Supplemental Results and Discussion

TpsA amino acid architecture. TpsA proteins are characterized by a common amino acid sequence architecture with minimal homology between family members[1]. Starting at the N-terminus, TpsA proteins contain a Sec signal peptide, a TPS domain, and one or more functional C-terminal domains (Figure S1A). Prior to secretion, the approximately 30 residue Sec signal peptide is proteolytically removed from the N-terminus[2], while the C-terminal functional domains provide the ability to classify TpsA members into functional families. The work presented within this study focuses on the TPS domain, an approximately 300 residue N-proximal segment juxtapositioned between the Sec signal and C-terminal functional domain (Figure S1).

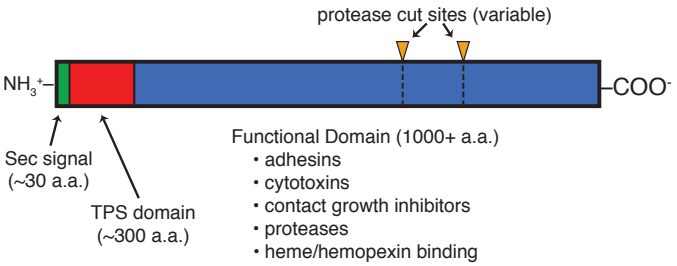

Figure S1) Sequence features of TpsA family members. TpsA sequence architecture. Shown are the Sec signal peptide (green), TPS domain (red), and functional domain(s) (blue). Approximate sizes in numbers of amino acids (a.a.) are shown for each region. Some TpsA members are proteolytically processed (orange triangles).

Past studies have shown that the TPS domain is essential for TpsB dependent translocation of related TpsA partners across the outer membrane[3]. Functionally, specific sequences on the unfolded TPS domain interact with POlypeptide-TRansport Associated (POTRA) domains on the TpsB protein to allow recognition and transport of TpsA proteins[4]. Although the POTRA domains are well characterized, the specific sequences within the TPS domain that interact with the POTRA domains are unknown and the events after POTRA domain binding remain poorly characterized[1,5,6].

Dependence of HpmA265 SD3 thermodynamic parameters on fitting model. The high stability of the polar core made it difficult to unambiguously determine the fitting parameters for the final transition in HpmA265. Two different models were used to try and determine the reliability of the parameters determined for this third transition. The first is as presented, the second is using a sloping baseline between the second and third transition (incorporated as a GdnHCl concentration dependent fraction change in CD for transition 2). As additional tests, we tested the effect on the fitting parameters of including the instrument provided errors in each data point or truncating the data at 7.5 M GdnHCl. These tests showed 1) that including a sloping baseline resulted in a higher $\Delta G°_{H_2O}$ of -69 ± 9 kJ/mol, without appreciable changes to the quality of the fit. These same changes resulted in a slight increase in $\Delta G°_{H_2O}$ for the second transition (-33 ± 2 kJ/mol) and essentially no change in the $\Delta G°_{H_2O}$ for the first transition (-11.5 ± 0.6 kJ/mol). Comparison of the effects of truncating or weighting the data showed that these procedures resulted in slight lower values of $\Delta G°_{H_2O}$ of -48 ± 3 kJ/mol and -49 ± 2 kJ/mol. Neither truncating nor weighting the data significantly affected the parameters for any of the other transitions. In all cases the C_m for the final transition remained constant to within 0.1 M, meaning that the changes in the $\Delta G°_{H_2O}$ were correlated with changes in the *m* value for the transitions. Regardless of the model used, the polar core subdomain (SD3) is the most stable subdomain. Although the trends are valid across the polar core mutations, care are should be taken in calculating $\Delta\Delta G°_{H_2O}$ values due to mutations in the polar core.

Correlation of *m* value on subdomain size. While the observed *m* values are within the range of values typically observed for a protein of this size[7,8], the loss of structure as measured by CD signal change and the *m* value for the transitions are uncorrelated for either the denaturation of SD2 ($I_1 \rightleftharpoons I_2$) or the denaturation of SD3 ($I_2 \rightleftharpoons D$). The *m* value for typical globular proteins is correlated with the change in solvent accessible surface area, and thus, the size of the structure that is denatured[9,10]. While the denaturation of SD1 ($N \rightleftharpoons I_1$) is accompanied by both the smallest change in CD signal (least amount of structure lost) and the smallest *m* value (smallest change in solvent accessible surface area), the *m* values determined for the other two subdomains do not follow the expected linear trend with the increase in fractional CD change compared to SD1 – either the *m* value associated with denaturation of SD2 is high, or that associated with SD3 is low given the values for the other two transitions. The polar core associated with SD3 may not provide the same *m* value as a typical protein of its size, which may help explain the noted discrepancy.

Denaturation at 7.8M GdnHCl. GdnHCl titration to 7.8 M results in a CD spectrum that is typical for a protein that lacks defined secondary structure (Figure S2). Likewise, this same CD spectrum is obtained for all of the mutants studied, suggesting that complete denaturation is obtained for all constructs.

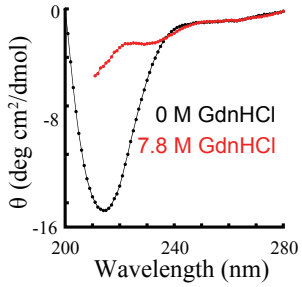

Figure S2) CD signal of native (0 M) and denatured (7.8M) HpmA265

Reversibility of individual transitions. In an effort to define equilibrium denaturation, reverse titrations were conducted on HpmA265. Identical thermodynamic parameters, within error, were obtained from unfolding and refolding titrations with up to 6.0 M GdnHCl (Figure S3A). Characterization of the kinetics of refolding from 6.0 M GdnHCl using manual dilution to any lower concentration of GdnHCl showed complete refolding in the approximately 10 second dead time of the experiment. Thus, the unfolding and folding pathways are fully reversible and at equilibrium for the region of the denaturation curve that encompasses T_1 and T_2.

To confirm that refolding resulted in a native-like structure we tested the ability of the refolded protein to function in a template assisted hemolytic activity (TAHA) assay. Dialysis of 6.0 M GdnHCl denatured HpmA265 with 10 mM Na_2HPO_4 pH 7.4 resulted in functionally active protein with similar kinetic parameters as non-denatured protein (Figure S3B). However, upon denaturation at 7.8 M GdnHCl and dilution or dialysis back to native conditions, no evidence of refolding is present; analysis of completely unfolded

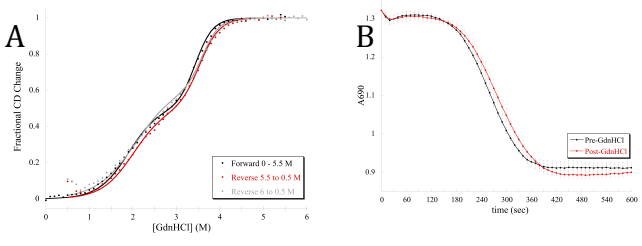

Figure S3. Reversibility of GdnHCl dependent unfolding. (A) Data and fit lines for the forward (0 to 5.5 M GdnHCl) and reverse titrations (6.0 to 0.5 M and 5.5 to 0.5 M GdnHCl) of HpmA265. (B) HpmA265 template-assisted hemolytic assays were conducted before (Pre-GdnHCl) and after 6.0 M GdnHCl addition and removal (Post-GdnHCl) via overnight dialysis. The average V_{max} for pre-GdnHCl and post-GdnHCl HpmA265 samples were 280 ± 30 and 310 ± 30 with lag times of 120 sec.

protein that has been returned to native conditions shows no evidence of β-structutre in CD nor is it functionally active in TAHA assays (data not shown). Furthermore, dialysis of HpmA265 equilibrated in 7.8 M GdnHCl, resulted in a visible precipitate.

Kinetics of Aggregation. Based on the observation that refolding could occur after incubation at 6 M GdnHCl but not after incubation at 7.8 M GdnHCl, a series of control experiments were conducted in order to confirm the presence of any off-pathway aggregates and determine the kinetics of their accrual. During these control experiments, aggregates were observed while investigating the irreversibility of HpmA265 incubated at 7.8 M GdnHCl. The kinetics of aggregation was determined to be GdnHCl concentration dependent, which allowed partial characterization of the aggregate forms. Incubation of HpmA265 at 6.0 M GdnHCl and subsequent dilution back to 0.5 M GdnHCl at timed intervals after the start of the incubation allowed partial aggregate characterization using DLS for particle size (Figure S4A), CD signal for β-strand content(Figure S4B), and GdnHCl titrations to judge amount of refolding (Figure S4C). DLS experiments showed that HpmA265 size shifts from 70 to 210 Å upon adjustment from 0 to 6.0 M GdnHCl, and this particle size is stable for 24 hrs without the appearance of larger particles (data not shown). However, after incubation at 6.0 M for 6 min to 48 hr and subsequent dilution 0.5 M GdnHCl, HpmA265 establishes DLS visible aggregate species within the dead time of the instrument. These aggregates had sizes ranging from 330 Å to 1900 Å compared to a crystallographic size of approximately 70 Å for the native state HpmA265 (Figure S4A). Additionally, these species are further characterized by maintenance of β-helix signal (Figure S4B). After incubation for 120 hrs at 6.0 M GdnHCl and subsequent dilution back to 0.5 M GdnHCl, the size of the aggregate that forms shifts to 680 Å and no longer maintains β-helix CD signal (Figure S4A, S4B). Thus, HpmA265 progresses through different states at 6.0M GdnHCl that results in distinct DLS and CD species being formed upon dilution to 0.5 M GdnHCl.

The first set of aggregates maintain β-strand-like signal albeit with a slightly different spectra than the CD signal observed with HpmA265 (Figure S2B). To further link the aggregation to the unfolding of the N-proximal polar core, we conducted GdnHCl titrations for the various timed single jump experimental samples (Figure S2C). As the length of time incubating in 6.0 M GdnHCl prior to dilution to 0.5 M GdnHCl increases, there is an observable loss of T_3, the transition linked to the N-proximal polar core subdomain.

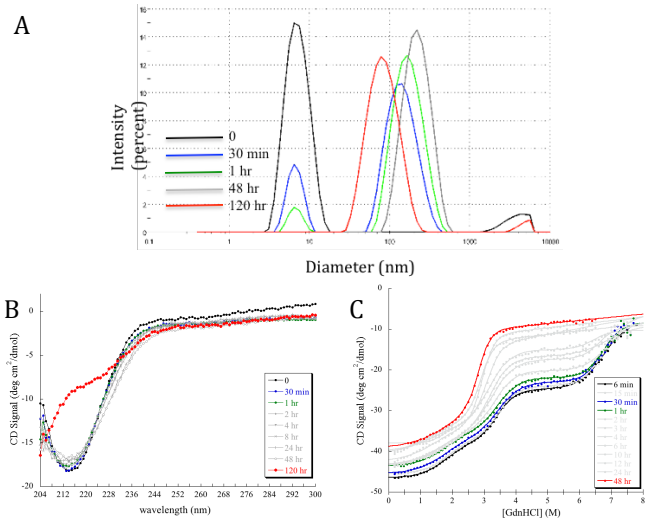

Figure S4) Aggregate characterization. (A) DLS results of samples diluted to 0.5M GdnHCl after specified time incubating in 6 M GdnHCl (B) CD wavelength scans post dilution to 0.5 M GdnHCl. (C) Re-titration of refolded samples.

Collectively, the data were interpreted as suggesting that an off-pathway aggregate develops during prolonged (more than 1 hour) exposure to ≥ 6.0 M GdnHCl. These timed incubations were repeated in the presence of 6.8 M GdnHCl (~C_m), which provided very similar results, though with faster kinetics (data not shown). Manual titration of HpmA265 at 0 M through 7.8 M GdnHCl followed by incubation for more than a month showed CD signals that recapitulated the automatic titrations at low (< 4 M) GdnHCl concentrations. Analogous refolding experiments starting with protein equilibrated at 7.8 M GdnHCl shows no evidence of β-helical structure at the same low GdnHCl concentrations (data not shown). This is consistent with the idea that during the refolding of HpmA265 from 7.8 M GdnHCl, an off pathway aggregate kinetically traps the protein and prevents refolding. We emphasize that the formation of aggregates is only significant at high GdnHCl concentrations, is kinetically slow to form, and is thus only apparent as the protein is refolded.

Stir Time Determination. To determine optimal parameters that would allow denaturation of the protein, but avoid significant contributions from off-pathway aggregate formation, stir times after GdnHCl addition were systematically tested. HpmA265 denaturation data collected with 1, 2, and 5 min stir times after GdnHCl addition showed essentially no effect on the fitted parameters (less than 1% difference which is less than the standard deviations of replicate measurements). This suggests that the samples had equilibrated during the 1 min stir time after GdnHCl addition. Significantly shorter stir times (36 sec) at each 0.1 M GdnHCl step failed to show systematic changes in T_1 and T_2, but did affect T_3. Similarly, longer stir times (10, 20 min) also only affected T_3 (data not shown). Given the apparent aggregation kinetics at high GdnHCl, we expect that 2 minutes of stirring after each addition of GdnHCl resulted in minimal aggregate formation during the approximately 20 minutes necessary to traverse T_3 with 0.1 M steps in GdnHCl concentration. Supporting this contention, larger steps between GdnHCl concentrations did not significantly change the thermodynamic parameters obtained from T_3 (data not shown). Thus, we have used 1 min stir times for GdnHCl concentrations encompassing the denaturation SD1 and SD2, and 2 min stir times for GdnHCl concentrations encompassing denaturation of SD3. We have analyzed the data assuming equilibrium and reversible denaturation for all of the transitions.

Sequence conservation supports different subdomain functions. Alignment of the HpmA265 sequence with other known and putative hemolysin members reports 23% identity, with a high degree of conservation of the internal polar core residues compared to the nonpolar core and C-terminal subdomains (33% identical residues versus 11% and 11% respectively). The identical residues include all of the internal hydrogen bonding partners within the polar core subdomain (Figure 2). The already low conservation in the nonpolar core (SD2) and C-terminal (SD1) subdomains degrades even further as more C-proximal residues are added from the full length protein. The fraction of identical residues is reduced to 4% when the entire ~1500 amino acid hemolysin proteins are aligned (data not shown). This analysis supports results by others showing high conservation of the internal polar core residues and the disulfide bond in the TPS domain of hemolysin proteins[2,11,12]. This higher degree of conservation in the TPS polar core subdomain, SD3, suggests that these residues are important for either interactions with the TpsB component and/or folding of the TPS domain. The discontinuity in sequence conservation between the polar versus the nonpolar cores among hemolysin TpsA members further suggests that these sequence regions are also under different selection, and thus may be different subdomains of the TPS domain.

Coincident transitions in independent and sequential unfolding models. The effect of destabilizing a single subdomain in a multisubdomain protein differs depending on if the unfolding of the

different subdomains are linked. Consider a two subdomain protein. For an independent unfolding model, the stability of one subdomain does not depend on the second subdomain. In this case, weakly destabilizing mutations would result in merger of the two transitions as seen by CD signal. Because each subdomain would still unfold as a separate cooperative transition, this merged transition would report a lower *m* value appropriate for the individual subdomains but less than expected for the cooperative unfolding for the entire structure. More destabilizing mutations would result in separate transitions, although the ordering of the subdomain denaturation would have flipped (Figure S5A).

A sequential model of unfolding leads to a different result. In this model unfolding of the second subdomain traps the first subdomain in the unfolded state. This means that regardless of how destabilizing of a mutation is made in the second subdomain, only a single transition will be observed. This single transition will have an increased *m* value compared to the WT protein. This increased *m* value reflects the increased amount of surface area exposed during the cooperative transition. Since the protein becomes trapped in the unfolded state once the second domain unfolds, as the degree of destabilization increases, the midpoint of the transition will decrease (Figure S5B).

The mutations L194N, and C144S-C147S demonstrate the sequential nature of folding the nonpolar core and C-terminal subdomains. For example, the increase in *m* value (Table S1) and change in CD (not shown) associated with T_1 can be interpreted as the L194N variant destabilizing SD2, the nonpolar subdomain, such that it now denatures cooperatively with the C-terminal subdomain.

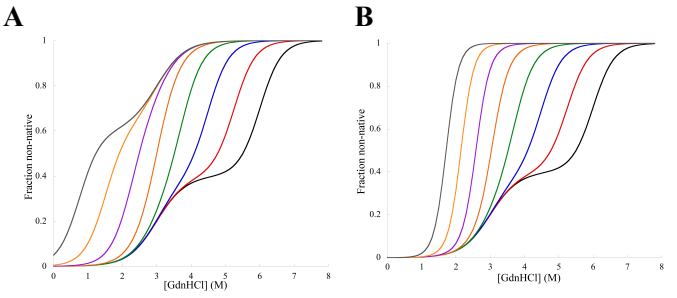

Figure S5. Comparison of the effect of destabilizing a subdomain on the titration profile for a three-state system for an (A) Independent or (B) Sequential model. The two subdomains are SDa and SDb, with associated free-energy changes (ΔG°), m-values, and change in CD signal. The wild type condition (far right black line) in both cases is the same: ΔG°a = -18 kJ/mol, m-value(a) = 6 kJ/mol, fraction CD change = 0.4, ΔGb = -48 kJ/mol, m-value(b) = 8 kJ/mol, fraction CD change = 0.6. Each successive line going from right to left represents an additional 6 kJ/mol destabilization for SDb. For the independent model, note that once the ΔGb is smaller than -24 kJ/mol (red-orange), the order of the denaturation changes and the larger loss of structure occurs first. For the sequential model, SDa has to denature first, which means that after ΔGb is smaller than -24 kJ/mol, the transitions become progressively steeper due to the cooperative denaturation of SDb as soon as SDa denatures. This single cooperative transition has a progressively lower and lower Cm even though the stability and m-value for SDa has not changed. At intermediate destabilizations of SDb, the two transitions can appear merged and will fit to a two-state model with a single transition, albeit with a lower *m* value than the total of the two subdomains.

Alternative divisions of C-terminal subdomain. The C-terminal subdomain described in the main text includes β-strands 23 - 28, which encompasses the second half of the flap and the final four strands in the β-helix. This arrangement results in the β-strands of the flap region being assigned to two separate subdomains. Two alternatives are considered that result in the entire flap remaining in a single subdomain. In the first, the residues in the flap are proposed to be entirely in SD1. Since the four anti-parallel β-strands comprising the flap region are interrupted by a two complete β-circuits including parallel β-strands 17 – 22 (Figure 2A) this would require that either F241 not be part of SD1 or that isolated β-strands remain folded after their neighbors denature. This alternative is therefore inconsistent with the F241K data and the cooperative nature of β-strand formation. In the second alternative, the flap is proposed to be entirely in SD2 or SD3. If the entire flap were part of these subdomains, then the entirety of β-circuit 4 would need to reside in SD1 to account for the change in CD that occurs during T_1. In this scenario, an alteration of I207 would lead to changes in T_1 but not in T_2. Instead, the I207N substitution affects both T_1 and T_2 due to a loss in stability for both SD2 and SD3 (Table S1). The simplest explanation that explains the results is thus that the C-terminal subdomain encompasses β-strands 23-28.

Discussion of differential affects on SD2 and SD3 for cysteine replacements. Loss of the disulfide bond has the largest affect on the stability of the polar core subdomain. However, increasing the polarity at these positions has the largest effect on the nonpolar subdomain stability. This apparent contradiction in results can be understood by assuming that the disulfide bond is at the interface between the two subdomains, and thus, the type of mutation can change the effect on the two subdomains. Placing polar serine residues next to the nonpolar core should greatly destabilize the nonpolar subdomain, in accord with the results. Destabilization of the polar core subdomain by the serine mutations could be through both the absence of the disulfide and/or the change in interactions involving the internal water molecule. The lack of change in stability of the nonpolar subdomain in the C144A-C147A double mutant suggests that polarity at these positions and not the disulfide bond itself is important for the stability of the nonpolar core. This is supported by the observation that hemolysin TPS domains that lack a disulfide at these positions maintain nonpolarity at position 144 and a glycine at position 147 (Figure 2A).

Interactions disrupted in polar core subdomain mutations. Phenylalanine 80, the only non polar residue targeted in the polar core subdomain forms an edge on ring interaction with F71 near the Q125-Y134 hydrogen bond. The semi-conserved F80 was replaced with a leucine residue (F80L). This is a conservative mutation based on sequence alignments. As expected, this conservative replacement leaves the thermodynamic values of all of the subdomains essentially unaltered in comparison to HpmA265.

Q125 has been previously shown to form a direct hydrogen bond to both Y134 and a water within the interior of HpmA265[13]. Substitutions at position 125 affect the stability of the polar core subdomain to differing degrees depending upon the size of the nonpolar residue replacing the polar glutamine (Table S1). Mutation of this glutamine to an alanine (Q125A) or phenylalanine (Q125F) destabilizes the polar core subdomain and lowers the C_m (Table 5). The Q125F substitution destabilizes the polar core subdomain to the point that denaturation of the nonpolar and polar core subdomains is nearly indistinguishable such that T_2 and T_3 form a broad second transition (Figure 5C). This new T_2 has a larger fractional CD change which encompasses the HpmA265 CD change for T_2 and T_3. However, the Q125F T_2 has a lower *m* value of 6 ± 1, compared to the HpmA265 values of 8.2 and 7.9 for T_2 and T_3 respectively. The discrepancy between the loss of structure as monitored by the change in CD signal and the *m* value suggests that the second transition in the Q125F denaturation curve is composed of multiple transitions.

Asparagine 66 and 111 are positioned within the central most portion of the polar core, where their associated side chain atoms form hydrogen bonds with other internal side chain and main chain atoms. Mutations of the asparagine residues to the nearly isosteric leucine were chosen to maintain much of the van der Waals contact area, while eliminating hydrogen bonding capability. Like the Q125F mutation, the denaturation data from the N111L protein is fit well by either a three-state or four-state model. As in Q125F, the native CD signal for the N111L protein is within error of the HpmA265 signal, suggesting that it is fully folded. Like Q125F, the *m* value for the new second transition is smaller even though the

Table S1. Denaturation Fit Values

Sample	T1			T2			T3		
	$\Delta G°_{H2O}$ (kJ/mol)	*m* value (kJ/mol/M)	C_m (M)	$\Delta G°_{H2O}$ (kJ/mol)	*m* value (kJ/mol/M)	C_m (M)	$\Delta G°_{H2O}$ (kJ/mol)	*m* value (kJ/mol/M)	C_m (M)
HpmA265	-11.7 ± 0.6 [a]	6.3 ± 0.2	1.87 ± 0.04	-29.9 ± 0.9	8.6 ± 0.2	3.50 ± 0.01	-52 ± 4	7.7 ± 0.5	6.8 ± 0.2
trHpmA249	-10 ± 3 [b]	7 ± 3 [b]	1.5 ± 0.2 [b]	-33 ± 3	9.5 ± 0.7	3.46 ± 0.05	-46 ± 5	6.9 ± 0.7	6.6 ± 0.2
F241K	-14 ± 5 [b]	9 ± 3 [b]	1.6 ± 0.2 [b]	-29 ± 3	8 ± 1	3.48 ± 0.07	WT [c]	WT	WT
I207N	-5.9 ± 0.3	8.2 ± 0.3	0.72 ± 0.03	-17.2 ± 0.7	8.5 ± 0.3	2.02 ± 0.01	WT	WT	WT
L194N	-13.4 ± 0.3	10.4 ± 0.2	1.29 ± 0.01	ND [d]	ND	ND	-64 ± 9	9 ± 1	6.9 ± 0.1
C144A-C147A	-15.4 ± 0.9	7.5 ± 0.4	2.05 ± 0.10	-32 ± 2	9.2 ± 0.5	3.44 ± 0.05	-45 ± 5	8.3 ± 0.9	5.4 ± 0.1
C144S-C147S	-26 ± 2	13.1 ± 0.5	1.99 ± 0.07	ND	ND	ND	-46 ± 2	8.1 ± 0.5	5.7 ± 0.1
Y134F	-12.4 ± 0.5	7.1 ± 0.2	1.75 ± 0.03	WT	WT	WT	-44 ± 7	8 ± 1	5.7 ± 0.1
Q125A	-15.9 ± 0.6	8.3 ± 0.3	1.91 ± 0.01	WT	WT	WT	-34 ± 1	7.0 ± 0.1	4.9 ± 0.1
Q125F	-14 ± 2 (-12 ± 1)	7.6 ± 0.9 (6.3 ± 0.4)	1.8 ± 0.10 (1.96 ± 0.06)	-23 ± 3 (WT)	6.1 ± 0.7 (WT)	3.82 ± 0.05 (WT)	ND (-28 ± 4)	ND (7 ± 1)	ND (4.2 ± 0.1)
N111L	-11 ± 1 (-9.5 ± 0.5)	6 ± 0.9 (4.2 ± 0.2)	1.84 ± 0.09 (2.26 ± 0.06)	-23 ± 2 (WT)	6 ± 0.5 (WT)	3.79 ± 0.04 (WT)	ND (-28 ± 1)	ND (6.9 ± 0.3)	ND (4.1 ± 0.1)
N66L	-12 ± 1	6.1 ± 0.8	1.91 ± 0.07	WT	WT	WT	-42 ± 5	9.1 ± 0.9	4.6 ± 0.1
F80L	-13.7 ± 0.3	7.3 ± 0.3	1.89 ± 0.04	-30 ± 2	8.7 ± 0.3	3.45 ± 0.08	-49 ± 2	7.6 ± 0.4	6.5 ± 0.1

[a] Errors represent standard deviation from at least three trials.
[b] trHpmA249 and F241K each had a small (~1%) change in CD signal at low GdnHCl concentrations. This change was modeled by a small transition as it corresponded roughly with the location of the HpmA265 T1. Inclusion of this transition improved the randomness of the residuals, but did not significantly affect the remaining fit parameters.
[c] To constrain the fits these values were held fixed at the values for the unmodified HpmA265 values.
[d] No separate transition was detected for this subdomain

fractional change in CD is larger. Fitting N111L to a four-state model with the parameters for SD2 fixed at the HpmA265 values allows estimation of a nearly 24 kJ/mol reduction in the $\Delta G°_{H_2O}$ due to the loss of the N111 side chain hydrogen bonds. Similarly, the effect of N66L is manifested as a destabilization of SD3. In both the N66L and the N111L mutations, the loss of the polar carboxamide group has resulted in a decrease in stability.

Q125F Controls. Q125F was assayed for native, full-length structure with native-like function in the TAHA assay. Q125F has a CD spectrum that recapitulates the folded HpmA265 within the resolution of the experiment, suggesting that the mutant forms a native state with similar amounts of secondary structure (data not shown). To ensure a degradative process had not eliminated the structure associated with T_3, N-terminal amino acid sequencing and MALD-TOF were conducted on Q125F. The results verified a fully intact, Sec pathway processed, protein with a molecular weight consistent with amino acids N30 through G265 (data not shown). Thus, the Q125F alteration has destabilized the N-terminal polar core such that T_2 and T_3 are visually indistinguishable.

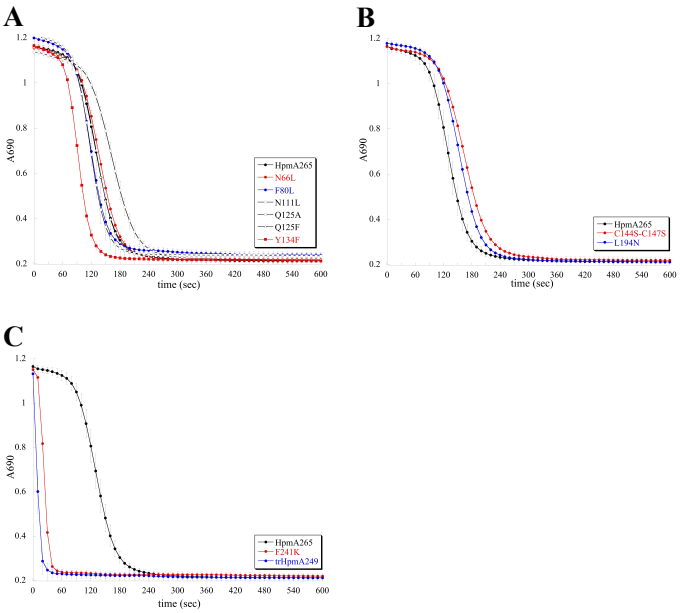

Figure S6. Template assisted hemolytic activity. HpmA265 template-assisted hemolysis activity has been compared to (A) polar, (B) nonpolar, and (C) C-terminal subdomain mutants.

Template-assisted hemolytic activity is maintained. To test that the introduced mutations still had HpmA265 like activity, and thus HpmA265-like structure, we performed template–assisted hemolytic assays. These assays depend on the ability of HpmA265 to function *in trans* as a template and activate a non-secreted, inactive, form of full-length HpmA[13]. The template assisted hemolytic activity of each protein construct was quantified in terms of the lag time and V_{max} (Figure S6). With the exception of F241K and trHpmA249, all of the mutants showed near HpmA265 lag times and V_{max} values (Figure S6 and Table S2). Because these times are related to the template-assisted buildup of folded full-length HpmA necessary to lyse a red blood cell, the similar values for the lag times suggest a HpmA265 like structure for the variants in this study. Intriguingly, both the trHpmA249 and F241K constructs virtually eliminated the lag time for full-length HpmA activation and increased the relative V_{max} by a factor of four. It appears the template-assisted function of HpmA265 resides in the C-terminal subdomain. These functional results were not the focus of this study, but are important in order to illustrate that the various site specific alterations did not, in general, affect ability of *in trans* template-assisted activation of full length HpmA, supporting the contention that the site specific alterations did not grossly affect the structure.

Table S2. Template-assisted hemolytic data

Sample	Relative V_{max} [a]	Relative Lag Time [c]
HpmA265	1.00 ± 0.02 [b]	1.00 ± 0.07 [b]
N66L	0.95 ± 0.06	1.04 ± 0.04
F80L	1.10 ± 0.05	0.92 ± 0.02
N111L	0.78 ± 0.04	1.25 ± 0.03
Q125A	1.20 ± 0.12	0.93 ± 0.02
Q125F	1.00 ± 0.04	1.02 ± 0.02
Y134F	1.50 ± 0.07	0.73 ± 0.03
DMS	0.81±0.02	1.27±0.02
L194N	0.93 ± 0.02	1.21 ± 0.02
F241K	4.50 ± 0.26	0.099 ± 0.001
trHpmA249	4.60 ± 0.77	ND [d]

[a] V_{max} for sample relative to V_{max} of HpmA265 experiment run in parallel.
[b] errors are relative standard deviations from at least three replicates.
[c] lag time for the sample relative to lag time for HpmA265 run in parallel.
[d] lag time was not determined (ND) under the experimental conditions.

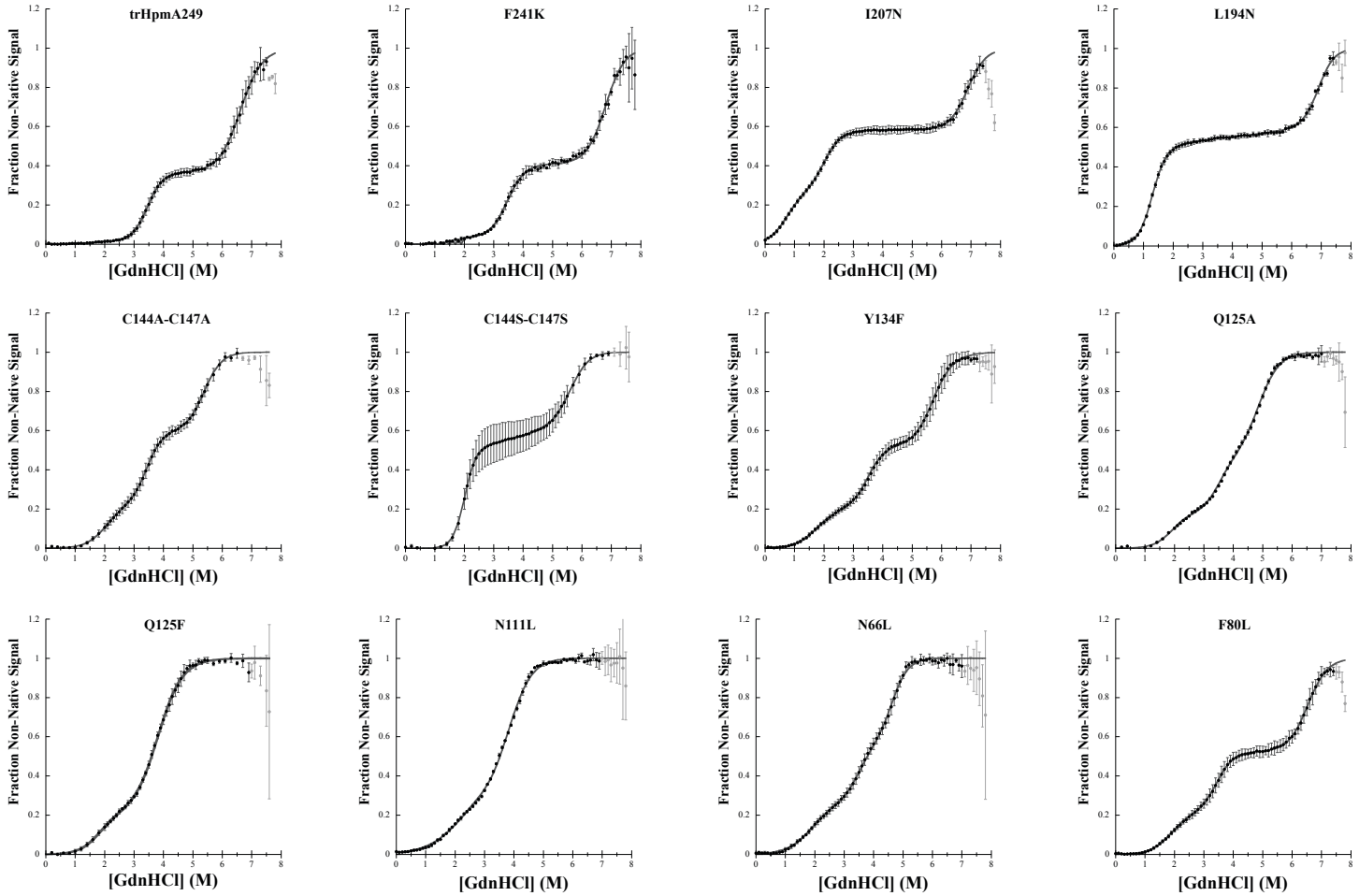

Figure S7. Average normalized data with fit line showing for each of the site directed mutants. Errors represent standard deviations from at least three measurements. Gray points were not fit.

Supplemental Materials and Methods

Assignment of secondary structure. Manual tabulation of residues based on dihedral angles was performed using the dihedrals output from the STRIDE program with the following two stipulations: 1) phi and psi were confined to be in the range of -45 to -180 or 170 to 180° for phi and -165 to -180 or 160 to 180° for psi; and, 2) residues were only tabulated if more than two residues in a row matched criteria (1). All of these methods assume that all of the β-structures provide the same contribution to the overall CD signal, which is not necessarily true[14-16]. Although the assigned secondary structure for the majority of the residues was consistent for the different models, the exact start and stop of any particular element differed between the different models.

Schemes for four-state unfolding of HpmA265. The three transitions observed for guanidine•HCl (GdnHCl) denaturation of HpmA265 are modeled in two ways as either independent unfolding transitions or as sequential unfolding transitions. Both models provide equilibrium free energy changes ($\Delta G°_{H_2O}$), GdnHCl dependence (*m* value) and fractional changes in CD signal for each transition. In addition, both models fit the data with low error and similar residuals, and they both provided similar values for the fit parameters. Since the two models provide nearly indistinguishable parameters and quality of fits, it is not possible to distinguish between them based on these results. This coincidence in the extracted parameters is due to the large differences in the C_m values for each transition. The schemes, the associated equilibrium constants and the equations used to fit the observed CD profiles are provided below.

Independent Subdomain Unfolding Scheme

In the context of independent subdomains, the native state (N) corresponds to $N_1N_2N_3$, the first intermediate (I_1) corresponds to $D_1N_2N_3$, the second intermediate (I_2) corresponds to $D_1D_2N_3$, and the denatured state corresponds to $D_1D_2D_3$.

Native Subdomain 1 (N_1) $\xrightleftharpoons{K_1}$ Denatured Subdomain 1 (D_1)

Native Subdomain 2 (N_2) $\xrightleftharpoons{K_2}$ Denatured Subdomain 2 (D_2)

Native Subdomain 3 (N_3) $\xrightleftharpoons{K_3}$ Denatured Subdomain 3 (D_3)

$$K_1 = \frac{[D_1]}{[N_1]}; \Delta G_1^o = -RT \ln K_1$$

$$K_2 = \frac{[D_2]}{[N_2]}; \Delta G_2^o = -RT \ln K_2$$

$$K_3 = \frac{[D_3]}{[N_3]}; \Delta G_3^o = -RT \ln K_3$$

The fraction unfolded (θ_D) for each of the independent transitions is given by the following formula.

$$\theta_D(x) = \theta_{Dx} = \frac{[D_x]}{[N_x]+[D_x]} = \frac{K_x}{1+K_x} = \frac{e^{\frac{-\Delta G°}{RT}}}{1+e^{\frac{-\Delta G°}{RT}}}$$

Assuming a linear dependence on the unfolding free energy[9] such that $\Delta G° = \Delta G°_{H2O} - m•[GdnHCl]$, and recognizing that the $\Delta G°(\text{unfolding}) = -\Delta G°(\text{folding})$.

$$\theta_{Dx} = \frac{e^{\frac{\Delta G^o_{H_2O}(\text{folding}) + m\cdot[GdnHCl]}{RT}}}{1 + e^{\frac{\Delta G^o_{H_2O}(\text{folding}) + m\cdot[GdnHCl]}{RT}}} \qquad (1)$$

The total signal is a linear combination of the signals due to each of the independent transitions.

$Y_{total} = Y_{N1}\theta_{N1} + Y_{D1}\theta_{D1} + Y_{N2}\theta_{N2} + Y_{D2}\theta_{D2} + Y_{N3}\theta_{N3} + Y_{D3}\theta_{D3}$

Recalling that $\theta_{N1} + \theta_{D1} = 1$.

$Y_{total} = Y_{N1}+(Y_{D1}-Y_{N1})\theta_{D1} + Y_{N2}+(Y_{D2}-Y_{N2})\theta_{D2} + Y_{N3}+(Y_{D3}-Y_{N3})\theta_{D3}$

Defining a ΔCD for each transition such that $\Delta CD = Y_x - Y_N$ and defining a total native signal Y_N.

$Y_{total} = Y_N + \Delta CD_1\bullet\ \theta_{D1} + \Delta CD_2\bullet\ \theta_{D2} + \Delta CD_3\bullet\ \theta_{D3}$

Where each fraction unfolded is given by equation (1). The signal observed during the titration with GdnHCl decreased due to dilution of the protein with GdnHCl. This raw observed signal was fit with an equation that accounted for the change in concentration.

$$Y_{obs} = Y_{total}\frac{[P_{x\ Gdn}]}{[P_o]}$$

Where $[P_o]$ is the initial protein concentration at the start of the titration and $[P_{x\ Gdn}]$ is the concentration at "x" concentration of GdnHCl. This allowed the errors to be treated as uniform throughout the range of the titration (See below for effect on the fit parameters of using instrument generated errors to weight data when fitting).

Sequential Subdomain Unfolding Scheme

$$N \underset{}{\overset{K_1}{\rightleftharpoons}} I_1 \overset{K_2}{\rightleftharpoons} I_2 \overset{K_3}{\rightleftharpoons} D$$

$$K_1 = \frac{[I_1]}{[N]}; \Delta G^o_1 = -RT\ln K_1$$
$$K_2 = \frac{[I_2]}{[I_1]}; \Delta G^o_2 = -RT\ln K_2$$
$$K_3 = \frac{[D]}{[I_2]}; \Delta G^o_3 = -RT\ln K_3$$

Following previously publish derivations[17-19] and similar logic to the independent pathway, the equation for the total signal can be derived (see equations 2-4 below). In this equation, the f_{CD} values represent the fractional change in CD signal for each transition.

Data Analysis. To consistently account for the poor baseline after T_3, data for F241K and I207N were fit assuming that the stability and GdnHCl dependence for denaturation of SD3 were the same as for HpmA265. These constraints resulted in little change for the parameters for SD1 from the unconstrained values. To obtain random residuals, L194N required a sloping baseline between the two transitions. To allow estimation of the effect of substitutions in the polar core for the highly destabilizing mutations, the SD3 mutants (N66L, N111L, Q125A, Q125F, and Y134F) were all fit assuming constant parameters for SD2. These constraints allowed separation of the second and third transition for Q125F and N11L. In the case of Q125F and N111L, results are provided for fits to both three and four-state models. For both SD1 and SD3 mutations, inspection of Figure 6A and 6C shows that the normalized data for the transitions that are held constant are nearly coincident with the HpmA265 data. Individual fits for each site specific mutation are presented in Figure S6. In all cases, the minimum number of constraints were applied provided that the residuals remained $\lesssim$ 1%. Transition mid-points (C_m) were calculated for each transition based on the $\Delta G°_{H2O}$ and m value for that transition using the formula $C_m = -\frac{\Delta G°_{H_2O}}{m\ value}$.

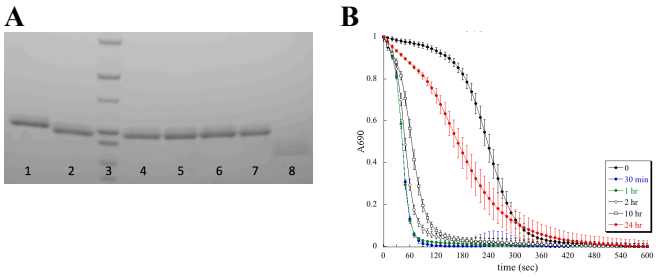

Figure S8. Trypsin digestion of HpmA265. (A) SDS-PAGE results from a 24-hr trypsin digest of HpmA265, where Lanes 1 – 8 represent 0, 0.5, 1, 2, 4, 10 and 24-hr incubation time points. (B) Template-assisted hemolytic activity results from various time points along a 24-hr trypsin digest of HpmA265. The 2-hr sample was subjected to MALDI-TOF (UW-Madison Biotechnology Center, Madison, WI) after heat inactivation trypsin at 95 °C for 4 hrs. A 23,050 Da fragment was determined, which maps most closely to a trypsin cut site at R249, leaving a LepB processed and HpmB secreted 23,053 Da HpmA265 fragment.

Guanidine Hydrochloride Timed Incubations. HpmA265 was adjusted 124 μM and 6 M GdnHCl and allowed to incubate at room temperature. In order to discern between slow denaturation and aggregate formation, aliquots were removed and diluted to 10 μM HpmA265 and 0.5 M GdnHCl at the following time points: 0, 15, 30 min, 1, 3, 6, 10, 12, 24, 48 and 120 hrs. Individual wavescans (300 – 200 nm) and forward titrations (0 to 7.8 M GdnHCl) were conducted via CD, while DLS was used to measure particle size and distribution.

Single-jump GdnHCl kinetics. Single-jump GdnHCl equilibration experiments were conducted in order to determine the reversibility of HpmA265 folding. Denatured stock solutions of HpmA265 at 7.8, 6.9, 6.0, 5.5, 4.0, 3.4, 3.0, 1.9 and 1.0 M, GdnHCl were mixed with buffer to obtain a final GdnHCl concentration of 0.5 M. Complete reversibility for samples incubated at less than or equal to 6.0 M GdnHCl was demonstrated from these single jump experiments via re-establishment of ellipticity at 220 nm (data not shown).

$$Y_{obs} = Y_N\left(1 - \frac{f_{CD1}K_1 + f_{CD2}K_1K_2 + f_{CD3}K_1K_2K_3}{1 + K_1 + K_1K_2 + K_1K_2K_3}\right)\frac{[P_{x\ Gdn}]}{[P_o]} \qquad (2)$$

$$f_{CD1} = 1 - \frac{Y_{I1}}{Y_N}; f_{CD2} = 1 - \frac{Y_{I2}}{Y_N}; f_{CD3} = 1 - \frac{Y_D}{Y_N} \qquad (3)$$

$$Y_{obs} = Y_N\left(1 - \frac{f_{CD1}e^{\frac{\Delta G^o_{H_2O,1} + m_1[GdnHCl]}{RT}} + f_{CD2}e^{\frac{(\Delta G^o_{H_2O,1} + \Delta G^o_{H_2O,2}) + (m_1+m_2)[GdnHCl]}{RT}} + f_{CD3}e^{\frac{(\Delta G^o_{H_2O,1} + \Delta G^o_{H_2O,2} + \Delta G^o_{H_2O,3}) + (m_1+m_2+m_3)[GdnHCl]}{RT}}}{1 + e^{\frac{\Delta G^o_{H_2O,1} + m_1[GdnHCl]}{RT}} + e^{\frac{(\Delta G^o_{H_2O,1} + \Delta G^o_{H_2O,2}) + (m_1+m_2)[GdnHCl]}{RT}} + e^{\frac{(\Delta G^o_{H_2O,1} + \Delta G^o_{H_2O,2} + \Delta G^o_{H_2O,3}) + (m_1+m_2+m_3)[GdnHCl]}{RT}}}\right)\frac{[P_{x\ Gdn}]}{[P_o]} \qquad (4)$$

Tryptic digestion. To confirm the stability and functionality of trHpmA249, trypsin treated HpmA265 aliquots (as described in the Materials and Methods) were removed for SDS-PAGE and template-assisted hemolytic activity at 0.5, 1, 2, 4, 10 and 24 hr time points. This trypsin resistant species was maintained up through 10 hours of digestion, and was more active in hemolytic assays compared to the full length HpmA265 (Figure S7). SEC-LS confirmed a size and mobility similar to HpmA265 (data not shown). Digestion for a full 24 hours resulted in further degradation and inactivation as evidenced by SDS-PAGE and hemolytic assays. Trypsin and the C-terminal fragment of HpmA265 were purified away from trHpmA249 via Ni2+-NTA chromatography.

Dynamic Light Scattering. The determination of HpmA265 size distribution profiles in the presence and absence of GdnHCl was conducted via dynamic light scattering using a Zetasizer Nano ZS (Malvern Instruments Ltd., Worcestershire, UK). Samples were prepared via ultracentrifugation at 100,000 rpm for 1 hr and 0.02 mM filtration (Sigma-Aldrich, St. Louis, MO) prior to DLS.

Template-Assisted Hemolysis. HpmA265 and all mutant forms of HpmA265 were analyzed for template-assisted hemolytic activity on a Synergy 2 microtiter plate reader (Biotek, Winooski, VT) operating in a continuous kinetic read mode as described previously[13]. Briefly, unfolded full length HpmA from *E. coli* lysate, red blood cells, 10 mM sodium phosphate, 150 mM NaCl pH 7.4 (PBS), and 786 nM of either native or mutant forms of HpmA265 were mixed; hemolysis was monitored by a decrease in optical density at 690 nm. Gene5.2 software (Biotek) was used to analyze each protein in terms of the maximum hemolytic velocity (V_{max}) and lag time. The V_{max} is the maximum slope of Absorbance at 690 nm versus time. The intercept between the extrapolated V_{max} fit and the absorbance at t = 0 defines the lag time. Both quantities depend on the concentration of the unfolded full length HpmA and on the quantity of red blood cells present in the assay. Since these quantities depend on the concentration of unfolded full length HpmA present in the assay, the extracted values are transformed to relative values by dividing by the values obtained for HpmA265 assayed in parallel with identical lysate and blood. Relative standard deviations are calculated as the standard deviation of the sample divided by the average value for HpmA265. Template-assisted hemolytic assays were also conducted on HpmA265 before (Pre-GdnHCl) and after GdnHCl addition and removal (Post-GdnHCl). HpmA265 (170 mM) was adjusted to 6.0 M GdnHCl and 100 mM via the addition of a solution containing 10 mM sodium phosphate (pH 7.4) and 8.04 M GdnHCl. This sample was allowed to equilibrate for 30 min at room temperature. This sample was diluted to 0.5 M GdnHCL (8 mM HpmA265) and dialyzed against 2 L of 10 mM sodium phosphate and 150 mM NaCl overnight at room temperature.

References

1. van Ulsen P, Rahman SU, Jong WSP, Daleke-Schermerhorn MH, Luirink J (2014) Type V secretion: from biogenesis to biotechnology. Biochim Biophys Acta 1843:1592–1611.
2. Jacob-Dubuisson F, Fernandez R, Coutte L (2004) Protein secretion through autotransporter and two-partner pathways. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research [Internet] 1694:235–257. Available from: http://www.sciencedirect.com/science/article/pii/S0167488904000886
3. Hodak H, Clantin B, Willery E, Villeret V, Locht C, Jacob-Dubuisson F (2006) Secretion signal of the filamentous haemagglutinin, a model two-partner secretion substrate. [Internet] 61:368–382. Available from: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=16771844&retmode=ref&cmd=prlinks
4. Ur Rahman S, Kajava AV, van Ulsen P, Baxa U, Steven AC (2013) System specificity of the TpsB transporters of coexpressed two-partner secretion systems of Neisseria meningitidis. J Bacteriol 195:788–797.
5. Mazar J, Cotter PA (2007) New insight into the molecular mechanisms of two-partner secretion. Trends Microbiol 15:508–515.
6. Noinaj N, Noinaj N, Buchanan SK, Buchanan SK (2014) FhaC takes a bow to FHA in the two-partner do-si-do. Mol Microbiol [Internet] 92:1155–1158. Available from: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=24798489&retmode=ref&cmd=prlinks
7. Hammack B, Attfield K, Clayton D, Dec E, Dong A, Sarisky C, Bowler BE (1998) The magnitude of changes in guanidine-HCl unfolding m-values in the protein, iso-1-cytochrome c, depends upon the substructure containing the mutation. Protein Sci 7:1789–1795.
8. Wrabl J, Shortle D (1999) A model of the changes in denatured state structure underlying m value effects in staphylococcal nuclease. Nat Struct Biol 6:876–883.
9. Street TO, Courtemanche N, Barrick D (2008) Protein folding and stability using denaturants. Methods Cell Biol 84:295–325.
10. Myers JK, Pace CN, Scholtz JM (1995) Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. Protein Sci 4:2138–2148.
11. Clantin B, Clantin B, Hodak H, Hodak H, Willery E, Willery E, Locht C, Locht C, Jacob-Dubuisson F, Jacob-Dubuisson F, et al. (2004) The crystal structure of filamentous hemagglutinin secretion domain and its implications for the two-partner secretion pathway. Proceedings of the National Academy of Sciences [Internet] 101:6194–6199. Available from: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=15079085&retmode=ref&cmd=prlinks
12. Swihart KG, Welch RA (1990) Cytotoxic activity of the Proteus hemolysin HpmA. [Internet] 58:1861–1869. Available from: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=2341182&retmode=ref&cmd=prlinks
13. Weaver TM, Smith JA, Hocking JM, Bailey LJ, Wawrzyn GT, Howard DR, Sikkink LA, Ramirez-Alvarado M, Thompson JR (2009) Structural and functional studies of truncated hemolysin A from Proteus mirabilis. J Biol Chem 284:22297–22309.
14. Sreerama N, Woody RW (2004) Computation and analysis of protein circular dichroism spectra. Meth Enzymol 383:318–351.
15. Martin SR, Schilstra MJ (2008) Circular dichroism and its application to the study of biomolecules. Methods Cell Biol 84:263–293.
16. Brahms S, Brahms J (1980) Determination of protein secondary structure in solution by vacuum ultraviolet circular dichroism. J Mol Biol 138:149–178.
17. Patel S, Chaffotte AF, Goubard F, Pauthe E (2004) Urea-induced sequential unfolding of fibronectin: a fluorescence spectroscopy and circular dichroism study. Biochemistry 43:1724–1735.
18. Hung H-C, Chen Y-H, Liu G-Y, Lee H-J, Chang G-G (2003) Equilibrium protein folding-unfolding process involving multiple intermediates. Bull. Math. Biol. 65:553–570.
19. Hung H-C, Chang G-G (2001) Multiple Unfolding Intermediates of Human Placental Alkaline Phosphatase in Equilibrium Urea Denaturation. Biophys J 81:3456–3471.