trypsin (Supplementary Fig. 1a). Cleavage in Pet^{Δ1-902}G¹⁰⁷⁶A in this position is consistent with the barrel being distorted: taken together, these data show that while the refolded barrel domains are processed autocatalytically between N¹⁰¹⁸ and N¹⁰¹⁹ in a manner that mimics the *in vivo* cleavage reaction, the mutant barrel domain tends to adopt a non-native, if near-native, conformation.

Many β -barrel proteins, once folded, are resistant to denaturation in SDS at low temperatures, with the folded form showing a faster electrophoretic mobility on SDS-PAGE than the unfolded counterpart³⁵, and this ‘heat-modifiability’ is a useful diagnostic tool. The refolded Pet^{Δ1-902} barrel displayed heat modifiability in SDS at 25 °C, whereas the refolded Pet^{Δ1-902}G¹⁰⁷⁶A barrel appeared completely denatured under these conditions (Supplementary Fig. 1f). We complemented these observations with thermal unfolding in CD by monitoring the changes in the β -sheet band at 218 nm (Fig. 2e). The data demonstrated that Pet^{Δ1-902}G¹⁰⁷⁶A is less stable than Pet^{Δ1-902}. Thus, while the Pet^{Δ1-902}G¹⁰⁷⁶A β -barrel domain acquires secondary and tertiary structure, it has a distorted fold, consistent with the observation that the first β -strand is trypsin accessible. We hypothesize that consequent misalignment of the α -linker segment delays the cleavage reaction resulting in inopportune initiation of passenger folding.

Translocation of Pet mutants measured *in vivo*. To assess whether the folding defect observed *in vitro* for PetG¹⁰⁷⁶A impacted on autotransporter assembly *in vivo*, cells harbouring a plasmid for the expression of Pet and PetG¹⁰⁷⁶A were induced in the presence of arabinose and, after homogenizing the cells, inner and outer membranes were separated on a discontinuous sucrose gradient. The cleaved Pet and PetG¹⁰⁷⁶A barrel domains, detected using antibodies that selectively recognize the barrel domain¹⁵, migrated in the gradient along with Bama, indicating that Pet and PetG¹⁰⁷⁶A had been integrated into the outer membrane (Fig. 3a).

We established a time course assay for the pulse-chase expression of Pet and PetG¹⁰⁷⁶A in whole cells, monitored by immunoblot analysis after SDS-PAGE using antibodies that selectively recognize the passenger domain³². This provided the means to assess the processing of the secreted form of the ~106 kDa passenger domain and residual ~30 kDa barrel domain, from the uncleaved ~136 kDa protein (Fig. 3b). In addition, at each time-point, an aliquot was treated with proteinase K to digest the population of molecules that were surface exposed: preliminary experiments were done to titrate the amount of proteinase K in use, so as to be sufficient to cleave the periplasmic protein SurA and the lipoprotein BamD that are exposed to the periplasm, but only when polymyxin B was also present (Fig. 3c); polymyxin B is an antibiotic that selectively permeabilizes the outer membrane^{36,37}.

A time course of pulsed-chase expression of Pet and PetG¹⁰⁷⁶A in intact cells demonstrated that Pet processing into distinct passenger domain and barrel domain was underway by 5 min and largely completed by 10 min (Fig. 3d, left panel, –PK). Proteinase K digestion of Pet was also observed after 5 min, indicating that uncleaved Pet was exposed on the cell surface by this time-point (Fig. 3d, left panel, +PK). In contrast, the cleavage of PetG¹⁰⁷⁶A occurred with much reduced efficiency, with only ~50% processing of the 136 kDa species to the 106 kDa mature species after 40 min (Fig. 3d, right panel, –PK). The 136 kDa form of PetG¹⁰⁷⁶A was only degraded by proteinase K in the presence of polymyxin B, indicating that a substantial part of the protein is exposed to the periplasm (Fig. 3e). A 70 kDa passenger domain fragment of PetG¹⁰⁶¹A behaves similarly under these conditions (Fig. 3e). To address what appeared to be a decrease in the steady-state level of the barrel domain of PetG¹⁰⁶¹A relative to Pet (Fig. 3d), we tested the thermal stability of the barrel lacking the key residues in the mortise and tenon motifs. The cleaved barrel domain of Pet was resistant to SDS denaturation, showing greater electrophoretic mobility (that is, the ~25 kDa form of the protein) in samples heated up to 70 °C and with total denaturation of the β -structure above 85 °C (Fig. 3f). In contrast, the PetG¹⁰⁷⁶A mutant migrated as completely denatured in SDS at all temperatures. The aromatic residue in the tenon motif also plays a role in barrel domain folding: a PetY¹¹⁰³A mutation also resulted in a defect in the thermal stability of the barrel domain, albeit less pronounced (Fig. 3g). Either the mortise–tenon structure maintains Pet resistance to SDS, or it is crucial for correct folding of the barrel *in vivo*, as *in vitro*.

Evolution has selected for mortise–tenon sequences. Multiple occurrences of each of these motifs were evident in the evaluation of the autotransporter protein family (Supplementary Fig. 2). This suggests that while motif searches are a useful sequence tool to detect the mortise–tenon structure, the searches underestimate the number of occurrences in each autotransporter sequence. To address whether other glycine residues might be conserved for similar structural reasons, an independent approach was taken using the Position Specific Scoring Matrix generated by the Conserved Domain Database^{38,39}. This tool uses multiple sequence alignments to calculate and depict the relative frequencies of amino acid residues at each position in the barrel domain of autotransporter sequences (Supplementary Fig. 3a). Inspection of the most frequently occurring residues at each position suggested several highly conserved residues including a tryptophan, two tyrosine residues and five glycine residues (Supplementary Fig. 3b).

On the topology model (Fig. 4a) of Pet, some of these residues map to neighbouring β -strands: strands 1 and 2 (W¹⁰³⁷ and G¹⁰⁶¹), strands 3 and 4 (G¹⁰⁷⁶ and Y¹¹⁰³), strands 5 and 6 (Y¹¹¹³ and G¹¹⁵¹) as summarized in Figure 4b, while the other two conserved residues G¹²⁰² and G¹²⁵⁸ on strands 8 and 10. Collectively, they form a ring positioned between the extracellular and periplasmic aromatic annuli of the β -barrel, but visual inspection of the barrel domain structure revealed G¹⁰⁶¹, G¹⁰⁷⁶ and G¹¹⁵¹ participate in mortise–tenon structures (Fig. 4c). To survey the impact of all conserved glycine residues on Pet biogenesis, we introduced an alanine residue at these positions and induced cells harbouring a plasmid for the expression of Pet, PetG¹⁰⁶¹A, PetG¹¹⁵¹A, PetG¹²⁰²A and PetG¹²⁵⁸A in response to the presence of the inducer arabinose.

There were no obvious differences in the insertion, assembly and cleavage of PetG¹¹⁵¹A, PetG¹²⁰²A and PetG¹²⁵⁸A relative to that observed in wild-type Pet (Fig. 5a). However, substitution of

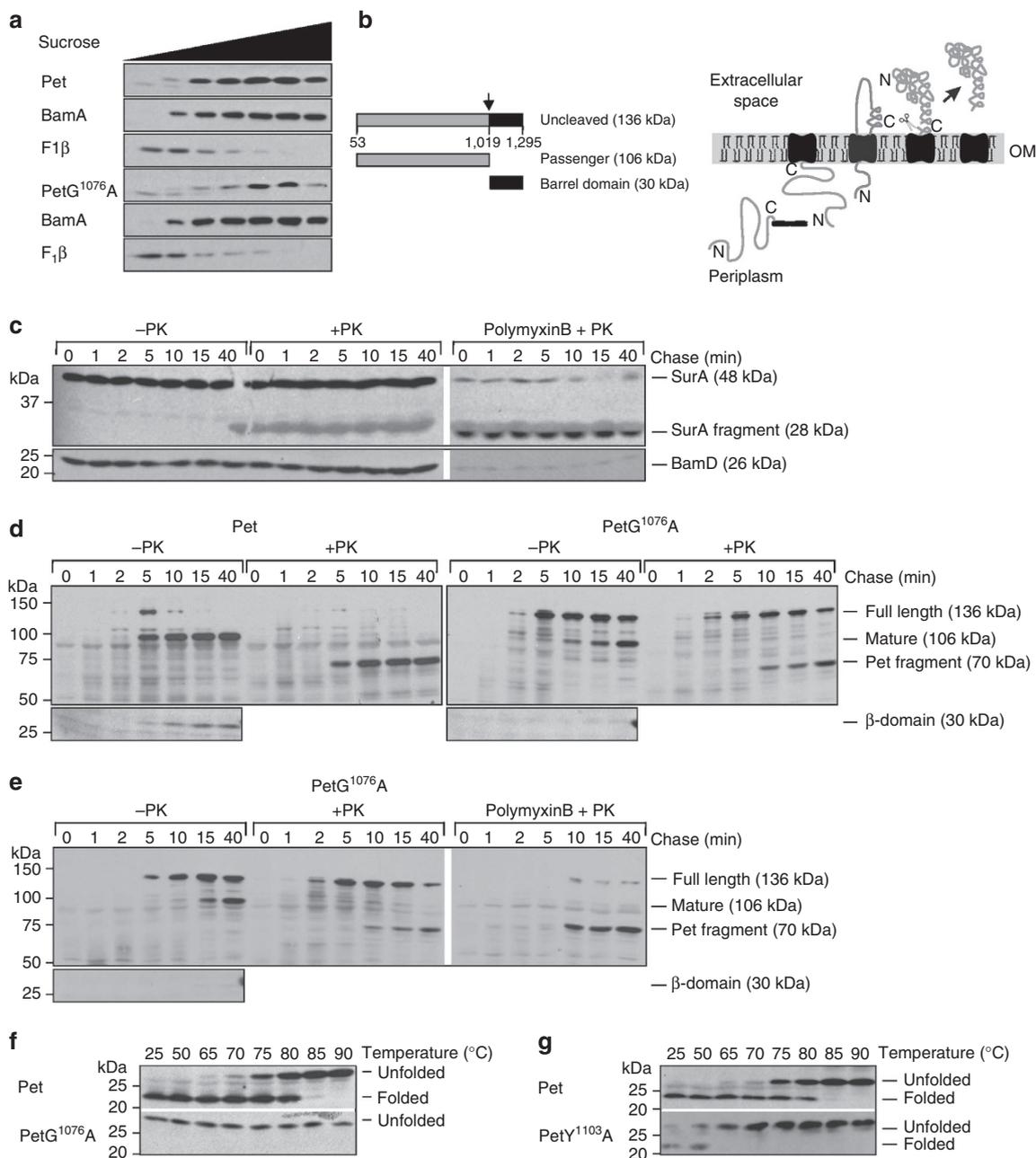


Figure 3 | Mutation of the mortise motif affects passenger domain translocation and autoproteolytic cleavage. (a) Subcellular location of Pet and PetG^{1076A} after pulse (5 min)-chase (40 min) expression by membrane separation on a discontinuous sucrose gradient. Gradient fractions were analysed by SDS-PAGE and immunoblots stained with anti-Pet barrel domain antibodies. Immunoblotting for the outer membrane protein BamA and the inner membrane F₁F₀-ATP synthase F₁β subunit served as controls for membrane separation. (b) The domain structure and schematic for biogenesis shows that after synthesis in the cytoplasm Pet is translocated across the inner membrane and processed by the signal peptidase to a 136 kDa polypeptide in the periplasm. Initiation of passenger domain translocation across the outer membrane continues until the N-terminus of the passenger domain traverses the barrel domain pore. Passenger domain folding at the bacterial cell surface results in the α-linker segment being positioned such that autocatalytic cleavage occurs within the lumen of the barrel domain, thereby liberating the 106 kDa passenger domain. (c) *E. coli* incubated for up to 40 min of a mock 'chase' period at 25 °C and then incubated on ice (– PK) or incubated on ice with the addition of 200 μg ml^{–1} proteinase K (+ PK) or incubated on ice with the addition of 200 μg ml^{–1} proteinase K and 200 μg ml^{–1} polymyxin B. After addition of Phenylmethanesulfonyl fluoride (see Methods), samples were TCA precipitated and analysed by SDS-PAGE and immunoblotting for SurA or BamD. The intact SurA protein (48 kDa) and a protease-resistant domain of 28 kDa were observed. A shadow corresponding to the migration position of proteinase K (28 kDa) is also observed. BamD is degraded to completion by proteinase K in the presence of Polymyxin B. (d) Pulse-chase Pet and PetG^{1076A} maturation and sensitivity to proteinase K, as indicated (+ or – PK), monitored by SDS-PAGE and immunoblotting with anti-passenger domain antibodies (upper panel) and anti-barrel domain antibodies (lower panel). All samples were TCA precipitated before SDS-PAGE. (e) Pulse-chase PetG^{1076A} maturation as above, but also in the presence of Polymyxin B before addition of proteinase K as indicated (Polymyxin B + PK). (f) Warming to the indicated temperatures of membranes after pulse expression of Pet or PetG^{1076A}, and (g) Warming at the temperatures indicated of membranes following pulse-chase Pet and PetY^{1103A} expression. Samples were analysed by SDS-PAGE and immunoblotting with anti-barrel domain antibodies.

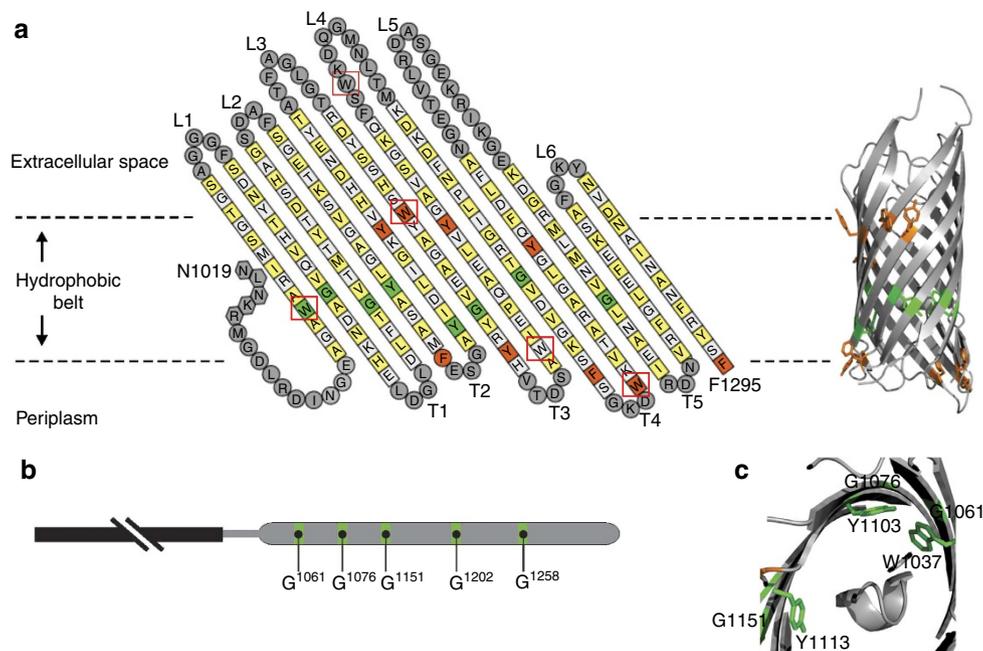


Figure 4 | Distribution of conserved glycine and aromatic residues around the barrel domain of autotransporters. (a) Topology model of Pet, detailing the location of conserved residues W¹⁰³⁷ (β -strand 1), G¹⁰⁶¹ (β -strand 2), G¹⁰⁷⁶ (β -strand 3), Y¹¹⁰³ (β -strand 4), Y¹¹¹³ (β -strand 5), G¹¹⁵¹ (β -strand 6), G¹²⁰² (β -strand 8) and G¹²⁵⁸ (β -strand 10) in green. In the topology diagram, white and yellow squares indicate residues with side chains that point toward the barrel lumen and the lipid bilayer, respectively. Extracellular loops (L1-L6) and periplasmic turns (T1-T5) are labelled. The residues that form the extracellular and periplasmic aromatic girdles are highlighted in orange. These are anticipated to align with the membrane interfaces and therefore define the positioning of the barrel within the plane of the outer membrane^{55,56}. The location of the tryptophan residues used to obtain information on the folding of the barrel domain using the intrinsic fluorescent properties of Pet are boxed: W¹⁰³⁷ (β -strand 1), W¹¹⁴⁴ (β -strand 6), W¹¹⁶¹ (β -strand 7), W¹¹⁸⁰ (L4), W¹²¹⁴ (β -strand 9). Also shown is a structural model of Pet, with conserved glycine and aromatic residues shown in green, and the residues that form the extracellular and periplasmic aromatic girdles highlighted in orange. (b) Schematic view of the positions of the residues subjected to mutagenesis in Fig. 5. (c) A cross-section view through the crystal structure of the EspP barrel domain demonstrates that like G¹⁰⁷⁶ in the mortise motif, G¹⁰⁶¹ and G¹¹⁵¹ also have a conserved aromatic residue that adopts an unusual position to bend tightly over them (W¹⁰³⁷ and Y¹¹¹³, respectively).

an alanine residue was more detrimental in PetG¹⁰⁶¹A, with both a delayed arrival at the cell surface and a slower processing into mature passenger and β -barrel domain (Fig. 5a). Like the PetG¹⁰⁷⁶A mutant, the PetG¹⁰⁶¹A protein appears to be exposed on the outer surface of the outer membrane, since in the absence of polymyxin B, the mature 106 kDa form of PetG¹⁰⁶¹A is proteinase K sensitive (Fig. 5b). In contrast, the 136 kDa form of PetG¹⁰⁶¹A was only degraded by proteinase K in the presence of polymyxin B, indicating that a substantial part of that form of the protein is exposed to the periplasm (Fig. 5b). PetG¹¹⁵¹A, PetG¹²⁰²A and PetG¹²⁵⁸A were only marginally destabilized relative to wild-type Pet (Fig. 5c). However, as was the case for PetG¹⁰⁷⁶A, the PetG¹⁰⁶¹A mutant protein is denatured in SDS at all temperatures (Fig. 5c), demonstrating that interfering with either of these conserved glycine residues impacts significantly on the folding pathway for the autotransporter.

Discussion

Recently, six sequence motifs were identified which, as a group, can be used to detect and classify autotransporters from genome sequence data²¹. The functional significance of two of these motifs has now been revealed: they represent conserved sequence signatures that enable a neighbouring pair of β -strands to anneal with each other such that the large side-chain of an aromatic residue in one strand (corresponding to the tenon motif) intercalates into the space made available by the specific positioning of a glycine residue in the neighbouring strand (in the mortise motif). We have also discovered that this joint

between neighbouring β -strands can also occur in positions which are not readily recognized by motif analysis, that is, only some of these mortise–tenon structures can be predicted with statistical significance. The Position Specific Scoring Matrix screen of autotransporter sequences in concert with visual inspection of the available crystal structures shows this common structural joinery between neighbouring strands is a prevalent feature in autotransporters.

Conceptually, the mortise–tenon structures can explain the molecular basis for mutant phenotypes seen in two previous studies. The autotransporter Hbp (also known as Tsh) has several conserved sequences⁴⁰ and mutations in these diminish its assembly in *E. coli*. In one case, HbpG¹¹⁵⁸Y, a specific defect was characterized by accumulation of an unprocessed form of Hbp in the periplasm⁴¹. G¹¹⁵⁸ sits in a mortise motif in the Hbp sequence, and analysis of the crystal structure of Hbp shows that the neighbouring β -strand carries a conserved aromatic residue (Y¹¹⁸⁵), where HbpG¹¹⁵⁸ and HbpY¹¹⁸⁵ are the equivalent of PetG¹⁰⁷⁶ and PetY¹¹⁰³ (Supplementary Fig. 1c). Conversely, mutational analyses of EspP show delayed passenger domain translocation in the mutant proteins EspPG¹⁰⁶⁶A and EspPG¹⁰⁸¹D, that did not affect interactions with the periplasmic chaperone Skp nor the BAM complex⁴². It was concluded that these mutations affect the insertion of the EspP barrel domain into the plane of the outer membrane⁴². Analysis of the crystal structure of EspP shows that these mutations, EspPG¹⁰⁶⁶A and EspPG¹⁰⁸¹D, would each affect a mortise–tenon structure (Supplementary Fig. 1c). Taken together with the delay of PetG¹⁰⁷⁶A to become surface exposed in intact *E. coli* and its

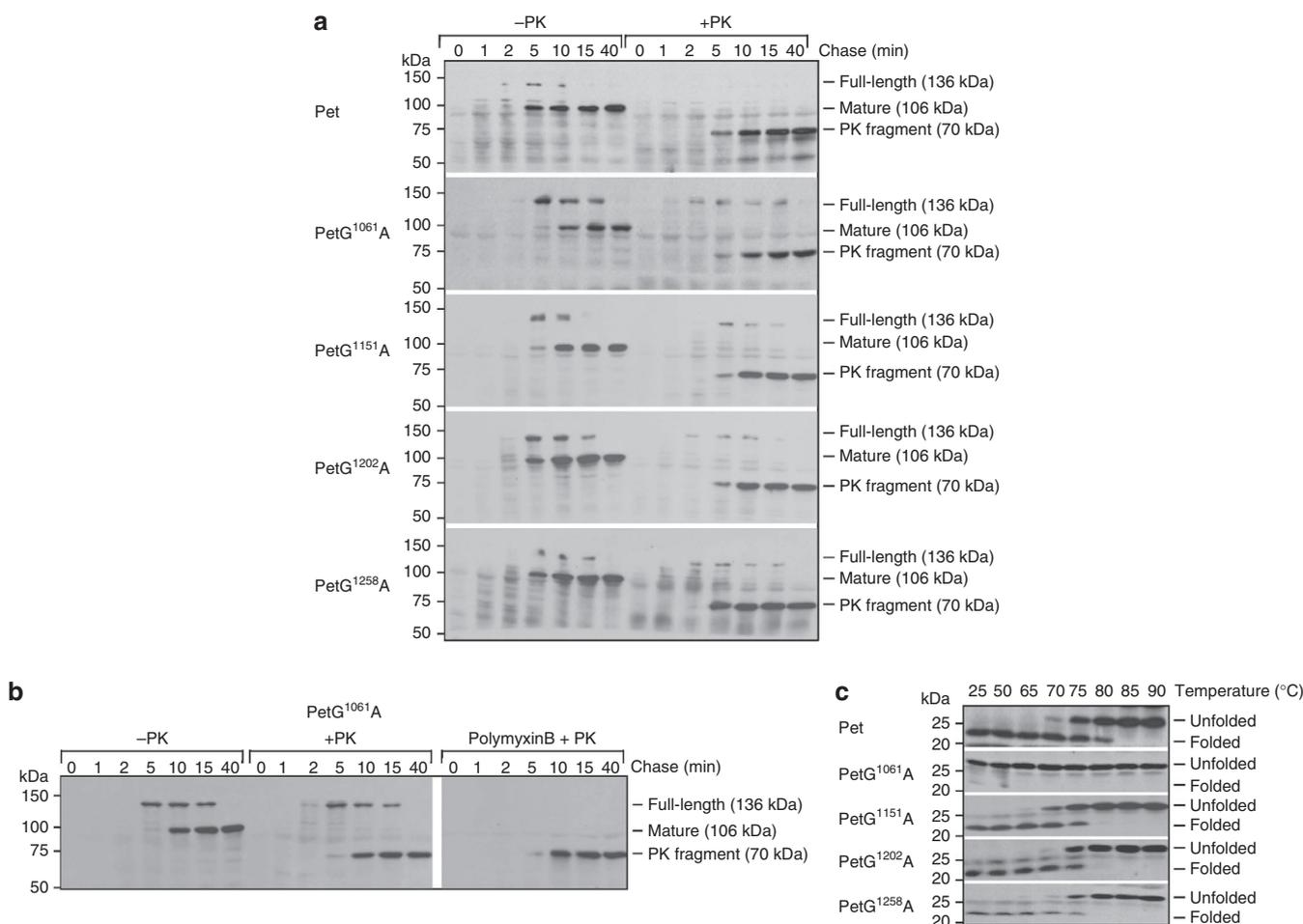


Figure 5 | Like G¹⁰⁷⁶ the conserved G¹⁰⁶¹ impacts on Pet translocation and barrel domain stability. (a) Pulse-chase maturation and proteinase K sensitivity, as indicated (+ or -PK), of Pet, PetG^{1061A}, PetG^{1151A}, PetG^{1202A} or PetG^{1258A}. **(b)** Pulse-chase maturation and proteinase K sensitivity of PetG^{1061A}, as indicated (+ or -PK, or Polymyxin B + PK). Samples were TCA precipitated (in **a** and **b**) before analysis by SDS-PAGE and immunoblotting with anti-passenger domain antibodies. **(c)** Warming to the indicated temperatures of membranes after pulse expression of Pet, PetG^{1061A}, PetG^{1151A}, PetG^{1202A} or PetG^{1258A}. Samples were analysed by SDS-PAGE and immunoblotting with anti-barrel domain antibodies.

slowed rate of autocatalytic processing on folding *in vitro*, we conclude that correctly formed interstrand joins are required in order to efficiently incorporate the passenger domain, thereby promoting efficient translocation through the lumen of the barrel domain.

Our results indicate that mortise-tenon structures promote efficient folding of an autotransporter, providing a rationale for the selection of these structures through the course of evolution: the conserved residues would assist in correctly registering neighbouring β -strand hairpins during the folding reaction. We hypothesize that their removal increases the ruggedness of the folding landscape, promoting misfolding as evidenced by the trypsin susceptibility of the first β -strand in a population of the molecules of the PetG^{1076A} protein, and the decreased thermal stability of the PetG^{1076A} and the PetY^{1103A} proteins. Related to this, a mortise-tenon structure allows, and is completed by, correct packing of the α -linker segment in the lumen of the autotransporter barrel domain, thereby further stabilizing the β -barrel structure^{30,31,43}.

Regulation over the folding reaction is essential to obtain a functional autotransporter. Secretion of autotransporter passenger domains across the outer membrane requires that, ultimately, the polypeptide has extended through an aqueous channel that is

formed by the barrel domain⁵. Thus, unlike other β -barrel proteins, the folding of an autotransporter barrel domain needs to proceed in a controlled fashion that also allows positioning of a segment of the passenger domain within the barrel domain lumen: in the correct structural context, and before the first and last β -strands anneal with each other to close the barrel domain. Cellular factors including the BAM and the translocation and assembly module have been identified which can act *in trans* to facilitate or retard aspects of protein folding, presumably to assist in coordinating the process as it occurs in the outer membrane¹²⁻¹⁵. It has been postulated that in order to facilitate positioning and translocation of the passenger domain through the lumen in the barrel domain, non-native 'barrel-like' conformations might be adopted within the outer membrane: perhaps including an additional β -strand, or a more rounded conformation, or an incomplete annealing of neighbouring β -strands⁴⁴. There are distinct views on whether the appropriate segment of polypeptide first positioned in the barrel would (i) encompass part of the α -linker segment or (ii) be a more distal region of the passenger domain, and only after considerable translocation would the α -linker segment take up its correct position. There is some experimental support for each view^{10,12,45}, but in any of these models, control over this complex

folding process holds the key to the success in initiating autotransporter assembly. Our findings indicate that the key residues facing into the barrel lumen have been under positive selective pressure through the course of evolution to achieve exactly that.

In summary, the mortise–tenon interaction favors the alignment of β -strands and incorporation of the passenger domain to improve translocation kinetics *in vivo*. Experiments with the purified proteins refolded *in vitro* verify that the conserved features in these neighbouring strands have a direct effect on protein stability. These aspects of structure would assist folding of the β -strands into a β -sheet, and coordinate this with the inward folding of the segment of passenger domain that first occupies the barrel lumen, ultimately creating a device that can efficiently promote passenger domain translocation onto the bacterial cell surface. The significance of the mortise–tenon structure is evident in the prevalence with which it has been selected for through evolution as a fundamental feature of autotransporters.

Methods

Reagents and bacterial strains. Unless otherwise stated, bacteria were grown at 37 °C in Luria-Bertani broth and where necessary, the growth medium was

supplemented with 100 $\mu\text{g ml}^{-1}$ ampicillin, 0.2% (v/v) D-glucose, 0.02% (v/v) L-arabinose, or 0.5 mM isopropylthio- β -galactoside. Details of the *E. coli* strains (TOP10 and BL21 (DE3)) and plasmids used in this study are presented in Table 1, and primers used for plasmid construction are presented in Table 2.

Plasmid construction. pBADPet has been described previously⁴⁶. To construct pBADPetG¹⁰⁶¹A, pBADPetG¹⁰⁷⁶A, pBADPetG¹¹⁵¹A and pBADPetY¹¹⁰³A, megaprimer PCR was performed as described previously^{47,48} with some variation. Briefly, all round 1 PCRs were performed on 500 ng of template DNA (pBADPet) with 1 μg of the appropriate mutagenesis primer always in combination with 1 μg of primer AatIIRv per 100 μl reaction mixture. Round 1 PCRs were then purified to remove residual primers from the megaprimer synthesized in this first round of amplification. Round 2 PCRs were performed with 4 μg of megaprimer and 1 μg of primer EagIFw on 500 ng of template DNA (pBADPet) per 100 μl reaction mixture. Round 2 amplicons and target vector (pBADPet) were then digested with EagI and AatII and ligated. Construct pBADPetG¹²⁰²A and pBADPetG¹²⁵⁸A were generated as described above with the exception that round 1 PCRs were performed with the appropriate mutagenesis primer and primers EcoRIIRv and HindIIIRv, respectively. Round 2 PCRs ensued using primers EagIFw and AatIIFw, respectively and the subsequent amplicons and target vector (pBADPet) were then digested with EagI and EcoRI (pBADPetG¹²⁰²A) and, AatII and HindIII (pBADPetG¹²⁵⁸A), and ligated. To construct pETPet ^{Δ 1–902}, the Pet barrel domain and the last 116 residues of the Pet passenger domain were amplified using 500 ng of template DNA (pBADPet) with 1 μg each of primers NdeIPetFw and XhoIPetRv. The subsequent amplicon and target vector (pET-22b+) were then digested with NdeI and XhoI, and ligated for an in frame C-terminal hexahistidine-tag fusion. pETPet ^{Δ 1–902} G¹⁰⁷⁶A was generated as above with the exception that PCRs were amplified using pBADPetG¹⁰⁷⁶A template DNA.

Table 1 | Strains and plasmids used in this study.

Strain/plasmid	Relevant description	Reference
<i>E. coli</i> TOP10	F– <i>mcrA</i> Δ (<i>mrr-hsdRMS-mcrBC</i>) Φ 80 <i>lacZ</i> Δ M15 Δ <i>lacX74</i> <i>recA1</i> <i>araD139</i> Δ (<i>ara leu</i>) 7697 <i>galU</i> <i>galK</i> <i>rpsL</i> (StrR) <i>endA1</i> <i>nupG</i>	Invitrogen
<i>E. coli</i> BL21 (DE3)	F– <i>ompT</i> <i>hsdSB</i> (rB–, mB–) <i>gal</i> <i>dcm</i> (DE3)	Invitrogen
pBADHisA	Arabinose-inducible expression vector, ampicillin resistant	Invitrogen 46,47
pBADPet	pBADHisA derivative expressing <i>de novo</i> synthesized Pet	This study
pBADPetG ¹⁰⁶¹ A	pBADPet derivative where G ¹⁰⁶¹ in the Pet barrel domain was mutated to A	This study
pBADPetG ¹⁰⁷⁶ A	pBADPet derivative where G ¹⁰⁷⁶ in the Pet barrel domain was mutated to A	This study
pBADPetG ¹¹⁵¹ A	pBADPet derivative where G ¹¹⁵¹ in the Pet barrel domain was mutated to A	This study
pBADPetG ¹²⁰² A	pBADPet derivative where G ¹²⁰² in the Pet barrel domain was mutated to A	This study
pBADPetG ¹²⁵⁸ A	pBADPet derivative where G ¹²⁵⁸ in the Pet barrel domain was mutated to A	This study
pBADPetY ¹¹⁰³ A	pBADPet derivative where Y ¹¹⁰³ in the Pet barrel domain was mutated to A	This study
pET-22b+	IPTG-inducible expression vector allowing a C-terminal hexahistidine-tag fusion, ampicillin resistant	Novagen
pETPet ^{Δ1–902}	Truncated pBADPet derivative containing a hexahistidine-tagged Pet barrel domain and the last 116 residues of the Pet passenger domain	This study
pETPet ^{Δ1–902} G ¹⁰⁷⁶ A	Truncated pBADPetG ¹⁰⁷⁶ A derivative containing a hexahistidine-tagged Pet barrel domain and the last 116 residues of the Pet passenger domain	This study

IPTG, isopropylthio- β -galactoside.

Note that the numbers next to the Glycine residues correspond to their position relative to the full-length Pet protein (from M¹ to F¹²⁹⁵).

Table 2 | Primers used in this study.

Primer	Sequence	Reference
EagIFw	5'-GCT CGGCCG GTGGTGGTTTCTCTG-3'	This study
AatIIRv	5'-TTACCG GACGTC AACGCCGGTACGAC-3'	This study
AatIIFw	5'-GCGTT GACGTC GGTAAATCTTTCTC-3'	This study
EcoRIIRv	5'-TC GAAATTC AGACCGAAACGCACGTTG-3'	This study
HindIIIRv	5'-CAGCC AAAGCTT TATCAATGATGATGAT-3'	This study
MPG ¹⁰⁶¹ A	5'-GACAACCTACACCCACGTTTACGGTT CGCG CGGACAACAAACCGAACTGGACG-3'	This study
MPG ¹⁰⁷⁶ A	5'-CTGGACGGTCTGGACCTGTTCACCG CGG TTACCATGACCTACACCGACTCTC-3'	This study
MPG ¹¹⁵¹ A	5'-CTTGGTACGCGGGTGCAGAAAGTT CGG TACCGTTACACGTTACCGACTC-3'	This study
MPG ¹²⁰² A	5'-CTTCAACCCGCTGATCGGTCGTA CCGCG GTTGACGTCGGTAAATCTTTCTC-3'	This study
MPG ¹²⁵⁸ A	5'-GACGGTCTGATGCTGATGACGTT CGCG CTGAACCGGAAATCCGTGACAACG-3'	This study
MPY ¹¹⁰³ A	5'-GAAACCAATCTGTTGGTGCGGGTCT CGCG CTCTGCGATGTTGAATCTGGTGCG-3'	This study
NdeIPetFw	5'-GGGAAT CCATATG TTATGCGTACCAACACG-3'	This study
XhoIPetRv	5'-CCG CTCGAG AGAGCCGAAAGAGTAAACGGAAGTTC-3'	This study

Restriction enzyme sequences are in bold font, site-directed mutations are underlined, and insertion sequences (Gly-Ser linker) are italicized. Note that the numbers next to the Glycine residues correspond to their position relative to the full-length Pet protein (from M¹ to F¹²⁹⁵).

Pulse-chase protein expression assay. The pBAD expression system⁴⁹ was used to produce protein. Bacterial cultures grown to mid-log phase in glucose to repress Pet synthesis were harvested (2,500 × g, 5 min, 4 °C), and then Pet expression was induced with arabinose for 5 min (pulse). Bacterial cells were recovered and resuspended in media containing glucose at 25 °C to ‘chase’ the protein produced during the ‘pulse’ phase of the experiment through the translocation pathway. Following the pulse-chase assay, four 1 ml aliquots were removed from each culture. The first aliquot was added to a tube containing a 10% (w/v) final concentration of trichloroacetic acid (TCA) and placed on ice. The second aliquot was treated with 200 µg ml⁻¹ proteinase K for 20 min on ice to digest Pet exposed on the cell surface and the third aliquot was treated with 200 µg ml⁻¹ polymyxin B for 10 min on ice as previously described^{36,37} before proteinase K treatment. Phenylmethanesulfonyl fluoride (PMSF) at 2 mM final concentration was used to stop the protease reaction before TCA precipitation. All TCA-precipitated samples were washed with acetone, dried and resuspended in 100 µl of SDS-PAGE loading buffer per 1 OD600 units of cells (fourth aliquot). Samples were analysed by SDS-PAGE on 3–14% gradient gels followed by immunoblotting using anti-passenger domain and anti-barrel domain antibodies both used at a 1:5,000 dilution. These antibodies have been described and validated previously^{15,32}. Briefly, to raise anti-barrel domain antibodies, the sequence corresponding to the Pet barrel domain was cloned into the pMAL-c2X expression vector, and the resulting maltose-binding protein-Pet barrel domain fusion was purified in accordance with the manufacturer’s instructions. For anti-passenger domain antibodies, the Pet passenger domain was isolated from culture supernatant of enteroaggregative *E. coli* O42 and purified using DEAE-cellulose and Sephadex columns, before immunization of rabbits and subsequent harvesting of serum.

Subcellular fractionation and heat modifiability assays. After pulse-chase protein expression, cells were harvested and prepared for sucrose density gradient centrifugation as described previously⁵⁰. In heat modifiability experiments, 40 µg of total membranes were mixed with SDS-PAGE loading buffer and then heated at temperatures between 25 °C and 90 °C. All samples were resolved by SDS-PAGE on 3–14% gradient gels and subjected to immunoblotting using anti-barrel domain antibodies as described above.

Protein expression and purification. Protein expression and refolding followed previously described methodology³³. Briefly, *E. coli* BL21 (DE3) cells transformed with pETPet^{Δ1-902} and pETPet^{Δ1-902}G¹⁰⁷⁶A were harvested after induction with 0.5 mM isopropylthio-β-galactoside, from which inclusion bodies were isolated for protein purification in 8 M urea. Pet^{Δ1-902} and Pet^{Δ1-902}G¹⁰⁷⁶A were then refolded from the urea extracts by rapid 10-fold dilution into 0.5% (w/v) lauryldimethylamine-oxide detergent micelles at 37 °C. After affinity purification using Ni-NTA resin, further purification by gel filtration chromatography (Superdex s200-10/300, GE HealthCare) resulted in purification of the folded ~30 kDa form. For N-terminal amino acid sequencing of the trypsin treated Pet^{Δ1-902} and Pet^{Δ1-902}G¹⁰⁷⁶A barrel domain fragments, proteins were first separated by 12% SDS-PAGE and blotted onto a polyvinylidene difluoride sequencing-grade membrane. After blotting, proteins were visualized by staining with Coomassie brilliant blue (R-250), and the appropriate bands were excised and sequenced by automated Edman degradation at the Biomedical Proteomics Facility (Monash University)⁵¹.

Biophysical analysis of protein structure. Folding kinetics of the Pet^{Δ1-902} and Pet^{Δ1-902}G¹⁰⁷⁶A barrel domains were measured after folding was initiated by rapid dilution of the unfolded proteins into detergent micelles at 37 °C as described above³³. Small aliquots of the reaction mixture were removed and quenched with SDS loading buffer (final concentration of SDS: 1% w/v) at the time points indicated, and immediately heated at 100 °C. The unfolded and folded forms were separated by SDS-12% PAGE and band intensities were integrated by densitometry to determine the fraction cleaved for each time point. The Pet^{Δ1-902} data fitted to a double exponential equation ($y = A_{burst} + A_{fast}(1 - \exp(-k_{fast}t)) + A_{slow}(1 - \exp(-k_{slow}t))$) and the Pet^{Δ1-902}G¹⁰⁷⁶A data to a single exponential equation ($y = A_{slow}(1 - \exp(-k_{slow}t))$) where y is the fraction cleaved at a given time (t), A is the amplitude and k is the rate constant. The stated values and error bars correspond to the average and s.d., respectively, of three independent experiments.

Kinetics of trypsin cleavage profiles during Pet^{Δ1-902} and Pet^{Δ1-902}G¹⁰⁷⁶A folding were initiated as described above. Limited treatment with trypsin was performed as described previously⁵². Aliquots of the reaction mixture were removed and treated with trypsin (20 µg ml⁻¹) for 20 min on ice at the time points indicated. Following trypsin treatment, digestions were quenched with 2 mM PMSF, SDS loading buffer was added to each sample and the resulting trypsin fragments were separated by SDS-PAGE and subjected to immunoblotting using anti-barrel domain antibodies as described above.

Far-UV CD spectra were collected from 190 to 260 nm using a Jasco J-815 spectropolarimeter. Folding of Pet^{Δ1-902} and Pet^{Δ1-902}G¹⁰⁷⁶A barrel domains was initiated as described above and monitored at 37 °C for 8 h in a 1-mm cuvette after removing aliquots of the reaction mixture containing 0.1 mg ml⁻¹ of protein. The bandwidth was set to 1.0 nm, integration time 1 s, scanning speed 20 nm/min, three

scans were averaged, and the spectrum was subtracted for buffer contribution. For thermal denaturation measurements, the change in CD signal at 218 nm was measured as a function of temperature from 25–100 °C at a temperature change of 1 °C min⁻¹. In both cases, the measured CD spectra were normalized to the mean residue molar ellipticity. Thermal unfolding was irreversible.

A Varian Cary Eclipse fluorimeter was used for all fluorescence measurements. Samples containing 0.01 mg ml⁻¹ of protein were collected at $t = 0$ min and $t = 8$ h after folding of Pet^{Δ1-902} and Pet^{Δ1-902}G¹⁰⁷⁶A barrel domains was initiated as described above, and measured at 25 °C, in a 1-cm cuvette. The emission spectra were collected between 300–400 nm on excitation of tryptophan at 295 nm, with an integration time of 0.1 s, a slit width of 5 nm and an increment of 1.0 nm. Three scans were averaged and the spectra were corrected for buffer contribution.

References

- Henderson, I. R., Navarro-Garcia, F., Desvaux, M., Fernandez, R. C. & Ala’Aldeen, D. Type V protein secretion pathway: the autotransporter story. *Microbiol. Mol. Biol. Rev.* **68**, 692–744 (2004).
- Dautin, N. Serine protease autotransporters of *Enterobacteriaceae* (SPATEs): biogenesis and function. *Toxins* **2**, 1179–1206 (2010).
- Nishimura, K., Tajima, N., Yoon, Y.-H., Park, S.-Y. & Tame, J. H. Autotransporter passenger proteins: virulence factors with common structural themes. *J. Mol. Med.* **88**, 451–458 (2010).
- Benz, I. & Schmidt, M. A. Structures and functions of autotransporter proteins in microbial pathogens. *Int. J. Med. Microbiol.* **301**, 461–468 (2011).
- Leyton, D. L., Rossiter, A. E. & Henderson, I. R. From self sufficiency to dependence: mechanisms and factors important for autotransporter biogenesis. *Nat. Rev. Microbiol.* **10**, 213–225 (2012).
- van Ulsem, P. in *Bacterial Adhesion* Vol. 715, 125–142 (Springer, 2011).
- Dautin, N. & Bernstein, H. D. Protein secretion in Gram-negative bacteria via the autotransporter pathway. *Annu. Rev. Microbiol.* **61**, 89–112 (2007).
- Renn, J. P. & Clark, P. L. in *Methods Enzymol.* Vol. 492, 233–251 (Academic Press (2011)).
- Braselmann, E. & Clark, P. L. Autotransporters: the cellular environment reshapes a folding mechanism to promote protein transport. *J. Phys. Chem. Lett.* **3**, 1063–1071 (2012).
- Junker, M., Besingi, R. N. & Clark, P. L. Vectorial transport and folding of an autotransporter virulence protein during outer membrane secretion. *Mol. Microbiol.* **71**, 1323–1332 (2009).
- Junker, M. *et al.* Pertactin β-helix folding mechanism suggests common themes for the secretion and folding of autotransporter proteins. *Proc. Natl Acad. Sci. USA* **103**, 4918–4923 (2006).
- Peterson, J. H., Tian, P., Ieva, R., Dautin, N. & Bernstein, H. D. Secretion of a bacterial virulence factor is driven by the folding of a C-terminal segment. *Proc. Natl Acad. Sci. USA* **107**, 17739–17744 (2010).
- Ieva, R. & Bernstein, H. D. Interaction of an autotransporter passenger domain with BamA during its translocation across the bacterial outer membrane. *Proc. Natl Acad. Sci. USA* **106**, 19120–19125 (2009).
- Ieva, R., Tian, P., Peterson, J. H. & Bernstein, H. D. Sequential and spatially restricted interactions of assembly factors with an autotransporter β domain. *Proc. Natl Acad. Sci. USA* **108**, E383–E391 (2011).
- Rossiter, A. E. *et al.* The essential β-barrel assembly machinery complex components BamD and BamA are required for autotransporter biogenesis. *J. Bacteriol.* **193**, 4250–4253 (2011).
- Ruiz-Perez, F. *et al.* Roles of periplasmic chaperone proteins in the biogenesis of serine protease autotransporters of *Enterobacteriaceae*. *J. Bacteriol.* **191**, 6571–6583 (2009).
- Ruiz-Perez, F., Henderson, I. R. & Nataro, J. P. Interaction of FkpA, a peptidyl-prolyl cis/trans isomerase with EspP autotransporter protein. *Gut. Microbes* **1**, 339–344 (2010).
- Selkrig, J. *et al.* Discovery of an archetypal protein transport system in bacterial outer membranes. *Nat. Struct. Mol. Biol.* **19**, 506–510 (2012).
- Saurí, A. *et al.* Autotransporter β-domains have a specific function in protein secretion beyond outer-membrane targeting. *J. Mol. Biol.* **412**, 553–567 (2011).
- Punta, M. *et al.* The Pfam protein families database. *Nucleic Acids Res.* **40**, D290–D301 (2012).
- Celik, N. *et al.* A bioinformatic strategy for the detection, classification and analysis of bacterial autotransporters. *PLoS ONE* **7**, e43245 (2012).
- Bitto, E. & McKay, D. B. The periplasmic molecular chaperone protein SurA binds a peptide motif that is characteristic of integral outer membrane proteins. *J. Biol. Chem.* **278**, 49316–49322 (2003).
- Robert, V. *et al.* Assembly factor Omp85 recognizes its outer membrane protein substrates by a species-specific C-terminal motif. *PLoS Biol.* **4**, e377 (2006).
- Sandoval, C. M., Baker, S. L., Jansen, K., Metzner, S. I. & Sousa, M. C. Crystal structure of BamD: an essential component of the β-barrel assembly machinery of Gram-negative bacteria. *J. Mol. Biol.* **409**, 348–357 (2011).

25. Reusch, R. N. Insights into the structure and assembly of *Escherichia coli* outer membrane protein A. *FEBS J.* **279**, 894–909 (2012).
26. Paramasivam, N., Habeck, M. & Linke, D. Is the C-terminal insertional signal in Gram-negative bacterial outer membrane proteins species-specific or not? *BMC genomics* **13**, 510 (2012).
27. Chamberlain, A. K. & Bowie, J. U. Analysis of side-chain rotamers in transmembrane proteins. *Biophys. J.* **87**, 3460–3469 (2004).
28. Chamberlain, A. K. & Bowie, J. U. Asymmetric amino acid compositions of transmembrane β -strands. *Protein Sci.* **13**, 2270–2274 (2004).
29. Barnard, T. J., Dautin, N., Lukacik, P., Bernstein, H. D. & Buchanan, S. K. Autotransporter structure reveals intra-barrel cleavage followed by conformational changes. *Nat. Struct. Mol. Biol.* **14**, 1214–1220 (2007).
30. Barnard, T. J. *et al.* Molecular basis for the activation of a catalytic asparagine residue in a self-cleaving bacterial autotransporter. *J. Mol. Biol.* **415**, 128–142 (2012).
31. Tajima, N., Kawai, F., Park, S.-Y. & Tame, J. R. H. A novel intein-like autoproteolytic mechanism in autotransporter proteins. *J. Mol. Biol.* **402**, 645–656 (2010).
32. Eslava, C. *et al.* Pet, an autotransporter enterotoxin from enteroaggregative *Escherichia coli*. *Infect. Immun.* **66**, 3155–3163 (1998).
33. Zhai, Y. *et al.* Autotransporter passenger domain secretion requires a hydrophobic cavity at the extracellular entrance of the β -domain pore. *Biochem. J.* **435**, 577–587 (2011).
34. Dautin, N., Barnard, T. J., Anderson, D. E. & Bernstein, H. D. Cleavage of a bacterial autotransporter by an evolutionarily convergent autocatalytic mechanism. *EMBO J.* **26**, 1942–1952 (2007).
35. Nakamura, K. & Mizushima, S. Effects of heating in dodecyl sulfate solution on the conformation and electrophoretic mobility of isolated major outer membrane proteins from *Escherichia coli* K-12. *J. Biochem.* **80**, 1411–1422 (1976).
36. Clements, A. *et al.* The reducible complexity of a mitochondrial molecular machine. *Proc. Natl Acad. Sci. USA* **106**, 15791–15795 (2009).
37. Webb, C. T. *et al.* Dynamic association of BAM complex modules includes surface exposure of the lipoprotein BamC. *J. Mol. Biol.* **422**, 545–555 (2012).
38. Marchler-Bauer, A. *et al.* CDD: conserved domains and protein three-dimensional structure. *Nucleic Acids Res.* **41**, D348–D352 (2013).
39. Marchler-Bauer, A. *et al.* CDD: a database of conserved domain alignments with links to domain three-dimensional structure. *Nucleic Acids Res.* **30**, 281–283 (2002).
40. Yen, Y. T., Kostakioti, M., Henderson, I. R. & Stathopoulos, C. Common themes and variations in serine protease autotransporters. *Trends Microbiol.* **16**, 370–379 (2008).
41. Yen, Y. T. *et al.* Importance of conserved residues of the serine protease autotransporter β -domain in passenger domain processing and β -barrel assembly. *Infect. Immun.* **78**, 3516–3528 (2010).
42. Pavlova, O., Peterson, J. H., Ieva, R. & Bernstein, H. D. Mechanistic link between β barrel assembly and the initiation of autotransporter secretion. *Proc. Natl Acad. Sci. USA* **110**, E938–E947 (2013).
43. Naveed, H. & Liang, J. Weakly stable regions and protein-protein interactions in beta-barrel membrane proteins. *Curr. Pharm. Des.* **20**, 1268–1273 (2013).
44. Tame Jeremy, R. H. Autotransporter protein secretion. *Biomol. Concepts* **2**, 525–536 (2011).
45. Ieva, R., Skillman, K. M. & Bernstein, H. D. Incorporation of a polypeptide segment into the β -domain pore during the assembly of a bacterial autotransporter. *Mol. Microbiol.* **67**, 188–201 (2008).
46. Leyton, D. L. *et al.* The unusual extended signal peptide region is not required for secretion and function of an *Escherichia coli* autotransporter. *FEMS Microbiol. Lett.* **311**, 133–139 (2010).
47. Leyton, D. L. *et al.* Size and conformation limits to secretion of disulfide-bonded loops in autotransporter proteins. *J. Biol. Chem.* **286**, 42283–42291 (2011).
48. Sarkar, G. & Sommer, S. S. The ‘megaprimer’ method of site-directed mutagenesis. *Biotechniques* **8**, 404–407 (1990).
49. Guzman, L. M., Belin, D., Carson, M. J. & Beckwith, J. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J. Bacteriol.* **177**, 4121–4130 (1995).
50. Dunstan, R. A. *et al.* Assembly of the type II secretion system such as found in *Vibrio cholerae* depends on the novel pilotin AspS. *PLoS Pathog.* **9**, e1003117 (2013).
51. Leyton, D. L., Sloan, J., Hill, R. E., Doughty, S. & Hartland, E. L. Transfer region of pO113 from enterohemorrhagic *Escherichia coli*: similarity with R64 and identification of a novel plasmid-encoded autotransporter, EpeA. *Infect. Immun.* **71**, 6307–6319 (2003).
52. Roussel-Jaz  d  , V. *et al.* Channel properties of the translocator domain of the autotransporter Hbp of *Escherichia coli*. *Mol. Membr. Biol.* **28**, 158–170 (2011).
53. van den Berg, B. Crystal structure of a full-length autotransporter. *J. Mol. Biol.* **396**, 627–633 (2010).
54. Oomen, C. J. *et al.* Structure of the translocator domain of a bacterial autotransporter. *EMBO J.* **23**, 1257–1266 (2004).
55. Domene, C., Bond, P. J., Deol, S. S. & Sansom, M. S. P. Lipid/protein interactions and the membrane/water interfacial region. *J. Am. Chem. Soc.* **125**, 14966–14967 (2003).
56. Weiss, M. S. *et al.* Molecular architecture and electrostatic properties of a bacterial porin. *Science* **254**, 1627–1630 (1991).

Acknowledgements

We thank Victoria Hewitt and Chris Stubenrauch for critical comments on the manuscript and Lillian Wong for technical assistance. We gratefully acknowledge the support of the Australian Research Council (ARC) for research funding through the ARC Super Science Fellowship grant FS110200015 (to T.L., J.R., R.A.S. and L.L.M.) and the National Health and Medical Research Council through a NHMRC Program Grant 606788 (to T.L. and R.A.S.). G.H.M.H. is a recipient of a Marie Curie Fellowship and an EMBO-Pasteur Fellowship (PIEF-GA-2010-272611; ALTF 1088-2010). M.J.B. is an NHMRC Postdoctoral Fellow, A.W.P. is an NHMRC Senior Research Fellow and J.R. an NHMRC Australia Fellow. D.L.L. and H.S. are ARC Super Science Fellows, T.L. is an ARC Australian Laureate Fellow.

Author contributions

D.L.L., G.H.M.H., M.D.J., R.A.D., N.C., R.T., H.-H.S., D.L. and T.L. designed and carried out analysis. D.L.L., G.H.M.H., M.J.B., I.R.H., T.B., J.R., L.L.M., T.L. and A.W.P. provided expertise to analyses. D.L.L., G.H.M.H. and T.L. supervised the experimental work and evaluated data. D.L.L., R.A.S. and T.L. evaluated results and wrote manuscript. All authors have reviewed the manuscript.

Additional information

Supplementary Information accompanies this paper at <http://www.nature.com/naturecommunications>

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

How to cite this article: Leyton, D. L. *et al.* A mortise–tenon joint in the transmembrane domain modulates autotransporter assembly into bacterial outer membranes. *Nat. Commun.* **5**:4239 doi: 10.1038/ncomms5239 (2014).