**Title**
Anthrax lethal toxin unfolding and translocation via a charge-state ratchet



**Author**
Brown, Michael Joseph

**Publication Date**
2013

Peer reviewed|Thesis/dissertation

# Anthrax lethal toxin unfolding and translocation via a charge-state ratchet

by

Michael Joseph Brown

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Molecular & Cell Biology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Bryan Krantz, Chair
Professor Susan Marqusee
Professor Andreas Martin
Professor David Wemmer

Spring 2013

Anthrax lethal toxin unfolding and translocation via a charge-state ratchet


Michael Joseph Brown

Abstract

Anthrax lethal toxin unfolding and translocation via a charge-state ratchet

by

Michael Joseph Brown

Doctor of Philosophy in Molecular & Cell Biology

University of California, Berkeley

Professor Bryan Krantz, Chair

A fundamental problem in biochemistry is how molecular machines convert chemical potential energy into mechanical work. Here, this problem is addressed in terms of how a proton gradient (ΔpH) drives anthrax toxin transmembrane protein translocation with a given directionality. One model to explain this phenomenon is a charge-state Brownian ratchet, wherein movement is biased by asymmetries across the membrane. Acidic residues in the substrate are protonated on the lower pH starting side, thus allowing them to pass through the cation-selective protective antigen (PA) channel. Movement can occur in either direction according to Brownian motion, but acidic residues that reach the higher pH of the destination are deprotonated, preventing retrotranslocation and resolving movement in one direction.

This model is probed through the use of planar lipid bilayer electrophysiology to analyze the charge requirements of the model substrate $LF_N$, the binding domain of the natural substrate lethal factor. Acidic residues are necessary and sufficient for ΔpH-driven translocation provided the starting side pH is low enough to sufficiently protonate the residues. Their position in the sequence, just before $LF_N$'s folded domain, plays a key role in substrate unfolding. Basic charges are important for initiation into the channel and to chaperone deprotonated acidic residues during translocation.

Further work by Sarah Wynia-Smith confirms the importance of acidic residues just before $LF_N$'s folded domain and points to the existence of electrostatic translocation barriers in the PA channel. I identify key residues in the upper portion of the channel's β barrel that contribute to its cation selectivity. These residues also prove to play a significant role in substrate initiation and translocation, supporting the role for a charge gate in the channel that prevents retrotranslocation.

Finally, secondary structure of the translocating substrate is analyzed. $LF_N$ translocates in a compact helical state nucleated by the channel's α clamp, a nonspecific helix-binding site recently identified by Geoffrey Feld. α clamp-induced helix formation lowers a previously identified but uncharacterized translocation barrier, and substrates that cannot form helices cannot translocate. This suggests a mechanism wherein helix melting, coupled to the electrostatic ratchet, drives translocation-coupled substrate unfolding.

Dedicated to Tania and my family.

# Table of Contents

# List of Figures

# List of Tables

# List of Abbreviations

| | |
|---|---|
| α clamp | nonspecific helix binding site in protective antigen channel |
| $\Delta\psi$ | membrane potential |
| $\Delta\psi_{rev}$ | reversal potential |
| $\Delta pH$ | proton gradient |
| $\Delta G^{\ddagger}$ | activation energy |
| $\Delta G_{NI}$ | thermodynamic difference between native and intermediate folding states |
| ϕ clamp | hydrophobic/aromatic binding clamp in protective antigen channel |
| ATP | adenosine-5'-triphosphate |
| BR | Brownian ratchet |
| CD | circular dichroism |
| cis | membrane side where channel and substrate are first added |
| EF | edema factor |
| $EF_N$ | edema factor amino-terminal domain |
| EM | electron microscopy |
| GdmCl | guanidinium chloride |
| $His_6$ | six-histidine affinity tag |
| LF | lethal factor |
| $LF_N$ | lethal factor amino-terminal domain |
| MES | 2-(*N*-morpholino)-ethanesulfonic acid |
| MS | mass spectrometry |
| MUT | mutant |
| PA | protective antigen |
| $PA_7$ | heptameric protective antigen oligomers |
| $PA_8$ | octameric protective antigen oligomers |
| $PA_xLF_y$ | holotoxin with x protective antigen and y lethal factor subunits |
| PCR | polymerase chain reaction |
| PDB | Protein Data Bank |
| PMF | proton motive force |
| PS | power stroke |
| S.D. | standard deviation |
| SUBB | supplemented universal bilayer buffer |
| $t_{1/2}$ | time required for half of the substrates to translocate during bulk translocation |
| trans | membrane side where substrate arrives after completed translocation |
| UBB | universal bilayer buffer |
| WT | wild-type |

# Acknowledgements

Thank you to everyone who helped me in my graduate career. A dissertation is by no means a solitary effort, and I have been especially lucky to have an excellent group of people supporting me. First, I would like to thank my advisor Professor Bryan Krantz, whose vision and expertise made this project possible and whose scientific training will continue to guide me throughout my career. I would also like to thank the other members of my thesis committee, Professors Susan Marqusee, Andreas Martin, and David Wemmer, for their time and guidance. Additionally, I wish to thank Professor Russell Vance who has helped me greatly in figuring out the next stage of my career.

I wish to thank all of the other Krantz Lab members for their training, assistance, helpful discussion, and friendship: Jen Colby, Geoff Feld, Shon Greenberg, Melinda Hanes, Alex Kintzer, Adam Schawel, Sam Stephenson, Yoki Tang, Katie Thoren, and Sarah Wynia-Smith. Special thanks also to Gina Chirichella and Claire Johns, two undergraduate students who each worked with me for a year and who helped me greatly. Additionally, I want to thank members of the Marqusee Lab, especially Diana Koulechova, Sabriya Rosemond, and Laura Rosen, for their friendship and scientific advice.

I would like to specifically thank the individuals who collected data that supported my own presented in this dissertation. Katie Thoren did single channel work in Fig. 2.3C and helped reproduce my results in Figs. 2.1 and 2.2. She also designed and made the $PA_{\alpha\text{-plug}}$ and polyproline $LF_N$ constructs in chapter 4, performed the binding assays in Figs. 4.1B and 4.4B, and did preliminary translocations that led to the work in Figs. 4.2A and 4.4C. Geoff Feld helped with the writing and figure design in the review that went on to form the basis for chapter 1. He also helped design many of the structural images in later chapters, and the work in chapter 4 would not have been possible without the crystal structure he discovered. Sarah Wynia-Smith performed all of the $LF_N/EF_N$ chimera work in chapter 3. Gina Chirichella helped in making the constructs used in Figs. 3.6 and 3.7, and performed preliminary assays that led to the work in Fig. 3.6B. Gigi Kemalyan performed the electron microscopy in Fig. 3.7B. Claire Johns analyzed the racemized substrate by circular dichroism in Fig. 4.5A. Tony Iavarone performed the mass spectrometry on the same substrate for Figs. 4.5C and 4.5D. Jen Colby analyzed peptide binding and translocation in Fig. 4.6B. And of course the advising of my mentor, Bryan Krantz, was instrumental in the designing, performance, and analysis of all my experiments.

Finally, this would not have been possible without the support of my loved ones. Thank you to my parents Andy and Sarah, my brother Rob, his wife Bev, and their three wonderful children, Alex, Jaime, and Bradley. My girlfriend, Tania Gonzalez, has been an incredible influence in my life during this time, affecting me both personally and scientifically.

This work was supported by the National Institutes of Health Research Grant R01-AI077703 (to Bryan Krantz) and National Institutes of Health Training Grants (to the department of Molecular & Cell Biology).

# Chapter 1

## Introduction to molecular ratcheting in protein translocation[1]

### 1.1 Significance of protein translocation

Transmembrane protein translocation is an essential process for all living things. An estimated 50% of all proteins must move across or into a membrane at some point in the course of their normal lifespan (1). This is unsurprising for eukaryotic cells, given the extent of compartmentalization into membrane-bound organelles, and particularly for multicellular organisms, where cell-cell communication is often mediated by exported proteins. Transmembrane translocation plays an essential role in a variety of processes, including protein trafficking, membrane and organelle biogenesis, endoplasmic reticulum-associated degradation of misfolded proteins, and antigen presentation (2, 3).

While these processes are unimportant for single-celled prokaryotes lacking complex organelles, over a quarter of all bacterial proteins are still thought to require translocation across the outer membrane (4). Of these proteins, effectors secreted by pathogens to manipulate and harm host cells are of particular interest to medical science. These toxins require passage through multiple membranes: one or two bacterial membranes (for gram-positive and gram-negative bacteria, respectively) and the host membrane. This can be accomplished using a single process, as occurs in type III (5) or type IV (6) secretion. Other times, secretion out of the bacterial cell (7) and toxin translocation into the host cell (8) require two distinct events.

While there are notable exceptions, including cotranslational translocation (9), nuclear pore transport (10), and twin-arginine transporters (11), generally the substrate protein must unfold, pass through a proteinaceous channel as an unfolded chain, and refold upon reaching its destination on the other side of the membrane. This process must be powered by ATP binding and hydrolysis (4), a chemical gradient (8), or a combination of the two (12). It should also be noted that, while not actually moving proteins across a membrane, there are cytoplasmic machines involved in protein degradation and disaggregation that share many of the properties of transmembrane translocases, including unfolding of the substrate, translocation of the unfolded substrate through a proteinaceous channel, and the coupling of an external energy source to power the process (13). Thus, studies of transmembrane translocation are likely to shed light on this analogous process as well.

### 1.2 Models of protein translocation

Generally, two competing mechanisms describing the functionality of nanoscale molecular machines have been presented (14) (Fig. 1.1). On one hand, nanomachines falling into the Brownian-ratchet (BR) classification do work by using external energy sources to harness Brownian thermal energy (15–18). A simplified system, based on the Smoluchowski-Feynman ratchet (15, 19), is depicted in Fig. 1.2. A particle is trapped in a stationary, anisotropic potential energy well. When an external energy source is used to momentarily *flash off* the potential, the particle is allowed to freely diffuse but is more likely to cross the closer energy barrier than the more distant one. Thus, when the potential is *flashed on*, there is a significantly higher chance of the particle progressing rather than regressing. Repeated cycle of such flashing will result in a net forward motion without an energy source ever being applied to the particle directly.

[1] Parts of this chapter are based on a previously published review, with permission from co-authors (124).

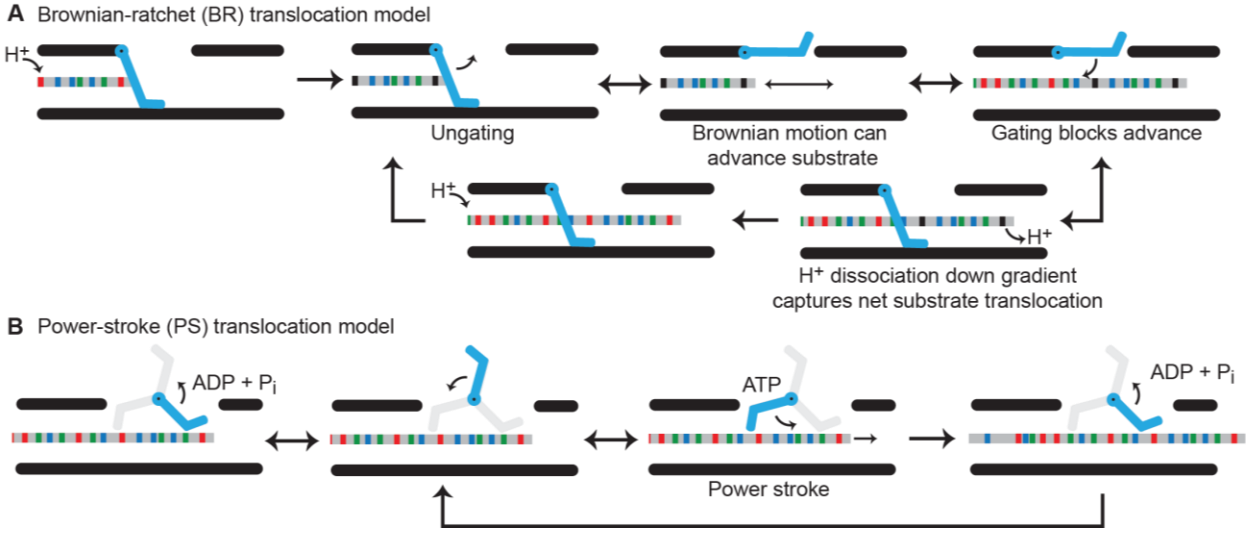


**Figure 1.1. Models of protein translocation.** (A) A ΔpH-dependent BR mechanism for protein translocation. The substrate polypeptide chain is depicted as a simplistic gray rod with different functional groups colored as follows acidic (red), basic (blue), and nonpolar (green). A gate (cyan) electrostatically excludes anionic charges on deprotonated acidic residues. In this cyclical mechanism, substrate acidic residues are protonated (black); the gate opens, allowing for Brownian motion to take place. The peptide can only advance up to the point where deprotonated acidic residues enter the channel. Closing of the gate traps the peptide in the channel, as the dissipation of $H^+$ ions down the gradient upon deprotonation prevents retrograde movements. Further protonation then allows the cycle to repeat. (B) An ATP-dependent PS mechanism for protein translocation. The substrate chain is colored as in (A) while in this case, the gate acts more like a paddle with active (cyan) and inactive (gray) states. In the ADP-bound state, the paddle has low affinity for peptide; the paddle exchanges ADP for ATP, and subsequent ATP hydrolysis triggers a conformational change, allowing the paddle to push the polypeptide chain forward (power stroke). Cycles of ADP release and ATP hydrolysis allow the mechanism to continue.

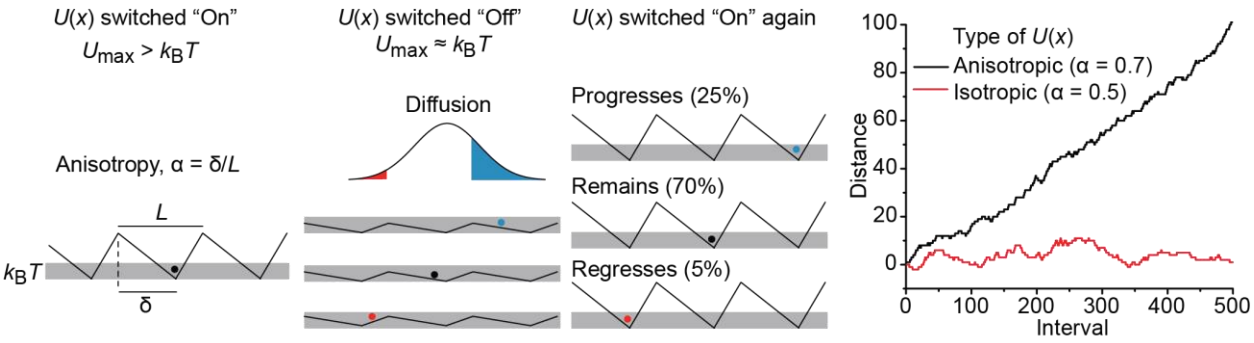


**Figure 1.2. Flashing BR model.** Far left, a saw-tooth potential function, $U(x)$, with respect to distance, $x$, is depicted, where the positional anisotropy, α, of $U(x)$ is related to the distance between the maximum and minimum, δ, of each period, $L$, such that $\alpha = \delta/L$. When $U(x)$ is switched *on*, the particle is trapped in a well, since $U_{max} > k_BT$. Middle left, when the potential is switched *off*, the particles diffuse freely according to Einstein's relation for the mean square displacement (20). Middle right, when $U(x)$ is switched back *on*, one interval is completed and the particle is trapped again. Thus the probability of progressing is greater than regressing. Far right, a Monte Carlo simulation of the flashing BR model plotting $x$ versus the number of switching intervals for an anisotropic (black, $\alpha = 0.7$) and an isotropic (red, $\alpha = 0.5$) $U(x)$.

Protein translocation can use an analogous system to harness Brownian thermal energy (Fig. 1.1A). Suppose an unfolded polypeptide chain has sites that can be modified to be in permissive and non-permissive states. The ratchet can switch between the permissive *on* and non-permissive *off* states through the modification of the substrate, for example by protonation. Analogously, a gate in the transporter switches to a permissive *open* or non-permissive *closed* state, depending on the protonation state of the polymer. Under the influence of a chemical gradient (e.g. a proton gradient, ΔpH) residues may be more likely to be in a permissive state on one side of the membrane than on the other. Thus translocation can proceed in the direction of non-permissive modification, since non-permissively modified portion of the substrate will be unable to retrotranslocate. A number of chemical modifications have been suggested to follow such a mechanism. Protonation (18, 21), chaperone binding (22, 23), glycosylation, disulfide bonding, and conditions that promote protein folding (16) are all capable of biasing movement through a translocase channel.

On the other hand, the power-stroke (PS) classification suggests that molecular machines use external energy sources to directly drive motion without the need for harnessing Brownian thermal fluctuations (Fig. 1.1B) (24). For example, the cycle of ATP binding, hydrolysis, and release of ADP and inorganic phosphate ($P_i$) induces a conformational change in the protein machine that performs useful work such as pulling or pushing a substrate through the translocase. The machine engages with the translocating polypeptide chain via a polypeptide clamp or other active site loop structure. Upon completion of the ATP binding, hydrolysis, and release cycle, the machine returns to its original conformation. However, the resetting of the polypeptide clamp or other active-site structure must occur via a different path back to the initial state, else it would essentially undo the work done in the ATP binding step.

The major difference between the BR and PS models is that the latter does not consider Brownian thermal fluctuations (or Langevin's random force) to be part of the mechanism. However, BR and PS mechanisms should not be held as mutually exclusive models. Instead, these two types of mechanisms may occur at distinct junctures in the transport cycle. For example, the two-headed motor kinesin uses an ATP-driven PS to nudge the trailing leg forward. However, this push provided by ATP binding is far too small to drive the leg the 8 nm span between binding sites on the microtubule, and a BR phase is required to bridge the remaining distance by allowing the head to diffuse and seek its next binding site, thus completing the transport cycle (25). Furthermore, as operating on polymers in a cell is akin to sailing in a hurricane, consider two readily available types of machinery for hoisting up a sail: a simple hand crank and a hand crank with a ratchet. Under ideal conditions, the crank driven by a PS may suffice in performing the task. However, under tropical cyclone conditions, a system that includes both a PS-driven crank and a ratchet may become necessary. While extremely windy and diffusive weather likely makes it difficult for the sailor to turn the simple hand crank in the proper direction, the ratchet ensures that sub-steps of a crank cycle are not lost to diffusive backsliding.

### 1.3 Anthrax toxin as a model system for protein translocation

The binary A/B bacterial toxins (26), including diphtheria, botulinum, and anthrax toxin, are a widely studied class of membrane transport systems, which use their own transport machinery to enter cells (27–29). These A/B toxins assemble into complexes on cell surfaces and then utilize host cell chemical potentials to unfold and translocate enzymatic factors into the host cell. Anthrax toxin, the binary $A_2B$ toxin (27) produced by *Bacillus anthracis* (8, 30) (Fig. 1.3A),

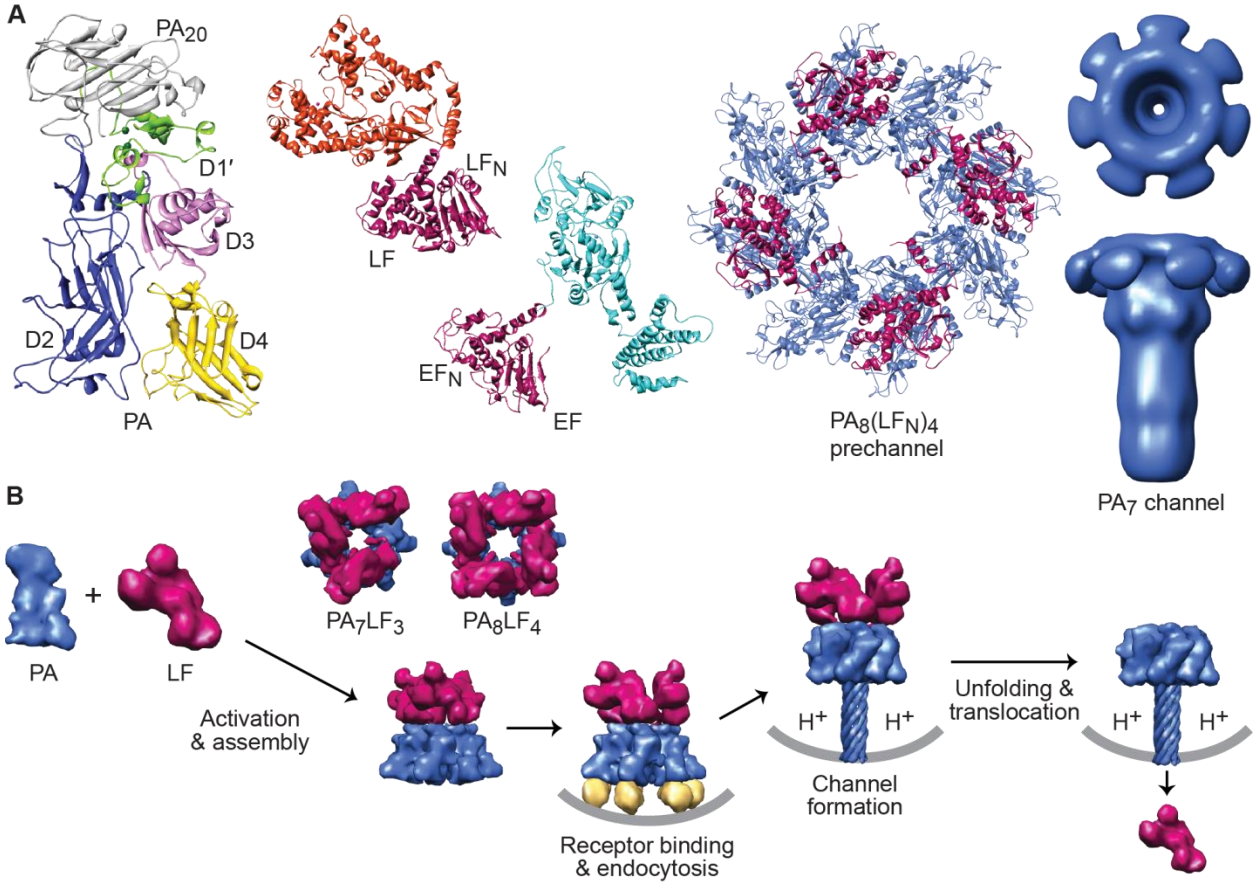


**Figure 1.3. An overview of the anthrax toxin protein translocation system.** (A) Components of anthrax toxin (left to right). Ribbons depiction of PA (3TEW) (31) colored by domain: $PA_{20}$ (gray), D1′ (green), Ca2+ ions (dark green), D2 (blue), D3 (magenta), and D4 (yellow). The enzymes, LF (1J7N) (32) and EF (1YOV) (33); their amino-terminal PA binding domains ($LF_N$ and $EF_N$, respectively) are colored red-violet and their catalytic domains colored orange and cyan, respectively. A representative PA prechannel complex, $PA_8(LF_N)_4$, (3KWV) (34); the PA oligomer and $LF_N$ colored denim and red-violet, respectively. Axial view (above) and side view (below) of a three-dimensional EM reconstruction of the $PA_7$ channel (35) (colored denim) (Prof. Mark Fisher graciously provided the EM density map). (B) Anthrax toxin assembly and transport. PA (denim) is proteolytically nicked and assembles with LF (red-violet) and forms $PA_8LF_4$ and $PA_7LF_3$ prechannel complexes (36, 37) (based on 3KWV (34) and ITZO (38), respectively). Prechannel complexes bind cellular receptors (gold; 1T6B (39) and 1TZN (40)) triggering endocytosis; acidic pH conditions in the endosome induce PA to form a transmembrane channel (35, 41, 42) (atomic model 1V36 (43)); the pH gradient that develops across the endosomal membrane destabilizes LF (44) and drives LF unfolding (21, 45) and translocation (18, 21) through the PA channel.

represents an attractive model system for probing the molecular mechanism of protein translocation across membrane bilayers for a variety of reasons.

First, for a membrane-protein system, structural studies using X-ray crystallography are tractable, because the translocase also exists in a soluble state (34, 36, 38, 46). The structures of the three anthrax toxin components are known (Fig. 1.3A). The protective antigen (PA) component, which forms the translocase channel, is secreted as an 83 kDa proprotein (46). The other two components that are transported by PA are ~90 kDa enzymes, lethal factor (LF; a protease that targets mitogen-activated protein kinase kinases) (32, 47), and edema factor (EF; a calmodulin-activated adenylyl cyclase) (48–50). In order for the toxin to function, PA must first be proteolytically nicked by a furin-family protease (51, 52), releasing an amino-terminal 20 kDa fragment, $PA_{20}$ (Fig. 1.3B). The resulting 63 kDa fragment can assemble into an active holotoxin complex comprised of multiple copies of LF and EF bound to a ring-shaped PA oligomer (Fig. 1.3B). There are two known PA "prechannel" oligomeric architectures, a homoheptamer ($PA_7$) (38, 46, 53) and a homooctamer ($PA_8$) (34, 36, 54). The assembly and binding interactions between the PA oligomer and its substrates are well characterized (34, 55–60). Currently only one crystal structure is known of a prechannel core complex, $PA_8(LF_N)_4$, where $LF_N$ is the PA-binding domain of LF (Fig. 1.3A) (34). These prechannel complexes either assemble on cells by binding a specific protein receptor (61, 62) or localize to cells after assembling in the plasma (54). Internalization and subsequent acidification within the endosomal compartment converts PA prechannels into membrane-embedded channels (41, 63, 64), which are strongly cation-selective (42). The ΔpH created between the endosome and the cytosol drives translocation of LF and EF into the cytoplasm where they can affect their targets (18).

Another useful feature of anthrax toxin is that its three protein components can be expressed recombinantly and studied independently. It is not necessary to even work with the full-length, catalytically-active versions of LF and EF, as their amino-terminal PA-binding domains, $LF_N$ and $EF_N$, are stable and sufficient for translocation. PA83 can be cleaved with trypsin or furin and oligomerized *in vitro* (36). The stable, soluble nature of the prechannel state means the oligomer can be stored and used when convenient.

Because the protein translocase and the individual substrates can be readily reconstituted from purified proteins, they are ideal for studies using planar bilayer electrophysiology at the ensemble (18, 21, 34, 45, 65–73) and single-molecule level (21, 36, 45, 66, 74). A schematic of the planar lipid bilayer setup and a typical trace of an ensemble translocation are depicted in Fig. 1.4. Two chambers, cis and trans, are connected by a 50-200-μm aperture (Fig. 1.4A). Ag/AgCl electrodes bathed in saturated 3 M KCl are linked to the chambers via 3 M KCl-agar salt bridges, and currents across the chambers are recorded with an Axoclamp 200B amplifier. Using a lipid solution to paint a bilayer over the aperture stops the flow of ions (Fig. 1.4B(i)). Addition of PA to the cis chamber results in spontaneous insertion of the channels into the bilayer (Fig. 1.4A). PA's cation-selective nature allows only positively charged ions to pass through, so there is an increase in current proportional to the number of channels that insert into the bilayer (Fig. 1.4B(ii)). When substrate is added to the cis chamber, it stably blocks the channels, reducing the current, but it does not begin to translocate until a driving force is applied (Fig. 1.4B(iii)). When a driving force is applied, either by increasing the $\Delta\psi$ or changing the pH of one of the chambers to create a ΔpH, translocation is initiated (Fig. 1.4B(iv)). Successful translocation of the substrate through PA leaves the channel open and can be measured by the return of current. In this way, both the efficiency (the ratio of the current after translocation to the current prior to addition of the substrate) and rate (the time required for half of the maximum post-translocation current to

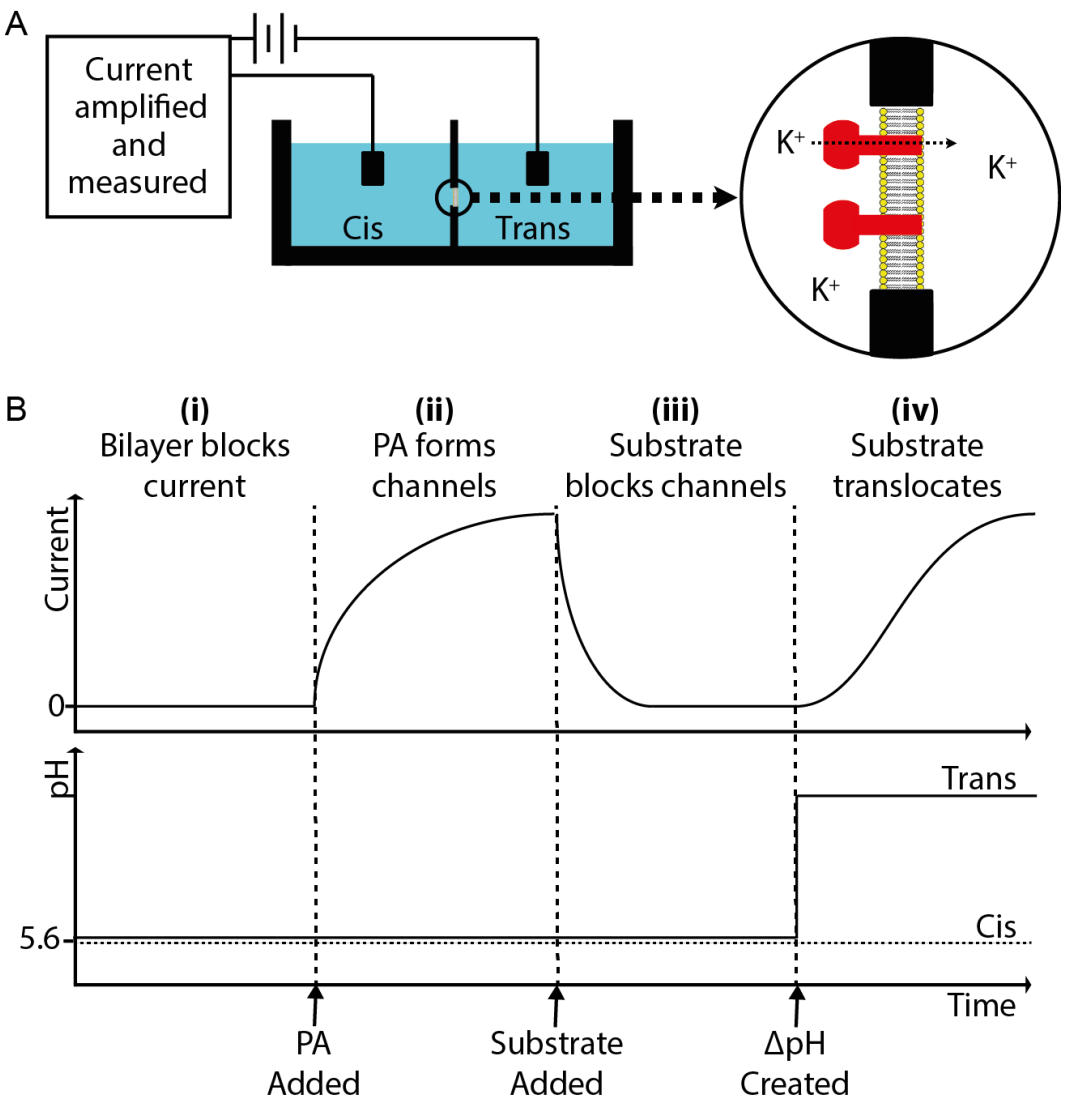


**Figure 1.4. Planar lipid bilayer electrophysiology as a tool for studying anthrax toxin translocation.** (A) A schematic of the instrument set up. Two chambers, called cis (where channel and substrate are added) and *trans*, are connected through an aperture, ranging from 50 to 200 μm. Electrodes linked to each chamber allow the current to be amplified and measured and provide for the application of a $\Delta\psi$. A lipid bilayer is painted across the aperture. Given proper pH conditions, PA channels will spontaneously insert into the bilayer, allowing the flow of positive ions. (B) Model recording of a translocation. (i) When a lipid bilayer is painted across the aperture separating the cis and trans chambers, no current can flow. (ii) The addition of PA channels and resulting flow of cations lead to an increase in current. (iii) Substrate will stably block channel conductance but will not begin translocating. (iv) Creation of a $\Delta$pH drives translocation, clearing the PA channels and allowing ions to flow again. An increase in $\Delta\psi$ may also function as a driving force. The rate and extent of translocation can be inferred by the increase in current.

return, $t_{1/2}$) can be measured. Single-channel studies can be done analogously, with channel insertion, blocking, and translocation appearing as distinct steps in current rather that smooth curves (21, 36, 45, 66, 74). The ability to separate the various steps in translocation makes it possible to assay oligomerization, prechannel stability, membrane insertion, substrate blocking, and movement through the channel individually (18, 21, 34, 45, 65–73, 75). Critical to this electrophysiological approach (also used in other systems (76–83)) is the ability to precisely control the driving force and solution conditions on either side of the membrane.

### 1.4 ΔpH-driven protein translocation

The acidic environment of the endosome is required for the action of the toxin (84), and it generates a proton motive force (PMF) comprised of both a chemical potential (ΔpH) and an electrical potential (Δψ) capable of driving LF/EF unfolding (21, 44, 45) and translocation (18, 21, 44, 45, 67). Translocation initiates from the amino terminus of LF (85), albeit a cationic unstructured amino terminus is sufficient for initiating translocation (34, 86). The PA translocase machine is powered by the ΔpH component of the PMF (18, 21). While a pure Δψ can drive the translocation of LF's amino-terminal domain, $LF_N$ (65), the ΔpH is required for the translocation of the full-length substrates, LF and EF (18). In one of my early experiments (see chapter 2), I show that the ΔpH component of the PMF is sufficient to unfold and translocate proteins (21).

A charge-state BR model was initially proposed for the molecular mechanism of ΔpH-driven translocation (Fig. 1.1A) (18). The model is based on the fact that the PA channel itself is cation-selective (or anion-repulsive) (42), and yet LF and EF are anionic substrates with isoelectric points of 5.4 and 6.6, respectively (more importantly, their amino-terminal domains, $LF_N$ and $EF_N$, have p*I* values of 4.9). Krantz *et al.* proposed that this paradox is resolved if acidic residues can be protonated during their translocation through the PA channel, thereby allowing their anionic charges to be neutralized (18). The ΔpH naturally favors substrate protonation on the endosomal side of the membrane, where the pH is ~5.5, but once the substrate reaches the higher pH side of the membrane (neutral cytosolic pH), these acidic residues would spontaneously deprotonate. Thus, while these protonated acidic residues could pass through the channel's cation-selective filter, they would be trapped on the opposite side of the membrane upon substrate deprotonation. In this mechanism, substrate motion is largely explained by Brownian motion, and the ratchet is an electrostatic trap created via cycles of acidic residue protonation and deprotonation on either side of the channel's charge-selectivity filter.

Several recent studies support the charge-state BR model of translocation. In one study, it was shown that sulfate groups attached to $LF_N$ via cysteine linkage inhibited Δψ-mediated translocation (71). Since the sulfate moiety essentially cannot be protonated under the experimental conditions, the authors concluded that only titratable negative charges could pass through the cation-selective channel. Another study reached a similar conclusion by incorporating cysteic acid residues in a semisynthetic $LF_N$ construct (69). Furthermore, semisynthetic $LF_N$ constructs lacking titratable acidic residues in the amino-terminal presequence display significant translocation defects (69).

In chapter 2, I probe the role of both positively- and negatively-charged residues in $LF_N$'s presequence. Interestingly, substrates lacking acidic residues could be translocated by a Δψ alone, yet they do not display ΔpH-dependent translocation. This evidence strongly supports the model that the ΔpH drives translocation by protonating acidic residues and is consistent with the charge-state BR model. Furthermore, I demonstrate that basic residues likely act to chaperone deprotonated acidic residues through the channel. That is, segments of sequence containing high

densities of negative charge will not enter the PA channel as efficiently as those containing positively-charged residues nearby. I observe that the ΔpH plays a role not only in translocation but also in the unfolding of the substrate. The most ΔpH-dependent step of translocation is associated with the protein folding barrier (45), which is consistent with the observation that a ΔpH alone is sufficient to unfold and translocate a fully folded substrate. Furthermore, acidic residues in $LF_N$ are naturally concentrated in an optimal position immediately before the substrate's folded domain and separating this region of optimal charge density from the folded domain greatly inhibits translocation. Taken together, these results imply a model for ΔpH-dependent unfolding. As the substrate presequence is translocated and captured on the high pH side of the membrane, the remainder of the polypeptide is caught in an extended state with more limited conformational options. This results in entropic tension that is only relieved via substrate unfolding.

### 1.5 The PA channel and its role in ΔpH-driven protein translocation

While acidic residues in the substrate make up the "teeth" in the metaphorical ratchet, the system is only complete if there is a "pawl", the component which allows the teeth to move in only one direction. With the teeth being a part of the substrate, the pawl must be a part of the channel. Using electron microscopy (EM), the structure of the $PA_7$ channel has been recently imaged (Fig. 1.3A) (35). The PA channel is mushroom-shaped and approximately 170 Å tall × 125 Å wide at its maximum dimensions. The wider, cap-shaped part of the structure is about 70 Å long and likely contains the LF/EF binding sites. Beneath the cap is a 100 Å long stem, which is likely a 14-strand β-barrel structure; the stem ultimately inserts into and spans the membrane bilayer (87). Earlier electrophysiology studies suggested the stem is an extended β-barrel structure (88). From basic modeling studies (44), the β-barrel stem is likely only able to accommodate structures as wide as an α helix (10–15 Å wide), confirming that LF and EF must unfold during translocation but leaving the possibility for secondary structure to be maintained.

The PA translocase channel can be divided into three sections (Fig. 1.5): the substrate docking surface in the cap (called the α-clamp site), a critical hydrophobic constriction point about one-third of the distance into the translocase (called the ϕ-clamp site), and the highly charged solvophilic β-barrel stem portion, which comprises the bottom two-thirds of the translocase. Interestingly, there appear to be separate polypeptide interaction sites, or clamps, associated with these different points in the PA translocase channel. These clamps also allow the channel to interact with the substrate nonspecifically. We consider the notion of a clamp to more closely resemble a dynamic binding site for polypeptide, where, for example, a clamp site can be modulated from a higher-affinity binding mode to a lower-affinity binding mode. This dynamic is critical to allowing the translocase to both favor unfolding and translocation while preventing tight binding interactions from occurring that would otherwise impede transport.

The α clamp was defined recently when the structure of $PA_8(LF_N)_4$ was reported, providing a molecular description of the interaction between the PA oligomer and $LF_N$ (34). LF binds to PA in two distinct subsites: a carboxy-terminal subsite comprised of specific interactions and an unconventional binding cleft formed at the interface of adjacent PA subunits, termed the "α clamp." $LF_N$ α1 helix and β1 sheet unfold from their native conformation and dock into the α-clamp subsite on the surface of the $PA_8$ oligomer (Figs. 1.5 and 1.6A). The manner in which the α clamp interacts with $LF_N$'s α1/β1 explains why the site is well equipped to recognize protein sequence in a nonspecific manner (34). Structurally, PA's EF-hand-type twin $Ca^{2+}$-binding sites frame the cleft and provide a structural scaffold consistent with how calmodulin

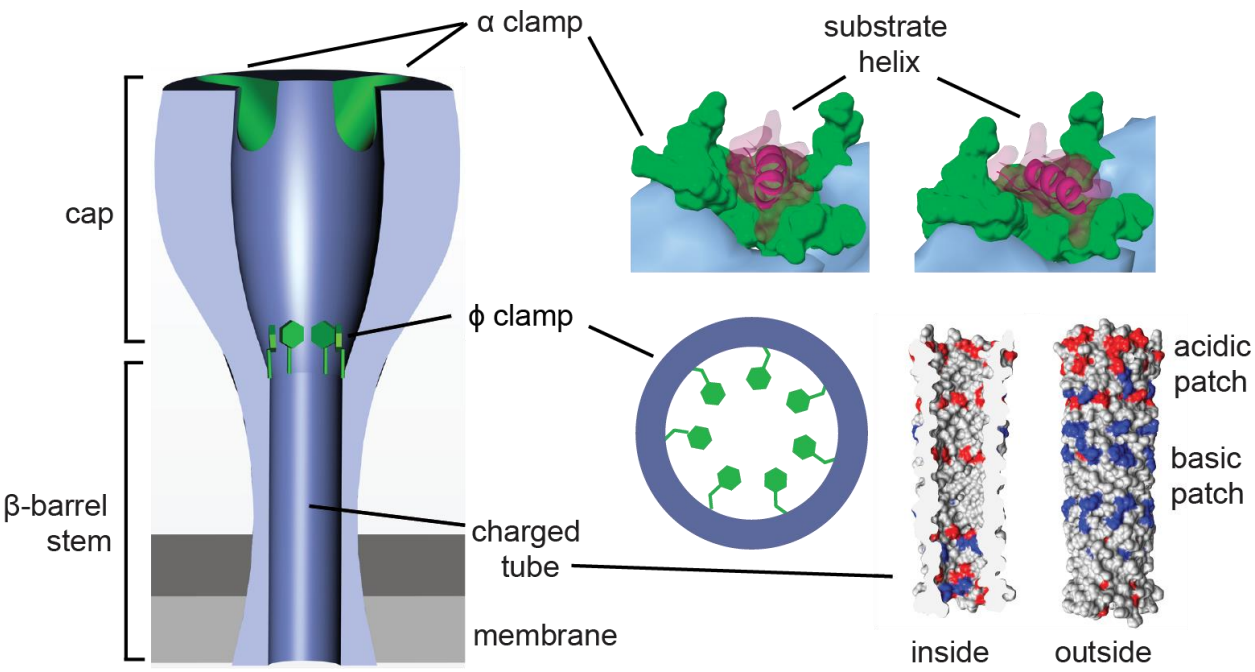


**Figure 1.5. Peptide clamps in the PA channel.** The PA channel (denim) contains clamping sites (green) and a β-barrel tube with positively (blue) and negatively (red) charged patches. Partially unfolded substrates (red-violet) first bind to the α-clamp site (34). Subsequently, the peptide threads into the ϕ clamp, which is comprised of a ring of F427 residues (67). Finally, the substrate polypeptide chain encounters the β-barrel tube (88, 89), which may act as a putative charge clamp, attracting cationic sequence and preventing the retrotranslocation of deprotonated acidic residues.

complexes bind to peptide helix substrates (Fig. 1.6B) (90, 91). Extensive mutagenesis in LF failed to identify specific side chain interactions in the α clamp. In fact, the α1 sequence in LF can be essentially replaced with other sequences in LF and EF and still maintain essentially wild-type binding and translocation activity (Fig. 1.6C) (34). Thus, the α-clamp site can garner significant nonspecific polypeptide binding activity using a general shape-complementarity binding mechanism, leading to the exciting prospect that helical structure in the substrate represents a key handle for the translocase to grip its substrate during translocation. The α clamp does not play a direct role in the charge-state BR aspect of translocation, but cooperates with the system is a significant way described below.

The next major site, the ϕ clamp, is formed by a ring of 7 or 8 phenylalanines (the conserved residue F427 from each PA monomer) and is critical for translocase function (67). F427 is the lone hydrophobic residue in an otherwise hydrophilic PA loop, $N^{422}$AQDD**F**SSTP. As most of the channel-lining residues are hydrophilic and/or anionic (88, 92), this aromatic site is unusual in terms of its chemistry and prominence in the interior of the PA channel. The ϕ clamp has broad substrate specificity with preference for hydrophobic aromatic compounds (67). Mutations at F427 manifest in increased substrate diffusion, backsliding, and retrotranslocation. The retrotranslocation phenotype of these ϕ-clamp mutants likely explains their >1000-fold losses in translocation activity (67). Analogous ϕ-clamp sites have been identified in numerous protein translocases, including soluble ones that hydrolyze ATP (93), revealing the general importance of the site to the mechanism of translocation. While the ratcheting mechanism of the ϕ clamp is doubtlessly of great importance for translocation and likely acts in concert with the charge-state BR (see model below), it is not strictly part of the BR and will not be addressed in this work.

Rather, my first studies of the channel dealt with its long, narrow, and highly-charged stem, an excellent candidate for the anion-exclusion site critical for the charge-state BR model of translocation. While an atomic-resolution structure of the PA channel is currently unavailable, a β-barrel model (43) has been proposed (88, 89). EM studies also show a tube-like stem consistent with a β-barrel architecture (35, 87). The β barrel's inner diameter is no wider than the width of an α helix (44) and may stabilize helical structure due to favorable van der Waals contacts and backbone desolvation. A number of charged residues populate both the inside and outside of the β barrel. These charges create patches of alternating electrostatic potential within the barrel (Fig. 1.5). The precise role of the charged groups in the barrel is unknown; however, the charge-state BR mechanism (Fig. 1.1A) proposes that differential electrostatic repulsion between the channel and substrate polypeptide is critical to the overall mechanism of ΔpH-driven translocation.

In chapter 3, I present studies of $LF_N/EF_N$ chimeras conducted by postdoctoral fellow Sarah Wynia Smith that confirm my observations in chapter 2 and, counter intuitively, demonstrate the importance of acidic residues in cis-positive Δψ-driven translocation, suggesting that the electrostatics of the channel override the applied Δψ. I go on to support this idea by neutralizing charged residues in PA's β-barrel stem and show that a ring of acidic residues along the top of the stem are important for the channel's cation selectivity. Loss of this ion selectivity damages the channel's ability to bind and translocate its substrate. This confirms the prediction of the charge-state BR model that the channel's repulsion of negative charges in the substrate prevents retrotranslocation and thereby is important for efficient productive translocation.

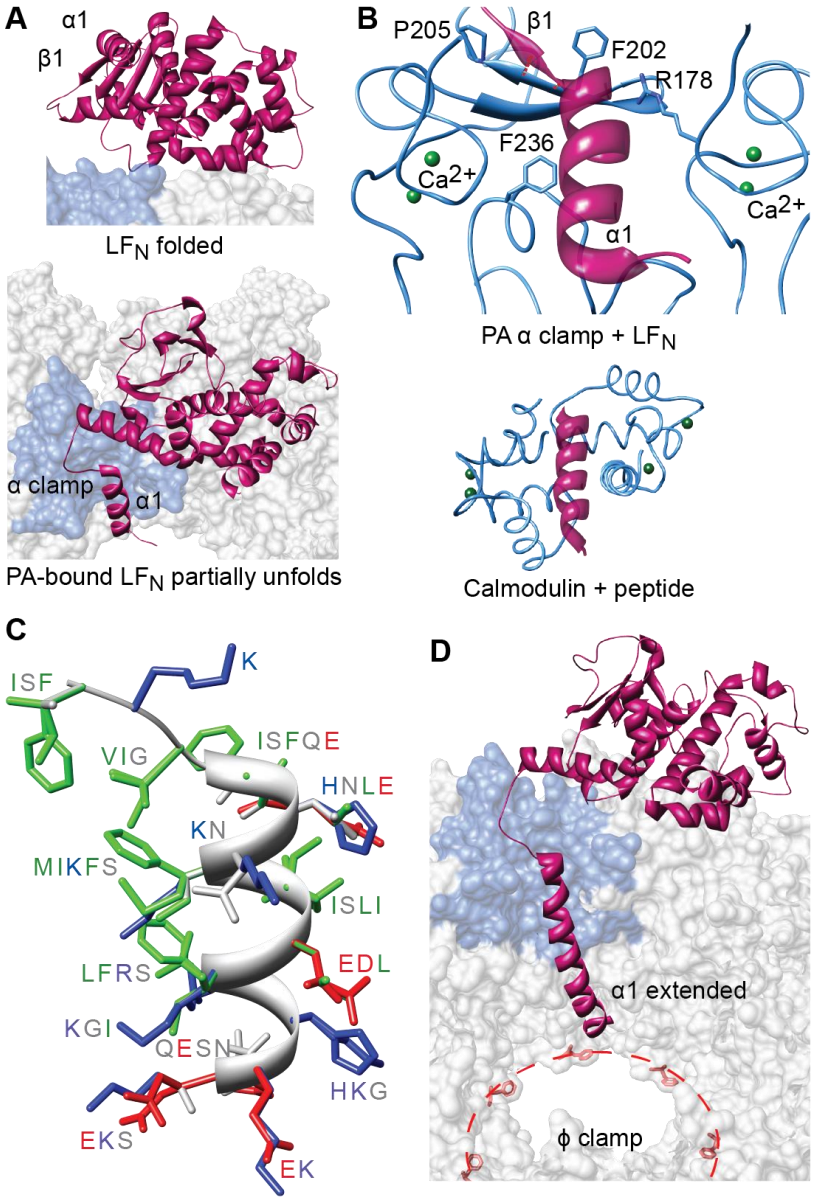


**Figure 1.6. The α clamp.** (A) $LF_N$ (red-violet) changes conformations from (top) a folded (1J7N) (32) to (bottom) a partially unfolded state (3KWV) (34), such that LF α1/β1 unfurl and bind into PA's α clamp (denim surface). (B) Top, a detailed view of the α clamp (denim ribbon) in complex with LF's α1 (red-violet ribbon) indicating the structural calcium ions (green) scaffolding the site. Bottom, calmodulin in complex with a peptide helix (1CDM) (90), where the latter is colored analogously to the α-clamp structure. (C) The α clamp can bind nonspecifically to and translocate a variety of sequences illustrated in a helical structural alignment, which is colored by residue chemistry: basic (blue), acidic (red), polar (gray), and hydrophobic (green) (34). (D) Considering the Zimm-Bragg formalism (94), the α clamp may act as an α-helix nucleation site (with additional residues modeled beyond the α1 helix). Elongated helical structure can then be fed into the ϕ clamp (red dotted line and F427 residues).

### 1.6 Ratchet-driven unfolding of LF during translocation

Since the mechanism described in our charge-state BR model can only bias a molecule's inherent thermal energy and does not involve the application of an external force, it may be difficult to imagine how such a system could lead to substrate unfolding. However, the thermal energy inherent in a polypeptide chain is significant. From statistical mechanics, an unstructured protein has many more degrees of freedom than a simple, rigid Brownian particle. A peptide backbone in the unfolded state can be assumed to have 3 degrees of freedom per amino acid, considering relevant ϕ/ψ-angle conformations. The multiplicity, $W$, of this peptide system, given as $W = 3^N$, translates into a configurational entropy, $S_{config}$, of $Nk_B \ln 3$. Therefore, a polypeptide with $N = 50$ residues has a thermal energy of ~33 kcal mol$^{-1}$ (~140 kJ mol$^{-1}$) at 25°C. So the energy required for unfolding is indeed present if it can be applied toward that purpose.

We envision that the necessary force is generated through the release of unfolded-state conformational entropy, $\Delta S_{config}$, upon substrate extrusion from the channel. If the substrate is constrained in an α-helical conformation during translocation and released into an unstructured coil upon extrusion from the channel, a great deal of force ($F$) can be generated. This is given by $F = T\Delta S_{config}/d$, where $T$ is the temperature and $d$ is the difference in length of peptide undergoing a α-helix-to-coil transition (2.2 Å per residue). A three-fold increase in the number of available conformations ($\Delta S_{config} = k_B \ln 3$) over this distance gives a force of $F = k_B T \ln 3 / 2.2 \times 10^{-10}$ m. This works out to ~20 pN at physiologically relevant temperatures, enough force to greatly accelerate protein unfolding reactions (95–98). The biasing action of electrostatic forces in the channel will ensure that the force is applied productively on the folded domains carboxy-terminal to the translocating chain. While this calculation is an upper-limit as there are undoubtedly entropic losses to factor in, these forces are significant and higher than those estimated for BR models, where the substrate is a simple Brownian particle with only three translational degrees of freedom. There is evidence suggesting that the substrate translocates in a compressed state. One lab used streptavidin-biotin intermediate-capture approach to determine the minimal length of peptide required to span PA's β-barrel stem (70). A 33-residue probe, which would have to exist as an extended chain in order span the 100-Å stem, could only be captured on time scales on the order of 1000 s. As a 263-residue substrate can translocate in less than 10 s, the fully extended state must be considered irrelevant to the process. Rather, the translocating chain must be in a more compact state >99% of the time. The absolute degree of helicity is not critical, only that the structure of polypeptide in the channel be more compact than that immediately following extrusion from the channel.

This leaves the question of how the translocating substrate can be loaded into the channel in the entropically-unfavorable compressed state. According to the Zimm–Bragg model for helix formation, α-helix stability, $K_N$, of an $N$-residue peptide is given as $K_N = \sigma s^N$, where $s$ is [helix]/[coil] (94). The initiation of α-helical structure, which is described by the factor σ, is the limiting step. The equilibrium constant, $K_N$, increases with each additional residue added to the nascent helix. Thus helix formation is a nucleation process, where the equilibrium stability of a helical substrate would be greatly enhanced by the presence of a structure that binds sequences non-specifically according to shape complementarity, in the spirit of PA's α clamp (Fig. 1.6D). Helical structure has an additional advantage of possessing reduced conformational entropy relative to unstructured peptide, minimizing the potential for dissipative losses and allowing for a more efficient utilization of the available electrical/chemical free energy source. We conclude that helix-nucleation machinery in translocases, embodied by the α clamp in PA, if oriented

properly along the axis of the translocase, would provide a significant kinetic benefit to the translocation mechanism.

In chapter 4, I build off the work of Feld, Thoren, et al. (34), and examine the evidence that the α clamp plays a role in nucleating helix formation in the substrate. I show that PA mutants with occluded α clamps (rendering them unable to nucleate helix formation) are impaired for substrate translocation. Further, the driving-force dependence of translocation is altered by occlusion of the α clamp in such a manner that suggests the a previously identified (18, 21, 45) yet uncharacterized energy barrier has been raised. We propose that helix nucleation induced by the α clamp serves to lower this barrier in WT channels. Disrupting the interaction from the other side (replacing $LF_N$'s α1/β1 region with a polyproline helix) similarly impairs translocation. Finally, I look at the system's ability to translocate substrates with a mix of L- and D-amino acids. Such constructs would be unable to form the compact structures our model indicates as necessary. Whether the D-amino acids are introduced through a chemical racemization process applied to WT $LF_N$ or by designing short peptides with alternating chiralities, the result is severely inhibited translocation. Interestingly, these substrates require no unfolding of tertiary structure prior to translocating, suggesting that the channel forces a preference for translocation as an α helix even in the absence of an unfolding step.

### 1.7 A model for substrate unfolding and translocation in anthrax toxin

In this work, I present evidence for a charge-state BR based on the role of differential protonation of acidic residues in the substrate and need for acidic residues in the channel to repel the deprotonated substrate and prevent retrotranslocation. I further show how linking this ratchet with helical translocation (nucleated by the α clamp) can produce a directed helix-to-coil transition that can generate the force necessary to unfold the substrate. Combined with the previously described ratcheting role of the ϕ clamp (67), an intricate model for anthrax translocation emerges (Fig. 1.7):

- *Step 1.* Anionic/deprotonated polypeptide loads into the upper vestibule of the PA channel by means of Brownian motion, binding into the α-clamp site as helix. The α clamp acts as a helix nucleating binding site, orienting, and feeding the growing helical chain toward the channel lumen.
- *Step 2.* The ϕ clamp switches to a *closed* position, tightly gripping the substrate polypeptide. Binding of unfolded chain at the ϕ clamp prevents backsliding. The ϕ clamp impedes $H^+$ flow, reducing $[H^+]$ below the ϕ clamp. Acidic substrate residues above the clamp are subsequently protonated, and consequently less anionic in charge. The stability of the interaction at the ϕ clamp site may be further strengthened by the reduction in negative charge of the substrate.
- *Step 3.* While the substrate polypeptide is tightly engaged at the ϕ clamp, the amino-terminal end of the peptide eventually begins to extrude from the end of the channel, where conformational space of the chain is less restricted. Translocation is thus thermodynamically favorable in the direction out of the channel due to the gain in $T\Delta S_{config}$. Importantly, the peptide can bypass the anionic charge repulsion site in the channel, since the peptide is now protonated at its Asp and Glu sites.

*Step 4.* The ΔpH at the ϕ clamp weakens as the substrate in the β barrel loses structure, solvent penetrates up the barrel, and $H^+$ dissipate out of the channel. The ϕ clamp switches to the *open* state, releasing bound peptide. Due to the charge selectivity of the channel, the peptide may only proceed through the channel until peptide and channel are no longer

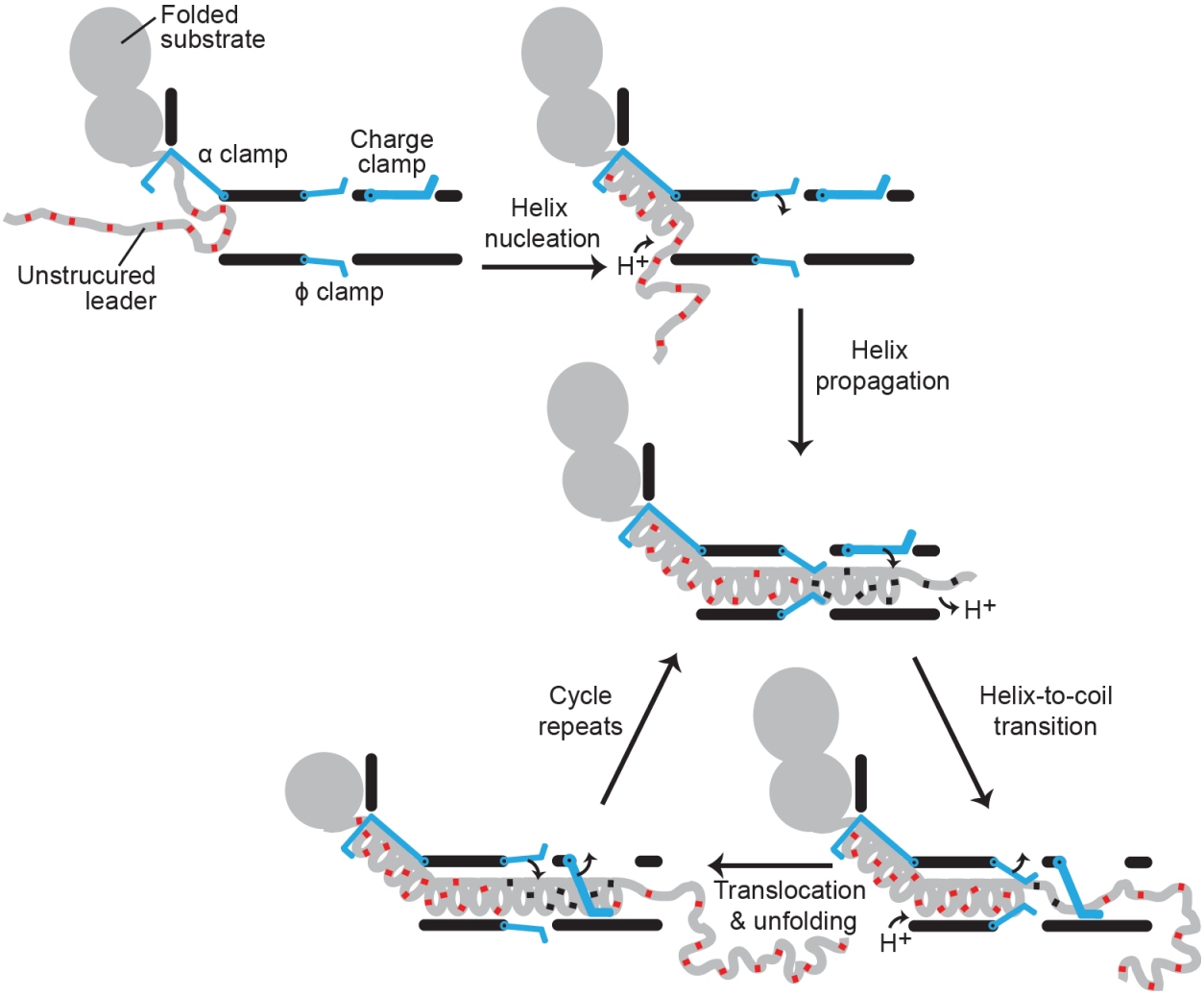


**Figure 1.7. Translocation by a proton-driven engine.** Folded substrate and unstructured leader sequence (gray) bind and dock to the PA channel. The α clamp nucleates and subsequently propagates helix formation. The ϕ clamp engages the compact translocating chain; deprotonated acidic residues (red) are then protonated (black). An increase in $T\Delta S_{config}$ outside the channel favors the transition of helix to unstructured random coil; acidic residues deprotonate in the higher pH of the cytosol. The charge clamp engages, permitting the passage of protonated acidic residues while preventing the retrotranslocation of deprotonated ones. Ungating of the ϕ clamp allows the chain to translocate, while the α clamp continues to stabilize and template unfolded polypeptide into helix. The cycle repeats until the substrate is fully unfolded and translocated.

electrostatically compatible. At this point, translocation pauses until peptide recompresses into helix, the ϕ-clamp site can reset, and the deprotonated section of chain is protonated. The cycle will repeat in a fashion analogous to the thermodynamic cycles of a heat engine.

I expect that many features of this model will be conserved across a wide variety of transmembrane and cytosolic protein transporters.

# Chapter 2

## Charge requirements for proton gradient-driven translocation of anthrax toxin[2]

### 2.1 Abstract

Anthrax lethal toxin is used as a model system to study protein translocation. The toxin is composed of a translocase channel, called protective antigen (PA), and an enzyme, called lethal factor (LF). A proton gradient (ΔpH) can drive LF unfolding and translocation through PA channels; however, the mechanism of ΔpH-mediated force generation, substrate unfolding, and the establishment of directionality are poorly understood. One recent hypothesis suggests that the ΔpH may act through changes in the protonation state of residues in the substrate. Here we report the charge requirements of LF's amino terminal binding domain ($LF_N$) using planar lipid bilayer electrophysiology. We find that acidic residues are required in $LF_N$ to utilize a proton gradient for translocation. Constructs lacking negative charges in $LF_N$'s unstructured presequence translocate independently of the ΔpH driving force. Acidic residues markedly increase the rate of ΔpH-driven translocation, and the presequence is optimized in its natural acidic-residue content for efficient ΔpH-driven unfolding and translocation. We discuss a ΔpH-driven charge-state Brownian-ratchet mechanism for translocation, where glutamic and aspartic acid residues in the substrate are the "molecular teeth" of the ratchet. Our Brownian-ratchet model includes a mechanism for unfolding and a novel role for positive charges, which we propose chaperone negative charges through the PA channel during ΔpH translocation.

### 2.2 Introduction

Transmembrane protein translocation (8, 18, 67, 99) and intracellular protein degradation (100) are essential processes that allow the cell to traffic protein, form new organelles, maintain protein quality control, and regulate the cell cycle. As folded proteins are thermodynamically stable under typical cellular conditions, protein translocation and degradation often require complex, energy-consuming molecular machines to catalyze the required series of unfolding and translocation steps in these pathways. Some of these protein transporters and degradation machines mechanically unfold their protein substrates via an ATP-dependent motor. However, a transmembrane proton gradient (ΔpH) may also be used to generate an appropriate driving force that can unfold (45) and translocate a substrate protein across a membrane (18).

A transmembrane proton gradient forms a proton motive force (PMF) comprised of two different types of available free energy: an electrical free energy ($\Delta G^{\Delta\psi}$) dependent on the membrane potential ($\Delta\psi$) and a chemical potential ($\Delta\mu^{\Delta pH}$) dependent on the proton concentration gradient (ΔpH). The former $\Delta G^{\Delta\psi}$ can be derived from the charge of the translocating protein, $z$, and Faraday's constant, $F$, according to the relation, $\Delta G^{\Delta\psi} = \Delta G^{\Delta\psi^\circ} + zF\Delta\psi$ (18, 45, 101). The latter $\Delta\mu^{\Delta pH}$ is expressed as a chemical potential, $\Delta\mu^{\Delta pH} = \Delta\mu^{\Delta pH^\circ} + 2.3\,nRT\,\Delta pH$ (where $n$ is the number of protons involved in the reaction, and $R$ and $T$ are the gas constant and temperature, respectively) (18, 45). Either of these energies can develop substantial force under physiological conditions, on the order of tens of pN (45, 102), and is sufficient to unfold a protein during translocation (45, 95). The molecular mechanism of ΔpH-driven

[2] This chapter is based on previously published work, with permission from co-authors (21).

translocation, in particular the enforcement of directionality and substrate unfolding, is not well understood.

Anthrax toxin (30) is a useful model system (8) to study ΔpH-driven protein unfolding (44) and translocation (18, 36, 45, 67). The toxin is a key virulence factor secreted by *Bacillus anthracis* (the causative agent of anthrax). It is comprised of three proteins: the translocase component, protective antigen (PA, 83 kDa); and two enzyme components, lethal factor (LF, 91 kDa) and edema factor (EF, 89 kDa). These secreted proteins must assemble into holotoxin complexes either on the surface of cells or in the blood serum (54). In each assembly mechanism, PA is initially cleaved by a protease, allowing a small 20-kDa portion to dissociate. The remaining 63-kDa PA fragment can then self-assemble into either a ring-shaped homoheptamer, $PA_7$ (38, 46, 53), or homooctamer, $PA_8$ (36, 67). PA oligomers then bind to LF and EF and form toxic complexes. Binding occurs at the top of the prechannel and in a deep amphipathic cleft, termed the α clamp, which also aids in unfolding the substrate (34).

The complex is then endocytosed by the target host cell, and subsequent acidification of the endosome causes the PA oligomer to convert into a transmembrane channel. Moreover, the ΔpH generated by the acidification of the endosome facilitates the unfolding (44) and translocation of LF and EF through PA (18) into the cytosol. ΔpH-driven translocation is aided by two catalytic active sites, the α clamp (34) and the ϕ clamp (67). The ϕ clamp is located inside the PA channel (18, 67). It is comprised of a narrowly-apposed ring of 7 or 8 Phe-427 residues, where one Phe is contributed by each PA monomer in the oligomer (67). The ϕ clamp prefers to make interactions with hydrophobic and aromatic groups, and its activity facilitates translocation (67) by helping to unfold the substrate (45). How the ϕ clamp stimulates (18) ΔpH-driven translocation is unclear.

While a ΔpH can drive translocation, the groups that are protonated and/or deprotonated in the substrate and/or channel are unknown. One clue is provided by the known cation selectivity of the PA channel (42). This selectivity is likely due to the ~100-150 glutamic and aspartic acid residues (depending upon PA's oligomeric state) known to be located inside the PA channel (67, 88). This cation selectivity is surprising given the acidic isoelectric points of both LF and EF (18). PA is a channel that only permits positive charges to pass but whose substrates are largely negative at neutral pH. Thus we previously proposed a ΔpH-driven charge-state Brownian ratchet translocation mechanism (18), which is based upon the theoretical work of Feynman (15), Oster (16), and Astumian (17). In this mechanism, the large available thermal energy of the translocating protein is biased in a useful and directed way by the ΔpH gradient, thereby supporting productive translocation.

Because the channel is cation-selective and, therefore, anion-repulsive, we expect that anionic residues within the portion of LF or EF passing through the channel must be protonated, thus making the translocating chain net cationic. The segment of LF within the channel is in fact net positive, as shown by Δψ-dependent measurements, which report positive *z* values (45). Once the Brownian thermal energy of the translocating protein causes the substrate to exit to the higher pH cytosolic side of the membrane, deprotonation to a more anionic state is favored. The resulting charge repulsion can then enforce directionality during translocation by preventing backward movement through the anionic channel. This process then repeats in multiple cycles toward completion. In this report, we investigate how ΔpH-driven translocation depends upon charged residues in the substrate.

## 2.3 Experimental Procedures

*Constructs and proteins*—Site-directed mutagenesis was performed using a Quikchange procedure. WT PA was expressed and purified as described (67). $PA_7$ oligomers were produced as described (36). $His_6$-$LF_N$ (residues 1-263 of LF) and mutants thereof were purified from overexpressing bacteria using standard $Ni^{2+}$-nitrilotriacetic acid affinity chromatography and Q-sepharose, anion-exchange chromatography, when required (GE Healthcare, USA) (18). When indicated, $His_6$ tags, which are an amino-terminal, 17-residue leader containing a hexahistidine sequence, MGSSHHHHHHSSGLVPR, were removed from $His_6$-$LF_N$ with bovine α-thrombin treatment (0.5 units/mg of protein) for 30 minutes at room temperature in 20 mM Tris, 150 mM NaCl, 2 mM $CaCl_2$ and 1 M glucose at pH 8.0. Note that after thrombin cleavage, an additional GSHM sequence is left at the amino terminus of all cleaved $LF_N$ constructs derived from the pET15b plasmid.

More complex synthetic $LF_N$ constructs were made using a three-step, gene-synthesis procedure, according to the following scheme:
HM$^{1}$AGGHGDVGM$^{10}$HVKEKEKNKD$^{20}$ENKRKDEERN$^{30}$KTQEEHLKEI$^{40}$MKHIVKIEVK$^{50}$
GEEAVKKEAA$^{60}$EKLLEKVPSD$^{70}$VLEMYKAIGG$^{80}$KIYIVD
The underlined pairs of amino acids on either end are encoded by the restriction sites, Nde I and a silent Sal I site (V84 and D85), respectively, which were used for cloning. Superscripted numbers indicate the numbering convention of LF residues in 1J7N (32). Overlapping oligonucleotides encoding the desired sequences with the amino-terminal replacement were synthesized (Elim Biopharmaceuticals, Inc., Hayward, CA) and amplified by two rounds of polymerase chain reaction (PCR). In Round I, 22 nM of nested oligonucleotides with consistent annealing temperatures of ~55 °C were amplified in a standard PCR reaction. In Round II, 1 μL of the PCR product made in Round I was amplified with the two outermost PCR primers (1 μM each) to make the synthetic double-stranded DNA fragment. These $LF_N$ synthetic DNA fragments were ligated via a 5' Nde I site and 3' Sal I site into the pET15b-$LF_N$(Sal I) construct, which contain an in-frame, silent Sal I restriction site in $LF_N$ at V84 and D85. The synthetic $LF_N$ constructs were overexpressed, purified, and their $His_6$ tags were subsequently removed as described above (when required).

*Electrophysiology*—Planar lipid bilayer currents were recorded using an Axopatch 200B amplifier (Molecular Devices Corp., Sunnyvale, CA) (36, 45). Membrane bilayers were painted onto a 100-μm aperture of a 1-mL polyethersulfone cup with 3% 1,2-diphytanoyl-*sn*-glycero-3-phosphocholine (Avanti Polar Lipids, Alabaster, AL) in neat *n*-decane. Cis (side to which the PA oligomer is added) and trans chambers were bathed in the indicated buffers as required. By convention, $\Delta\psi \equiv \psi_{cis} - \psi_{trans}$ ($\psi_{trans} \equiv 0$ V), and $\Delta pH \equiv pH_{trans} - pH_{cis}$.

*Translocation assays*—Translocation experiments were carried out as described (36, 45) generally using a universal pH bilayer buffer system (UBB: 10 mM oxalic acid, 10 mM phosphoric acid, 10 mM MES, 1 mM EDTA). However, for translocations requiring a pH > 7.5, we used an altered UBB (6 mM oxalic acid, 6 mM phosphoric acid, 6 mM MES, 6 mM boric acid, 6 mM TAPS, 1 mM EDTA), which is better at buffering in the 7.5-9 range. We found that these two types of buffers produced consistent translocation results. Generally, an additional 100 mM equivalent of KCl was added to the UBB and maintained symmetrically in most translocation experiments, excepting when translocation kinetics were monitored in the absence of a $\Delta\psi$. In that case, additional KCl was only added to the cis-side buffer. Our electrophysiology setup can maintain a $\Delta\psi$ of 0 mV even with the asymmetric ion concentrations across the membrane. The $pH_{cis}$ was generally 5.6 unless indicated otherwise.

In a typical experiment, we first inserted $PA_7$ channels into a planar bilayer by adding the prechannel oligomeric form to the cis side of the membrane under a small $\Delta\psi$ of 20 mV. Insertion was observed as an increase in current. Then WT or a mutant $LF_N$ was added also to the cis side of the membrane, and a decrease in current was observed. Excess $LF_N$ in the cis compartment was removed by perfusion, and the translocation process was initiated by either raising the $\Delta\psi$ or by creating a $\Delta$pH by raising the pH of the trans chamber. When translocation was initiated by raising the pH of the trans chamber, we set the time when 1% of maximum translocation occurs as t = 0 s in order to control for different mixing times.

An alternative method of creating a proton gradient was also used and provided similar results: the $pH_{cis}$ and $pH_{trans}$ buffer were preadjusted to their final conditions; after substrate binding to the channel was complete, then the cis buffer was perfused at ~0 mV at pH 5.6 to maintain the $\Delta$pH; then the voltage was raised to the final voltage upon the completion of perfusion (after ~30 s).

*Single-channel recordings*—Single-channel $LF_N$-docking experiments were performed as described (36, 45). Planar lipid bilayers were bathed in symmetric buffer (100 mM KCl, 1 mM EDTA, 10 mM succinic acid, pH 5.6). PA oligomer was applied directly to membranes at ~$10^{-16}$ M. Single channel insertion was observed by a discrete step in current under an applied voltage. Once a single channel inserted into the membrane, $LF_N$ was added to the cis side of the membrane at 200 pM. Blocking events were recorded at a constant $\Delta\psi$ of 20 mV for various $LF_N$ mutants with and without their $His_6$ tag. Data were acquired at 400 Hz using a low-pass filter of 200 Hz. For $LF_N$ that stably blocked a PA channel, the voltage was reversed to -80 mV after several minutes of conductance block in order to clear the substrate from the channel; this procedure verified that the channel is still present. Histograms of the current versus time data were fit to a two or three Gaussian function to obtain the relative percentages of time spent in the open, blocked, and partly-blocked states.

**2.4 Results**

*Chemical potential component of PMF is sufficient to drive translocation*—To study translocation, we use planar lipid bilayer electrophysiology (18, 45, 67) and a model substrate, $LF_N$, the amino-terminal, 263-residue, PA-binding domain of LF (57). In this assay, a planar lipid bilayer separates two aqueous chambers (called cis and trans). We first insert PA channels into a planar bilayer by adding the prechannel heptameric form of PA to the cis side of the membrane under a $\Delta\psi$ of 20 mV ($\Delta\psi \equiv \psi_{cis} - \psi_{trans}$, where $\psi_{trans} \equiv 0$ V). Channel formation, observed as an increase in current, then stabilizes, and either wild type (WT) $LF_N$ or an $LF_N$ mutant (MUT) is added also to the cis side of the membrane. Subsequently, an exponential decrease in current is observed as $LF_N$'s amino-terminal presequence inserts into the ion-conducting channel and blocks ion flow (85). Excess $LF_N$ in the cis compartment is removed by perfusion, and the translocation process is initiated by either changing the $\Delta\psi$ and/or $\Delta$pH. The ensuing current increase generally observed results from $LF_N$ translocation to the trans side of the membrane, as inferred by control experiments (18, 65). Thus two types of parameters can be obtained from translocation records: the time for half of the protein to translocate ($t_{1/2}$, measured in seconds) and the efficiency of translocation, which is equivalent to the fraction of $LF_N$ that successfully translocates.

While a $\Delta$pH can drive substrate translocation in the presence of a small, positive $\Delta\psi$ (18), it has not yet been shown if a $\Delta$pH alone is sufficient. To test this possibility, we set up the bilayer with a potassium chloride gradient; this procedure allows current to flow when $\Delta\psi$ is 0

mV. It should be noted that by design, the voltage-clamp amplifier maintains the system at 0 mV even under a ΔpH, which we define as $\Delta pH \equiv pH_{trans} - pH_{cis}$. At a Δψ of 0 mV and positive ΔpH values, we found that both the rate and efficiency of $LF_N$ translocation increase with larger positive ΔpH gradients (Fig. 2.1 inset). We estimate the activation energy of translocation ($\Delta G^{\ddagger}$) using the following relationship, $\Delta G^{\ddagger} = RT \ln(t_{1/2}/c)$, where $c$ is an arbitrary constant of 1 s. The ΔpH dependency of the $\Delta G^{\ddagger}$ for $LF_N$ translocation is biphasic with two different limiting slopes (Fig. 2.1). This dependency is similar to what has been observed under varying ΔpH conditions with a constant, nonzero Δψ (18, 45). The relationship fits to a two-barrier model,

$$\Delta G^{\ddagger}(\Delta pH) = RT\ln[\exp((\Delta G^{\ddagger\circ}_1 - 2.3n_1\Delta pH)/RT) + \exp((\Delta G^{\ddagger\circ}_2 - 2.3n_2\Delta pH)/RT)] \quad \text{(Eq. 1)}$$

where the indexed $\Delta G^{\ddagger\circ}$ and $n$ values are the activation energy and number of protons required to cross each barrier, respectively (45). The steeper ΔpH dependency of the $\Delta G^{\ddagger}$ on the left-hand side of the graph (less than 1.2 ΔpH, $n_1$ value of 4.0 (±0.5)) corresponds to a barrier limited by $LF_N$ unfolding, (45) whereas the shallower ΔpH dependency on the right-hand side (at higher ΔpH values, $n_2$ value of 0.22 (±0.03)) may correspond to a barrier limited by a yet uncharacterized translocation process (45) (Fig. 2.1). We conclude that the chemical potential component of the PMF is sufficient to drive $LF_N$ translocation.

*ΔpH-driven translocation depends on charged residues in $LF_N$*—The charge-state Brownian ratchet model (18) predicts that negatively-charged acidic residues in the substrate are critical to ΔpH-driven translocation. To test this model, we replaced the unstructured amino-terminal presequence of $LF_N$ (residues 1 to 27) with a randomized neutral/uncharged background comprised of Gly, Ser and Thr residues (Fig. 2.2A). We call this neutral mutant background $LF_N$ Syn°. This sequence is still polar and flexible but lacks charged residues. We found that $LF_N$ Syn° translocates slower than $LF_N$ WT under conditions of ΔpH = -0.6 to 1.0, $pH_{cis}$ = 5.6, and Δψ of 60 mV (Fig. 2.2B). Additionally, the construct possessed a significantly reduced ΔpH dependence ($LF_N$ WT: $n_1$ = 1.9 (±0.2); $LF_N$ Syn°: $n$ = 0.59 (±0.06); Fig. 2.2B). (It should be noted that we included the amino-terminal $His_6$ tag in these $LF_N$ constructs due to defects in channel blocking for certain constructs—an effect that we explore more fully below. To our knowledge, the $His_6$ tag does not obfuscate our results or interpretations, since we are looking at effects relative to similarly $His_6$-tagged control constructs.) Initially, we predicted that acidic residues alone would be sufficient to restore $LF_N$ Syn°'s defect in translocation rate and its weaker ΔpH dependence; however, we found this was not the case. $LF_N$ $Syn^-$, which has acidic residues restored at their WT positions, further reduced the rate and efficiency of translocation relative to $LF_N$ Syn° and had essentially no ΔpH dependence ($n$ value of 0.0 (±0.1);Fig. 2.2B). Restoring only basic residues in the $LF_N$ $Syn^+$ construct increased translocation rate relative to $LF_N$ Syn° but further reduced the ΔpH dependence ($n$ value of 0.14 (±0.07); Fig. 2.2B). Only when both the acidic and basic residues are restored in $LF_N$ $Syn^{\pm}$ do we see translocation that is identical to the WT substrate (Fig. 2.2B). Moreover, since $LF_N$ WT and $LF_N$ $Syn^{\pm}$ translocate with identical ΔpH dependencies, we can safely assume that the noncharged residues in the WT leader sequence are unimportant to ΔpH-driven translocation, as these have been substituted *en masse* with Gly, Ser and Thr residues in $LF_N$ $Syn^{\pm}$. Thus acidic residues alone in $LF_N$'s presequence are not sufficient for efficient ΔpH-driven translocation under these conditions; instead, a balance of acidic- and basic-charged residues is required.

To further explore the differences between the Syn constructs, we translocated them under more physiologically relevant conditions (ΔpH = 2, Δψ of 0 mV). Suspecting that

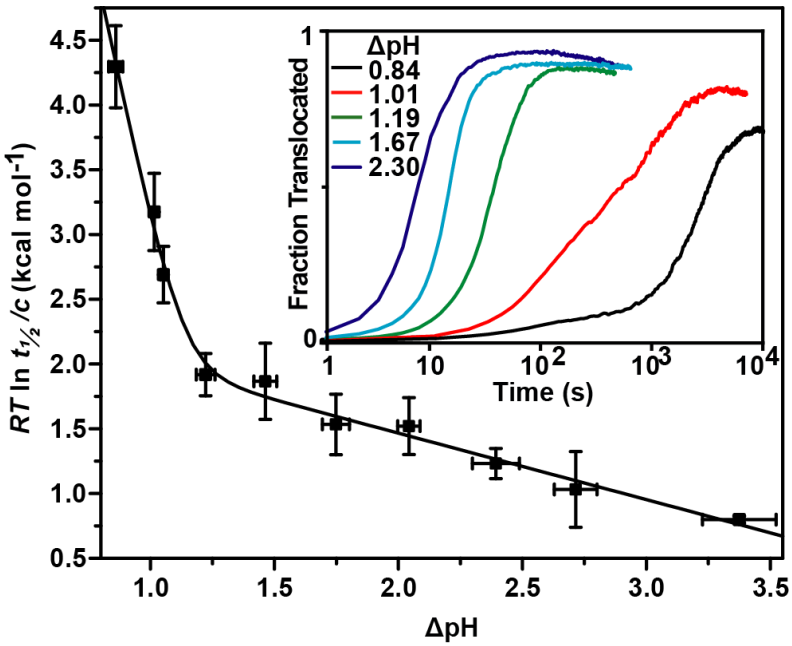


**Figure 2.1. The chemical potential of the PMF is sufficient to drive $LF_N$ translocation.** Analysis of $LF_N$ translocations driven purely by a ΔpH. Activation free energies ($\Delta G^{\ddagger}$; expressed as $RT\ln t_{1/2}/c$) for individual translocations of $His_6$-$LF_N$ are plotted against their respective ΔpH values and fit to Eq. 1. The fit parameters are: $\Delta G^{\ddagger\circ}{}_1$ = 12.2 (± 1.1) kcal $mol^{-1}$, $\Delta G^{\ddagger\circ}{}_2$ = 2.5 (± 0.2) kcal $mol^{-1}$, $n_1$ = 4.0 (± 0.5), and $n_2$ = 0.22 (± 0.03). (inset) representative $LF_N$ translocation records normalized as a fraction of the theoretical maximum of translocation under the following ΔpH values at a Δψ of 0 mV: 0.84 (black), 1.01 (red), 1.19 (green), 1.67 (blue), and 2.30 (purple). The universal bilayer buffer was consistently at a $pH_{cis}$ of 5.6, and $pH_{trans}$ was adjusted to form the indicated ΔpH values. To control for buffer mixing lag times, $t$ = 0 was set as the time when 1% of maximum translocation occurred. The error bars are the means ± S.D. ($n$ = 2–5).

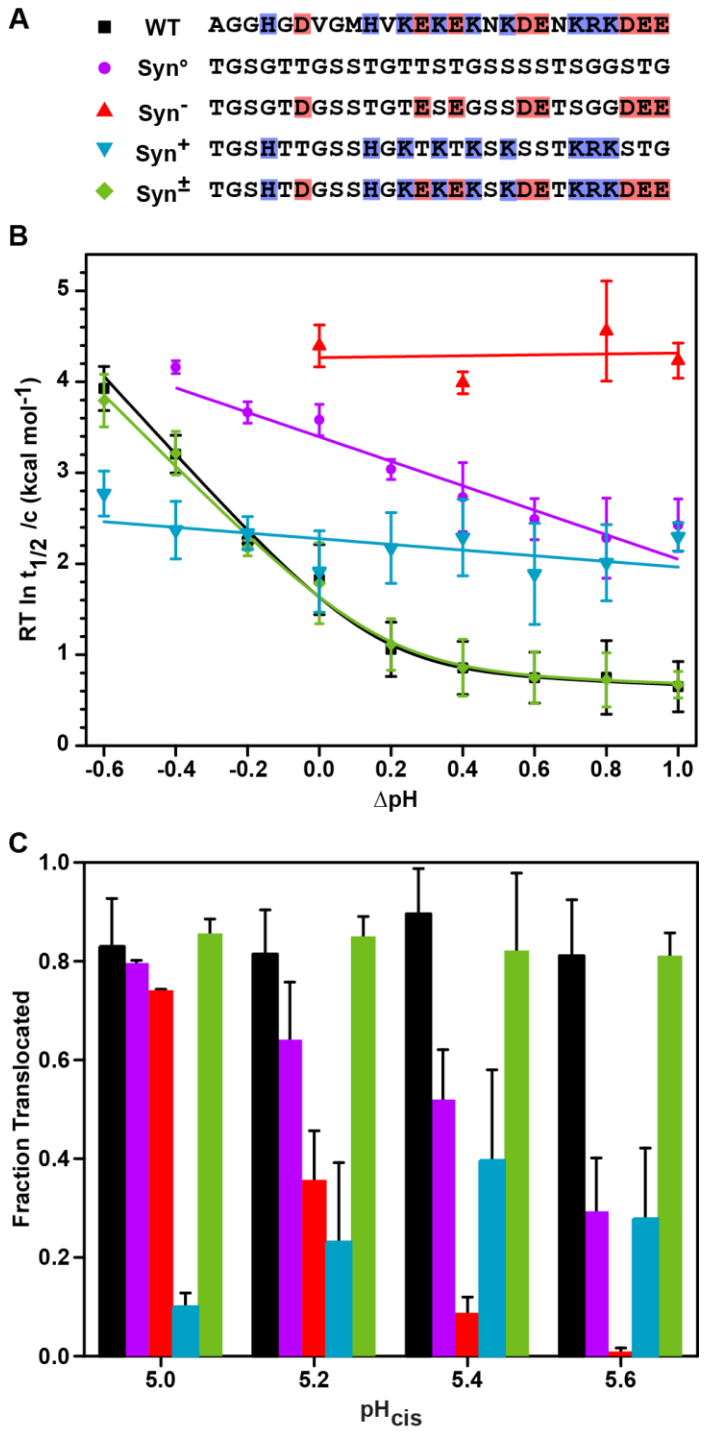


**Figure 2.2. Efficient ΔpH-driven translocation requires both acidic and basic charged residues.** (A) Design of synthetic $LF_N$ constructs ($LF_N$ Syn) in which the first 27 amino acids are replaced using a sequence containing both acidic and basic residues ($Syn^{\pm}$), neither acidic nor basic residues (Syn°), only acidic residues ($Syn^{-}$), and only basic residues ($Syn^{+}$). The exact sequence compositions of these constructs are shown alongside the $LF_N$ WT sequence. (B) Translocation activation energy versus ΔpH results for the $His_6$-$LF_N$ Syn constructs shown in (A): $His_6$-WT (black square), $His_6$-Syn° (purple circle), $His_6$-$Syn^{-}$ (red triangle), $His_6$-$Syn^{+}$ (blue inverted triangle), and $His_6$-$Syn^{\pm}$ (green diamond). These ΔpH-driven translocation experiments were conducted at a ΔpH of −0.6 to 1.0, $pH_{cis}$ of 5.6, and a $\Delta\psi$ of 60 mV. The error bars are the means ± S.D. ($n$ = 2–5). The fit parameters for $His_6$-WT and $His_6$-$Syn^{\pm}$ using Eq. 1: $His_6$-WT, $\Delta G^{\ddagger\circ}_1$ = 1.5 (± 0.2) kcal $mol^{-1}$, $\Delta G^{\ddagger\circ}_2$ = 0.8 (± 0.3) kcal $mol^{-1}$, $n_1$ = 1.9 (±0.2), $n_2$ = 0.1 (± 0.2); $His_6$-$Syn^{\pm}$, $\Delta G^{\ddagger\circ}_1$ = 1.5 (±0.2) kcal $mol^{-1}$, $\Delta G^{\ddagger\circ}_2$ = 0.8 (± 0.3) kcal $mol^{-1}$, $n_1$ = 1.7 (± 0.2), $n_2$ = 0.1 (± 0.2). $His_6$-Syn°, $His_6$-$Syn^{-}$, and $His_6$-$Syn^{+}$ were fit to a single-barrier model ($\Delta G^{\ddagger} = \Delta G^{\ddagger\circ} - 2.3n\Delta pH$), with the parameters: $His_6$-Syn°, $\Delta G^{\ddagger\circ}$= 3.40 (± 0.08) kcal $mol^{-1}$, $n$ = 0.59 (± 0.06); $His_6$-$Syn^{-}$, $\Delta G^{\ddagger\circ}$ = 4.3 (± 0.2) kcal $mol^{-1}$, $n$ = 0.0 (± 0.1); $His_6$-$Syn^{+}$, $\Delta G^{\ddagger\circ}$ = 2.28 (± 0.09) kcal $mol^{-1}$, $n$ = 0.14 (± 0.07). (C) Maximum translocation efficiency achieved within 5 min at varying $pH_{cis}$ values for the $His_6$-$LF_N$ Syn constructs. These ΔpH-driven translocation experiments were conducted at a ΔpH of 2.0, $pH_{cis}$ ranging from 5.0 to 5.6, and a $\Delta\psi$ of 0 mV. To control for buffer mixing lag times, $t$ = 0 was set as the time when 1% of maximum translocation occurred. The legend colors are identical to those in (B). The error bars are the means ± S.D. ($n$ = 2–8).

difficulties in protonating all of the acidic residues in $LF_N$ $Syn^-$ account for its poor translocation, we also varied the cis pH to see if lower pHs would increase its translocation speed relative to the chargeless $LF_N$ Syn°. Here, we report the fraction of substrate translocated after five minutes. This is a better method for comparing substrates that display large differences in both translocation efficiencies and rate because kinetic values alone would not tell the full story. Unsurprisingly, $LF_N$ WT and $LF_N$ $Syn^{\pm}$ show no cis pH dependence (Fig. 2.2C). $LF_N$ Syn°, however, displays a moderate, linear cis pH dependence, likely due to the increased substrate destabilization at lower pHs. As predicted, $LF_N$ $Syn^-$ has a much stronger cis pH dependence. It displays little translocation within five minutes at cis pH 5.6 (though it does eventually reach 15-20% translocation). By cis pH 5.0, $LF_N$ Syn° and $LF_N$ $Syn^-$ translocate with efficiency similar to $LF_N$ WT and $LF_N$ $Syn^{\pm}$. When the membrane potential is 0 mV, $LF_N$ $Syn^+$ translocates poorly at all cis pHs. We assume that $LF_N$ $Syn^+$'s comparatively better performance relative to $LF_N$ $Syn^-$ under higher voltage conditions (Fig. 2.2B) is an artifact of the supplemental $\Delta\psi$-driving force. Thus we conclude that $LF_N$ $Syn^-$ is greatly stimulated by a PMF dominated by a $\Delta$pH with lower cis pH conditions, i.e., below ~pH 5.2.

*Cationic groups in presequence allow anionic residues to penetrate channel*—The translocation defect observed for $LF_N$ $Syn^-$ at pH 5.6 was surprising and led us to hypothesize that a presequence dominated by acidic groups was unable to traverse the cation-selective PA channel, likely due to electrostatic repulsion, which impeded entry into the channel. To test this possibility, we examined the ability of the $His_6$-tagged and untagged versions of the $LF_N$ Syn constructs to bind inside the PA channel. In this assay, when $LF_N$'s presequence binds the PA channel it blocks ion conductance. In ensemble channel-binding assays, we found that all of the $His_6$-Syn constructs could fully engage and dock into the PA channel, as evidenced by the complete block the of channel's conductance (Fig. 2.3A). Upon removing the $His_6$-tag, we found that $LF_N$ Syn° and $LF_N$ $Syn^-$ are defective in forming a stable interaction with the PA channel. Moreover, when we tested the ability of these $LF_N$ Syn constructs to remain stably engaged in a blocked-conductance state with the PA channel, we found that $LF_N$ Syn° and $LF_N$ $Syn^-$ rapidly dissociated from the channel with dissociation lifetimes of ~30 and ~10 s, respectively (Fig. 2.3B). All other Syn constructs showed little dissociation; and the $His_6$-tagged versions of $LF_N$ Syn° and $LF_N$ $Syn^-$ maintained a stable complex with the PA channel. We conclude that uncharged and purely anionic leader presequences are unable to stably engage the PA channel; however, the defects in these constructs may be complemented by a 17-residue $His_6$-tag.

We then examined the channel-blocking activity of the $LF_N$ Syn constructs at symmetrical pH 5.6 in single-channel experiments. The $His_6$-tagged Syn constructs all formed stable, fully-blocked complexes with the PA channel, where few opening (or unblocking) events were observed (Fig. 2.3C). However, for the untagged $LF_N$ Syn° and $LF_N$ $Syn^-$ constructs, we again found noticeable defects in channel blocking activity. The two untagged constructs differ notably in their behavior. Untagged $LF_N$ Syn° is able to form transient blockades of the PA channel (which fully block the conductance) for durations of 5-10 s. Untagged $LF_N$ $Syn^-$ is not able to form lasting blockades, and only blocked 29 (±2)% of the channel's conductance. These partly-blocked flickering events lasted 10-30 s (Fig. 2.3C). From these single-channel studies, we conclude that the two defective mutant $LF_N$ presequences (Syn° and $Syn^-$) possess distinguishable characteristics: the purely anionic version ($Syn^-$) cannot penetrate the PA channel and only produces a partial flickering 30% blocked state, whereas the neutral Syn° presequence can completely penetrate and block the channel, albeit ~70% less often than WT, $Syn^+$, and $Syn^{\pm}$.

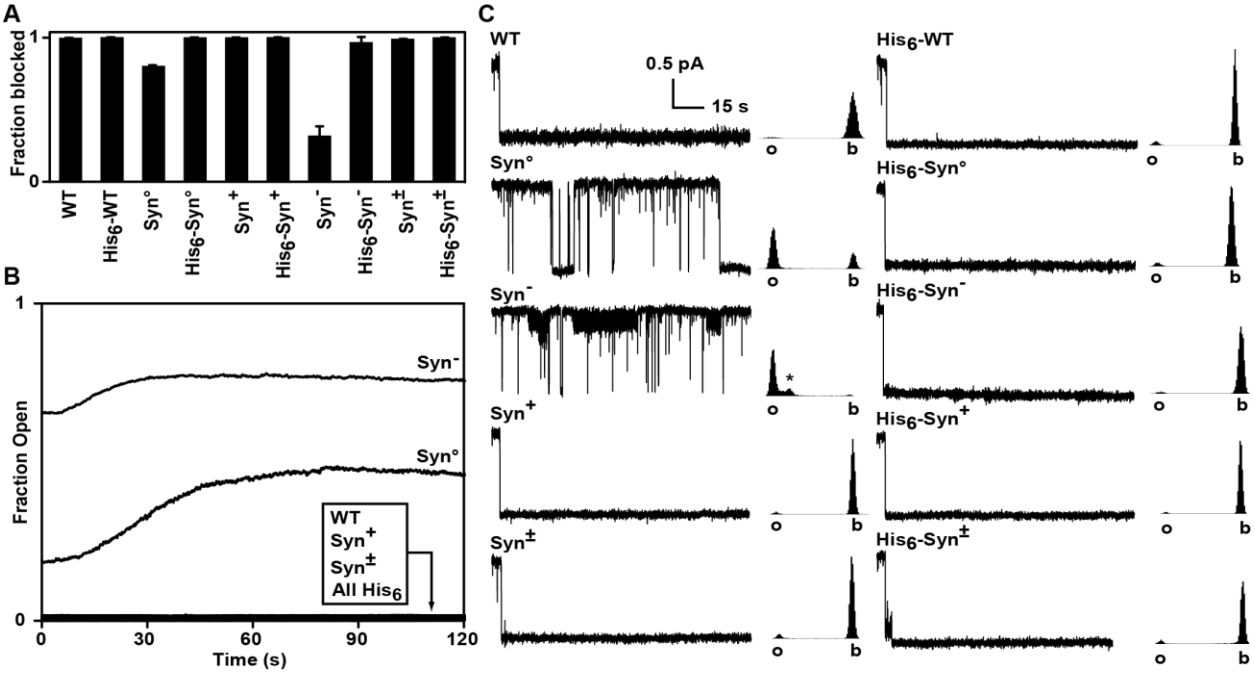


**Figure 2.3. Charge requirements for $LF_N$ docking.** (A) Ensemble bilayer recordings of PA channel conductance block by $LF_N$ WT and $LF_N$ Syn mutants with and without their $His_6$ tags were obtained at symmetrical pH 5.6 and saturating concentrations of $LF_N$ (100 nM). The relative fraction of conductance block is given for each Syn mutant construct relative to $LF_N$ WT. The error bars represent the means ± S.D. ($n$ = 2). (B) Ensemble bilayer recordings of PA channel conductance block by $LF_N$ WT and $LF_N$ Syn mutants at symmetrical pH 5.6 following perfusion of the cis side of the membrane. Perfusion removes excess $LF_N$, allowing the channel dissociation kinetics to be recorded. Significant rapid dissociation was only observed with $LF_N$ Syn° and $Syn^-$; all $His_6$-tagged $LF_N$ and the non-$His_6$-tagged $LF_N$ (including WT and mutants, $Syn^+$ and $Syn^{\pm}$) did not dissociate during the recording. (C) Single-channel blocking events recorded for $LF_N$ WT and $LF_N$ Syn mutants with and without their $His_6$ tags at symmetrical pH 5.6. Once a single channel inserts into the membrane, $LF_N$ is added under a $\Delta\psi$ of 20 mV. Approximately 2 min of a typical blocking event is shown for each $LF_N$. The data are acquired at 400 Hz under a low pass filter of 200 Hz. For clarity, the displayed single-channel traces were downsampled by a factor of 10. To the right of each trace is a histogram of the current level for each recording. Gaussian functions fit to these histograms assess the percentage of the time the $LF_N$-channel complexes spend in the open (o), blocked (b), and partly blocked states (*). The percentages of time in the blocked and partly blocked states are given as follows: all $His_6$-tagged $LF_N$ constructs as well as untagged $LF_N$ WT, $LF_N$ $Syn^+$ and $LF_N$ $Syn^{\pm}$ were 100% blocked; untagged $LF_N$ Syn° was 28% blocked; and $LF_N$ $Syn^-$ was 1.6% blocked and 3% partly blocked (*). The errors for percentage of the time spent in the block states are all better than ±1%. The partly blocked (*) state is 29 (± 2)% less conducting than the fully open state.

These results, taken together, imply that anionic residues are effectively repulsed by the channel; however, cationic residues can effectively chaperone anionic residues through the channel.

*Acidic-residue dependence of ΔpH translocation extends beyond presequence*—Based upon our initial characterization of the sequence dependence of ΔpH-driven translocation, we determined that acidic residues are responsible for the ΔpH dependence, but counter-balancing positively-charged residues can assist passage through the cation-selective PA channel. It should be noted that the requirement of acidic residues on the ΔpH dependence of $LF_N$ translocation exists when all the negative charges are mutated to serine (Fig. 2.4A) and when the $His_6$ tag is cleaved and the negative residues mutated to their uncharged analogs (i.e. Glu→Gln and Asp→Asn; Fig. 2.4B). Here we produced a series of amino-terminal $LF_N$ mutants in which successively larger amino-terminal stretches of acidic residues are replaced with serine residues (Fig. 2.5A). We call these anionic-residue replacement constructs $LF_N$ $des^{(-)}_{x\text{-}y}$, where *x* and *y* designate the region of sequence in which the natural Asp and Glu residues were replaced with Ser. Such extensive mutations in $LF_N$ destabilized the substrates to varying degrees, as shown in equilibrium protein-folding stability measurements (Fig. 2.6A). In fact, our analysis is limited to the first 68 residues of $LF_N$ because removing acidic residues beyond this point resulted in too substantial a loss in protein solubility and stability.

To normalize for these differences in protein stability, we translocated each $His_6$-tagged $LF_N$ $des^{(-)}$ mutant in the presence (i and ii) or absence (iii) of a ΔpH driving force:

(i) lower Δψ and higher ΔpH, ΔpH = 1, $pH_{cis}$ = 5.6, Δψ of 20 mV (Fig. 2.6B);

(ii) higher Δψ and lower ΔpH, ΔpH = 0.6, $pH_{cis}$ = 5.6, Δψ of 40 mV (Fig. 2.6C);

(iii) higher Δψ and no ΔpH, ΔpH = 0, $pH_{cis}$ = 5.6, Δψ of 60 mV (Fig. 2.6D).

By comparing the translocation rates of a single mutant under two different conditions, we can calculate a $\Delta\Delta G^{\ddagger}$ value by $\Delta\Delta G^{\ddagger} = \Delta G^{\ddagger}(\Delta pH>0) - \Delta G^{\ddagger}(\Delta pH=0)$, which normalizes for stability. We then referenced the $\Delta\Delta G^{\ddagger}$ values for each mutant to that of $LF_N$ WT by determining a $\Delta\Delta\Delta G^{\ddagger}$ value, where $\Delta\Delta\Delta G^{\ddagger} = \Delta\Delta G^{\ddagger}(MUT) - \Delta\Delta G^{\ddagger}(WT)$. Here, larger $\Delta\Delta\Delta G^{\ddagger}$ values indicate a reduced capacity to translocate using a pH gradient compared to $LF_N$ WT. These comparisons revealed a clear trend showing that the more acidic residues that are removed from $LF_N$, the less it was able to effectively use a pH gradient for translocation (Fig. 2.5B). Also, we generally found that condition (i), which contains a higher ΔpH, increases $\Delta\Delta\Delta G^{\ddagger}$ more so than the lower ΔpH condition (ii), indicating that the effect was specifically due to an inability to use the pH gradient (Fig. 2.5B). Based on these results, we conclude that acidic residues are critical to the mechanism of ΔpH-dependent translocation both within the presequence region and well into the folded structure of the substrate protein.

*Optimal positions for acidic residues in the presequence*—We then asked which acidic residues in the leader presequence were most critical to ΔpH-driven translocation. Using the $LF_N$ $des^{(-)}_{1\text{-}46}$ construct as our background, we introduced single glutamates at various positions and measured the translocation kinetics of these new constructs under the influence of a ΔpH. We found that little or no gain (and perhaps a small decrease in rate) is observed when introducing a single Glu at positions 1 to 13 (inclusive, Fig. 2.7B). A small but reproducible ~2-fold increase in translocation rate (relative to the $LF_N$ $des^{(-)}_{1\text{-}46}$ background) is observed when introducing a Glu into positions 15 to 20 (inclusive, Fig. 2.7B). The optimal region for the introduction of a Glu residue appears to occur between residues 23 and 31 (inclusive), where the most optimal site is position 23 (increase of 5-fold; Fig. 2.7B). More modest increases in the translocation rates are observed from residues 34 to 43 (the extent of the analysis in this mutant background, Fig. 2.7B). Similar results were obtained in the $LF_N$ $des^{(-)}_{1\text{-}32}$ background, and multisite replacements with as

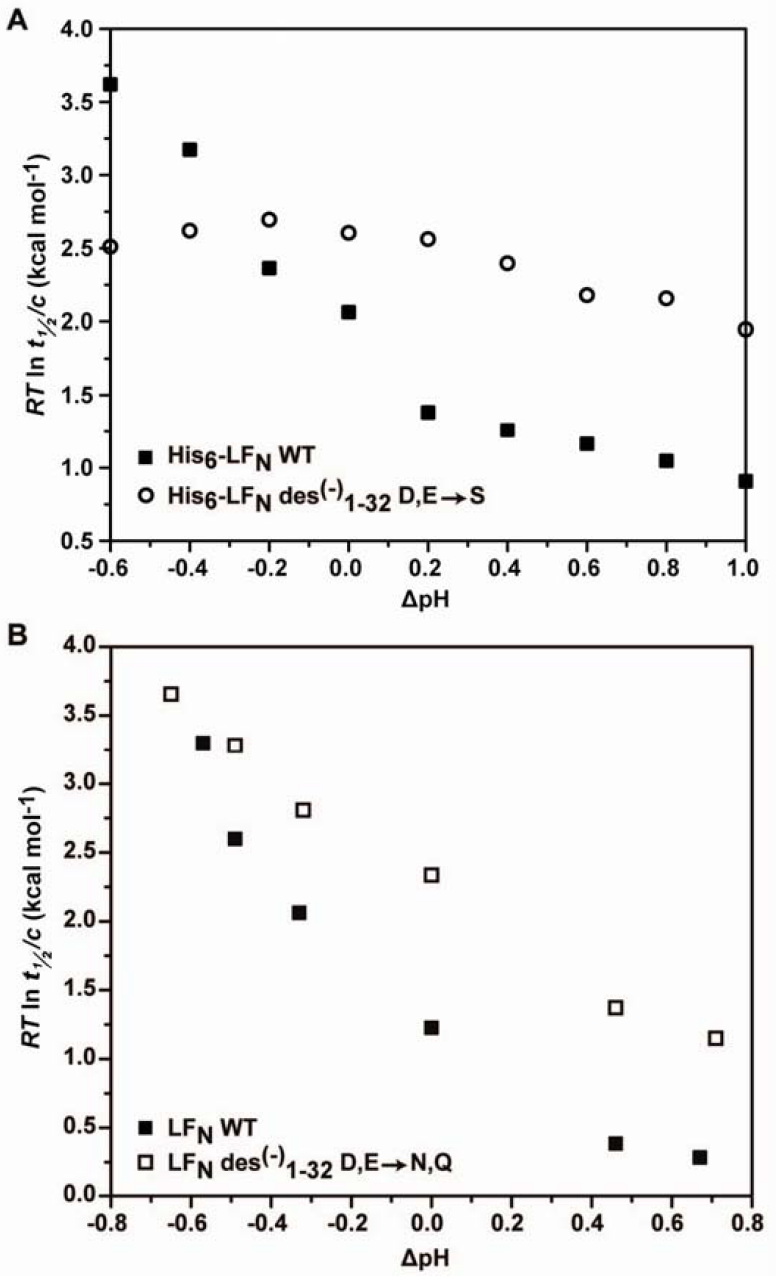


**Figure 2.4. $LF_N$ $des^{(-)}_{1-32}$ has reduced ΔpH dependence relative to $LF_N$ WT regardless of replacement residue or $His_6$ tag.** The $LF_N$ $des^{(-)}_{1-32}$ construct changes the aspartates and glutamates in $LF_N$'s presequence to (A) serines or (B) their respective nonpolar analogs asparagines and glutamines. Additionally, in (B) the $His_6$ tags have been removed from the proteins. $\Delta G^{\ddagger}$ versus ΔpH profiles are shown for (A) $His_6$-$LF_N$ WT (■) and $His_6$-LFN $des^{(-)}_{1-32}{}^{D,E\rightarrow S}$ (○) or (B) $LF_N$ WT (■) and $LF_N$ $des^{(-)}_{1-32}{}^{D,E\rightarrow N,Q}$. In both cases, conditions are $\Delta\psi$ of 60 mV and $pH_{cis}$ = 5.6.

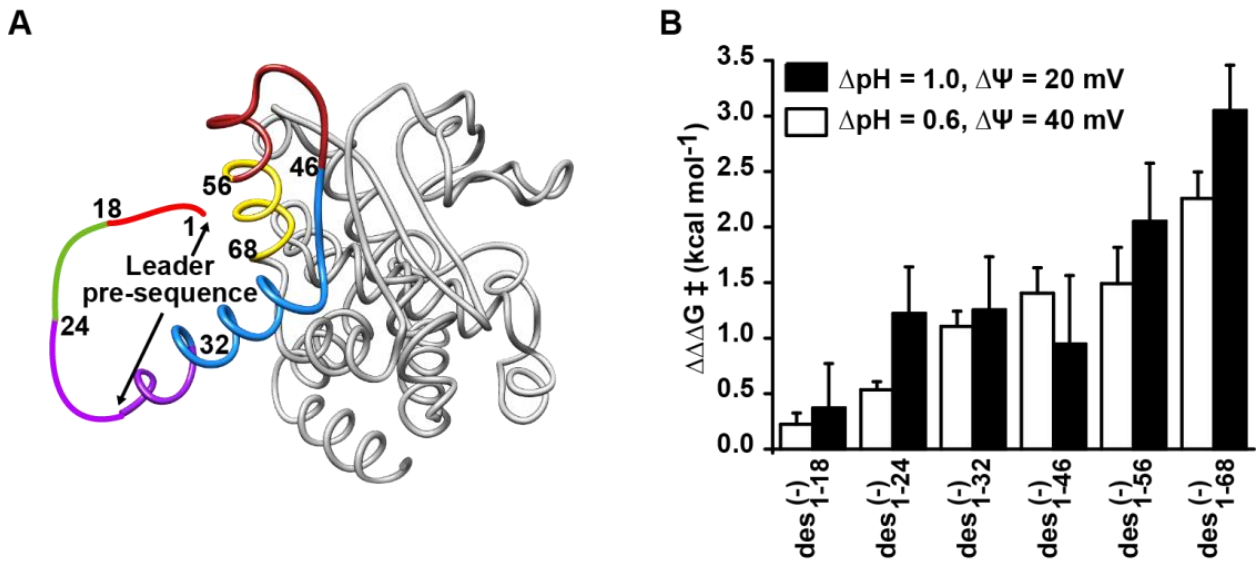


**Figure 2.5. Acidic residues within the folded domain of $LF_N$ are also critical to ΔpH-driven translocation.** (A) The ribbon depiction of the structure of $LF_N$ is from PDB ID 1J7N (32), where regions of sequence are colored by sequence position. The unstructured leader presequence (residues 1–32) is indicated as a thick colored line. Regions in which acidic residues were replaced with serine residues are colored as follows: residues 1–18 (red), residues 19–24 (green), residues 25–32 (purple), residues 33–46 (blue), residues 47–56 (*brown*), and residues 57–64 (gold). (B) The difference in activation energy $\Delta\Delta G^{\ddagger}$ for each $His_6$-$LF_N$ $des^{(-)}$ series mutant are obtained using two ΔpH-driven conditions (ΔpH = 1.0 and Δψ of 20 mV; ΔpH = 0.6 and Δψ of 40 mV) and one condition in the absence of a ΔpH (ΔpH = 0 and Δψ of 60 mV), where $\Delta\Delta G^{\ddagger} = \Delta G^{\ddagger}(\Delta pH>0) - \Delta G^{\ddagger}(\Delta pH = 0)$. $\Delta\Delta G^{\ddagger}$ values for each mutant (MUT) are then referenced to that of $LF_N$ WT to give the reported $\Delta\Delta\Delta G^{\ddagger}$ values plotted above. $\Delta\Delta\Delta G^{\ddagger} = \Delta\Delta G^{\ddagger}(MUT) - \Delta\Delta G^{\ddagger}(WT)$. In all cases, $pH_{cis} = 5.6$. The error bars are the means ± S.D. ($n = 2$).

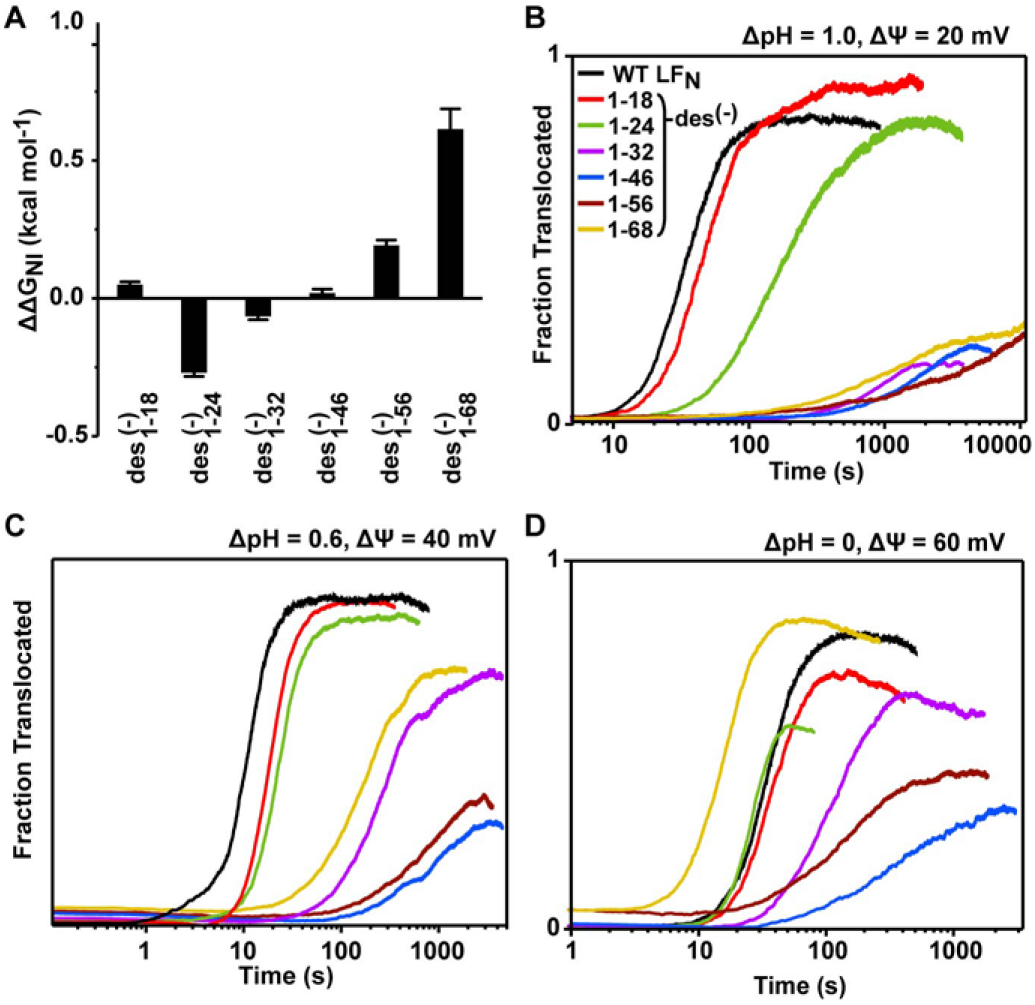


**Figure 2.6. Acidic residues within $LF_N$'s folded domain are also critical to ΔpH-driven translocation.** (A) The change in protein stability ($\Delta\Delta G_{NI}$) for the $His_6$-$LF_N$ $des^{(-)}$ series relative to $His_6$-$LF_N$ WT estimated from guanidinium chloride denaturant melts probed by circular dichroism. The thermodynamic quantity, $\Delta\Delta G_{NI}$, compares difference in the *N* and *I* state free energies of the mutant (MUT) to WT as follows: $\Delta\Delta G_{NI} = \Delta G_{NI}(MUT) - \Delta G_{NI}(WT)$. Error bars are the mean ± S.D. ($n$ = 2). (B-D) Translocation records for $His_6$-$LF_N$ WT (black), $His_6$-$LF_N$ $des^{(-)}_{1\text{-}18}$ (red), $His_6$-$LF_N$ $des^{(-)}_{1\text{-}24}$ (green), $His_6$-$LF_N$ $des^{(-)}_{1\text{-}32}$ (purple), $His_6$-$LF_N$ $des^{(-)}_{1\text{-}46}$ (blue), $His_6$-$LF_N$ $des^{(-)}_{1\text{-}56}$ (brown), and $His_6$-$LF_N$ $des^{(-)}_{1\text{-}68}$ (gold) under different ΔpH and Δψ driving force conditions: (B) ΔpH = 1.0, $pH_{cis}$ = 5.6, Δψ of 20 mV; (C) ΔpH = 0.6, $pH_{cis}$ = 5.6, Δψ of 40 mV; and (D) ΔpH = 0, $pH_{cis}$ = 5.6, and Δψ of 60 mV, as summarized in Fig. 2.5B.

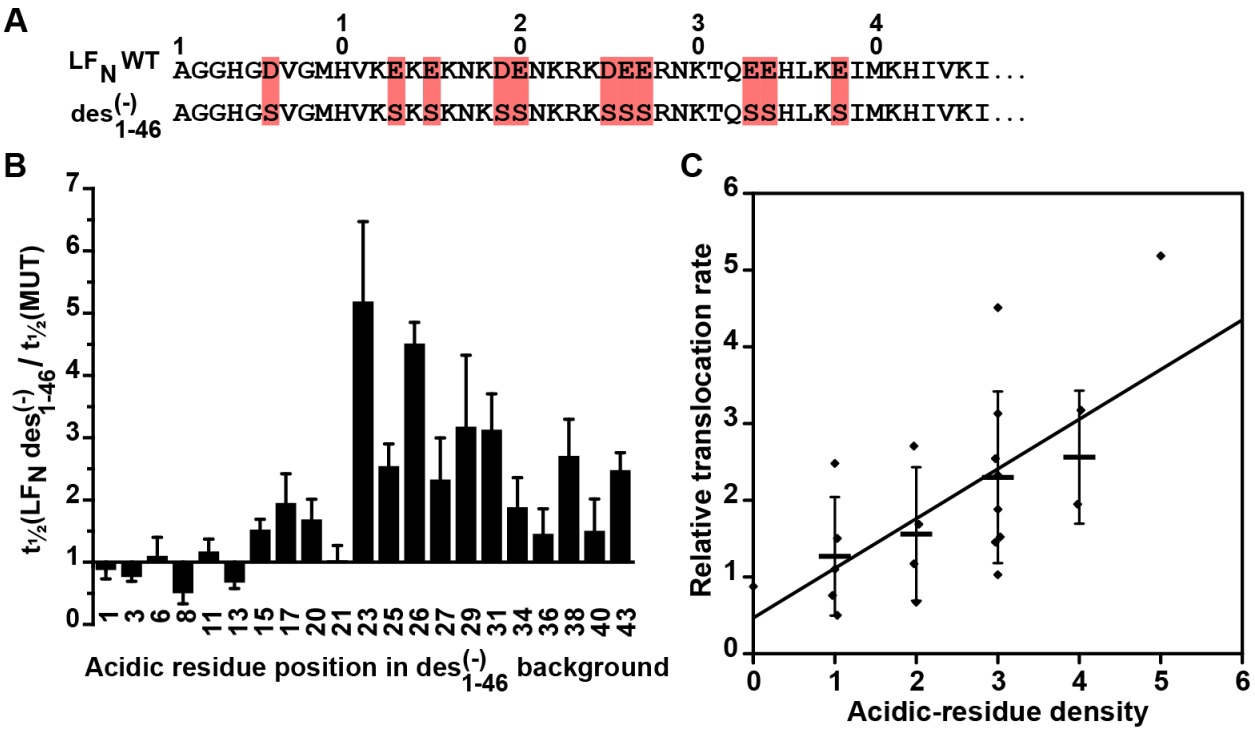


**Figure 2.7. Acidic residue positions in the presequence of $LF_N$ are most critical to ΔpH-driven translocation.** Acidic residues were added back into the $His_6$-$LF_N$ $des^{(-)}_{1–46}$construct background. (A) The sequences of the first 46 residues of $LF_N$ WT and $LF_N$ $des^{(-)}_{1–46}$ are shown, where acidic residues in the WT sequence are shaded red. (B) The relative translocation $t_{1/2}$ times for acidic residue introductions into the $His_6$-$LF_N$ $des^{(-)}_{1–46}$ mutant backgrounds are given as the ratio $t_{1/2}(des^{(-)}_{1–46})/t_{1/2}(MUT)$ for ΔpH-driven translocation. The numbers on the *x* axis indicate the positions in which acidic residues are reintroduced into the $His_6$-$LF_N$ $des^{(-)}_{1–46}$ mutant background. ΔpH-driven translocation conditions were ΔpH = 0.8, $pH_{cis}$ = 5.6, and Δψ of 50 mV. The error bars are the means ± S.D. (*n* = 2–4). (C) A correlation of relative translocation rate (given as the ratio $t_{1/2}(des^{(-)}_{1–46})/t_{1/2}(MUT)$) for the mutants in *B* versus the density of acidic residues normally found in the WT sequence. Acidic residue density in this instance is the total number of acidic residues found in the four residues amino- and carboxyl-terminal to the probed site. The linear regression fit to all of the individual measurements (filled diamonds) is significant with a *p* value of 0.001 for the fit function, $y = a + bx$, where $a$ = 0.5 (± 0.5) and $b$ = 0.6 (± 0.2). Because multiple observations of particular acidic residue densities were obtained in certain instances, a heavy horizontal bar (mean) and error bars (± S.D.) are also given.

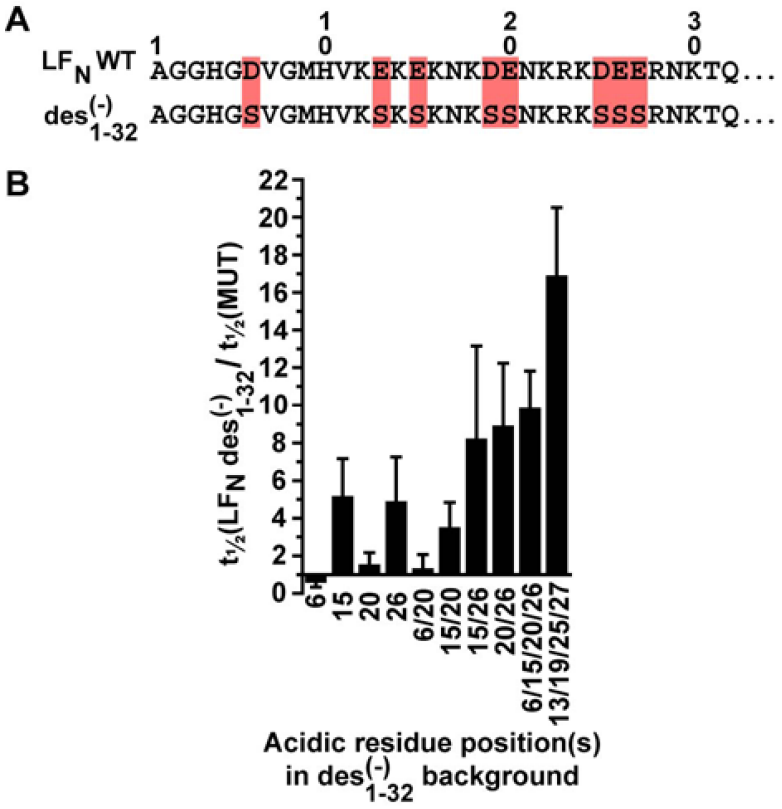


**Figure 2.8. Acidic-residue positions in $LF_N$'s presequence are most critical to ΔpH-driven translocation.** Acidic residues were reintroduced in their wild-type positions into the $His_6$-$LF_N$ $des^{(-)}{}_{1\text{-}32}$ background. (A) The sequences of the first 32 residues of $LF_N$ WT and $His_6$-$LF_N$ $des^{(-)}{}_{1\text{-}32}$ are shown, where acidic residues in the WT sequence are shaded red. (B) The relative translocation $t_{1/2}$ times for ΔpH-driven translocation are given as the ratio $t_{1/2}(des^{(-)}{}_{1\text{-}32})/t_{1/2}(\text{MUT})$, where the mutant (MUT) is the construct with the reintroduced acidic residue(s). The numbers on the *x*-axis indicate the position(s) in which acidic residues are reintroduced into the $His_6$-$LF_N$ $des^{(-)}{}_{1\text{-}32}$ backgrounds. Multisite acidic-residue reintroductions are indicated with slashes separating the residue number. ΔpH-driven translocation conditions were ΔpH = 0.6, $pH_{cis}$ = 5.6, and Δψ of 35 mV. Error bars are the mean ± S.D. ($n$ = 2). For reference, the relative translocation time for $LF_N$ WT compared to $His_6$-$LF_N$ $des^{(-)}{}_{1\text{-}32}$ , $t_{1/2}$ $(des^{(-)}{}_{1\text{-}32})$ / $t_{1/2}$ (WT), is 14.8 (±3.0).

few as four of the eight wild-type charges could fully restore the translocation rate to that of $LF_N$ WT (Fig. 2.8). We conclude that residues 23 to 31 in $LF_N$'s presequence are the most optimal positions to place acidic residues for ΔpH-mediated translocation.

*Anionic-charge density in presequence is optimized for ΔpH-driven translocation*—We then determined whether the natural sequence of $LF_N$ is optimized in terms of the placement of the acidic residues in the presequence. First, we determined the density of various classifications of amino acids: acidic, basic, acidic plus basic, and hydrophobic. Our density metric is calculated as the number of residues in each classification that are ±4 positions, or one turn of an α helix, from the probed site. We then performed linear regression analysis to determine the degree to which the density of these various types of residue classifications correlated with the relative change in translocation rate resulting from the introduction of Glu residues into the $LF_N$ $des^{(-)}_{1\text{-}46}$ background (Fig. 2.7B). We find that acidic residue density provides the strongest correlation with a significant *p* value of 0.001 (Fig. 2.7C). While the density of total charges, acidic plus basic, is a fairly good predictor ($p = 0.003$), this correlation likely reflects the contribution of the negative-charge density and not the positive-charge density, since the latter correlation is not significant ($p = 0.3$). The degree of hydrophobicity also poorly correlated ($p = 0.09$) with the most optimal positions for acidic residues in ΔpH-mediated translocation. We conclude that the natural placement of acidic residues in $LF_N$'s presequence is optimized for ΔpH-dependent translocation.

*Charge-dense regions are optimally positioned proximal to folded structure*—Since we find that when translocating under a ΔpH driving force that the region of highest charge density in $LF_N$'s presequence (residues ~15 to 30) is also the optimal region for introducing Glu residues into the $LF_N$ $des^{(-)}_{1\text{-}46}$ mutant background (Fig. 2.7), we hypothesized that the location of this cluster of charged residues is only critical in relation to $LF_N$'s folded domain. That is, if the charged cluster were moved further away from the folded structure of $LF_N$, then the force generated by the applied ΔpH gradient would not be as fully realized on the folded structure of $LF_N$. Thus acidic residues more proximal to the substrate's folded region would better facilitate proper $LF_N$ unfolding. To test this hypothesis, we created the $LF_N$ 27Ins° construct, which inserts a series of 16 random Gly, Ser, and Thr residues between positions 27 and 28 of $LF_N$ WT (Fig. 2.9A). Despite possessing all of the acidic residues of $LF_N$ WT and only extending the substrate's overall length by ~6%, translocation is slowed approximately 50-fold under a 1.06 unit ΔpH and no Δψ (Fig. 2.9B). An identical insertion placed earlier ($LF_N$ 0Ins° and $LF_N$ 11Ins°) in the $LF_N$'s amino terminus does not appreciably effect ΔpH-dependent translocation. Restoring charges to the inserted region in a pattern matching that of $LF_N$ residues 12 to 27 ($LF_N$ $27Ins^{\pm}$) returns translocation speed to that of $LF_N$ WT. Interestingly, all constructs with inserts (regardless of position, charge, or background) displayed a reduction in efficiency of translocation: approximately 50-65% that of $LF_N$ WT.

To test whether the reduction of the translocation rate in the $LF_N$ 27Ins° construct is due to an effective stabilization of the folded substrate, we introduced the destabilizing mutation, L145A (45), into the $LF_N$ WT and 27Ins° backgrounds (Fig. 2.9C). Under the same conditions described above, the translocation rate of $LF_N$ WT was increased by a factor of 2.3 (±0.7) by the L145A mutation while 27Ins° L145A's rate increased by a factor of 6.4 (±1.9). We posit that 27Ins° weakens the mechanical connection between the forces derived from the anionic charge cassette in the unstructured leader and the folded domain, effectively increasing the unfolding barrier's height (Fig. 2.9D). When the L145A mutation is introduced in $LF_N$ WT, a second barrier becomes limiting, but in 27Ins° the unfolding barrier remains limiting. Thus despite

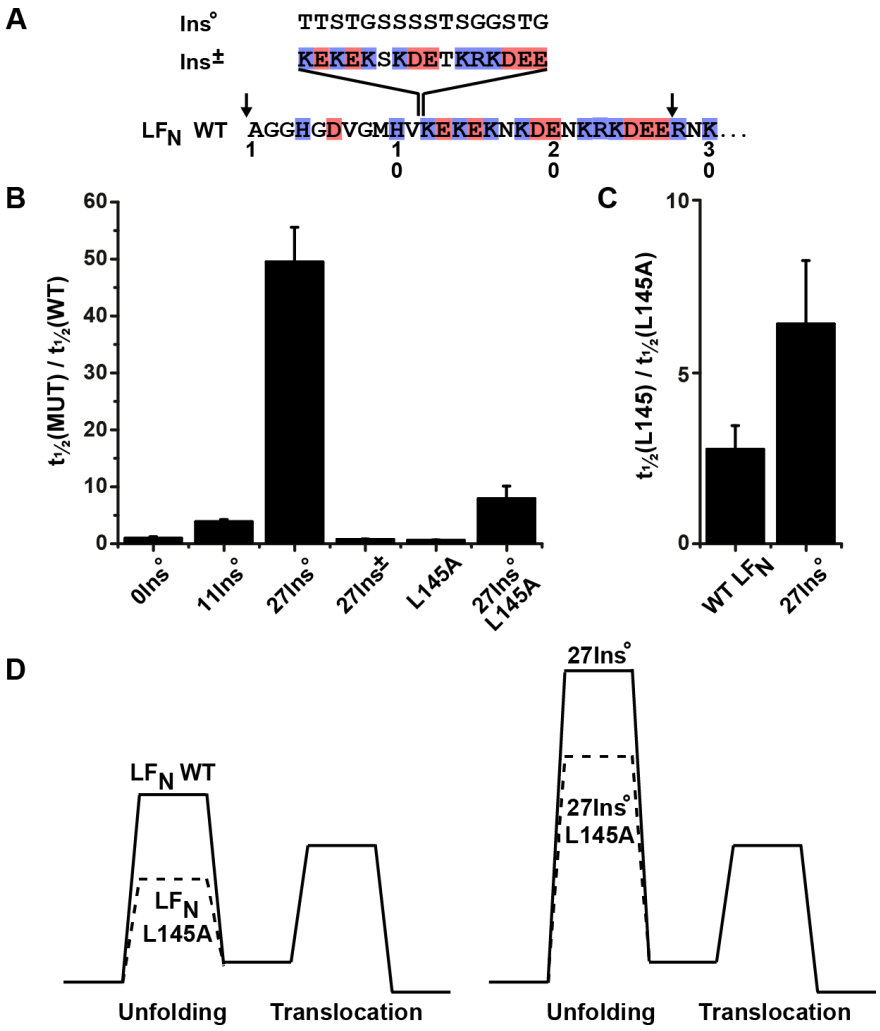


**Figure 2.9. Charged residues must be located immediately before the folded domain of the substrate for efficient translocation.** (A) Sixteen-residue inserts with either no charge (Ins°) or a mix of positive and negative charges (Ins$^{\pm}$) were inserted into $LF_N$WT or $LF_N$ L145A. The resulting constructs are called *y*Ins$^{x}$, where *y* denotes the last WT residue prior to the insert, and the superscript *x* represents the charge of the insert. The arrows indicate two other positions (0 and 27) where the Ins° sequence was inserted. (B) The relative translocation $t_{1/2}$ times of insertion and $LF_N$ L145 mutants (MUT) are given as the ratio $t_{1/2}$(MUT)/$t_{1/2}$(WT). ΔpH-driven translocation conditions were: ΔpH = 1.06, $pH_{cis}$ = 5.6, and Δψ of 0 mV. The error bars are the means ± S.D. (*n* = 2). (C) The relative translocation $t_{1/2}$ times of the $LF_N$ L145A and $LF_N$ 27Ins° L145A mutants compared with their respective (L145) counterpart given as the ratio $t_{1/2}$(L145)/$t_{1/2}$($LF_N$ L145A). Translocation conditions are as in (B). (D) Model energy diagrams depicting the changes in energy barriers caused by the L145A and 27Ins° mutations, wherein we interpret 27Ins° as greatly increasing the unfolding barrier. The L145A mutation reduces the unfolding barrier by the same extent in the WT and 27Ins° backgrounds. However, in the WT background, the rate becomes limited by the translocation barrier, so the relative increase in speed is not as large as that observed for the 27Ins° background.

reducing the height of the unfolding barrier to the same extent, the L145A mutation has a greater effect in 27Ins°. We conclude that the highly charged region must be present immediately prior to the folded domain for efficient protein unfolding to take place under a ΔpH.

*Intermixed cationic and anionic groups in $LF_N$'s presequence are optimal for ΔpH translocation*—So far we have reported that basic residues are important for the initial insertion into the channel and acidic residues are most important just before $LF_N$'s folded domain. From these observations, we might predict that translocation would not be inhibited when $LF_N$'s presequence is rearranged so that all the basic residues are at the beginning and all the acidic residues are at the end. We made such a construct ($LF_N$ PosNeg) along with its inverse ($LF_N$ NegPos) and assayed their abilities to translocate with a proton gradient, ΔpH = 1.06 and Δψ of 0 (Fig. 2.10A). Both constructs were essentially unable to translocate (Fig. 2.10B). However, a construct with a randomized amino-terminal presequence ($LF_N$ Mix) is able to translocate much more similarly to $LF_N$ WT (Fig. 2.10B). Interestingly, none of the positions in this construct share the charge of their WT counterparts. The effectiveness of the $LF_N$ Mix construct suggests the specific WT arrangement of positive and negative charges is less important than simply maintaining an intermixed arrangement of positively- and negatively-charged residues. We propose that positively-charged residues act locally as ionic chaperones for negatively-charged acidic residues.

**2.5 Discussion**

To perform work, molecular machines and protein translocases in the cell require a source of energy, generally in the form of chemical energy (ATP) or potential energy stored as either a PMF or other ionic gradient. These sources of energy are interchangeable as evidenced by the highly homologous (103) bacterial flagellum, (104, 105) $F_1F_o$-ATP synthase, (106, 107) and the vacuolar $V_oV_1$-ATPase membrane proton pump (108–110). In the former two cases, a PMF is used to drive the rotational motion of the flagellum (105) or the cycles of enzyme reactions (111) that produce ATP in the transmembrane $F_1F_o$-ATP synthase (112). In the latter system, ATP hydrolysis powers a rotary proton pump to generate a PMF, which is essentially the ATP synthase in reverse. The PMF is also an important driving force for transmembrane protein translocation in mitochondria (113), chloroplasts (11), and the endosomal compartment (18). Proton gradients also exist across the endoplasmic reticulum and other compartments in the cell (114) and in theory may comprise an important driving force for other types of translocases in the cell.

*Active pushing/pulling translocation models*—Protein translocation is basically a series of reactions that convert a source of free energy to mechanical work used to drive protein unfolding and unidirectional transport. This process is involved in transmembrane transport, protein degradation, and chaperone activities. Here, either ATP hydrolysis cycles (100, 115–117) or utilization of the PMF (18) can propel polypeptides across membranes or into hydrolytic compartments for degradation. Often the molecular machine that carries out translocation processes contains loop regions with critical aromatic groups at their tips. Also, additional accessory proteins outside the channel may be utilized to engage the polypeptide during translocation. It is thought that cycles of ATP binding, ATP hydrolysis, and ADP and inorganic phosphate release cause these loops/proteins to move like simple linear actuators, which effectively pushes the translocating polypeptide in a particular direction (24, 115, 117, 118). This mechanism is supported in part because these ϕ-clamp-type structures are found in various conformations for different nucleotide bound states. For anthrax toxin, the utilization of the

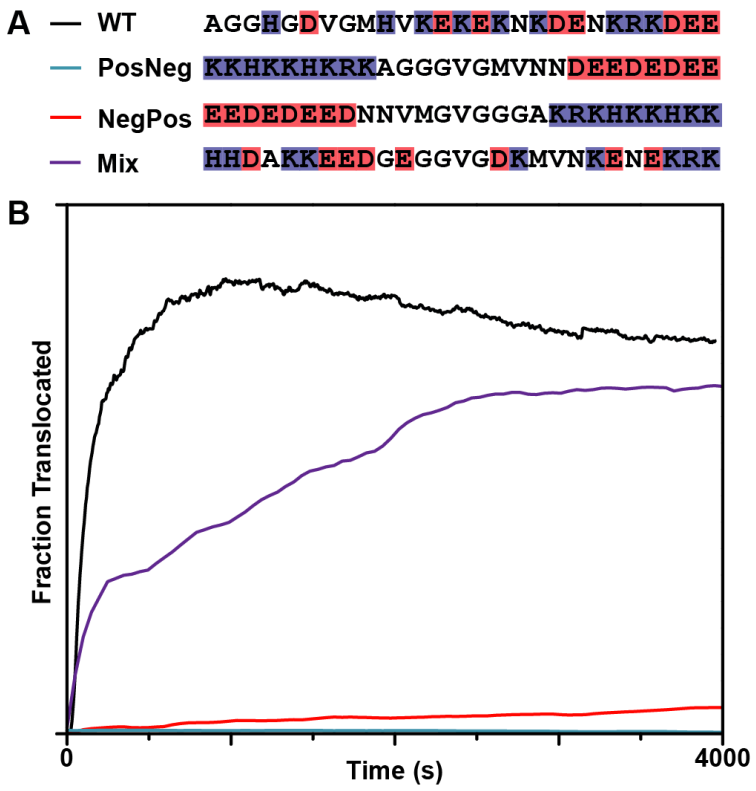


**Figure 2.10. Proton gradient-driven translocation requires that acidic and basic residues be intermixed in the substrate.** (A) Design of constructs that separate the acidic and basic amino acids in the first 27 residues of $LF_N$ to opposite ends of the presequence. PosNeg has a contiguous stretch of basic residues at the amino-terminal end and a contiguous stretch of acidic residues at the carboxyl-terminal end; NegPos is the inverse sequence of PosNeg. A randomized construct was also prepared (called Mix), in which the acidic and basic residues are intermixed, but no position has the same charge as its WT counterpart. (B) Translocation records for $His_6$-$LF_N$ WT (black), $His_6$-$LF_N$ PosNeg (blue), $His_6$-$LF_N$ NegPos (red), and $His_6$-$LF_N$ Mix (purple). ΔpH-driven translocation conditions were: ΔpH = 1.06, $pH_{cis}$ = 5.6, and Δψ of 0 mV.

proton gradient occurs most optimally with an intact aromatic ϕ-clamp structure, (18, 67) suggesting that these aromatic internal pore loops may push the translocating polypeptide in a similar manner. At present, this active pushing mechanism of translocation is favored in many systems, but it cannot by itself explain all mechano-chemical coupled mechanisms.

*Brownian-ratchet models for translocation*—An alternative but potentially complementary view of mechano-chemical coupling should be considered in addition to the active pushing model. The Brownian ratchet model (15–17, 119, 120) suggests that non-equilibrium fluctuations, namely in the form of chemical reactions or thermal gradients, can bias random thermal motion in a productive and directional way. In Feynman's initial conception (15), stochastic fluctuations may be harnessed for useful work using a temperature gradient. While a thermal gradient is difficult to achieve across a biological membrane, chemical asymmetries and membrane gradients, such as a ΔpH or those created by ATP hydrolysis, are readily available to drive biological Brownian motors.

*Ratchet model of ΔpH-driven anthrax toxin translocation*—Anthrax toxin activity is optimal when the endosomal compartment is allowed to naturally acidify (84). Acidification produces a transmembrane PMF that stimulates translocation (18, 45, 67). The chemical potential component of the gradient, $\Delta\mu^{H+}$, is sufficient by itself to drive translocation (Fig. 2.1). How might ΔpH-driven translocation follow a Brownian-ratchet mechanism? Brownian fluctuations in the substrate polypeptide chain may be coupled to protonation state changes in the substrate chain or channel (Fig. 2.11). While the substrates have net acidic isoelectric points, we infer that their acidic residues are protonated upon entering the PA channel, because the charge-dependence, or *z* value, we observe for translocation is positive. This fact implies that the substrate chain is positively charged over the span of the potential drop in the channel (18, 45). The positive-charge dependence also reflects the known fact that the PA channel is cation-selective (42), and a highly anionic polypeptide would be repulsed by the channel as we observe with our $LF_N$ $Syn^-$ construct (Fig. 2.2B). A similar effect has been observed when $LF_N$ is modified with a sulfate group (71). Therefore, a segment of substrate chain with protonated acidic residues will more readily achieve a positive charge enabling its passage through the PA channel under Brownian motion. It can then subsequently release protons into the trans side of the membrane (as favored in the higher pH condition), thus leading to a build-up of anionic charge in the trans-side substrate. The anionic charge in the trans substrate should effectively impede retrograde efflux (under Brownian motion) back into the anion-repulsive channel and enforce proper directionality. Cycles of this charge-state Brownian-ratchet mechanism would then translocate the remainder of the polypeptide across the membrane.

*Acidic residues in the substrate are the teeth in the ΔpH-driven Brownian ratchet*—One obvious and testable feature of this charge-state Brownian-ratchet mechanism is the role of acidic residues in the substrate. A previous study using a synthetic peptide attached to the folded domain of $LF_N$ via native chemical ligation showed that chargeless or positive-only presequences were defective in translocation (69). Here we more extensively explore charged-residue involvement in ΔpH translocation. By replacing the residues in the presequence *en masse* with neutral residues and measuring translocation under a variety of ΔpH conditions, we determined the ΔpH dependency of translocation, which is the best indicator of the ability to use a proton gradient as a driving force (Fig. 2.2). We found that sequences lacking acidic residues have comparatively weak ΔpH dependencies (Fig. 2.2), and the ability to translocate using a ΔpH decreases as more acidic residues are removed (Fig. 2.5). A more detailed scan, using an anion-less variant, $LF_N$ $des^{(-)}_{1\text{-}46}$, showed that the region spanning residues 23-31 received the largest

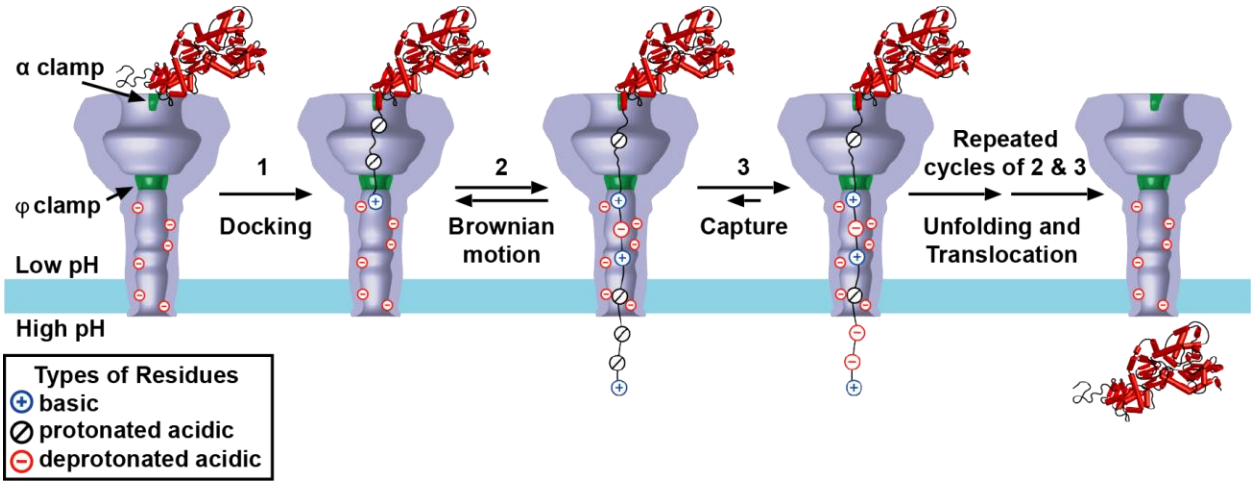


**Figure 2.11. The charge-state Brownian ratchet model for ΔpH-driven translocation of anthrax toxin.** At the time of endosome acidification, the substrate is bound to the top of the channel, with its α1 helix in the α clamp. Upon channel conversion, the unstructured amino terminus docks in the ϕ clamp. The low pH of the endosomal compartment will protonate most of the acidic residues, whereas positive charges chaperone any remaining deprotonated aspartic or glutamic acids. This ensures that the translocating polypeptide will have a net positive charge, allowing it to move freely through the cation-selective channel. Because of Brownian motion, a portion of the substrate will eventually emerge in the cytosol. There, the higher pH of this region will result in frequently deprotonated acidic residues, thereby giving the emerged portion of the polypeptide a net negative charge and capturing it on the cytosolic side of the membrane. The chaperoning positive charges would also lower the $pK_a$s of neighboring acidic residues, possibly making the capture events easier. Repeated cycles of emergence from the channel through Brownian motion and capture via deprotonation allow the remaining portion of the substrate to translocate across the membrane.

gain in translocation kinetics when a single acidic residue was introduced (Fig. 2.7B). Furthermore, the natural sequence appears to maximize its inherent acidic-residue density in this very region to best take advantage of the ΔpH driving force (Fig. 2.7C). Thus based on this model, we conclude that acidic residues in the substrate form the molecular teeth in ΔpH-driven Brownian-ratchet mechanism.

*Novel role for positive charges in the charge-state ratchet*—We identified one other complicating factor in our analysis when we tested for the role of positive charges in the presequence. Using the synthetic presequence Syn series, we find that a $Syn^-$ presequence dominated by acidic residues and lacking basic residues is defective in translocation (Fig. 2.2B). However, this result does not mean that ΔpH-driven translocation requires only positive charges, as presequences engineered to only contain cationic residues tend to have flat ΔpH dependencies (Fig. 2.2B). Various lines of evidence indicate that acidic residues require neutralization either by direct protonation (Fig. 2.2C) or by proximal positively-charged residues (Figs. 2.3 and 2.10). This requirement demonstrates that the channel has robust charge selectivity. We propose this selectivity enforces directionality through electrostatic repulsion, thereby making it an integral feature of the charge-state Brownian ratchet translocation mechanism.

*Can ΔpH-driven Brownian ratchets generate robust unfoldase activity?*—Some formulations, (16, 24) but not all, (121) of the Brownian-ratchet mechanism imply that unfolding must occur prior to the engagement of the ratcheting process. In the unfolding-limited conceptualization of the Brownian-ratchet model, it is suggested that the channel must wait for the protein to unfold and the channel itself is not a participant in the denaturation. In this view, it is thought that the Brownian-ratchet mechanism cannot fully explain the large translocation rate accelerations observed for stable substrates, such as dihydrofolate reductase (122), barnase (123), or cytochrome $b_2$ (12).

This formulation of the Brownian-ratchet model requires some updating with more current reports (18, 34, 45, 67). The PA channel, for example, utilizes two principle unfoldase or denaturation sites within the channel, namely the **ϕ** clamp (45, 67) and α clamp (34). The former site is comprised of a ring of 7 or 8 Phe residues (depending on the oligomerization state of the PA complex), which can clamp onto the amino terminus of LF (67) and drive unfolding (45), presumably by binding to hydrophobic moieties in the substrate in a nonspecific manner (67). The latter α-clamp site is a deep cleft, situated between the twin $Ca^{2+}$-ion binding sites on the surface of the PA oligomer; this clamp is capable of binding nonspecifically to a ~10 residue α helix and short ~5 residue β strand (34). The initial characterization of the α clamp revealed an interesting capability of the site; that is, upon binding to the surface of the PA oligomer, $LF_N$ is partially unfolded. In this unfolding, LF's first α helix and β strand unfurl and dock into the α-clamp cleft. The α clamp also has broad binding specificity, allowing the site to recognize many types of sequences and making it a general denaturation site on the translocase. The second factor to consider is that ΔpH-driven translocation facilitates unfolding (45). Two different types of barriers have been identified in $LF_N$ translocation. The more force-dependent barrier (as determined by the *n* value, Eq. 1) corresponds with protein unfolding (45). Thus substrate unfolding may not be out of the realm of possibility for a Brownian-ratchet powered translocase, since the closely-spaced denaturation sites in the channel (α and ϕ clamps) allow for small displacements of sequence (small fluctuations) to be captured and stabilized via interactions with the channel.

*How might denatured-protein binding sites on the channel operate in conjunction with a ΔpH-driven ratchet?*—Brownian thermal fluctuations are a significant source of energy that, if

partially harnessed, could serve to denature a protein. For a polypeptide presequence (30 residues in length), Brownian thermal fluctuations can provide an upper limit of ~12 kcal $mol^{-1}$ of thermal energy (assuming 2 degrees of freedom per backbone Φ and Ψ angle per residue). Thus while the α-clamp structure maintains contact with the sequence carboxy-terminal to the presequence, the amino-terminal end may move through the channel purely motivated by Brownian thermal fluctuations. Successful excursions require that the presequence region is electrostatically compatible with the cation-selective lumen of the channel. We believe that this is ensured either through the lowered pH of the cis side of the membrane or by adjacent positive charges, which may act to help chaperone the acidic groups across the membrane. Successful excursions are then anchored or captured on the trans side of the membrane once the acidic groups further deprotonate in the higher pH medium. The high degree of correlation between the best sites to introduce a Glu residue in $LF_N$ $des^{(-)}_{1\text{-}46}$ and the naturally high density of negative charges in LF's presequence (Fig. 2.7C) suggests that this anchoring step is more favorable with higher acidic-residue densities.

Interestingly, we find that the acidic-residue-rich sequence is located immediately adjacent to the folded structure. Thus after successful anchoring of the presequence takes place, an entropic tension may be effectively applied to the remaining folded structure. This entropic tension is derived from the fact that a more extended polypeptide chain would have fewer possible conformations; and to relieve this entropic tension, the protein would be driven to unfold. We tested the positional dependence of this acidic residue cluster by creating an artificial gap between this naturally dense region and the folded structure of $LF_N$. When the gap contains neutral residues, the translocation rate decreases 50-fold, but when the gap contains charged groups, the translocation rate is similar to WT (Fig. 2.9B). Furthermore, the decrease in the translocation rate coincides with an observed stabilization of the substrate (Fig. 2.9C). Previous studies show that a large nucleus of structure unfolds during translocation in a large cooperative unfolding step (45); however, this unfolding occurs after a smaller portion of the amino-terminal structure is unfolded and docked into the α-clamp site (34). Therefore, we expect that the α- and ϕ-clamp protein denaturation sites in the channel can reduce the overall unfolding barrier by allowing for small incremental unfolding steps to be stabilized. Effective trapping of Brownian fluctuations may force partially unfolded intermediates to then disengage from the channel, leading ultimately to the larger-scale cooperative unfolding event (45) we observe under a ΔpH.

Finally, the charge requirements we identify for the substrate strongly favor a Brownian-ratchet model. However, we cannot rule out the possibility that ΔpH-driven transport also involves coordinated, proton-dependent movement of loops within the PA channel, which help to push the substrate chain during translocation. A combination of the active-pushing and Brownian-ratchet models may apply to this system since these models are not mutually exclusive. In fact, other translocases already known to push proteins using ATP-driven loop movements may also use a Brownian ratchet to further drive transport.

# Chapter 3

## Electrostatic ratchet in the protective antigen channel promotes anthrax toxin translocation[3]

### 3.1 Abstract

Central to the power-stroke and Brownian-ratchet mechanisms of protein translocation is the process through which nonequilibrium fluctuations are rectified or ratcheted by the molecular motor to transport substrate proteins along a specific axis. We investigated the ratchet mechanism using anthrax toxin as a model. Anthrax toxin is a tripartite toxin comprised of the protective antigen (PA) component, a homooligomeric transmembrane translocase, which translocates two other enzyme components, lethal factor (LF) and edema factor (EF), into the host cell's cytosol under the proton motive force (PMF). The PA-binding domains of LF and EF ($LF_N$ and $EF_N$) possess identical folds and similar solution stabilities; however, $EF_N$ translocates ~10- to 200-fold slower than $LF_N$, depending on the electrical-potential ($\Delta\psi$) and chemical-potential ($\Delta$pH) compositions of the PMF. From an analysis of $LF_N/EF_N$ chimera proteins, we identified two 10-residue cassettes comprised of charged sequence that were responsible for the impaired translocation kinetics of $EF_N$. These cassettes have nonspecific electrostatic requirements: one cassette surprisingly prefers acidic residues when driven by either a $\Delta\psi$ or a $\Delta$pH; the second requires basic residues only when driven by a $\Delta\psi$. Residue identity or specific position in the cassettes are not important as long as these basic charge requirements are met. Through modeling and experiment, we identified a charged surface in the PA channel responsible for charge selectivity. The charged surface latches the substrate and promotes PMF-driven transport. We propose an electrostatic ratchet in the channel, comprised of opposing rings of charged residues, enforces directionality by interacting with charged cassettes in the substrate, thereby generating forces sufficient to drive unfolding.

### 3.2 Introduction

Protein translocation is a fundamental molecular process required to transport proteins across membranes and to disassemble, denature, renature and/or degrade proteins within the cell (8, 124). Many biological events depend upon protein translocation (99), namely microbial toxin translocation into host cells (8, 18, 21, 34, 36, 45, 67, 124), toxin secretion (5), antigen presentation (3), membrane and organelle biogenesis (2), and retrograde transport of misprocessed proteins from the endoplasmic reticulum (3). The translocase machinery is generally, but not always (9), comprised of proteinaceous components. While diverse in mechanism, translocases utilize common driving forces such as ATP hydrolysis or the proton motive force (PMF) to provide the necessary energy for unfolding and translocation (8).

Brownian-ratchet (BR) (Fig. 3.1A) and power-stroke (PS) (Fig. 3.1B) mechanisms have been invoked to describe how molecular machines convert potential energy (the PMF or ATP) into useful work, such as unfolding and translocating proteins (124). The PS mechanism is believed to do work via a direct chemomechanical coupling of the energy source, whereas the BR mechanism does work by rectifying Brownian motion. In each case, the PS and BR mechanisms function via a cyclical dissipation of the potential energy source, creating repeated nonequilibrium fluctuations in the system. The substrate polymer is then directed to move in a

[3] This chapter is based on previously published work, with permission from co-authors (134).

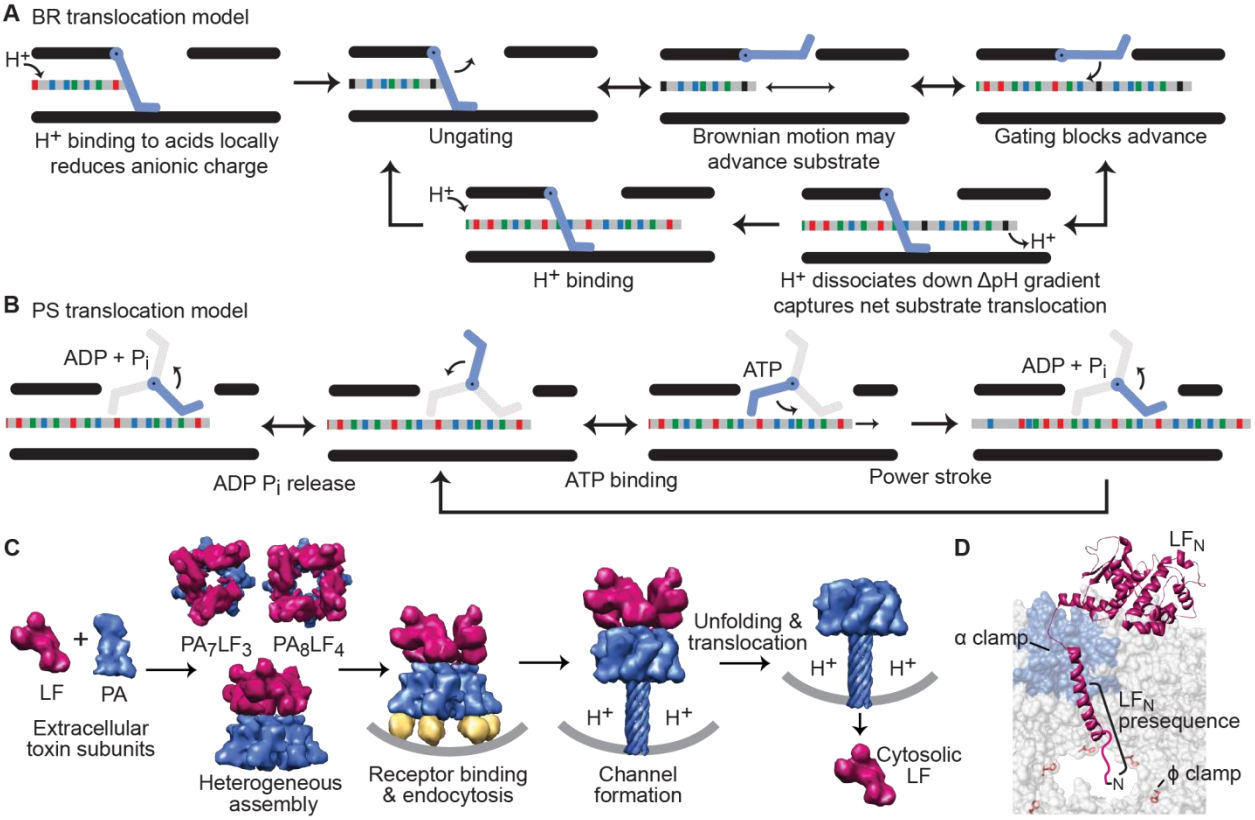


**Figure 3.1. Models of translocation.** Shown are (A) $H^+$-powered BR and (B) ATP-driven PS protein-translocation models. The translocating peptide has its residue chemistries colored: deprotonated acidic (red), protonated acidic (black), basic (blue), and hydrophobic (green). Dynamic gates and clamps that cycle in these systems are shown as steel blue. See text for details. (C) Assembly and translocation mechanism of anthrax lethal toxin. The components of lethal toxin, PA (steel blue) (31, 46) and LF (magenta) (32), assemble into heterogeneous oligomeric complexes, $PA_8LF_4$ (34) and $PA_7LF_3$, which are then endocytosed upon binding a receptor (gold). Acidification triggers the PA oligomer to form a translocase channel (35, 54, 87), and the ΔpH component of the PMF drives LF unfolding and translocation into the cytosol (18, 21, 45). (D) The PA oligomer (gray surface) facilitates LF (magenta) unfolding and translocation with several known polypeptide clamps. The α clamp (light blue surface) (31) binds nonspecifically to peptide helices and initiates $LF_N$ (magenta) translocation by binding to its first helix, α1, which is just carboxy-terminal to the modeled amino-terminal leader sequence leading into the central lumen. The ϕ clamp, a ring of 7 or 8 Phe427 residues (red sticks) in the PA oligomer, which is depicted here in the prechannel conformation to show its approximate location, then engages the amino-terminal leader sequence again through nonspecific interactions (67). These clamps may work in concert to bind and release substrate promoting unfolding and translocation (124).

unidirectional manner by means of some type of rectification or ratchet mechanism. The ratchet can be thought of energetically as an asymmetrical potential energy barrier that fluctuates as the energy source dissipates, or structurally, as a loop that forcefully pushes the peptide in one direction and/or biases against retrotranslocation. The molecular bases of these ratchet features are not well understood.

Anthrax toxin (8, 29, 124), the tripartite virulence factor secreted by *Bacillus anthracis* (the etiologic agent of anthrax), is ideally suited for biophysical studies probing the molecular mechanism of PMF-driven protein translocation (8, 18, 21, 34, 36, 45, 65, 67–69, 71) (Fig. 3.1C). Using electrophysiology, the electrical-potential ($\Delta\psi$) and chemical-potential ($\Delta$pH) compositions of the PMF can be externally controlled (18, 21, 34, 36, 45, 67–69). Lethal factor (LF) and edema factor (EF) are the two different ~90-kDa enzyme components of the toxin, which are translocated by the oligomeric channel formed by a third component, protective antigen (PA, 83 kDa).

To function, PA, LF, and EF must assemble into holotoxin complexes (Fig. 3.1C). PA is initially cleaved by a furin-type protease. The resulting 63-kDa PA subunits assemble into either heptameric ($PA_7$) (35, 38, 46) or octameric ($PA_8$) (34, 36, 54, 125) oligomers, or prechannels. $PA_7$ and $PA_8$ can bind up to three and four EF/LF moieties, respectively (34, 36). Crystal structures of LF (32), EF (33, 50), PA (46), the $PA_7$ (38) and $PA_8$ (36) prechannel oligomers, and the core of a $PA_8LF_4$ holotoxin complex (34) have been described. Once assembled, toxin complexes are endocytosed and trafficked to an acidic compartment in the cell, where PA converts to a cation-selective channel (42). The channel structure as resolved by electron microscopy (EM) (35) has a putative extended tubular β-barrel architecture (88, 89), analogous to the *Staphylococcus aureus* α-hemolysin toxin pore (126).

The narrowness of the PA channel requires that LF and EF unfold during translocation. Some destabilization of these proteins is imparted by the acidic conditions of the endosome (44). Interestingly, some unfolding occurs when LF and EF initially form a complex with the PA oligomer. In a recent crystal structure of the core of the $PA_8LF_4$ holotoxin, it was determined that the first α helix and β strand of the amino-terminal PA-binding domain of LF ($LF_N$) are unfolded and docked into a cleft, called the α clamp (34) (Fig. 3.1D). The α clamp is created at the interface of adjacent PA subunits, such that the deep cleft is framed by twin $Ca^{2+}$-ion binding sites (34). The α clamp is also a highly nonspecific binding site, and it can interact with diverse sequence chemistries, binding amphipathic and nonamphipathic helices with similar affinities (34). Detailed mutagenesis studies have shown that the most force-dependent step of the translocation mechanism coincides with the unfolding of the remaining structure of $LF_N$ (45). In fact, to cross the rate-limiting barrier, a significant portion of the amino-terminal β-sheet subdomain of LF is required to unfold (45). The unfolding process appears to also require another unfoldase active site, called the ϕ clamp (45, 67). The ϕ clamp is a ring of Phe-427 residues, which also bind nonspecifically to substrates that are dense in aromatic, hydrophobic, and cationic functional groups (67) (Fig. 3.1D). These two unique protein-denaturation sites in the PA channel (α and ϕ clamps) together favor the unfolding process. Though the mechanism is uncertain, these protein-denaturation sites are not thought to be traditional protein-binding sites; rather they are believed to be dynamic, coordinated, and ratchet-like, switching between high- and low-affinity states to promote directional motion, where binding at one clamp site can allosterically control binding at the other clamp site (8).

While translocation can be driven by either the $\Delta\psi$ (65) or $\Delta$pH (18), the $\Delta$pH is sufficient (21) and critical to the efficient translocation of the full-length enzymes, LF and EF

(18). A consensus picture is emerging that the underlying mechanism of ΔpH-driven translocation involves a charge-state BR (18, 21, 45, 68, 69, 71). Differences in the relative extent of protonation on either side of the membrane are believed to be able to bias Brownian fluctuations and impart directionality in the translocation mechanism. I have shown that acidic residues in a protein substrate are required for ΔpH-driven translocation (21). These residues are effectively the molecular teeth upon which an electrostatic-ratchet feature within the channel acts to produce forces during translocation.

An anionic charge requirement for ΔpH-dependent protein translocation may seem unusual, as the PA channel itself is strongly cation-selective (or anion-repulsive) (42). However, the protonation of acidic residues is likely required to make a portion of the translocating chain within the channel near neutral or slightly cationic. Doing so allows the protein to pass through the anion-rejection site of the channel by means of Brownian motion (Fig. 3.1A). Once the protonated portion of the translocating protein reaches the higher pH of the cytosol, these sites are more frequently deprotonated, becoming electrostatically incompatible with the channel. The same electrostatic feature that repels anion flux into the channel may then also act to ratchet and exclude retrograde efflux back into the channel. This rectification/ratchet feature is a critical aspect of BR- and PS-type molecular machines because it can bias nonequilibrium substrate fluctuations by limiting retrograde efflux.

Cycles of substrate protonation, Brownian motion, and deprotonation are likely required to pull the protein across the membrane. Analogously, with ATP-dependent systems, 100s of cycles of ATP binding and hydrolysis are required to unfold and transport a substrate protein. Several critical questions remain unanswered as to how this mechanism applies to protein translocation. What substrate sequence features allow for rapid translocation? What feature in the channel rectifies or ratchets Brownian motion and nonequilibrium fluctuations? How does the proposed charge-state BR mechanism develop forces sufficient to unfold substrate proteins? To address these questions, we investigated electrostatic requirements of the substrate and channel in PMF-driven anthrax toxin translocation. Our results and modeling studies are consistent with an electrostatic-ratchet translocation model.

### 3.3 Experimental Procedures

*Proteins*—Recombinant wild-type (WT) PA, $LF_N$, the amino-terminal PA-binding domain of EF ($EF_N$), and resulting chimeras and mutants were expressed and purified as described (36, 45). Assembly PCR was used to construct $LF_N/EF_N$ chimeras (21, 34). The amino-terminal six-histidine affinity tags ($His_6$) were removed from $LF_N/EF_N$ chimeras using bovine α thrombin (45). $PA_7$ prechannel oligomers were assembled as described (36). For the PA mutants $PA_{top}$ (containing the substitutions D276S, D335S, and E343S) and $PA_{bot}$ (containing the substitutions E302T, H304T, E308T, and H310T), and a WT PA control, 10 μg of each PA monomer were proteolyzed by 0.4 U furin (New England Biolabs) in 20 mM Tris-Cl pH 9, 150 mM NaCl and 1 mM $CaCl_2$ at room temperature. After 30 minutes, $LF_N$ was added at a 1:1 molar ratio, and following another 30-minute incubation at 25 °C, Fos-choline-14 was introduced to a final concentration of 2 mM to stabilize the PA oligomers in the channel form (127). Proper PA assembly was verified by native-PAGE, SDS-PAGE, and negative-stain EM.

*Electrophysiology*—Planar lipid bilayers were formed by painting (128) a membrane-forming solution (3% 1,2-diphytanoyl-*sn*-glycerol-3-phosphocholine in *n*-decane) across a 100 μm aperture in a 1-mL white-Delrin or polysulfone cup (34, 36, 45). A capacitance test confirmed the quality of the membrane. The membrane separated the cis and trans chambers,

each containing 1 mL of universal bilayer buffer (UBB; 100 mM KCl, 1 mM EDTA, 10 mM oxalic acid, 10 mM MES, 10 mM phosphoric acid). Ag/AgCl electrodes bathed in saturated 3 M KCl were linked to the chambers via 3 M KCl-agar salt bridges. PA currents were recorded with an Axoclamp 200B amplifier in CLAMPEX10.

*Translocation assays*—Bilayers were bathed in symmetrical UBB. $PA_7$ prechannels were added to the cis chamber (held at 20 mV), and conductance was blocked by the addition of substrate ($LF_N$, $EF_N$, or chimera) to the cis side (held at 20 mV in symmetric pH 5.6 experiments). The substrate blockade was >95% of the original current. Excess substrate was perfused by a hand-cranked, push-pull perfusion system. In Δψ-driven translocation assays, substrate translocation was initiated by increasing the Δψ. $\Delta\psi \equiv \psi_{cis} - \psi_{trans}$ ($\psi_{trans} \equiv 0$). Translocation activation energy ($\Delta G^{\ddagger}$) was computed by $RT \ln t_{1/2}/c$ (45). The $t_{1/2}$ value is the time for half the substrate to translocate; $c$ is a 1-sec reference; $R$ is the gas constant; and $T$ is the temperature. In ΔpH-driven experiments, the cis and trans chambers were bathed in UBB differing only in pH ($pH_{cis} = 5.6$; $pH_{trans} = 6.6$), where $\Delta pH \equiv pH_{trans} - pH_{cis}$. The Δψ was -1 mV during substrate blockade and perfusion. Translocation was initiated by increasing Δψ to 20 mV. Translocation records in either case were acquired across a range of Δψ values and repeated several times across multiple membranes ($n$ = 6 to 30).

*Equilibrium stability measurements*—Guanidinium chloride (GdmCl) titrations of $LF_N$, $EF_N$, and chimeras were carried out as described (44, 45) in 10 mM sodium phosphate, 1 M glucose, pH 7.5, at 20 °C. The stabilizing glucose additive was used to define the native-state baseline. Each titration point was monitored after reaching equilibrium by circular dichroism (CD) spectroscopy at 222 (±2) nm using a Jasco J-810 spectropolarimeter. The CD-probed curves fit to a four-state thermodynamic model ($N \leftrightarrow I \leftrightarrow J \leftrightarrow U$), where native ($N$), two intermediates ($I$ and $J$), and an unfolded ($U$) state are populated (44). We used the thermodynamic difference between the $N$ and $I$ states ($\Delta G_{NI}$) to assess the stability of the protein.

*Reversal potential ($\Delta\psi_{rev}$) measurements*—A planar bilayer was formed with the cis chamber bathed in 5 mM potassium phosphate, 100 mM KCl, pH 6.6 and the trans chamber bathed in unbuffered saline consisting of 100 mM KCl, pH 5.8. Assembled mutant and WT PA oligomer-$LF_N$ prechannel complexes were added to the cis side. Following channel insertion, the cis chamber was perfused thoroughly with fresh 100 mM KCl, making the system symmetrical, unbuffered KCl, pH 5.8. Residual $LF_N$ was then removed by applying a strong 100 mV Δψ to translocate it through the channel; and in some cases, a 1-unit ΔpH was established to aid in channel clearance of residual $LF_N$. Upon stabilization, a series of 50 µL aliquots of 3 M KCl were added to the cis side, and $\Delta\psi_{rev}$ was recorded as the Δψ required to drop the current to zero. All given KCl ratios of the two sides of bilayer have been corrected for activity in water (129), and following the experiment, the chambers were weighed to confirm their volume.

*Ensemble channel blocking*—A planar bilayer was formed with both chambers in 10 mM potassium phosphate, pH 6.6. The cis chamber alone had an additional 100 mM equivalent of KCl. Assembled mutant and WT PA oligomer-$LF_N$ prechannel complexes were added to the cis side, and the chamber was perfused following insertion. To remove remaining $LF_N$, 10 µL of 0.4 M phosphoric acid was added to the cis chamber to lower the pH to ~4.4, and a Δψ of 20 mV was applied. Afterward, the cis chamber was perfused with fresh pH 6.6 buffer, and the Δψ was returned to 0 mV. $LF_N$ was added to a given concentration and allowed equilibrate. The percent blockade was determined by the equilibrium drop in current following the addition of $LF_N$.

*Electron microscopy*—Preparations of $PA_{top}$, $PA_{bot}$, and a WT PA control were purified by anion exchange chromatography to remove residual PA monomer and excess $LF_N$. Fos-

choline-14 was only added to a concentration of 0.05 mM to avoid reaching the critical micelle concentration. All samples were diluted to an estimated final concentration of 70 nM (based on absorbance at 280 nm). Diluted complexes were incubated for 30 s on 400 mesh copper grids (Electron Microscopy Sciences) coated with continuous carbon on nitrocellulose, stained with 1% uranyl formate, and imaged with a Tecnai 12 TEM operated at 120 kV and at 49,000× magnification. Single particles were selected automatically using boxer (EMAN). Total numbers of particles (*n*) analyzed: PA WT (*n* = 4847), $PA_{top}$ (*n* = 4577), and $PA_{bot}$ (*n* = 4971). Class averages were determined iteratively using 10 successive cycles of Adapt, an automated classification program (written in house, see http://cryoem.berkeley.edu/adapt.html) and 2D multi-reference alignment in IMAGIC.

*Molecular models*—$EF_N$ and $LF_N$ domains from EF (PDB 1Y0V (33)) and LF (PDB 1J7N (32)), respectively, were α-carbon-($C_\alpha$)-aligned in CHIMERA (130). A three-dimensional model of the 14-stranded β-barrel region of the PA channel (residues 275 to 352) was made by coaxially stacking multiple copies of the heptameric β-barrel from α hemolysin (PDB 7AHL (126)). Peptide bonds were formed and residues were repopulated using COOT (131). The model was aligned to the *z* axis in CHIMERA (130). To obtain an electrostatic energy $U(z)$ as a function of the distance moved axially through the barrel *z* axis, we computed the sum of all pair-wise electrostatic energies in a PERL script (zforce.pl, which is available on request), using a 1-unit elementary point charge, $q_{test}$, moved along the center of the barrel in 0.1-Å increments, $U(z) = q_{test}\, b\, \Sigma\, q_i \cos\theta_i / d_i$, where $d_i$ is the distance between the $C_\alpha$ of the $i^{th}$ charged site within the channel of elementary charge, $q_i$, and $q_{test}$; $\theta_i$ is the angle between the charges and the *z* axis; and *b* is an electrostatic-energy conversion constant of 1390 kJ Å $mol^{-1}$.

**3.4 Results**

*$EF_N$ translocates slower than $LF_N$*—$LF_N$ and $EF_N$ share high levels of sequence (57) and structural homology (32, 33); however, the most divergent sequence homology occurs on the amino terminus (Fig. 3.2A). In planar lipid bilayer electrophysiology experiments, $LF_N$ and $EF_N$ translocate through the PA channel at remarkably different rates. While $LF_N$ translocates with a $t_{1/2}$ value of ~10 s at symmetrical pH 5.6 and a Δψ of 60 mV (18, 45), $His_6$-$EF_N$ translocates with a $t_{1/2}$ of ~140 s under identical conditions (36). The $His_6$ tag used in affinity purification tends to have modest effects on the translocation $t_{1/2}$ (21), and so we re-examined these translocation differences under two different driving-force extremes, a pure Δψ and a strong ΔpH, using the constructs in which the $His_6$ tag was removed by a protease. In our electrophysiological assay (18, 45, 67), a planar lipid bilayer separates two aqueous chambers (cis and trans). We first insert $PA_7$ channels into the bilayer. Either WT $LF_N$ or $EF_N$ is added to the cis side of the membrane (side to which $PA_7$ was added). Generally, an exponential decrease in current is observed as the substrate's amino-terminal presequence inserts into the ion-conducting PA channel (85). A brief perfusion removes excess substrate from the cis chamber, and translocation is initiated by changing the Δψ and/or ΔpH. The subsequent current increase results from substrate translocation to the trans side of the membrane, as determined by control experiments (18, 65). Two parameters are obtained from these "single-turnover" translocation records: the $t_{1/2}$ and the efficiency of translocation, which is equivalent to the fraction of substrate that successfully translocates. We note that there are multiple $LF_N$ or $EF_N$ bound to each PA complex so these translocation records likely represent the turnover of several substrates. Therefore, "single-turnover" kinetics refers to a single loaded PA complex that has translocated all of its substrates. We analyzed $LF_N$ and $EF_N$ translocation under identical conditions. Under a pure Δψ driving

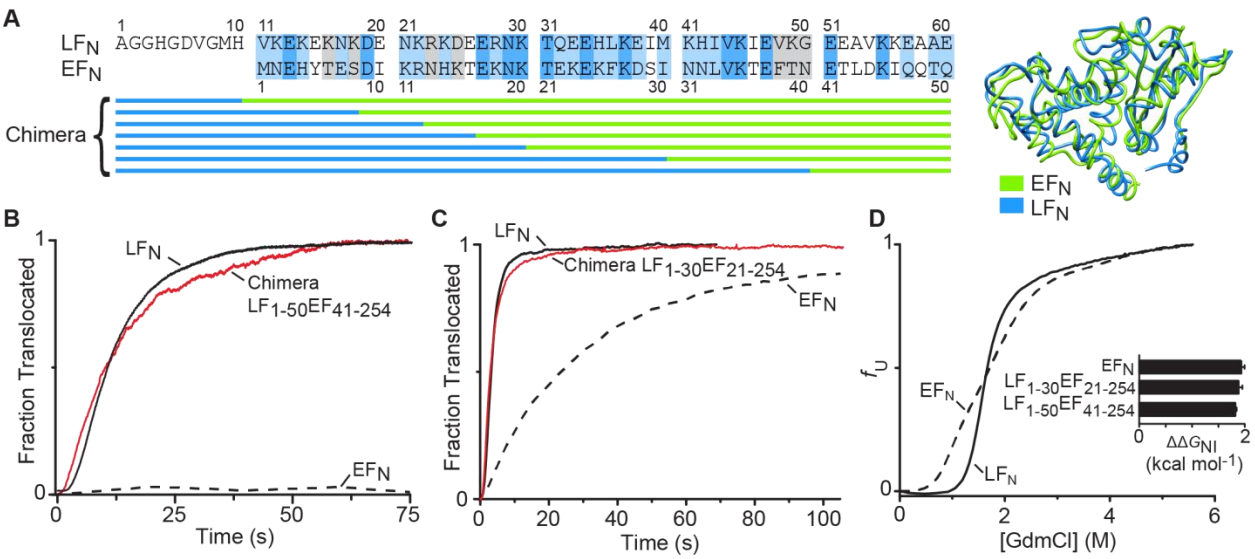


**Figure 3.2. $LF_N/EF_N$ chimeras are sufficient to mimic $LF_N$-like translocation kinetics.** (A) (left) Sequence alignment of the first 50 amino acids of $LF_N$ and $EF_N$. Residue pairs are shaded as follows: identity (blue), similarity (light blue), and weak similarity (gray). $LF_N/EF_N$ chimera constructs are shown below where the increasing amounts of amino-terminal sequence from $LF_N$ (blue) appended to the $EF_N$ carboxy-terminal folded domain (green). (right) $C_\alpha$-backbone alignment of $EF_N$ (1Y0V, green) and $LF_N$ (1J7N, blue) computed in CHIMERA (130). (B) Representative translocation recordings of $LF_N$ (black), $EF_N$ (dashed), and $LF_{1\text{-}50}EF_{41\text{-}254}$ (red) under a $\Delta\psi$ driving force (at symmetric pH 5.6, $\Delta\psi$ of 50 mV). (C) Representative translocation records of $LF_N$ (black), $EF_N$ (dashed), and $LF_{1\text{-}30}EF_{21\text{-}254}$ (red) under a 1-unit $\Delta$pH driving force ($pH_{cis}$ = 5.6, $pH_{trans}$ = 6.6, $\Delta\psi$ of 20 mV). Records in panels B and C are normalized to maximal expected fraction translocated. (D) Representative equilibrium denaturant titrations comparing $LF_N$ (solid) and $EF_N$ (dashed) in GdmCl (1 M glucose, pH 7.5, 20 °C) probed by CD at 222 nm and normalized to fraction unfolded ($f_U$). (inset) Equilibrium stability differences ($\Delta\Delta G_{NI}$) are referenced to WT $LF_N$ (where $\Delta\Delta G_{NI}$ compares $EF_N$ and chimeras to $LF_N$). For other chimeras, see Table 3.1. Errors are the mean (±S.D.) for $n$ = 3.

**Table 3.1. Equilibrium chemical denaturation stability fit parameters.**[1]

| Protein | $\Delta G_{NI}$ (kcal mol$^{-1}$) | $\Delta G_{IJ}$ (kcal mol$^{-1}$) | $\Delta G_{JU}$ (kcal mol$^{-1}$) | $\Delta\Delta G_{NI}$[2] (kcal mol$^{-1}$) |
|---|---|---|---|---|
| $LF_N$ | -5.68 (±0.01) | -2.46 (±0.06) | -5.00 (±0.03) | 0 |
| $EF_N$ | -3.24 (±0.08) | -2.62 (±0.01) | -5.02 (±0.03) | 2.44 (±0.08) |
| $LF_{1-10}EF_{1-254}$ | -3.03 (±0.04) | -2.46 (±0.01) | -5.03 (±0.03) | 2.65 (±0.04) |
| $LF_{1-18}EF_{9-254}$ | -3.22 (±0.05) | -2.66 (±0.02) | -4.66 (±0.05) | 2.46 (±0.05) |
| $LF_{1-22}EF_{13-254}$ | -3.00 (±0.05) | -2.47 (±0.02 | -4.97 (±0.04) | 2.68 (±0.05) |
| $LF_{1-26}EF_{17-254}$ | -3.38 (±0.08) | -2.85 (±0.03) | -4.29 (±0.04) | 2.3 (±0.08) |
| $LF_{1-30}EF_{21-254}$ | -3.30 (±0.05) | -2.55 (±0.02) | -4.65 (±0.04) | 2.38 (±0.05) |
| $LF_{1-40}EF_{1-254}$ | -3.54 (±0.03) | -2.43 (±0.02) | -4.69 (±0.04) | 2.14 (±0.03) |
| $LF_{1-50}EF_{41-254}$ | -3.39 (±0.05) | -2.64 (±0.02) | -4.92 (±0.03) | 2.29 (±0.05) |
| $LF_{1-22}EF_{13-254}$ N23D K25D T26E | -2.47 (±0.33) | -3.10 (±0.03) | -5.11 (±0.05) | 3.21 (±0.33) |
| $LF_{1-40}EF_{31-254}$ N41K T49K | -3.24 (±0.24) | -2.83 (±0.03) | -4.95 (±0.05) | 2.44 (±0.24) |
| $LF_{1-40}EF_{31-254}$ N41E N42D T49E | -3.4 (±0.3) | -2.86 (±0.03) | -5.06 (±0.05) | 2.30 (±0.32) |

[1]The four-state ($N \leftrightarrow I \leftrightarrow J \leftrightarrow U$) equilibrium free energy parameters, $\Delta G_{NI}$, $\Delta G_{IJ}$, $\Delta G_{JU}$, are generally obtained from fitting equilibrium denaturation experiments probed using CD as shown in Fig. 3.2D and described elsewhere. *N*, native; *I*, intermediate 1; *J*, intermediate 2; and *U*, unfolded state. The denaturant sensitivity for each thermodynamic transition, or *m* value, was generally fit using values consistent with previously published data, where $m_{NI}$, $m_{IJ}$, $m_{JU}$ are 3.59, 1.37, and 1.05, respectively.

[2]The equilibrium free energy differences between *N* and *I* ($\Delta\Delta G_{NI}$) are computed as $\Delta\Delta G_{NI} = \Delta G_{NI}(\text{mutant}) - \Delta G_{NI}(\text{WT } LF_N)$.

force, $EF_N$ translocated ~200-fold slower than $LF_N$ (Fig. 3.2B). Likewise, under a one-unit ΔpH, $EF_N$ translocated ~10-fold slower than $LF_N$ (Fig. 3.2C). Interestingly, previous studies (44) and our more recent thermodynamic analysis (Fig. 3.2D; Table 3.1) show that the equilibrium stability of $EF_N$, $\Delta G_{NI}$, is ~2.4 kcal mol$^{-1}$ less than $LF_N$ (44). As destabilization should in the most extreme case increase the rate of translocation due to the lowered unfolding barrier (45), it is unlikely that the weakened solution thermodynamic stability of $EF_N$ explains the observed increase in the activation energy of translocation relative to $LF_N$.

*Amino-terminal chimeras with $LF_N$ complement slow $EF_N$ translocation*—To determine the sequence differences responsible for the relatively slow translocation of $EF_N$, we created a series of chimera constructs (Fig. 3.2A). In these, we used the bulk of the $EF_N$ domain and only replaced the amino-terminal peptide with corresponding sequence from $LF_N$, where specifically 10, 18, 22, 26, 30, 40, or 50 $LF_N$ residues replaced equivalent positions in the $EF_N$ construct. (In our scheme, $LF_{1\text{-}a}EF_{b\text{-}254}$, *a* and *b* inclusively delimit the last residue of $LF_N$ and starting residue of $EF_N$, respectively.) We found that the $LF_{1\text{-}50}EF_{41\text{-}254}$ and $LF_{1\text{-}30}EF_{21\text{-}254}$ chimeras represented the minimal chimera constructs (Fig. 3.2B,C) of all tested chimeras (Fig. 3.3A,B) to exhibit $LF_N$-like translocation under a pure Δψ and a one-unit ΔpH, respectively. The sequence determinants that define the relatively slow translocation kinetics of $EF_N$ are found on its amino terminus. Thus the translocation kinetic stabilization we observe with $EF_N$ relative to $LF_N$ cannot be attributed to a phenomenon that occurs in solution (in isolation), but rather this difference manifests only in the context of the unfolding machine—the PA channel (Fig. 3.2D).

We then further explored the translocation differences of these chimeras under a variety of driving-force conditions. Under pure Δψ-driven translocation at symmetric pH, we found that the more $LF_N$ sequence introduced into the chimera, the faster the rate of translocation (Fig. 3.3A). Due to the complex nature of these ensemble translocation kinetics, a rate constant for translocation, *k*, was estimated using the $t_{1/2}$ for translocation, as $k \propto 1/t_{1/2}$, and from this we compute the $\Delta G^{\ddagger}$. Interestingly, we found the $LF_{1\text{-}10}EF_{1\text{-}254}$ chimera had similar to slightly slower translocation rates than $EF_N$ across many Δψ values (Fig. 3.3C), indicating that these additional 10 residues in $LF_N$ are not responsible for the observed differences in translocation. To effectively recapitulate the $LF_N$ Δψ-dependence curve, the $LF_{1\text{-}50}EF_{41\text{-}254}$ chimera was sufficient.

We then examined the set of chimeras under a 1-unit ΔpH gradient (Fig. 3.3B). Interestingly, the $LF_{1\text{-}10}EF_{1\text{-}254}$ and $LF_{1\text{-}18}EF_{9\text{-}254}$ chimeras showed slower translocation than $EF_N$ (Fig. 3.3D), indicating potentially that these sequences, which have more densely hydrophobic amino termini (Fig. 3.2A), may impede translocation due to the formation of an unusually tight binding interaction at the ϕ-clamp site. We found that the $LF_{1\text{-}30}EF_{21\text{-}254}$ chimera, however, was sufficient to completely restore $LF_N$-like translocation (Figs. 3.2C and 3.3D); and in contrast to purely Δψ driving forces, the sequence determinant for this restoration was concentrated between $LF_N$ residues 20 to 30.

*Two sequence cassettes modulate the translocation stability of $EF_N$ and $LF_N$*—A summary of the Δψ- and ΔpH-driven translocation results (Fig. 3.4A) identified two sequence regions of interest, or "cassettes": (i) the 20s cassette (residues 19-30); and (ii) the 40s cassette (residues 41-50) (Fig. 3.4B). (Note that because $EF_N$ is 10 residues shorter than $LF_N$ on the amino-terminal end, we are applying the $LF_N$-numbering scheme to $EF_N$.) Under symmetric pH conditions and a Δψ driving force, there is a ~1.3 kcal mol$^{-1}$ difference in $\Delta G^{\ddagger}$ between $LF_{1\text{-}18}EF_{9\text{-}254}$ and $LF_{1\text{-}26}EF_{17\text{-}254}$ in the 20s cassette (Fig. 3.4A). Under a one-unit ΔpH gradient, there is a ~2 kcal mol$^{-1}$ difference between the same chimeras (Fig. 3.4A). Also notable is the ~1.5 kcal mol$^{-1}$ $\Delta G^{\ddagger}$ difference between the $LF_{1\text{-}40}EF_{31\text{-}254}$ and $LF_{1\text{-}50}EF_{41\text{-}254}$ chimeras (Fig. 3.4A);

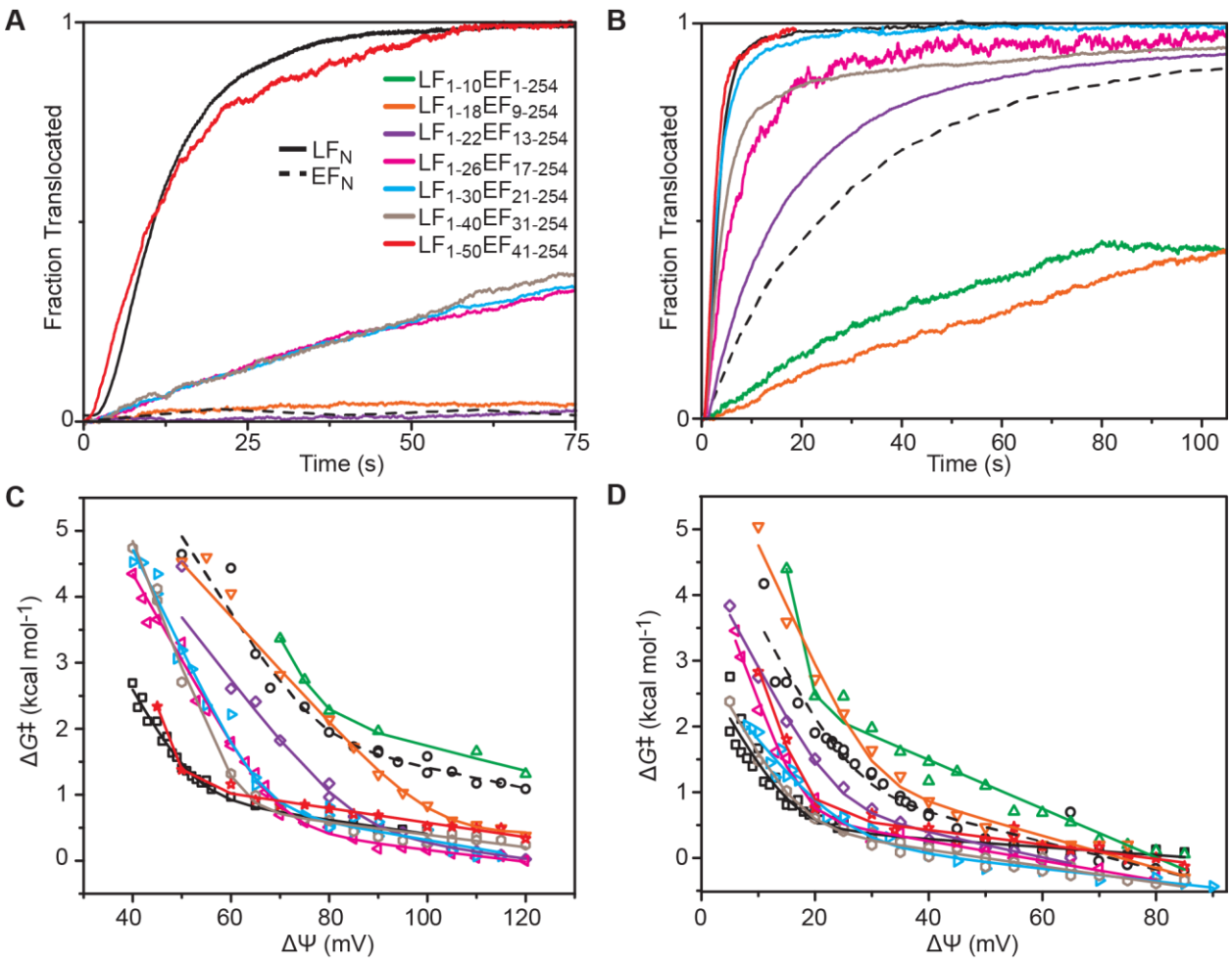


**Figure 3.3. Translocation Δψ and ΔpH driving-force dependencies for $LF_N/EF_N$ chimeras.** (A) Translocation records for $LF_N$, $EF_N$, and various chimeras at symmetric pH 5.6 and a Δψ of 50 mV. (B) Translocation records for $LF_N$, $EF_N$, and various chimeras under a 1-unit ΔpH ($pH_{cis}$ = 5.6; $pH_{trans}$ = 6.6) and a Δψ of 20 mV. (C) Translocation $\Delta G^{\ddagger}$ versus Δψ for the chimeric constructs shown in Fig. 3.2A. The pH was symmetric at 5.6. Data were fit to a two barrier model as described in Table 3.2, where the fit parameters are given. (D) Translocation $\Delta G^{\ddagger}$ versus a variety of Δψ values for the chimeric constructs under a 1-unit ΔpH ($pH_{cis}$ = 5.6; $pH_{trans}$ = 6.6). See Table 3.3 for fit parameters. Note the legend in panel A is used throughout.

**Table 3.2. Two-barrier fit parameters[1] for $\Delta G^{\ddagger}$[2] under a $\Delta\psi$ driving force[3].**

| Protein | $\Delta G^{\ddagger\circ}_1$ (kcal mol$^{-1}$) | $-z_1$ | $\Delta G^{\ddagger\circ}_2$ (kcal mol$^{-1}$) | $-z_2$ |
|---|---|---|---|---|
| $LF_N$ | 8.1 (±0.4) | 5.9 (±0.4) | 1.5 (±0.2) | 0.5 (±0.1) |
| $EF_N$ | 10.8 (±0.9) | 5.1 (±0.7) | 2.9 (±0.8) | 0.7 (±0.3) |
| $LF_{1\text{-}10}EF_{1\text{-}254}$ | 15.3 (±2.5) | 7.5 (±1.5) | 3.6 (±0.4) | 0.8 (±0.2) |
| $LF_{1\text{-}18}EF_{9\text{-}254}$ | 9.0 (±0.4) | 3.8 (±0.3) | 0.9 (±2.1) | 0.2 (±0.8) |
| $LF_{1\text{-}22}EF_{13\text{-}254}$ | 11.6 (±0.8) | 6.2 (±0.7) | 2.6 (±0.8) | 1.0 (±0.3) |
| $LF_{1\text{-}26}EF_{17\text{-}254}$ | 9.6 (±0.3) | 5.7 (±0.3) | 1.1 (±0.4) | 0.4 (±0.2) |
| $LF_{1\text{-}30}EF_{21\text{-}254}$ | 10.8 (±0.4) | 6.6 (±0.3) | 1.6 (±0.3) | 0.6 (±0.1) |
| $LF_{1\text{-}40}EF_{1\text{-}254}$ | 12.6 (±0.4) | 8.4 (±0.4) | 1.3 (±0.2) | 0.4 (±0.1) |
| $LF_{1\text{-}50}EF_{41\text{-}254}$ | 9.8 (±0.8) | 6.9 (±0.7) | 1.6 (±0.2) | 0.4 (±0.1) |

[1]Translocation data (from Fig. 3.3A) were fit to a two-barrier model, $\Delta G^{\ddagger} = RT \ln[\exp((\Delta G^{\ddagger\circ}_1 + z_1F\Delta\psi)/RT) + \exp((\Delta G^{\ddagger\circ}_2 + z_2F\Delta\psi)/RT)]$, where $R$, $T$, and $F$ are the gas constant, temperature, and Faraday's constant. $\Delta G^{\ddagger\circ}_1$ and $\Delta G^{\ddagger\circ}_2$ are the respective activation energies of the two barriers at a $\Delta\psi$ of 0 mV; and $z_1$ and $z_2$ are related to the amount of formal charge required to cross each respective barrier.

[2]$\Delta G^{\ddagger}$ values obtained from protein translocation experiments in an ensemble planar lipid bilayer assay using $\Delta G^{\ddagger} = RT \ln t_{1/2}/c$, where $R$, $T$, $t_{1/2}$, and $c$ are the gas constant, temperature, time for half of the protein to translocate, and an arbitrary constant of 1 s.

[3]The driving force was a $\Delta\psi$ with no $\Delta$pH (i.e. $\Delta$pH = 0). By convention, $\Delta\psi \equiv \psi_{cis} - \psi_{trans}$ ($\psi_{trans} \equiv 0$).

**Table 3.3. Two-barrier fit parameters[1] for $\Delta G^{\ddagger}$ under a $\Delta\psi$ driving force and a constant 1-unit $\Delta pH$[2].**

| Protein | $\Delta G^{\ddagger\circ}_1$ (kcal mol$^{-1}$) | $-z_1$ | $\Delta G^{\ddagger\circ}_2$ (kcal mol$^{-1}$) | $-z_2$ |
|---|---|---|---|---|
| $LF_N$ | 2.9 (±0.2) | 6.6 (±0.8) | 0.5 (±0.1) | 0.3 (±0.1) |
| $EF_N$ | 5.1 (±0.3) | 6.8 (±0.8) | 1.5 (±0.2) | 1.0 (±0.2) |
| $LF_{1\text{-}10}EF_{1\text{-}254}$ | 11.9 (±1.9) | 22 (±5) | 3.0 (±0.2) | 1.6 (±0.1) |
| $LF_{1\text{-}18}EF_{9\text{-}254}$ | 6.6 (±0.3) | 8.0 (±0.7) | 1.8 (±0.2) | 1.6 (±0.2) |
| $LF_{1\text{-}22}EF_{13\text{-}254}$ | 4.5 (±0.1) | 7.0 (±0.4) | 1.1 (±0.2) | 0.8 (±0.1) |
| $LF_{1\text{-}26}EF_{17\text{-}254}$ | 4.7 (±0.2) | 10.1 (±0.7) | 0.2 (±0.1) | 0.6 (±0.1) |
| $LF_{1\text{-}30}EF_{21\text{-}254}$ | 2.8 (±0.1) | 4.8 (±0.5) | 0.4 (±0.1) | 0.4 (±0.1) |
| $LF_{1\text{-}40}EF_{1\text{-}254}$ | 3.1 (±0.1) | 6.7 (±0.6) | 0.6 (±0.1) | 0.5 (±0.1) |
| $LF_{1\text{-}50}EF_{41\text{-}254}$ | 5.5 (±0.4) | 11.2 (±1.2) | 0.9 (±0.1) | 0.5 (±0.1) |

[1]Two-barrier model, $\Delta G^{\ddagger} = RT \ln[\exp((\Delta G^{\ddagger\circ}_1 + z_1 F\Delta\psi)/RT) + \exp((\Delta G^{\ddagger\circ}_2 + z_2 F\Delta\psi)/RT)]$, fit to data shown in Fig. 3.3B.

[2]The driving force was a variable $\Delta\psi$ and a constant 1-unit $\Delta pH$ (i.e. $\Delta pH = 1$).

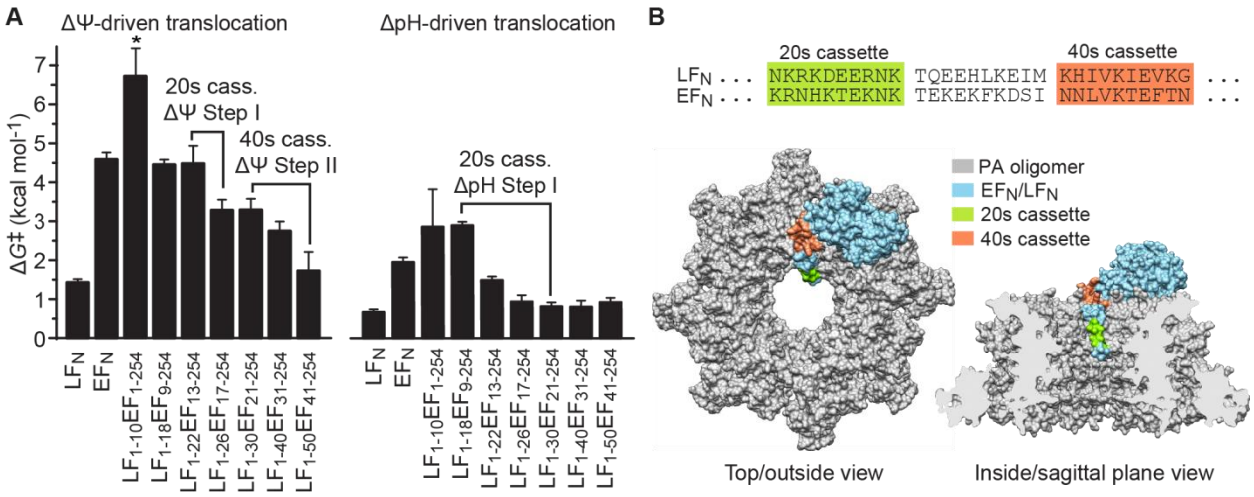


**Figure 3.4. Charged residues in the 20s and 40s cassettes utilize the Δψ and ΔpH driving forces to promote unfolding and translocation.** (A) (left) Translocation activation energy for chimeric constructs at symmetric pH 5.6 and Δψ of 50 mV. One value was estimated by extrapolation (*) based on a larger Δψ-dependent dataset (Fig. 3.3C) and associated fit parameters (Table 3.2). (right) Translocation $\Delta G^{\ddagger}$ for $LF_N$, $EF_N$, and indicated $LF_N/EF_N$ chimeras under a 1-unit ΔpH ($pH_{cis}$ = 5.6, $pH_{trans}$ = 6.6, Δψ of 20 mV). Brackets indicate significant differences (or "steps") in $\Delta G^{\ddagger}$ due to inclusion of the intervening $LF_N$ sequence cassette (cass.) Additional Δψ-dependent data at a 1-unit ΔpH are given in Fig. 3.3D, where associated fit parameters are given in Table 3.3. (B) (above) Amino-terminal 20s (green) and 40s (orange) cassette peptides are highlighted and the residue sequences in $LF_N$ and $EF_N$ are shown. (below) Top/outside and inside/sagittal plane vantages of a molecular model of $LF_N$ (blue) in complex with the $PA_8$ oligomer (gray) (3KWV (34)).

however, this difference was only observed under a $\Delta\psi$ driving force. Therefore, we hypothesize that sequence divergences in the 20s and 40s cassettes are responsible for the slow translocation kinetics of $EF_N$.

Mutations in these two sequence cassettes may have destabilized the chimera, thereby altering the height of the unfolding barrier for the translocation mechanism, as has been previously observed (21, 45). To test this possibility, we measured $\Delta G_{NI}$ of the base and most highly internally mutagenized chimera constructs using standard solution unfolding procedures (44, 45). We generally found no significant differences in $\Delta G_{NI}$ between these chimera and $EF_N$ (Fig. 3.2D; Table 3.1). As the bulk of the folded domain is from $EF_N$, this result was expected. The residues differing between the chimeras are contained in the amino-terminal unstructured region and first α helix and β strand, which are highly solvent-accessible. Thus we ruled out protein destabilization for these chimeras, and the amino-terminal sequence divergence in $EF_N$ likely affects the mechanisms of PA channel-dependent unfolding and translocation.

*Charge content of cassettes controls driving-force dependence of translocation*—In order to identify sequence features in the two cassettes contributing to the observed translocation $\Delta G^{\ddagger}$ differences, we introduced several point mutations within the existing chimera constructs (Fig. 3.5A,B). These mutations were made given the variation in net charge ($z$) observed within the cassettes. Net charge was estimated by $z = n_{basic} - n_{acidic}$, where $n_{basic}$ and $n_{acidic}$ are the number of basic and acidic residues, respectively. For the 20s cassette, we found that $EF_N$ and $LF_N$ had fairly different $z$ values of +5 and 0, respectively. Likewise, for the 40s cassette, $EF_N$ and $LF_N$ had $z$ values of 0 and +3, respectively. Upon our examination of their translocation kinetics, we found that correlations emerged between $z$ values within the cassettes and their translocation $\Delta G^{\ddagger}$ values (Fig. 3.5C,D). Thus as expected, the subtraction of positive charge in the 20s cassette and addition of positive charge in the 40s cassette, tended to generally increase the rate of translocation for $EF_N$-based chimeras.

We also examined the residue identity and position dependence of these effects. When we separately introduced an Asp at position 23 and 28 of $LF_{1\text{-}22}EF_{13\text{-}254}$ ($LF_{1\text{-}22}EF_{13\text{-}254}$ N23D, $z$ = +1; $LF_{1\text{-}22}EF_{13\text{-}254}$ K28D, $z$ = 0) (Fig. 3.5A), the rate of translocation increased relative to the parent construct ($z$ = +2) (Fig. 3.5C). Furthermore, both $LF_{1\text{-}22}EF_{13\text{-}254}$ K25D and $LF_{1\text{-}22}EF_{13\text{-}254}$ K25E ($z$ = 0) increased the translocation rate similarly, indicating that there is a general requirement for negative charge, but residue identity is not critical. In general when examining all the data, translocation rates were only affected by changes in $z$ values and not by changes in the position of the charges (Fig. 3.5C). The rate of translocation is similar for the $LF_{1\text{-}22}EF_{13\text{-}254}$ K25D T26E and $LF_{1\text{-}22}EF_{13\text{-}254}$ H24D K25N chimeras ($z$ = -1). Finally, the negative charge-neutralizing mutation $LF_{1\text{-}25}EF_{16\text{-}254}$ D25N ($z$ = +2) showed slowed translocation compared to its parent construct $LF_{1\text{-}25}EF_{16\text{-}254}$ ($z$ = +1). A similar but opposite effect can be seen in the 40s cassette, where there is a general requirement for positive charges independent of the specific positions (Fig. 3.5D). For example, $LF_{1\text{-}40}EF_{31\text{-}254}$ N41E and $LF_{1\text{-}40}EF_{31\text{-}254}$ T49E ($z$ = -1) had similarly decreased translocation rates relative to their parent chimera ($z$ = 0). Thus we conclude that the 20s and 40s cassettes indeed have particular anionic and cationic charge requirements, respectively, but these requirements are highly nonspecific in terms of both position and residue identity.

While most of the charge-dependent $\Delta G^{\ddagger}$ data for the 20s cassette is linear with respect to charge, the presence of outlier data at higher negative charge density led to the hypothesis that there may be two barriers in the charge-dependent transport mechanism. Increasing negative charge can lower one barrier; however, the second barrier is either charge insensitive or

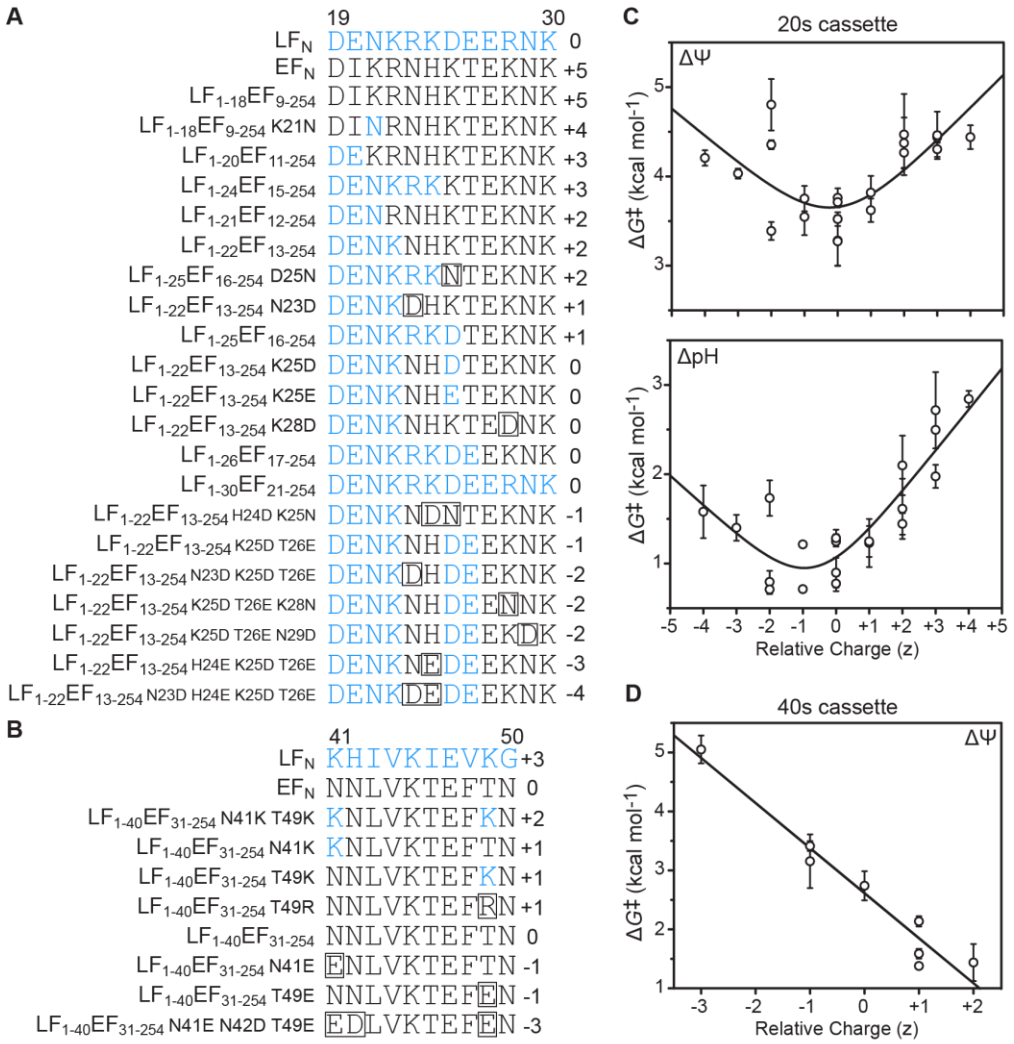


**Figure 3.5. Charged cassettes are nonspecific.** (A) Construct design for chimeras and derivative mutants in the 20s cassette (residues 19-30) are arranged from the most positive to the most negative. Net charge given to the right of each sequence is computed using the following scoring system: D, E = -1; H, K, R = +1. Residues from native $LF_N$ (blue) and native $EF_N$ (black) are shown alongside non-native mutations (boxed) to either $LF_N$ or $EF_N$. Residue-numbering scheme is according to $LF_N$ (32). (B) Constructs altering the 40s cassette (residues 41-50). Net charge is computed as in panel A. (C) (top) $\Delta G^{\ddagger}$ versus $z$ at symmetric pH 5.6, $\Delta\psi$ of 50 mV for $LF_N/EF_N$ chimeras and related mutants affecting the 20s cassette (residues 19-30 inclusive). Two-barrier model fit (Eq. 1): $\Delta G^{\ddagger\circ}_1 = 3.2$ (±0.5), $\delta_1 = -0.3$ (±0.2), $\Delta G^{\ddagger\circ}_2 = 3.2$ (±0.5), and $\delta_2 = 0.3$ (±0.1) ($n = 21$, $p < 0.001$). (bottom) $\Delta G^{\ddagger}$ versus $z$ at a $\Delta\psi$ of 20 mV, 1-unit $\Delta$pH ($pH_{cis} = 5.6$, $pH_{trans} = 6.6$) for the same 20s-cassette variants. Two-barrier fit parameters: $\Delta G^{\ddagger\circ}_1 = 0.3$ (±0.5), $\delta_1 = -0.7$ (±0.4), $\Delta G^{\ddagger\circ}_2 = 0.9$ (±0.3), and $\delta_2 = 1$ (±0.2) ($n = 21$, $p < 0.001$). (D) $\Delta G^{\ddagger}$ versus $z$ at symmetric pH 5.6, $\Delta\psi$ of 50 mV for $LF_N/EF_N$ chimeras and related mutants affecting the 40s-cassette region (residues 41-50 inclusive). Single-barrier model (Eq. 2) fit parameters: $\Delta G^{\ddagger\circ} = 2.7$ (±0.1) and $\delta = -0.58$ (±0.07) ($p < 0.001$). Error bars are the mean (±S.D.) ($n \geq 3$).

somewhat inversely dependent on negative charge. To allow for partial-charge character (δ) during each respective barrier crossing (132), we used the following model,

$$\Delta G^{\ddagger}(z) = RT \ln[\exp((\Delta G^{\ddagger\circ}_1 + \delta_1 zF\Delta\psi)/RT) + \exp((\Delta G^{\ddagger\circ}_2 + \delta_2 zF\Delta\psi)/RT)] \quad \text{(Eq. 1)}$$

where $F$ is Faraday's constant. For the ΔpH-dependent data ($n$ = 21), the fit to Eq. 1 was significant ($p < 0.001$) (Fig. 3.5C). The δ parameter was obtained for each barrier as $\delta_1$ = -0.7 (±0.4) and $\delta_2$ = 1.0 (±0.2). The corresponding activation energies, $\Delta G^{\ddagger\circ}_1$ and $\Delta G^{\ddagger\circ}_2$, in the absence of net charge were 0.3 (±0.5) and 0.9 (±0.3), respectively. For the Δψ-dependent translocation ($n$ = 21), the fit was also significant ($p < 0.001$) with $\delta_1$ = -0.3 (±0.2) and $\delta_2$ = 0.3 (±0.1) and $\Delta G^{\ddagger\circ}_1$ = 3.2 (±0.5) and $\Delta G^{\ddagger\circ}_2$ 3.2 (±0.5) (Fig. 3.5C). Typically, δ values are challenging to interpret: residues may be partially charged due to p$K_a$ shifts; metal ions may bind to the translocating peptide and alter net charge; and finally, only part of the charged region in the substrate may be required to cross the rate-limiting barrier. Nevertheless, the goodness of the fit suggests that indeed two unique charge-dependent barriers with inverse charge requirements are present in the translocation mechanism.

In the 40s cassette region (Fig. 3.5B), $LF_N$ possesses additional positive charge comparative to $EF_N$ at positions 41, 42, and 49. We created several point mutations in the existing chimeras to determine the effects of increasing or decreasing charge of the 40s cassette and investigated the charge-based differences in this region via translocation assays (Fig. 3.5D). Starting with a sequence similar to $EF_N$ and increasing its positive charge to that of $LF_N$, we again observe a direct relationship between charge and the translocation $\Delta G^{\ddagger}$, where increasing positive charge leads to faster translocation. Also the charge dependence was again nonspecific (Fig. 3.5D) where the position and identity of the residues did not appear to matter as much as the overall $z$ value (Fig. 3.5B). These charge-dependent data ($n$ = 8) for the 40s cassette were best fit by a single-barrier model (132),

$$\Delta G^{\ddagger}(z) = \Delta G^{\ddagger\circ} + \delta zF\Delta\psi \quad \text{(Eq. 2)}$$

The fit was significant ($p < 0.001$) with a δ of -0.58 (±0.07) and $\Delta G^{\ddagger\circ}$ of 2.7 (±0.1) (Fig. 3.5D). The type of cationic-charge preference in the 40s cassette is classical in the sense that it coincides with the direction of the electric field created by the applied membrane potential (i.e. the field is cis-positive).

*Electrostatic analysis of the PA β barrel*—Given the unusual preference for anionic residues in the 20s cassette when driven by a Δψ (which is exactly opposite of the result expected for a cis-positive membrane potential), we hypothesized that the local electrostatic field produced by features within the channel, $E_{chan}$, may override the electrical potential applied across the membrane, $E_m$. The overall electric field, $E$, is a vector, where $E = E_{chan} + E_m$. The force applied upon the translocating chain is related to the sign and magnitude of the charge, $q$, of groups in the translocating chain and $E$ by $E \times q$. Since the electrical field contributed by the membrane potential relates to Δψ as $E_m = \Delta\psi/d$, where $d$ is the distance over which the potential drops, we can assume that the membrane potential will contribute unproductively to a negatively-charged substrate if the membrane potential is positive in polarity. Therefore, $E_{chan}$ likely provides an oppositely-oriented electrical field component that can apply a productive force on the substrate that aligns with the productive direction of translocation. Our hypothesis is also supported by the fact that the anion-charge preference in the 20s cassette appears

independent of the makeup of the driving force; both Δψ-driven and ΔpH-driven kinetics can be accelerated by including additional negative charge in the 20s cassette of $EF_N$.

To characterize the electrostatic features within the PA channel, we initially built a model of the β-barrel portion of the PA channel using the coordinates of α hemolysin (126) (Fig. 3.6A). From this β-barrel model, we calculated the sum of all pairwise electrostatic potentials for a point charge translocated along the central axis of the channel (Experimental Procedures). Our analysis revealed two prominent and oppositely charged electrostatic features, which were juxtaposed in the β barrel. One is a strongly anion-repulsive feature (PA residue ranges 275-283 and 343-352, generally localized to the top of the β barrel), and the other is a strongly cation-repulsive feature (PA residue ranges 287-299 and 328-340, generally localized to the middle of the β barrel) (Fig. 3.6A). The PA residues contributing to these two features were located both inside and outside of the β barrel. Based on the same analytical model, we produced two β-barrel mutants, one that would disrupt the anionic feature and one that would not. $PA_{top}$ disrupted the upper, cis-most portion of the β barrel, targeting its negatively-charged residues by substituting them with isosteric Ser residues (D276S, D335S, and E343S). We chose Ser or Thr substitutions because the inside of the channel is hydrophilic and composed of mostly of Ser and Thr residues (44). $PA_{bot}$ disrupted the lower trans-most portion of the β barrel and channel via the similar isosteric Thr substitutions (E302T, H304T, E308T, and H310T). The modeled electrostatic effects of these two mutant PA β barrels are shown in Fig. 3.6A.

*The ion-selectivity filter of the PA channel is critical for Δψ- and ΔpH-driven translocation*—In order to characterize $PA_{top}$ and $PA_{bot}$, however, we first needed to properly assemble the monomeric PA into oligomers. The multisite mutations would not assemble using the traditional ion-exchange approach (42). Hence we developed a modified assembly procedure. We nicked the PA monomers at pH 9 with furin instead of trypsin (to avoid nonspecific tryptic degradation), co-assembled the PA at pH 9 by adding $LF_N$ (36, 54), and finally added Fos-choline-14 detergent to convert the prechannel oligomers into stable, detergent-solubilized channels (127). As a control, we also assembled WT PA by the same procedure. Native and SDS-PAGE (Fig. 3.7A) and negative-stain EM (Fig. 3.7B) verified the proper assembly of these samples. To monitor channel formation by planar bilayer electrophysiology, we had to remove the $LF_N$ *in situ* by perfusing the cis chamber and translocating the residual $LF_N$ through the channels. We found that the Fos-choline-14 favorably weakened the interaction of $LF_N$ with the channel, making its removal rapid and complete. In conclusion, the three preparations had reasonable insertion activities, albeit WT PA was optimal.

To determine whether these mutations change the ion selectivity of the PA channel, we first measured $\Delta\psi_{rev}$ for WT PA, $PA_{top}$, and $PA_{bot}$. ($\Delta\psi_{rev}$ is the voltage required to reduce the ionic current to zero under asymmetrical KCl gradients.) Each of these complexes was applied to planar bilayer membranes to form stable populations of channels following the removal of excess $LF_N$ by perfusion and translocation. The removal of residual $LF_N$ was judged to be complete by the stabilization of the current. Over a range of tested KCl gradients (in unbuffered saline, pH 5.8), WT PA and $PA_{bot}$ possessed similar $\Delta\psi_{rev}$ values and, therefore, possessed similar ion selectivity (Fig. 3.6B). However, $PA_{top}$ showed a reduced magnitude of $\Delta\psi_{rev}$ relative to WT PA (Fig. 3.6B). Thus $PA_{top}$ disrupts a portion of the channel's ion-selectivity filter, presumably by reducing its anionic-charge character (Fig. 3.6A).

$PA_{top}$ and $PA_{bot}$ were then assayed for their ability to translocate $LF_N$ under either a Δψ or a ΔpH. We found strong translocation deficiencies for $PA_{top}$ with either type of driving force (Fig. 3.6C). Under a 1-unit ΔpH ($pH_{cis}$ 5.6 to $pH_{trans}$ 6.6) with Δψ of 20 mV, translocation of $LF_N$

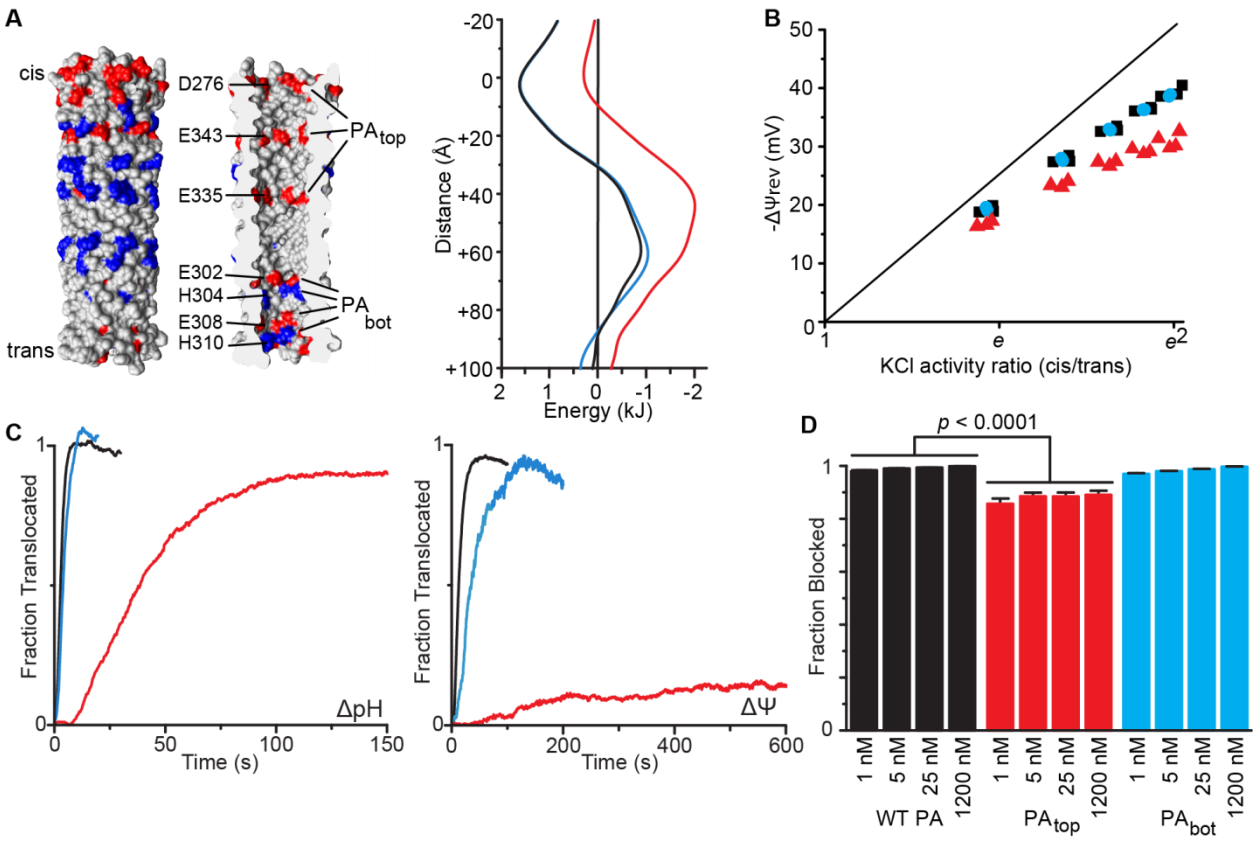


**Figure 3.6. Charge-selectivity filter in PA β barrel is required for efficient translocation.** (A) (left) Molecular model of the PA channel β barrel (gray), where acidic (red) and basic (blue) residues are highlighted. The outside and a sagittal section of the inside of the β-barrel structure are depicted. (right) The electrostatic energy for a negative point charge moved down the central axis of the β barrel of the channel. The origin on the distance axis is at the cis-most end of the β barrel, and increasing positive values indicates productive translocation. The potential was computed as described in the Experimental Procedures. (B) Relative differences in ion selectivity for WT PA (black squares), $PA_{top}$ (red triangles), and $PA_{bot}$ (blue circles) determined by $-\Delta\psi_{rev}$ versus the KCl activity ratio (cis:trans). The *x* axis is plotted as a natural-log scale marked by factors of *e*. The ideal cation-selective Nernstian relationship (*e*-fold activity ratio per 25.2 mV at 20 °C) is indicated with a solid line. Three independent measurements assessed on three different membranes were corrected for membrane and electronics offsets. (C) Representative protein translocation records for WT $LF_N$ under a (left) ΔpH and (right) Δψ using WT PA (black), $PA_{top}$ (red), and $PA_{bot}$ (blue). The Δψ and ΔpH conditions are identical to those applied in Fig. 3.2B,C. Results shown are consistent with replicates obtained on at least two separate membranes. (D) Ensemble bilayer recordings of WT PA (black), $PA_{top}$ (red), and $PA_{bot}$ (blue) channel conductance block by WT $LF_N$ at 1, 5, 25, and 1200 nM were obtained at symmetrical pH 6.6 and no Δψ. Error bars are the mean (±S.D.) ($n$ = 2). WT and $PA_{top}$ were tested for significance using an unpaired *t* test ($p < 0.0001$) for all observations ($n$ = 16) at each set of conditions.

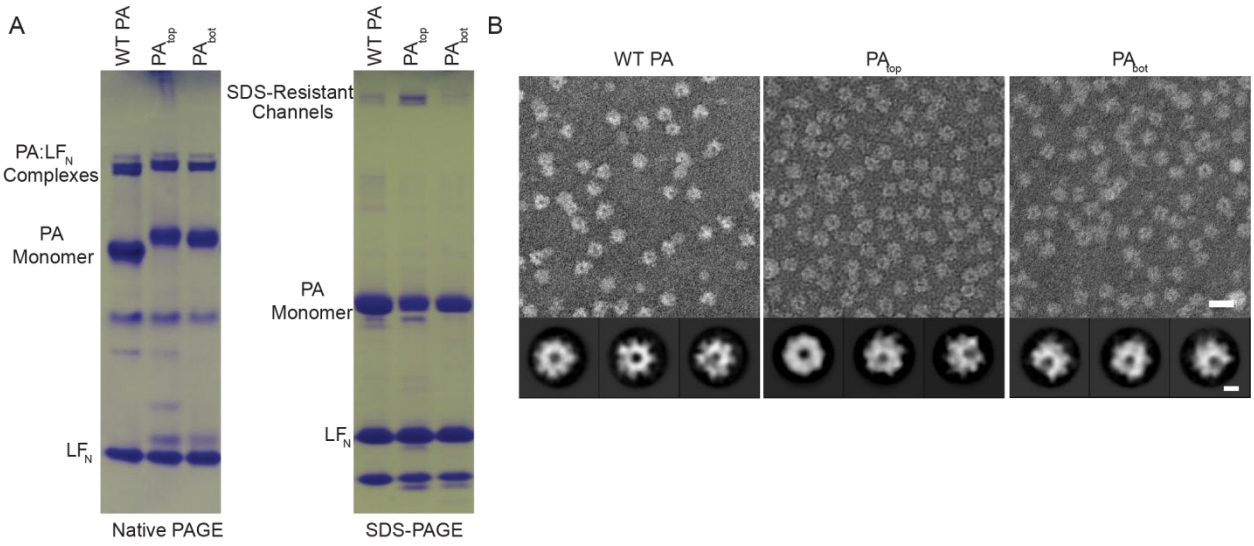


**Figure 3.7. $PA_{top}$ and $PA_{bot}$ form WT oligomeric structures.** (A) Native-PAGE (left) and SDS-PAGE (right) gels of oligomerized WT PA and mutants, $PA_{top}$ and $PA_{bot}$. Native-PAGE reveals consistent oligomerization and formation of PA-$LF_N$ complexes; and SDS-PAGE shows the formation of the SDS-resistant channel state. (B) Fields (top) and class averages (bottom) of negative stain EM images collected on WT PA (left), $PA_{top}$ (middle), and $PA_{bot}$ (right). All three complexes are capable of forming $PA_7$ and $PA_8$ complexes. White scale bars (lower right) are 20 nm and 10 nm for the fields and the class averages, respectively. All EM work was performed and analyzed by Gigi Kemalyan.

through $PA_{top}$ is slowed more than 10-fold compared to WT PA, while $PA_{bot}$ is unaffected (Fig. 3.6C left). With a 50 mV $\Delta\psi$ at symmetrical pH 5.6, $PA_{top}$ was also less able to translocate $LF_N$ relative to WT PA (Fig. 3.6C right). Under these conditions, the rate and efficiency of translocation were affected. While WT PA and $PA_{bot}$ are fully translocated within two minutes, $PA_{top}$ achieved less than 20% efficiency after ten minutes. Thus $PA_{top}$ reveals significant translocation deficiencies under either a $\Delta\psi$ or $\Delta$pH driving force.

Finally, $LF_N$ was assayed for its ability to block $PA_{top}$ and $PA_{bot}$ channels. In this experiment, we added 5 nM of $LF_N$ to the channels bathed in an asymmetrical KCl gradient at symmetrical pH 6.6 and a $\Delta\psi$ of 0 mV. Under these conditions, we found 99.0 (±0.1)% of WT PA channel current was blocked (Fig. 3.6D). For $PA_{bot}$, we observed 98.0 (±0.1)% conductance blockade; however, for $PA_{top}$, 88 (±1)% of the conductance was blocked by $LF_N$. The binding defect observed with $PA_{top}$ may indicate that the charge disruption in that region affects the ability of the amino terminus of $LF_N$ to properly dock inside the pore and block conductance. In this model (Eq. 3), we expect two different stages of binding. In stage one, $LF_N$ binds to the top surface of the channel, forming the ($PA \cdot LF_N$) complex; and in stage 2, the amino terminus docks into the channel to block conductance, forming the $(PA \cdot LF_N)^*$ complex.

$$PA + LF_N \leftrightarrow (PA \cdot LF_N) \leftrightarrow (PA \cdot LF_N)^* \qquad \text{(Eq. 3)}$$

To test whether stage 1 or stage 2 was affected by the $PA_{top}$ mutation, we determined the percent blockade as a function of $LF_N$ concentration. While the concentration of $LF_N$ should affect the equilibrium of stage 1, the equilibrium describing stage 2 is, of course, concentration-independent. To test for these two possibilities, we altered the $LF_N$ concentration. Reducing the concentration to 1 nM resulted in small changes in channel blockade (PA WT: 98.4 (±0.1)%; $PA_{bot}$: 97.0 (±0.3)%; $PA_{top}$: 86 (±2)%). However, increasing the concentration five-fold to 25 nM did not appreciably change the blockade (PA WT: 99.3 (±0.0)%; $PA_{bot}$: 98.7 (±0.2)%; $PA_{top}$: 88 (±1)%), indicating that the system is at saturating levels of $LF_N$. Indeed, even increasing the concentration to 1.2 μM did not appreciably affect the percent block (Fig. 3.6D). The inability of $LF_N$ to fully saturate channel conductance blockade in the $PA_{top}$ mutant over a 1000-fold concentration range demonstrates that channel docking (stage two) is impaired, and the $PA_{top}$ mutation likely disrupts a latching or ratcheting feature within the PA channel.

### 3.5 Discussion

*General substrate charge requirements*—To address the molecular mechanism of PMF-driven translocation, we traced the source of the differences in the translocation kinetics between $LF_N$ and $EF_N$. Previous translocation studies (36, 45) and our more-controlled re-examination here show that $EF_N$ translocates ~200-fold slower than $LF_N$ under a $\Delta\psi$ alone and ~10-fold slower than $LF_N$ under a combined $\Delta\psi$ and $\Delta$pH (Fig. 3.2B,C). This phenomenon occurs despite the fact that $LF_N$ and $EF_N$ have ~55% sequence similarity, adopt identical folds (32, 33), possess similar solution stabilities (Fig. 3.2D) (44), and bind to the same location on the PA channel (34, 57). Interestingly, while LF and EF initiate translocation starting from the amino termini of their homologous $LF_N$ and $EF_N$ domains, the amino-terminal initiation sequence of these domains is the most divergent sequence in the domain. We anticipated that this region of the sequence was responsible for the differences we observed in their translocation kinetics. Swapping the 40-residue amino terminus of $EF_N$ with the homologous 50-residue amino terminus from $LF_N$ allows the chimera to translocate as rapidly as $LF_N$. The inability of $EF_N$ to utilize the PMF as well as

$LF_N$ is hence due to sequence differences in the amino-terminal presequence, and therefore, the charged presequence is critical to allowing the substrate to best capture the PMF to drive unfolding and translocation.

Within the presequence, we were then able to locate two sequence cassettes, or motifs, required for efficient translocation (Fig. 3.4B). When additional acidic residues are added within the 20s cassette of $EF_N$, its translocation becomes more $LF_N$-like. My previous studies have shown that under a ΔpH driving force, acidic residues are needed in the 20s cassette for efficient translocation (21), and while the studies here support prior observations, they also show that higher acidic-residue content in the 20s cassette is favorable under a pure Δψ. Hence the acidic-residue-dependent mechanism we observe is independent of the nature of the driving force. This dependence, at first glance, is most unusual since it is opposite to the effect expected for a cis-positive Δψ, and we will expand on this point in detail below. But from this unusual charge requirement, we expect that the electrostatics of the channel itself govern the overall mechanism. We also identified a 40s cassette in the presequence and found it prefers cationic residues. This preference in the 40s cassette is only observed under a pure Δψ driving force, and while the 40s cassette is a novel sequence feature, it was expected to exist since a productive Δψ driving force is cis-positive.

*Broad sequence specificity in protein translocases*—The broad sequence specificity we observe for these charged cassettes (Fig. 3.5C,D) is similar to the binding preferences of other polypeptide-clamping sites in the PA channel and in other systems (124). This observation is the case during translocation for several reasons that invite analogy to Levinthal's paradox. Levinthal originally stated that a folding protein would be unable to sample all the possible configurations of the unfolded state in a reasonable timescale, and instead, proteins must fold via a specific pathway (133). Similarly, a translocating chain cannot fully explore its vast conformational space during the timescale of translocation, and the translocase cannot individually recognize the enormous combinatorial sequence space of the substrate's primary amino acid sequence. As the hydrophobic effect, for example, is likely a key feature that guides many folding pathways, the PA channel interacts with general charge patterns in these cassettes rather than their specific sequences. The broad sequence specificity observed for these two particularly important charge cassettes is a testament to the channel's overall ability to recognize a broad range of sequences. It is reminiscent of the ϕ clamp's broad recognition of hydrophobic and aromatic sequences (67) and the α clamp's general shape complementarity to α helices (34). Hence, we propose that broad sequence specificity is key for a protein translocase, since it must process unfolded protein, which may otherwise occupy too many possible states.

The charged cassettes we report here again have general electrostatic requirements, but the specific details are far less critical. Other examples of these nonspecific clamping sites in the PA channel include the α clamp and the ϕ clamp. The ϕ clamp prefers hydrophobic and aromatic substrates (67), whereas the α clamp binds most optimally to α-helical structure with minimal sequence specificity (34). Each clamping site binds broadly to a different type of chemical handle in the translocating chain, where specific hydrogen bonds and salt bridges are non-critical. Polypeptide clamps are critical because forces cannot be applied to the unfolding substrate protein without a fulcrum. Also, competing diffusive and entropic forces in the system scale with the size of the unfolded state; i.e., for an *N*-residue-long unfolded chain where each residue can sample an average of *C* conformations, the total number of potential configurations scales as $C^N$. Limiting the size of the unfolded chain that is freely diffusible through nonspecific clamping allows the force-generating apparatus to focus more efficiently on producing

directional motion and mechanically unfolding the substrate. The downside to nonspecifically clamping the chain becomes immediately apparent; because when interactions are too tight translocation should become impeded. However, we have proposed instead that clamping sites are dynamic, and the chain is continually bound and released during translocation, and hence such events would reduce diffusive entropic costs, improve energy transduction and force generation, and lower the overall barriers to translocation.

*Role of channel electrostatics in translocation*—My previous work has shown that sites within the 20s cassette of $LF_N$ were optimal for the placement of acidic residues when translocation is driven by a ΔpH (21). The key finding in this report is that $EF_N$ chimeras also require additional acidic-residue density in the 20s cassette; however, this requirement for more rapid translocation kinetics holds even under a pure Δψ driving force. The requirement is counterintuitive because the relationship expected between a purely Δψ-driven process and charge should rather be a preference for cationic residues. Since the acidic-residue requirement in the 20s cassette is driving-force independent, we surmised that the electric field acting on the negatively-charged region is not purely derived from the Δψ (as that would create forces opposite in sign to productive translocation) but rather from charged residues residing inside the PA channel.

Simplified electrostatic modeling of the PA channel β barrel reveals two strong oppositely charged electrostatic barriers/wells are present depending upon the identity of the test charge used (Fig. 3.6A). We started with the β barrel because the structure is well supported by numerous studies (35, 88, 89). The electrostatic features we identified in the β barrel are produced by residues pointing into the lumen of the barrel and residues on the outside of the barrel. We mutated various residues in the β barrel in clusters to investigate their role in the translocation mechanism. The contribution of these charged-residue mutations is, of course, amplified by the 7- to 8-fold nature of the oligomer. Based on our electrostatic modeling, $PA_{bot}$ (which removed 4 charges per monomer, 2 positive and 2 negative) will have very modest effects on the electrostatic energy landscape; however, $PA_{top}$ (which removed 3 negative charges) is expected to diminish the anion-repulsive barrier (Fig. 3.6A). We hypothesized that this would shift the ion selectivity and confirmed this to be true by measuring a reduction in $\Delta\psi_{rev}$ for $PA_{top}$ relative to WT PA and $PA_{bot}$ (Fig. 3.6B). This result implies that this region is part of the ion-selectivity filter. It should also be stated that other reports have implicated the ϕ-clamp site as a key electrostatic filter central to ΔpH translocation, albeit it is unclear what charged residue comprises the ϕ-clamp filter itself (71). We report here that when the charge-selective filter is removed from the $PA_{top}$ mutant, both substrate docking and translocation are defective (Fig. 3.6C,D). The inability to properly dock $LF_N$ argues that a clamping or latching feature in the channel is disrupted in the $PA_{top}$ mutation, and we suspect this element in the top of the PA β barrel is a key piece of the electrostatic ratchet expected in our BR model.

*Model*—Our BR model (Fig. 3.8) suggests that ion selectivity plays an important role in PMF-driven translocation (18, 21). We expect that a polypeptide chain can pass through the anion-repulsive charge filter once it is partially protonated by the lower cis pH. As this chain moves through the charge filter, the chain may deprotonate in the higher trans pH and become net repulsive to the charge filter. Such changes in protonation state may also occur in the channel itself, since the residues we have identified in the $PA_{top}$ mutant are also acidic, and this change would favor the proposed model. At this stage, the filter acts like a ratchet and holds the chain in way that limits retrotranslocation. An entropic tension develops in the leading sequence and favors further substrate unfolding of the lagging folded domain (21). In our current model, based

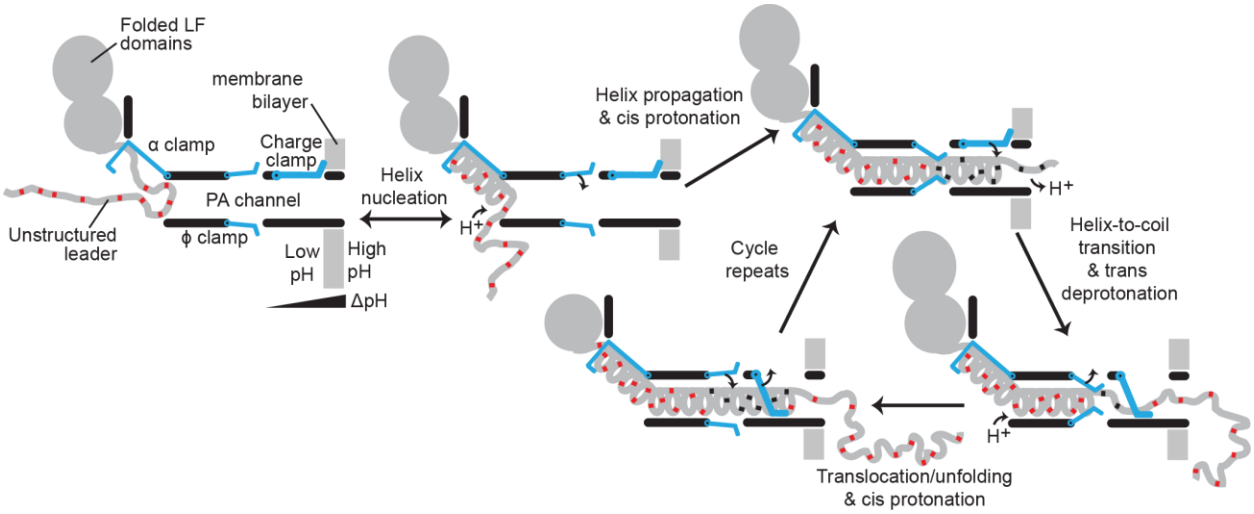


**Figure 3.8. Electrostatic ratchet model.** A cartoon model of the PA channel (black outline) with indicated α-, ϕ-, and charge-clamp sites (blue moveable gates) based on results described here and elsewhere (18, 21, 34, 45, 67, 124). The folded substrate domains from LF are indicated as gray circles on the top surface of the channel, where its amino-terminal leader sequence is shown as a thick gray line. The α clamp may nucleate helical structure into the channel, where the ϕ clamp can grip the amino-terminal leader. Protonation of the peptide on the lower pH side (cis protonation) converts acidic, charged residues (red squares) to neutral ones (black squares), allowing for the leader to move past the charge-clamp site via Brownian motion. Deprotonation of these acidic residues on the higher pH side (trans deprotonation) and an accompanying helix-to-coil transition in the leader are thermodynamically favorable and result in further translocation. The deprotonated sequence is, however, unable to retrotranslocate due to the charge-clamp site. Entropic tension in the upstream folded substrate maintained by the clamp sites leads to domain unfolding. Further cycles complete the translocation of the remaining domains.

upon the recent discovery of the helix stabilizing cleft, the α clamp, we propose that helical structure can be stabilized inside of the channel. The transition from helix → random coil is highly favorable entropically, and thus could generate an entropic force that pulls the upstream peptide through the channel. Some coordination with the ϕ clamp site is evident in prior studies, and hence dynamics at the ϕ clamp site may be required for coordinated peptide movement or protonation state changes in the system (18). Brownian motion likely underlies the transitions in this system, especially when particular electrostatic barriers are lowered upon protonation/deprotonation cycles. Such diffusive motion is critical to driving the overall helix-to-coil transition we have proposed. This process can repeat in subsequent sequences and domains until translocation is complete. In fact, LF has cassettes in the 350s and 370s with similar charges and spacing as the 20s and 40s cassettes. Further investigation into the role of these cassettes is warranted.

It is tantalizing to point out that there is also a cation-repulsive site downstream of the anion-repulsive site in the β barrel. This cation-repulsive site will be stabilizing, however, to the formation of deprotonated Glu and Asp residues, favoring their deprotonation effectively. Such an activity would reinforce our BR model. The energy landscape we have computed is consistent with the biphasic nature of the $\Delta G^{\ddagger}$-versus-charge relationship observed in the 20s cassette (Fig. 3.5C). One barrier prefers negative charge and the other prefers positive charge in the region. Based on these electrostatic features, the channel may hold amino-terminal poly-cationic substrates, such as $His_6$ tags, (86) at low driving forces in a peptide-clamped or conductance-blocked stage indefinitely without actually translocating the substrate until a higher cis-positive potential is applied (18, 21, 67, 85). Many phenomena involving the amino-terminal presequences of LF, EF, and other heterologous substrates likely derive their origins from their interactions with the highly-charged β barrel.

# Chapter 4

## Role of α-helical secondary structure in the protein translocation mechanism of anthrax toxin

### 4.1 Abstract

Anthrax toxin, a model for studying transmembrane protein translocation, also can serve to investigate translocation-coupled unfolding. The toxin's channel-forming component, protective antigen (PA), forms a homooligomer that assembles into toxic complexes with the two enzyme substrates, lethal factor (LF) and edema factor (EF). Upon endocytosis into host cells, acidification of the endosome induces PA to form a membrane-inserted channel, which unfolds LF and EF and translocates them into the host cytosol. Translocation is driven by the proton motive force, comprised of the chemical potential, the proton-gradient (ΔpH), and the electrical potential (Δψ). The ΔpH is critical for translocating the full length enzyme substrates, LF and EF. While the proposed mechanism for utilizing the ΔpH is consistent with an electrostatic ratchet, it remains unclear whether the substrate must fully unfold to a random coil or whether it maintains secondary structure during translocation. A recent crystal structure of the lethal toxin core complex revealed an "α clamp" structure that binds to substrate helices nonspecifically. One hypothesis is that the α clamp nucleates helix formation in the substrate's amino terminus, and the substrate moves through the channel more efficiently when it is able to form a helical intermediate state. We report that the α clamp more stably binds substrates that can form helices than those (such as polyproline) that cannot, and hence the α clamp recognizes substrates by a general shape-complementarity mechanism. To further investigate the function of the α clamp, we produced a synthetic PA mutant in which a helix was crosslinked into the α clamp to block its function. This synthetic construct impairs translocation by raising a previously identified, yet uncharacterized, energy barrier shown to be much less force dependent than the known major unfolding barrier. Also, substrates that are incapable of forming compact secondary structure (due to the introduction of either a polyproline II helix or D-amino acids) are severely deficient for translocation. Therefore, polypeptides do not efficiently translocate in an extended and unstructured state, suggesting their ability to form helical structure may be critical to the overall translocation mechanism.

### 4.2 Introduction

The secretion (99) and degradation (100) of proteins are essential for a variety of processes including protein trafficking (99), membrane and organelle biogenesis (2), microbial toxin secretion (5) and subsequent entrance into host cells (3, 124), antigen presentation (3), and destruction of damaged proteins (3). These cellular processes require large and complex molecular machines that are typically composed of multiprotein complexes, though this is not always the case (9). These multiprotein complexes, called translocases, form aqueous pores in lipid bilayers thereby allowing other substrate proteins and peptides to be translocated across the membrane. Of course, translocases require substantial energy inputs, usually in the form of ATP binding and hydrolysis or the proton motive force (PMF) (3, 8, 100).

Anthrax toxin, one of two key virulence factors produced *Bacillus anthracis* (30), the causative agent of anthrax disease, is an ideal model for studying transmembrane translocation (18, 21, 34, 36, 45, 65, 67–69, 71, 85, 134). It consists of three separately nontoxic proteins that must associate to make toxic complexes. The three proteins include an 83 kDa homooligomeric

channel-forming protein, protective antigen (PA), and two ~90 kDa enzyme effectors, lethal factor (LF) and edema factor (EF). First, PA is proteolytically activated to form a 63 kDa activated form, which assembles into heptameric or octameric prechannels that bind respectively three or four moieties of LF or EF (34, 36). These toxic complexes are endocytosed, and the acidification of the compartment drives prechannel-to-channel conversion, substrate unfolding (44, 45), and translocation into the host cytosol (18, 21).

The emerging consensus is that the proton gradient (ΔpH) component of the PMF created by the endosomal acidification drives substrate translocation via a charge-state Brownian ratchet (18, 21, 68, 69, 71, 134). In brief, acidic residues in the substrate are more frequently protonated on the low pH endosomal side of the membrane, resulting in a net positive stretch of peptide that is allowed to move bidirectionally, according to Brownian motion, through the cation-selective channel. When the translocating chain reaches the higher pH cytosol, the acidic residues are more frequently deprotonated, yielding an anionic region no longer capable of retrotranslocation. In this way, movement is resolved in one direction, and productive translocation occurs until the entire substrate reaches the cytosol. However, it is not immediately clear how a system that does not involve the application of an outside force can drive protein unfolding.

Our understanding of how the substrate unfolds is complicated by our lack of knowledge about the degree of folding, i.e. is secondary structure maintained in the translocating chain? Recently, the crystal structure of the octameric prechannel bound to LF's PA-binding domain ($LF_N$) was reported (34). It revealed that the α1 helix and β1 strand of $LF_N$ unfold and bind a hydrophobic grove near the top of the PA channel, termed the α clamp. Mutational analysis of the substrate revealed that the α clamp will bind non-specifically to helices with a wide variety of side chain size, charge, and polarity. Additionally, the clamp is positioned in such a way that extending the substrate's α helix would orient it toward the channel lumen. We propose that the α clamp functions as a helix nucleation site, in accordance with the Zimm-Bragg formalism (94). In this report, we show how occlusion of the α clamp and using substrates incapable of forming compact states alter translocation rates and efficiency, consistent with obligate helical translocation. We present a model for how helical translocation could be a necessary step in driving tertiary unfolding.

### 4.3 Experimental Procedures

*Proteins*—Recombinant PA, LF, and mutants were expressed and purified as described (36, 45). Assembly PCR and/or QuikChange were used to construct mutants (21, 34). Where indicated, the $His_6$ tags were removed from substrates using bovine α thrombin (45). $PA_7$ prechannel oligomers were assembled as described (36). Peptides were synthesized and purified to >95% (Elim Biopharmeceuticals). The 10mers had the sequence KKKKKWWSWW, where either all residues were L or alternating with D (odd-numbered residues) and L (even).

*Racemization*—Chemical racemization was carried out as described (135). In brief, 200 μM $His_6$-$LF_N$ was incubated with 100 mM NaOH, 1 mM sodium bisulfite, and approximately 5 M guanidinium chloride at 85°C in for various times up to 90 minutes. Time courses for racemization were monitored by measuring CD spectra on a Jasco J810 spectropolarimeter from 200 to 260 nm in a 1 cm pathlength cell. For each time point, 10 μL of sample was diluted and quenched in 2 mL of 10 mM sodium phosphate buffer before analysis. Since maximum signal loss occurred within 60 min, further large-scale (1 μmol of $LF_N$ in the same conditions described above) sample racemization was done for 60 min. Afterward, the sample was dialyzed against ultra-pure water, with any resulting precipitate removed by centrifugation. The resulting protein

was analyzed by LC-MS and SDS-PAGE, revealing low molecular weight degradation products. Extent of racemization was determined by $CD_{205}$ readings of acid hydrolyzed sample compared to non-racemized controls. The acid hydrolysis was performed by incubation with 6M HCl overnight at 100°C, after which the samples were dried with a $N_2$ gas stream and then re-dissolved in ultra-pure water. To check for modifications, the protein was digested with trypsin and analyzed by LC-MS/MS. Before assaying the treated protein for translocation activity, the sample was lyophilized, dissolved in 5% acetonitrile 0.1 % trifluoroacetic acid, purified using reverse phase HPLC to isolate the full-length protein, lyophilized again, and dissolved in an aqueous solution.

*Electrophysiology*—Planar lipid bilayers were formed as described (45) by painting (128) a membrane-forming solution (3% 1,2-diphytanoyl-*sn*-glycerol-3-phosphocholine in *n*-decane; Avanti Polar Lipids) across a 100 μm aperture in a 1-mL, white-Delrin or polysulfone cup (34, 36, 45). A capacitance test confirmed the quality of the membrane. The membrane separates the cis and trans chambers, which each contained 1 mL of universal bilayer buffer (UBB; 10 mM oxalic acid, 10 mM MES, 10 mM phosphoric acid, 1mM EDTA) or a supplemented UBB (SUBB; 6 mM oxalic acid, 6 mM MES, 6 mM phosphoric acid, 6 mM TAPS, 6 mM boric acid, 1mM EDTA), which is a better at buffering at higher pH. Generally, an equivalent of 100 mM KCl was added to each buffer, except when indicated. Ag/AgCl electrodes bathed in saturated 3 M KCl were linked to the chambers via 3 M KCl-agar salt bridges. PA currents were obtained via an Axopatch 200B amplifier and recorded using AXOCLAMP10 (Molecular Devices).

*Translocation assays*—Bilayers were bathed in symmetrical UBB or SUBB. $\Delta\psi \equiv \psi_{cis} - \psi_{trans}$ ($\psi_{trans} \equiv 0$). $PA_7$ prechannel was added to the cis chamber (held at 20 mV), and conductance was blocked by the addition of substrate (WT or mutated $LF_N$) to the cis side (still held at 20 mV in symmetric pH 5.6 experiments). The substrate blockade was >95% of the original current, except where indicated. Excess substrate was perfused by a custom hand-cranked, push-pull perfusion system. In $\Delta\psi$-driven translocation assays, substrate translocation was driven by increasing the $\Delta\psi$. An approximate translocation activation energy ($\Delta G^{\ddagger}$) was determined using $RT \ln t_{1/2}/c$ (45). The $t_{1/2}$ value is the time for half the substrate to translocate; $c$ is a 1-sec reference.

*Binding assays*—Bilayers were bathed in asymmetric KCl solutions buffered in 10 mM potassium phosphate ($[\text{added KCl}]_{cis}$ = 100 mM, $[\text{added KCl}]_{trans}$ = 0 mM, $pH_{cis}$ = 6.5, $pH_{trans}$ = 7.40). Once PA channel insertion was complete the cis buffer was perfused and exchanged to pH 7.40, 100 mM KCl. (The pH of the cis and trans buffers were matched to 0.01 units.) $LF_N$ was then added to the cis side of the membrane at small concentration increments, allowing for binding equilibrium to be maintained. Final current ($I$) levels were recorded, and the equilibrium current-block versus ligand concentration, [$L$], curves were fit to a simple single-binding site model, $I/I_o = 1 - a/(1 + K_d/[L])$, to obtain the equilibrium dissociation constant, $K_d$, where $I_o$ is the current amplitude with no substrate and $a$ estimates the value of $1 - I/I_o$ under saturating concentrations of substrate.

*Peptide blocking and translocation*—Bilayers were bathed in symmetric buffer (100 mM KCl, 1 mM EDTA, 10 mM succinic acid, pH 5.6). $PA_7$ prechannel was added to the cis chamber (held at 20 mV), and single-channel insertion was observed by a discrete increase in current. Once a single channel had inserted into the membrane, peptide was added to the cis side of the membrane at 18-20 nM. Data were recorded at 400 or 600 Hz using a 200-Hz low-pass filter.

## 4.4 Results

*A mechanism-based α clamp inhibitor construct disrupts substrate binding*—Previously, we have shown that interactions between the α clamp and LF's α1/β1 region provide 2.5 to 4 kcal $mol^{-1}$ of stabilization for the bound state (34). Interestingly, replacements of LF's α1/ β1 region with other peptides from LF and EF do not readily alter the interaction with the α clamp, revealing its large degree of nonspecificity. The consequence of this property of the α clamp is that it is difficult to completely eliminate binding to the site with point mutations, and ultimately its role in the translocation mechanism is unknown. In point of fact, the clamp cannot be disrupted by mutating the structural twin-calcium ion binding sites, as they are too integral to the overall structure and stability of PA oligomer. To develop an approach that probes α clamp function yet maintains structural integrity of the oligomer, we attached a sequence corresponding to LF's α1/β1 sequence (residues 26-49) to $PA_{63}$'s amino terminus (i.e. directly after $PA_{83}$'s cleavage site and preceding the first residue of post-cleavage $PA_{63}$; Fig. 4.1A). The four known Lys residues in the α1/β1 sequence were mutated to Ala to prevent unwanted proteolytic cleavage when $PA_{83}$ is treated with trypsin immediately prior to oligomerization. These Lys residues are not important for binding the α clamp (34). Furthermore, PA's P173 residue, one of six unstructured amino acids at $PA_{63}$'s amino terminus was mutated to a Gly to increase the region's flexibility and allow the appended sequence to occupy the α clamp and act as a "plug"; as such, we termed the construct "$PA_{\alpha\text{-plug}}$." Additionally, we made several constructs in which this interaction was stabilized by cysteine crosslinking: $^{236\text{-}40}PA_{\alpha\text{-plug}}$, $^{464\text{-}32}PA_{\alpha\text{-plug}}$, and $^{465\text{-}30}PA_{\alpha\text{-plug}}$, where the first superscripted number is the position in PA mutated to Cys and the second is the equivalent LF residue on the appended helix also mutated to Cys.

To measure the substrate $LF_N$-binding thermodynamics, we used a planar lipid bilayer electrophysiology binding assay (34). Here, two aqueous chambers, cis and trans, are separated by a planar lipid bilayer. The PA prechannel oligomer is added to the cis side of the membrane under an asymmetric KCl gradient ($[KCl]_{cis}$ = 100 mM, $[KCl]_{trans}$ = 0 mM, $pH_{cis}$ = 6.5, $pH_{trans}$ = 7.40), thereby allowing for detection of PA channel insertion and current increase at a $\Delta\psi$ of 0 mV ($\Delta\psi \equiv \Delta\psi_{cis} - \Delta\psi_{trans}$; $\Delta\psi_{trans} \equiv 0$ mV). Once the current stabilizes, excess prechannel is removed by perfusion with $pH_{cis}$ 7.40 buffer to maintain a KCl gradient simultaneously while removing the pH gradient. Substrate $LF_N$ is added at a variety of concentrations. Conductance blockade is observed in response to each addition of $LF_N$. The fraction of unblocked channels as a function of substrate concentration is well fit using a single-site binding model.

The equilibrium dissociation constant, $K_d$, for WT $LF_N$ with WT PA channels is 120 (±30) pM (Fig. 4.1B). With the non-crosslinked $PA_{\alpha\text{-plug}}$ construct, that value increases modestly to 530 (±60) pM. The $^{236\text{-}40}PA_{\alpha\text{-plug}}$ channel slightly increases the extent of binding disrupting, bringing the $K_d$ to 860 (±150) pM. However, reactions with Ellman's reagent suggest poor crosslinking in this construct. The other constructs, $^{464\text{-}32}PA_{\alpha\text{-plug}}$, and $^{465\text{-}30}PA_{\alpha\text{-plug}}$, had more substantial effects, increasing the $K_d$ to 3.3 (±0.6) nM and 7.4 (±1.7) nM, respectively. For the most severe mutant, $^{465\text{-}30}PA_{\alpha\text{-plug}}$, this corresponds to a loss of 2.4 (±0.2) kcal $mol^{-1}$ of stabilization for the bound state, which was the previously reported value for stability imparted by the alpha clamp as determined by α1/β1 truncation (34).

Furthermore, the baseline indicating the maximum amount of current blocked under saturating concentrations of WT $LF_N$ is shifted when the α clamp is rendered inaccessible (Fig. 4.1B). For WT PA, this baseline is 93 (±4)%. The non-crosslinked $PA_{\alpha\text{-plug}}$ and unsuccessfully crosslinked $^{236\text{-}40}PA_{\alpha\text{-plug}}$ only shift this value to 88 (±2)% and 83 (±3)%, respectively. Once again, the largest effects come from $^{464\text{-}32}PA_{\alpha\text{-plug}}$ and $^{465\text{-}30}PA_{\alpha\text{-plug}}$, where the maximum block

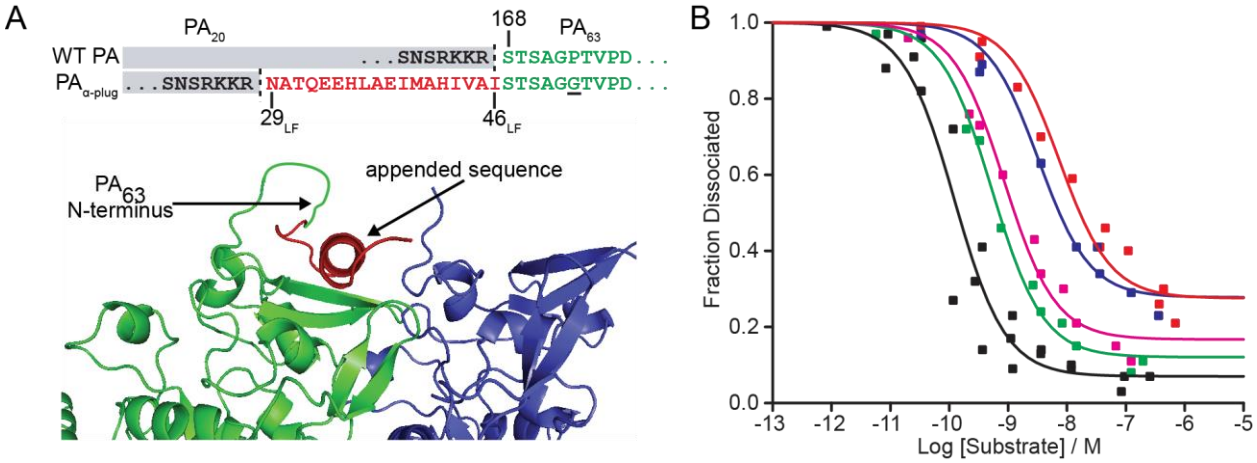


**Figure 4.1. α clamp occlusion inhibits substrate binding.** (A) (top) Sequences of WT PA and $PA_{\alpha\text{-plug}}$. $PA_{20}$, which is cleaved prior to oligomerization, is highlighted in gray. $PA_{63}$ is shown in green, with the P173G mutation, added for linker flexibility, underlined. The point of cleavage is indicated with a dashed line. The appended sequence, corresponding to LF residues 29-46, is in red. (bottom) A model of $PA_{\alpha\text{-plug}}$, based on the crystal structure of $LF_N$ bound to the octameric PA prechannel (3KWV) (34). Adjacent PA63 subunits are in green and blue with the appended LF sequence in red. Six unstructured residues of PA's N terminus not seen in the crystal structure are drawn in as the flexible linker connecting PA to the appended LF helix. (B) Fraction of PA bound to substrate, inferred from the ratio of current after and before substrate addition ($I/I_o$), as a function of WT $LF_N$ ligand concentration, [$L$], fit to a single-state binding model: $I/I_o = 1 - a/(1 + K_d/[L])$. The parameter $a$ estimates the value of $1 - I/I_o$ under saturating concentrations of substrate. The equilibrium dissociation constant, $K_d$, is shifted from 120 (±30) pM for WT PA (black) to 530 (±60) pM for $PA_{\alpha\text{-plug}}$ (green), 860 (±150) pM for $^{236\text{-}40}PA_{\alpha\text{-plug}}$ (magenta), 3.3 (±0.6) nM for $^{464\text{-}32}PA_{\alpha\text{-plug}}$ (blue), and 7.4 (±1.7) nM for $^{465\text{-}30}PA_{\alpha\text{-plug}}$ (red). The baseline for conductance block at saturating concentrations of $LF_N$ is 93 (±4)% for WT PA, 88 (±2)% for $PA_{\alpha\text{-plug}}$, 83 (±3)% for $^{236\text{-}40}PA_{\alpha\text{-plug}}$, 72 (±2)% for $^{464\text{-}32}PA_{\alpha\text{-plug}}$, and 72 (±3)% for $^{465\text{-}30}PA_{\alpha\text{-plug}}$.

baselines are 72 (±2)% and 72 (±3)%, respectively. Hence for the $PA_{\alpha\text{-plug}}$ mutations, while substrate binding to the α clamp was disrupted by up to 2.5 kcal $mol^{-1}$, the ability of substrate to initiate and dock into the channel's conductance-blocking site, the ϕ clamp was perturbed but not fully disrupted.

*$PA_{\alpha\text{-plug}}$ mutant possesses defective translocase activity*—Because the substrate could still initiate into the channel, we then measured the translocation kinetics in the α-plug mutant background. As it was the more severe mutant, we specifically focused on $^{465\text{-}30}PA_{\alpha\text{-plug}}$. Planar lipid bilayer electrophysiology was again used to make translocation kinetics measurements (18, 21, 34, 36, 45, 68, 69, 71, 134). Here the planar bilayer separated symmetrical aqueous chambers of 100 mM KCl, pH 5.6. PA prechannel oligomer was added to the cis side of the membrane under a Δψ of 20 mV. An increase in current followed by stabilization indicated successful channel formation, and excess prechannel was removed by perfusion. Next, a saturating concentration of substrate was added and allowed to fully block the channel, as inferred by the decrease in ion flow, before a second round of perfusion. Translocation was then initiated by increasing the Δψ. The observed translocation kinetic records are complex and multi-exponential. To estimate the rate of translocation, the time for half of the substrate to translocate ($t_{1/2}$) is measured. This approximation allows the rate-limiting step of the translocation reaction to be monitored according to the established theory of transit times used in enzymology. With this latter parameter, we can estimate the activation energy of translocation ($\Delta G^{\ddagger}$) by $\Delta G^{\ddagger} = RT \ln (t_{1/2}/c)$, where $c$ is a constant of 1 s.

Typically as a function of the driving force, $\Delta G^{\ddagger}$ does not vary linearly as expected for a simple system with only one major driving force-dependent rate-limiting barrier. Rather, there is a steep slope at lower driving forces and a much shallower slope at higher driving forces. This is consistent whether the driving force is a Δψ, ΔpH, or a combination thereof (18, 21, 45). The two slopes correspond to two major barriers for translocation, a highly driving force-dependent barrier that dominates with lower driving forces and a largely driving force-independent barrier that is rate limiting with greater driving forces. Earlier work has extensively characterized the driving force-dependent barrier and has shown it to be limited by large-scale substrate unfolding, specifically the unfolding of $LF_N$'s β-sheet subdomain region (45). The second barrier has yet to be identified. Interestingly, the $^{465\text{-}30}PA_{\alpha\text{-plug}}$ mutant has a different barrier profile that WT PA (Fig 4.2A). In addition to shifting the curve upward, the driving force-independent slope dominates over a wider range of driving force strengths than it does in the case of WT PA. While there does appear to be increased driving force-dependence only between 40 and 50 mV for $^{465\text{-}30}PA_{\alpha\text{-plug}}$, this range normally extends until 60 or 70 mV for WT PA.

*$PA_{\alpha\text{-plug}}$ mutant disrupts the more force-independent translocation step*—The upward shift in activation energies shows that the $^{465\text{-}30}PA_{\alpha\text{-plug}}$ mutant is defective in translocating WT $LF_N$ (Fig. 4.2A). However, this translocation defect is more pronounced with higher driving forces. At 40 mV, our lowest driving force assayed, the change in activation energy for translocation ($\Delta\Delta G^{\ddagger} = \Delta G^{\ddagger}_{\alpha\text{-Plug}} - \Delta G^{\ddagger}_{WT}$) was increased less than 0.5 kcal $mol^{-1}$ (Fig. 4.2B). As the magnitude of the driving force is increased, $\Delta\Delta G^{\ddagger}$ rises as well, until it eventually levels off slightly above 1 kcal $mol^{-1}$ at the highest Δψ values we measured. Our model suggests that the α clamp plays a role in unfolding, so we investigated the ability of the $PA_{\alpha\text{-plug}}$ mutant to translocate a construct with a mutation previously demonstrated to destabilize it, $LF_N$ L145A (45). Once again, translocation was slowed at all voltages, but in this case the extent of the defect did not vary with the magnitude of the driving force (Fig. 4.3A). The $\Delta\Delta G^{\ddagger}$ values for $LF_N$ L145A remained just

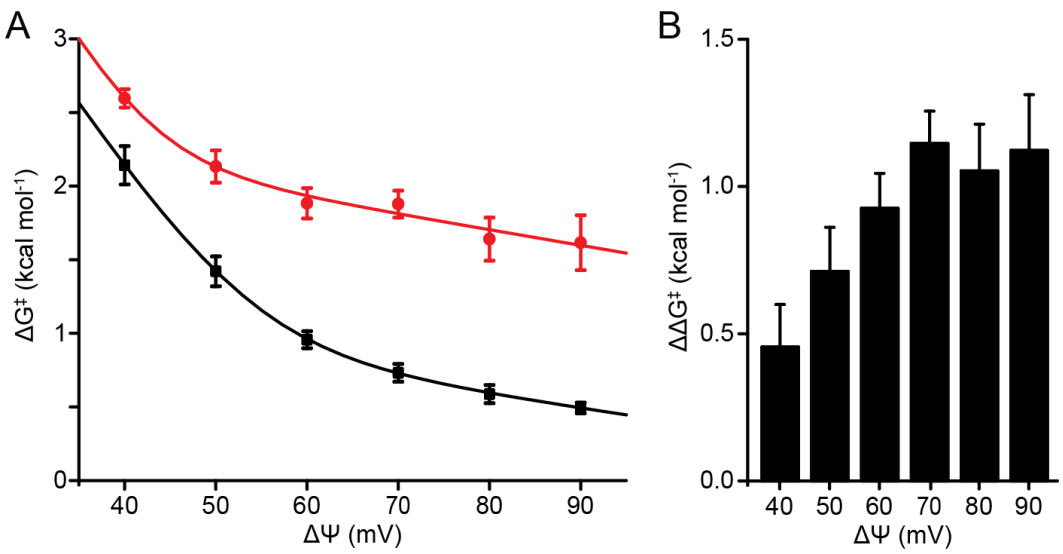


**Figure 4.2. Blocking the α clamp inhibits WT $LF_N$ translocation more severely with greater driving forces.** (A) Activation energy versus Δψ results for the translocation of WT $LF_N$ with either WT PA (black) or $PA_{α\text{-plug}}$ (red). These translocations were conducted at a symmetrical pH of 5.6. The error bars are the means ± S.D. ($n$ = 2–4). (B) Values of $\Delta\Delta G^{‡} = \Delta G^{‡}_{α\text{-plug}} - \Delta G^{‡}_{WT}$ for the driving force magnitudes displayed in (A).

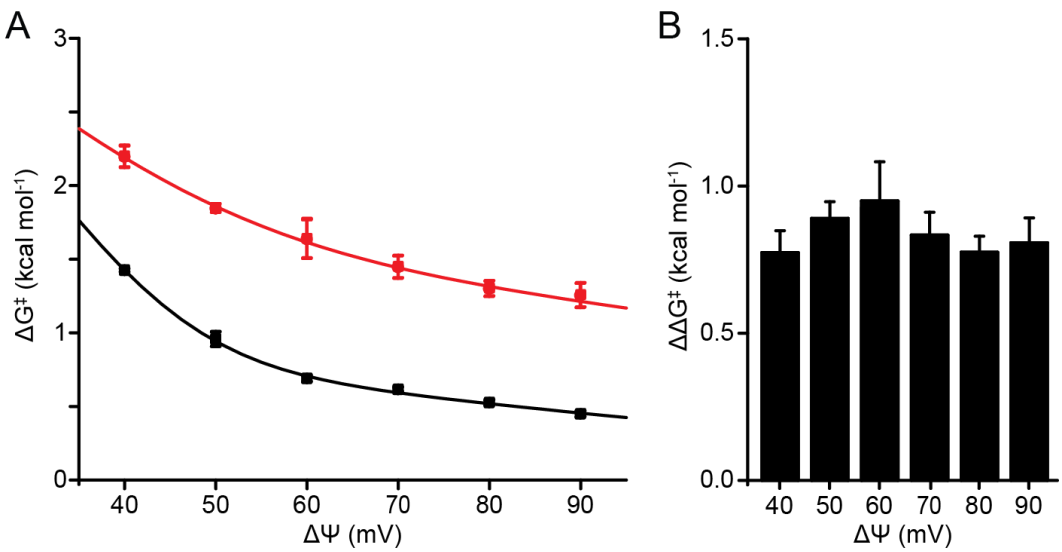


**Figure 4.3. Blocking the α clamp inhibits translocation of a destabilized mutant $LF_N$ regardless of the driving force magnitude.** (A) Activation energy versus Δψ results for the translocation of $LF_N$ 145A with either WT PA (black) or $PA_{α\text{-plug}}$ (red). These translocations were conducted at a symmetrical pH 5.6. The error bars are the means ± S.D. ($n$ = 2–3). (B) Values of $\Delta\Delta G^{‡} = \Delta G^{‡}_{α\text{-plug}} - \Delta G^{‡}_{WT}$ for the driving force magnitudes displayed in (A).

below 1 kcal mol$^{-1}$ across the same range of driving forces under which the $\Delta\Delta G^{\ddagger}$ doubled with WT $LF_N$ (Fig. 4.3B).

Finally, it should be noted that $^{465\text{-}30}PA_{\alpha\text{-plug}}$ cannot be fully blocked by either WT or mutant substrate. Even with our low pH and small $\Delta\psi$ pre-translocation conditions, there are still unblocked channels, and we are reporting the kinetics only for the substrates that are bound. If some channels bind no substrate, it stands to reason that a portion of the blocked channels have only a single substrate bound, whereas WT PA heptamers can bind up to three. It is possible that clearance of WT PA channels requires three translocation events while $^{465\text{-}30}PA_{\alpha\text{-plug}}$ only requires one. Thus, we may be in fact be underestimating the extent to which the $^{465\text{-}30}PA_{\alpha\text{-plug}}$ is inhibiting translocation.

*Altering the shape of $LF_N$'s α1/β1 sequence disrupts channel binding and translocation*— A complementary approach to obstructing the α clamp is to alter the α1/β1 sequence (residues 30-47) that typically binds there. We previously showed that the α clamp is a nonspecific site, and it stably binds a variety of polypeptide sequences (34). The current hypothesis is that the α clamp may recognize substrates via a general steric shape-complementarity mechanism. To test this idea, we replaced $LF_N$ α1/β1 (positions 30-47) with varying densities of Pro, which would result in a drastic change in their backbone configuration. Pro residues are highly disruptive to α-helix formation due to steric interference and the lack of amide hydrogens for hydrogen binding. Consecutive prolines do form a helical structure, the left-handed polyproline II helix, which in steric terms is narrower and consequently longer than a typical α helix (136) (Fig. 4.4A).

In our initial constructs, we replaced either the entire α1/β1 region ($LF_N$ $Pro_{30\text{-}47}$), a portion of the α1 helix and the β1 strand ($LF_N$ $Pro_{36\text{-}47}$) or just the β1 region ($LF_N$ $Pro_{43\text{-}47}$) with consecutive Pro residues. If the prolination causes the α1/β1 region to form a shape the α clamp cannot recognize, this substrate should be deficient in its ability to block the channel, similar to what was observed with an occluded α clamp. Indeed, $LF_N$ $Pro_{30\text{-}47}$ bound WT PA ~100 times weaker than WT $LF_N$, with the $K_d$ increasing to 2.1 (±0.5) nM for the mutant from 120 (±30) pM for the WT substrate (Fig. 4.4B). This is a loss of 1.7 (±0.2) kcal mol$^{-1}$ of stabilization for the bound state. While substantial, that is value is somewhat lower than ~2.5 kcal mol$^{-1}$ observed by occlusion of the α clamp or full truncation of the α1/β1 region (34). Further, this mutant can only achieve a maximum block of 78 (±3)% compared to WT $LF_N$'s 93 (±4)%, again a defect similar to though slightly smaller in magnitude compared with what was observed when the α clamp was made inaccessible. Blocking only a portion of the α1 helix and β1 sheet with $LF_N$ $Pro_{36\text{-}47}$ yielded similar results with a $K_d$ of 1.2 (±0.3) nM and a maximum block parameter of 73 (±3)%. However, $LF_N$ $Pro_{43\text{-}47}$ behaved much more like WT $LF_N$, with $K_d$ of 310 (±50) pM and maximum block of 90 (±2)%. This indicates that disrupting the β1 sheet is tolerated, but the α1 helix cannot be replaced with a polyproline sequence.

*Proline-substituted $LF_N$ substrates are deficient in translocation*—While the polyproline sequence imparts binding deficiencies, the $LF_N$ $Pro_{30\text{-}47}$ construct can block PA channels under our normal pre-translocation conditions ($pH_{cis}$ = 5.6; $\Delta\psi$ of 20 mV). This allows us to explore whether the α clamp plays a role in translocation after successful initiation. If the α clamp plays no role post-initiation, there should be no difference in translocation between WT $LF_N$ and $LF_N$ $Pro_{30\text{-}47}$ under these conditions. Differences would only appear if interaction with the α clamp and/or the shape of the translocating chain matter in downstream translocation steps.

Indeed, there appear to be great post-initiation translocation differences between WT $LF_N$ and $LF_N$ $Pro_{30\text{-}47}$. $LF_N$ $Pro_{30\text{-}47}$ was unable to translocate at all with a 50 mV membrane potential, conditions under which a large fraction of WT $LF_N$ can translocate (Fig. 4.4C). We next looked

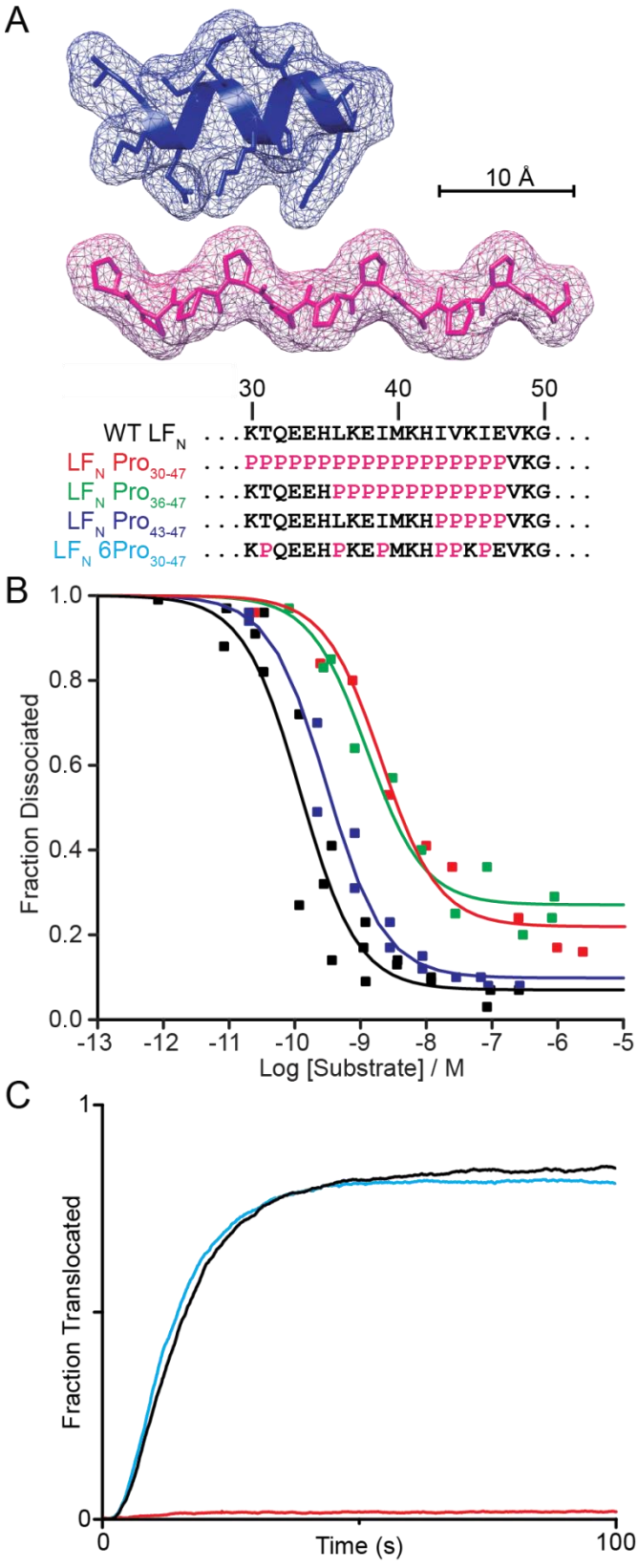


**Figure 4.4. A polyproline helix disrupts binding of the α clamp and translocation through the PA channel while less severe modifications are tolerated.** (A) Comparison of an α helix (top; based on LF's α1 helix) and a polyproline helix (middle) with an equal number of residues. (bottom) The sequences of the WT $LF_N$ substrate and mutants with part or all of the α1/β1 region replaced with Pro residues. (B) Fraction of WT PA bound to substrate, inferred from the ratio of current after and before substrate addition ($I/I_o$), as a function of ligand concentration, [$L$], fit to a single-state binding model: $I/I_o = 1 - a/(1 + K_d/[L])$. The parameter $a$ estimates the value of $1 - I/I_o$ under saturating concentrations of substrate. The equilibrium dissociation constant, $K_d$, is shifted from 120 (±30) pM for WT $LF_N$ (black) to 7.8 (±0.2) nM for $LF_N$ $Pro_{30-47}$ (red), 1.2 (±0.3) nM for $LF_N$ $Pro_{36-47}$ (green), and 310 (±50) pM for $LF_N$ $Pro_{43-47}$ (blue). The baseline for conductance block at saturating concentrations of $LF_N$ is 93 (±4)% for WT PA, 78 (±3)% for $LF_N$ $Pro_{30-47}$, 73 (±3)% for $LF_N$ $Pro_{36-47}$, and 90 (±2)% for $LF_N$ $Pro_{43-47}$. (C) Translocations of WT $LF_N$ (black) and $LF_N$ $Pro_{30-47}$ (red), and $LF_N$ $6Pro_{30-47}$ (cyan) under a $\Delta\psi$ (symmetrical pH 5.6 and $\Delta\psi$ of 50 mV). Each trace is representative of translocations performed on multiple membranes.

to see if we could sufficiently disrupt the shape the α1/β1 region using six helix-disrupting proline substitutions spaced throughout this sequence (Fig. 4.4A). This construct, $LF_N$ $6Pro_{30-47}$, would be predicted to disrupt α helix formation but would not form the unique polyproline II helix. Surprisingly, $LF_N$ $6Pro_{30-47}$ was not deficient in translocation under the same driving forces (Fig. 4.4B). If these mutants are indeed interacting with the α clamp despite the helix-disrupting presence of prolines, the site's specificity may be even broader than we initially expected. Further characterization of the $LF_N$ $6Pro_{30-47}$ substitution will be required, however, to better understand these phenomena.

*Racemized substrates are deficient in translocation*—Another approach for disrupting secondary structure in the substrate is the introduction of D-amino acids into a predominantly L-amino acid protein sequence. A polypeptide chain with a mixture of L- and D-amino acids is incapable of forming any secondary structural features. To introduce D-amino acids into $LF_N$, we used a chemical racemization procedure described by Schwass and Finley (135). $His_6$-tagged $LF_N$ was incubated with 100 mM NaOH and approximately 5 M guanidinium chloride at 85°C. 1 mM sodium bisulfite was added to prevent lysinoalanine crosslinking that may occur at low frequency during this treatment (137). A time course shows that about 50% of the $CD_{222}$ signal is lost with a 60 min treatment (Fig. 4.5A). To analyze the extent of racemization, the sample and an untreated control were hydrolyzed to individual amino acids by treatment with 6 M HCl at 100°C overnight. After drying and redissolving the amino acids in water, they were analyzed by CD. When normalized for concentration, we see a 20 (±5)% decrease in the CD band at 205 nm (Fig. 4.5B). While this is not complete racemization, the amino acid conversion and thus disruption of secondary structure is extensive according to CD spectra taken of the untreated and treated proteins (Fig. 4.5A). To confirm that there was no lysinoalanine crosslinking, we performed LC-MS/MS analysis of trypsinized protein (Fig. 4.5C). While no crosslinking was detected, this assay revealed certain peptides with Asn and Gln residues experienced side chain deamidation. However, only one residue was deamidated in excess of 50% (N29: 2 of 3 hits). Finally, analysis by intact protein LC-MS (Fig. 4.5D) and SDS-PAGE (Fig. 4.5E) revealed the existence of truncated polypeptide products from racemization procedure, but these smaller products were largely removed using reverse-phase HPLC (Fig. 4.5E).

Interestingly, under pre-translocation conditions (pH 5.6; Δψ of 20 mV), the racemized $LF_N$ can fully block WT PA. This is likely due to the intact $His_6$ tag, which has previously been shown to be sufficient for initiation into the PA channel (86). However, even though initiation into the channel has occurred and unfolding of the racemized substrate is unnecessary, it is completely unable to translocate under a pure Δψ (Fig. 4.6A).

Since the LC-MS/MS analysis revealed some deamidation (Fig. 4.5C), this raises the possibility that this chemical modification rather than the racemization itself is responsible for the deficiency. Indeed, previous work has shown that high sequence densities of negative charge can prevent translocation (21, 134). However, the sequences stalled by substantially acidic sequence density were recovered by lowering the pH to 5.0 during translocation. We find that the racemized substrate remains unable to translocate at symmetrical pH 5.0 (Fig. 4.6A). Additionally, the presence of many examples of peptides without modified Asn and Gln (Fig. 4.5C) suggests that not every substrate has been modified. Even if some of the substrates had these putatively deleterious modifications, a subset of protein should still translocate, especially under pH 5.0 conditions. The complete absence of translocation suggests that the inhibition is due to a more efficient modification, such as the racemization.

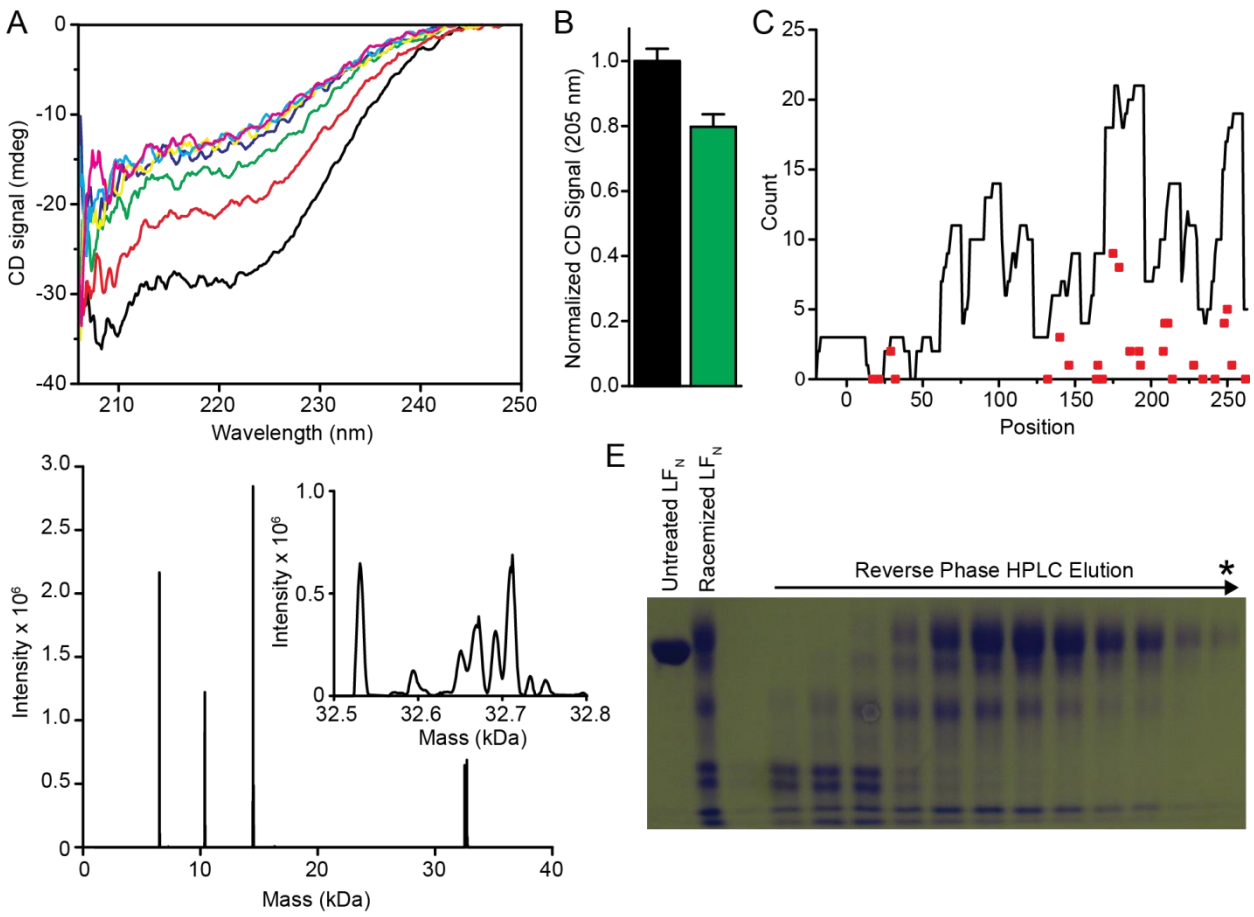


**Figure 4.5. D-amino acids can be introduced into $LF_N$ through a chemical treatment.** (A) CD spectra of racemization time course at 0 (black), 15 (red), 30 (green), 45 (blue), 60 (cyan), 75 (magenta), and 90 min (yellow). (B) Normalized $CD_{205}$ signals for the amino acid hydrolysates untreated (black) and racemization treated (green) $His_6$-$LF_N$ samples. The error bars are the means ± S.D. (*n* = 3). (C) LC-MS/MS analysis of the racemized protein included 95% sequence coverage, including all but two of the potentially deamidated Asn and Gln residues. The number of peptides that each residue occurred in is shown with the black line. For each Asn and Gln residue, the number of peptides featuring a deamidiation at that site is shown a red box. The numbering convention is that used for LF in PDB ID 1J7N (32), with the added N-terminal tag numbered -20 through 0. (D) Intact protein LC-MS spectrum of $His_6$-$LF_N$ treated with the racemization procedure to show major breakdown products, with a zoomed in view around the predicted mass of the full-length substrate. (E) SDS-PAGE gel showing the difference between untreated and treated $His_6$-$LF_N$ and our ability to purify the full-length treated protein using reverse phase HPLC. The (*) indicates the fraction that was used for later translocation analysis.

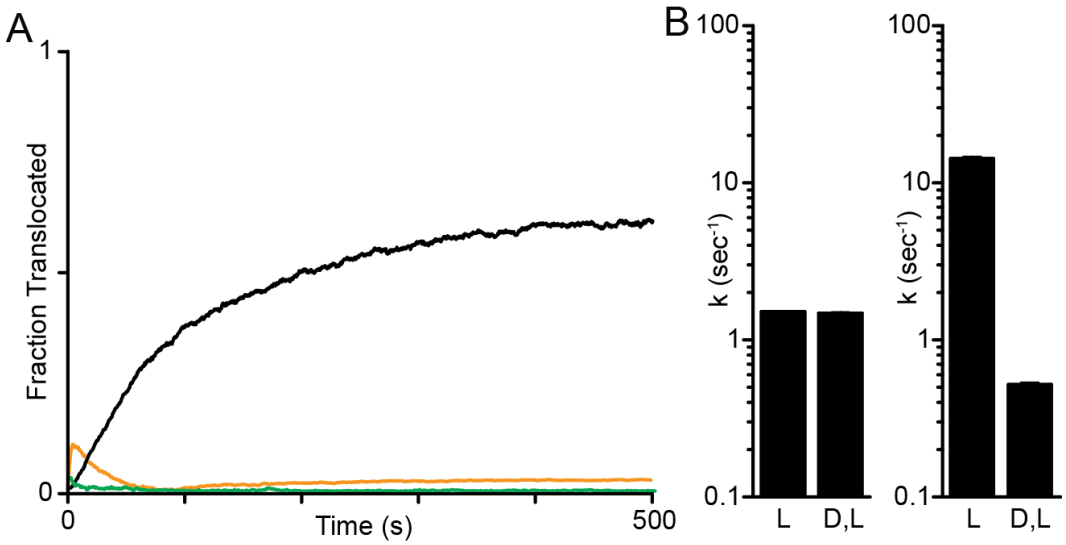


**Figure 4.6. Racemized substrates are impaired in translocation through the PA channel.** (A) Translocations of untreated (black) and racemization treated (green) WT $His_6$-$LF_N$ under a $\Delta\psi$ (symmetrical pH 5.6, $\Delta\psi$ of 50 mV). Racemized substrate translocation is only 2.2 (±1.1)% efficient compared to the non-racemized sample that reaches 66 (±1)% efficiency. A control of racemized WT $His_6$-$LF_N$ under symmetrical pH 5.0, $\Delta\psi$ of 50 mV is shown in orange. Each trace is representative of translocations performed on multiple membranes. (B) Binding (left) and translocation (right) kinetics of L- and D,L-peptides with the WT PA channel (symmetrical pH 5.6, $\Delta\psi$ of 20 mV). Peptides have the sequence KKKKKWWSWW. In the D,L construct, odd-numbered residues have the D configuration and even-numbered residues are L.

To look at the effects of introducing D-amino acids in a more controlled manner, we used synthesized 10mer oligopeptides with all L-residues or alternating D,L (D: odd; L: even). Using a single-channel analysis, we assayed the kinetics of channel blocking and translocation for these peptides. Both species are capable of blocking the channel conductance at equivalent rates: 1.51 (±0.01) $s^{-1}$ for L and 1.48 (±0.00) $s^{-1}$ for D,L. Such short peptides should not engage the α clamp while blocking the channel, so this suggests that blocking (when separated from α clamp binding) does not require helix formation. However, the translocation rate is significantly slower for the D,L version at 0.52 (±0.01) $s^{-1}$ relative to the pure L-amino acid version at 14.2 (±0.3) $s^{-1}$. Taken together, these results indicate that there is a step in translocation separate from α clamp binding and channel blocking that seems to be facilitated by the substrate's ability to form an α helix.

## 4.5 Discussion

*Disrupting the substrate's interaction with the α clamp impairs channel blocking and translocation*—Previous work extensively demonstrated the importance for the α clamp in substrate binding (34). By occluding the α clamp with the $PA_{\alpha\text{-plug}}$ mutations (Fig. 4.1B) and disrupting $LF_N$'s α1/β1 region with $LF_N$ $Pro_{30\text{-}47}$'s polyproline II helix (Fig. 4.4B) we have found further support for this role. However, it was surprising to note that $LF_N$ $Pro_{30\text{-}47}$ did not disrupt binding as extensively as other approaches, suggesting some interaction between the α clamp and a polyproline II helix is possible. While these interactions with polyproline are generally weaker, the expected activity of the α clamp may be to recognize polypeptide by a shape-complementarity mechanism, where α helix is preferred relative to narrower helices, such as polyproline II. Although earlier work determined the binding stability imparted by the α clamp based on α1/β1 truncations, those experiments could not address the α clamp's role in channel blocking. With the $PA_{\alpha\text{-plug}}$ and $LF_N$ $Pro_{30\text{-}47}$ mutants, we can disrupt the α clamp-α1/β1 interaction while still maintaining a WT N terminus in $LF_N$ and a WT ion constriction site in PA. At our highest substrate concentrations, $LF_N$ should dock to PA's second binding site on the channel's top surface regardless of the presence of the α clamp interaction. Neither $PA_{\alpha\text{-plug}}$ nor $LF_N$ $Pro_{30\text{-}47}$ have mutations that can interrupt that interaction. That full blockade cannot be achieved under saturating conditions, however, indicates that some percentage of bound substrates cannot block conductance fully and suggests a role for the α clamp in guiding the substrate's N terminus into the channel. Further single-channel studies will be required to understand the partial blockade phenomenon we observe for polyproline-containing sequences.

Our main interest in these investigations of the α-clamp site is its role in substrate translocation post-initiation. Its nonspecific helix-binding nature led us to hypothesize a role in nucleating helix formation. This led to two main predictions: (i) blocking access to the α clamp could inhibit substrate initiation at the ϕ clamp, and (ii) disrupting the substrate's ability to bind the α clamp and form an α helix may generally impede the largely force-independent translocation step. Indeed, our data provides some insight on these possible models. Based on prior work, substrate initiation is defined as conductance blockade at the ϕ-clamp site, which is the major conductance-blocking site in PA (67). We found that occlusion of the α-clamp site with the $PA_{\alpha\text{-plug}}$ mutations, introduction of polyproline in substrate sequences, and substrate racemization did not block initiation and conductance blockade. However, we did find that substitution of a polyproline sequence in place of $LF_N$'s α1/β1 sequence greatly inhibited translocation (Fig. 4.4), as did the introduction of D-amino acids through chemical racemization (Fig. 4.6). Occlusion of the α clamp gave a more nuanced phenotype, revealing an increase in the

force-independent translocation barrier. This led us to some surprising conclusions about the potential role of substrate secondary structure in the mechanism of translocation.

*The α clamp lowers a major barrier for substrate translocation*—Earlier work indicated that there are two major barriers that can be rate-limiting during for translocation (18, 21, 45). One dominates when the driving force is weaker and is heavily force-dependent. This barrier has been well-characterized and corresponds to substrate unfolding, specifically at the amino-terminal β sheet subdomain (45). Conversely, the relatively force-independent barrier that dominates at higher driving forces has remained elusive. However, our results are consistent with a scenario where this barrier corresponds to the substrate's adoption of compact secondary structure, such as α helix. The kinetics of forming this secondary structure would not be predicted to be significantly altered under different driving force strengths, making it compatible with the barrier's force-independent nature. This would also explain the difficulty in translocating a polyproline sequence (Fig. 4.4C) or substrate with mixed chirality (Fig. 4.6A). In these cases of defective substrates, forming a compact α-helical secondary structure is impossible, and so the barrier is raised substantially.

When we blocked the substrate's access to the α clamp with the $^{265\text{-}30}PA_{\alpha\text{-plug}}$ mutation, translocation was more modestly inhibited at lower driving forces but greatly shifted with higher driving forces (Fig. 4.2). This observation is consistent with a situation where the force-independent barrier has been raised (Fig. 4.7). Earlier, we hypothesized that the α clamp plays a role in helix nucleation, acting as a Zimm-Bragg nucleation site for the substrate helix (94). According to our model, the α clamp normally lowers the force-independent translocation barrier by reducing the free energy required for the substrate to form a compact secondary structure. Under a lower driving force, when substrate unfolding is normally the larger barrier, the effect of losing the α clamp is minimal. However, with higher driving forces, when the putative secondary-structure formation barrier dominates, the full effect of the α clamp's loss can be felt (Fig. 4.7).

If we are correct that plugging the α clamp is in effect raising the force-independent translocation barrier, then we would predict a different set of results if we translocated a highly destabilized substrate. Such a substrate would have a reduced unfolding barrier, and therefore the secondary structure barrier would dominate at all driving forces. Raising this latter barrier should effect translocation equally regardless of driving-force strength and to an extent comparable to the maximum change observed with the WT substrate (Fig. 4.7). Indeed, we see this to be the case (Fig. 4.3). The α clamp appears to lower the force-independent translocation barrier by ~1 kcal $mol^{-1}$. We suggest that this barrier corresponds to helix formation in the substrate, and the α clamp lowers the activation energy for helix formation by serving as an upstream helix-nucleation site.

*Model*—Given the results described here and prior work described elsewhere (21, 34, 45, 67, 134), we can formulate a model for translocation that takes substrate unfolding into account (Fig. 4.8). First, the substrate binds the top of the PA channel, and the substrate's α1/β1 region binds into the channel's α-clamp site. This latter interaction serves to nucleate helix and may then propagate further helix formation into the channel. However, the deprotonated acidic residues in the substrate will not readily form structure past the channel's anionic repulsion barrier until they are protonated and become neutral in charge. Once protonated, the acidic residues can pass this barrier, and the now net-positive substrate chain will actually be pulled toward that formerly repulsive barrier (which also serves as a cation-attractive well). The helical dipole moment should further increase the net-positive charge and attraction toward the cation-

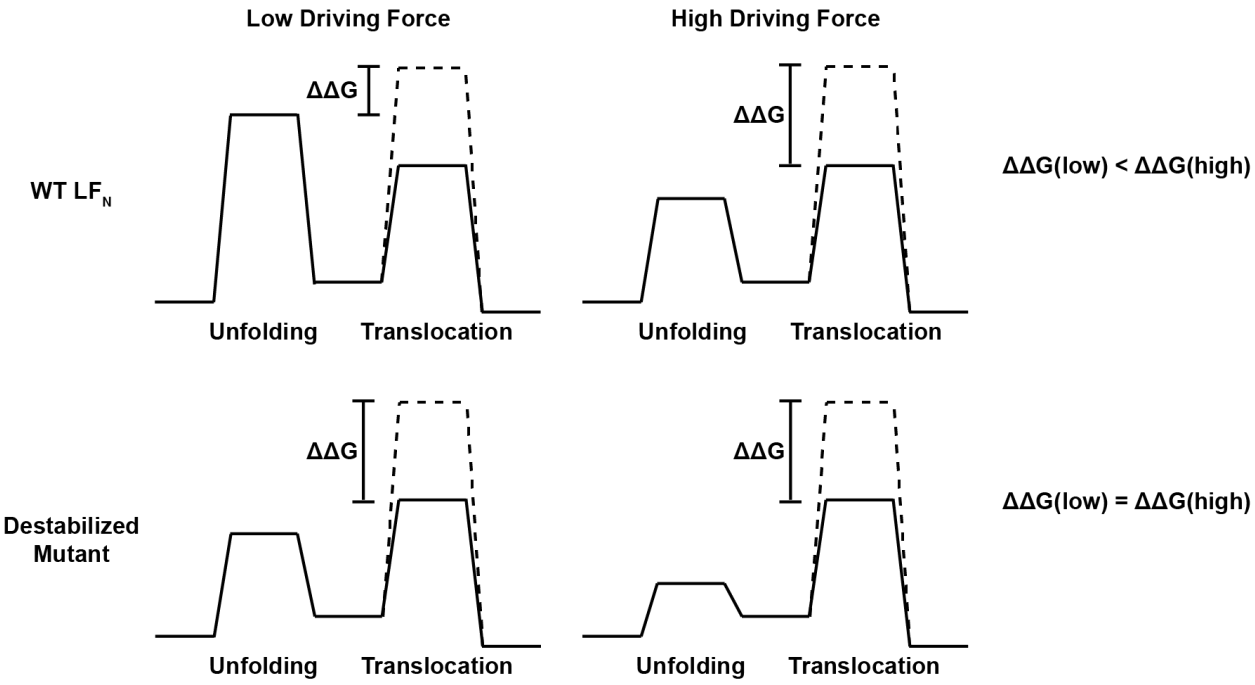


**Figure 4.7. Occluding PA's α clamp raises anthrax toxin's previously uncharacterized translocation barrier.** Earlier work showed two major barriers for anthrax toxin translocation: one that dominates at low driving forces and is strongly force dependent, and one that dominates at high driving forces and is largely force independent (18, 21, 45). The energy diagrams shown here explain the changes in translocation activation energy ($\Delta\Delta G^{\ddagger} = \Delta G^{\ddagger}_{MUT} - \Delta G^{\ddagger}_{WT}$) that would be predicted with a mutated channel (MUT) that increases the latter barrier. With a WT substrate (top), the $\Delta\Delta G^{\ddagger}$ would be predicted to increase as the driving force increases. However, with a destabilized substrate, the second barrier would always be rate limiting and the $\Delta\Delta G^{\ddagger}$ should be consistent across a wider range of driving forces. This is consistent with the data collected in Figs. 4.2 and 4.3. Given our hypothesis that the α clamp's role is to nucleate helix formation, we expand our model to suggest that helix formation is in fact the previously uncharacterized translocation barrier.

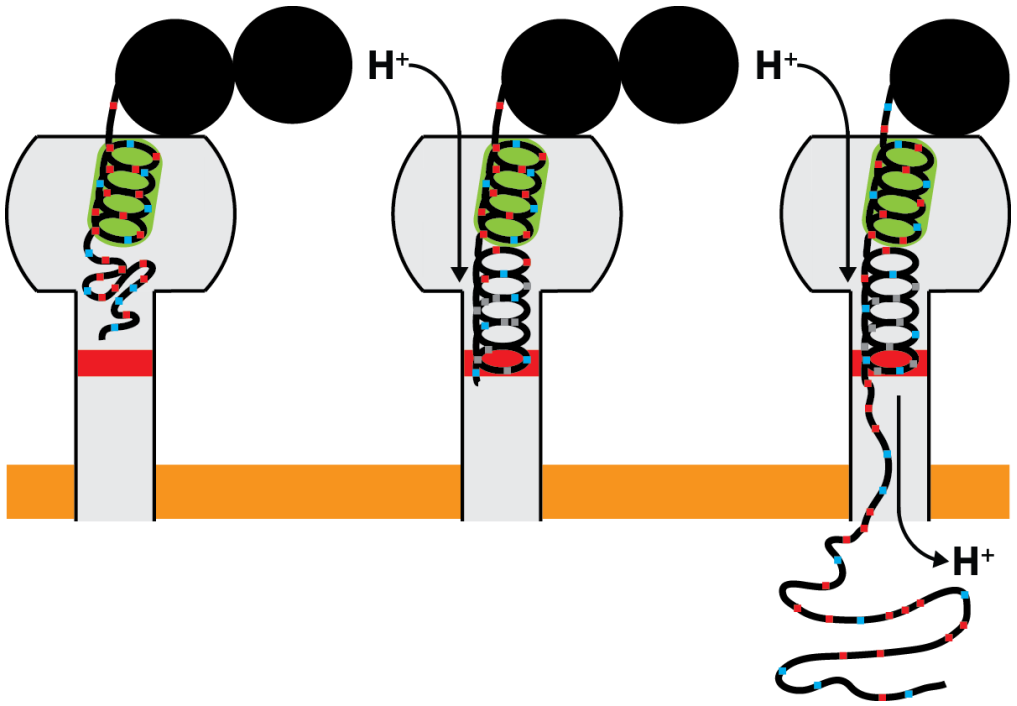


**Figure 4.8. Current model for a proton-engine helix-to-coil translocation model.** Based on results described herein and those described elsewhere (21, 34, 45, 67, 134), a model explaining substrate unfolding and translocation emerges. First, the substrate (black) binds the top of the PA channel (light gray). The substrate's α1/β1 region binds into the channel's α-clamp site (green); this interaction serves to nucleate helix and may then propagate further helix formation into the channel. However, the deprotonated acidic residues in the substrate (red squares), will not readily form structure past the channel's anionic repulsion barrier (red zone in channel) until these substrate residues are protonated and become neutral in charge (gray squares). The acidic residues, once protonated, can pass the channel's anionic repulsion barrier, and the now net-positive substrate chain will actually be pulled toward that formerly repulsive barrier (which also serves as a cation-attractive well). The helical dipole moment should further increase the net-positive charge and attraction toward the cation-interacting well. This helix-formation and proton-binding coupled attraction to the well may be considered a "power-stroke" phase of the polypeptide transport cycle. The subsequent reaction of the transport cycle may reflect an entropic expansion more resembling a "Brownian-ratchet" phase. Here, additional force may be applied to the substrate as Brownian motion in the helix allows it to move past the anionic-repulsion barrier. As acidic residues reach the cytosolic side of the acidic-repulsion barrier, they will become more frequently deprotonated, making the peptide net negative and significantly less capable of retrotranslocation. This deprotonation will be coupled to the helix-to-coil transition in the substrate helix, and an entropic tension may develop in the extended random-coil chain. The resolution of this tension is likely the reformation of helical structure from translocation of upstream chain. This restoring force will be accompanied by a new round of substrate protonation and helix formation that allows new helical structure to pass through the anionic charge repulsion site. Repeated cycles of these processes lead to the unfolding of each domain and complete translocation.

interacting well. This attraction to the well from helix formation and substrate protonation may be considered a “power-stroke” phase of the polypeptide transport cycle, while subsequent steps may reflect an entropic expansion more resembling a “Brownian-ratchet” phase. Here, additional force may be applied to the substrate as Brownian motion in the helix allows it to move past the anionic-repulsion barrier. As acidic residues reach the higher pH of the cytosolic side of the barrier, they will become more frequently deprotonated, making the peptide net negative and significantly less capable of retrotranslocation. This deprotonation will be coupled to the helix-to-coil transition in the substrate helix, and an entropic tension may develop in the extended random-coil chain. The resolution of this tension is likely the reformation of helical structure via translocation of upstream chain. This restoring force will be accompanied by a new round of substrate protonation and helix formation that allows new helical structure to pass through the anionic charge repulsion site. Repeated cycles of these processes lead to the unfolding of each domain and complete translocation.

# References


1. Schatz, G., and Dobberstein, B. (1996) Common principles of protein translocation across membranes. *Science* **271**, 1519–1526

2. Ruiz, N., Kahne, D., and Silhavy, T. J. (2006) Advances in understanding bacterial outer-membrane biogenesis. *Nat. Rev. Microbiol.* **4**, 57–66

3. Tsai, B., Ye, Y., and Rapoport, T. A. (2002) Retro-translocation of proteins from the endoplasmic reticulum into the cytosol. *Nat. Rev. Mol. Cell Biol.* **3**, 246–255

4. Driessen, A. J. M., and Nouwen, N. (2008) Protein translocation across the bacterial cytoplasmic membrane. *Annu. Rev. Biochem.* **77**, 643–667

5. Galán, J. E., and Wolf-Watz, H. (2006) Protein delivery into eukaryotic cells by type III secretion machines. *Nature* **444**, 567–573

6. Cascales, E., and Christie, P. J. (2003) The versatile bacterial type IV secretion systems. *Nat. Rev. Microbiol.* **1**, 137–149

7. Lycklama A Nijeholt, J. A., and Driessen, A. J. M. (2012) The bacterial Sec-translocase: structure and mechanism. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* **367**, 1016–1028

8. Thoren, K. L., and Krantz, B. A. (2011) The unfolding story of anthrax toxin translocation. *Mol. Microbiol.* **80**, 588–595

9. Egea, P. F., Stroud, R. M., and Walter, P. (2005) Targeting proteins to membranes: structure of the signal recognition particle. *Curr. Opin. Struct. Biol.* **15**, 213–220

10. Hetzer, M. W., Walther, T. C., and Mattaj, I. W. (2005) Pushing the envelope: structure, function, and dynamics of the nuclear periphery. *Annu. Rev. Cell Dev. Biol.* **21**, 347–380

11. Sargent, F. (2007) The twin-arginine transport system: moving folded proteins across membranes. *Biochem. Soc. Trans.* **35**, 835–847

12. Glick, B. S., Wachter, C., Reid, G. A., and Schatz, G. (1993) Import of cytochrome b2 to the mitochondrial intermembrane space: the tightly folded heme-binding domain makes import dependent upon matrix ATP. *Protein Sci.* **2**, 1901–1917

13. Sauer, R. T., and Baker, T. A. (2011) AAA+ proteases: ATP-fueled machines of protein destruction. *Annu. Rev. Biochem.* **80**, 587–612

14. Tomkiewicz, D., Nouwen, N., and Driessen, A. J. M. (2007) Pushing, pulling and trapping--modes of motor protein supported protein translocation. *FEBS Lett.* **581**, 2820–2828

15. Feynman, R. P., Leighton, L. B., and Sands, M. (1963) *The Feynman Lectures on Physics*, Addison-Wesley, Reading, MA

16. Simon, S. M., Peskin, C. S., and Oster, G. F. (1992) What drives the translocation of proteins? *Proc. Natl. Acad. Sci. U.S.A.* **89**, 3770–3774

17. Astumian, R. D. (1997) Thermodynamics and kinetics of a Brownian motor. *Science* **276**, 917–922

18. Krantz, B. A., Finkelstein, A., and Collier, R. J. (2006) Protein translocation through the anthrax toxin transmembrane pore is driven by a proton gradient. *J. Mol. Biol.* **355**, 968–979

19. Smoluchowski, M. (1912) Experimentell nachweisbare, der Ublichen Thermodynamik widersprechende Molekularphenomene. *Phys. Zeitshur.* **13**, 1069–1079

20. Einstein, A. (1905) Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Annalen der Physik* **17**, 548–560

21. Brown, M. J., Thoren, K. L., and Krantz, B. A. (2011) Charge requirements for proton gradient-driven translocation of anthrax toxin. *J. Biol. Chem.* **286**, 23189–23199

22. Matlack, K. E., Misselwitz, B., Plath, K., and Rapoport, T. A. (1999) BiP acts as a molecular ratchet during posttranslational transport of prepro-alpha factor across the ER membrane. *Cell* **97**, 553–564

23. Liebermeister, W., Rapoport, T. A., and Heinrich, R. (2001) Ratcheting in post-translational protein translocation: a mathematical model. *J. Mol. Biol.* **305**, 643–656

24. Glick, B. S. (1995) Can Hsp70 proteins act as force-generating motors? *Cell* **80**, 11–14

25. Bier, M. (2003) Processive motor protein as an overdamped Brownian stepper. *Phys. Rev. Lett.* **91**, 1–4

26. Drazin, R., Kandel, J., and Collier, R. J. (1971) Structure and activity of diphtheria toxin. II. Attack by trypsin at a specific site within the intact toxin molecule. *J. Biol. Chem.* **246**, 1504–1510

27. Barth, H., Aktories, K., Popoff, M. R., and Stiles, B. G. (2004) Binary bacterial toxins: biochemistry, biology, and applications of common Clostridium and Bacillus proteins. *Microbiol. Mol. Biol. Rev.* **68**, 373–402

28. Falnes, P. O., and Sandvig, K. (2000) Penetration of protein toxins into cells. *Curr. Opin. Cell Biol.* **12**, 407–413

29. Collier, R. J. (2009) Membrane translocation by anthrax toxin. *Mol. Aspects Med.* **30**, 413–422

30. Young, J. A. T., and Collier, R. J. (2007) Anthrax toxin: receptor binding, internalization, pore formation, and translocation. *Annu. Rev. Biochem.* **76**, 243–265

31. Feld, G. K., Kintzer, A. F., Tang, I. I., Thoren, K. L., and Krantz, B. A. (2012) Domain flexibility modulates the heterogeneous assembly mechanism of anthrax toxin protective antigen. *J. Mol. Biol.* **415**, 159–174

32. Pannifer, A. D., Wong, T. Y., Schwarzenbacher, R., Renatus, M., Petosa, C., Bienkowska, J., Lacy, D. B., Collier, R. J., Park, S., Leppla, S. H., Hanna, P., and Liddington, R. C. (2001) Crystal structure of the anthrax lethal factor. *Nature* **414**, 229–233

33. Shen, Y., Zhukovskaya, N. L., Guo, Q., Florián, J., and Tang, W. J. (2005) Calcium-independent calmodulin binding and two-metal–ion catalytic mechanism of anthrax edema factor. *EMBO J.* **24**, 929–941

34. Feld, G. K., Thoren, K. L., Kintzer, A. F., Sterling, H. J., Tang, I. I., Greenberg, S. G., Williams, E. R., and Krantz, B. A. (2010) Structural basis for the unfolding of anthrax lethal factor by protective antigen oligomers. *Nat. Struct. Mol. Biol.* **17**, 1383–1390

35. Katayama, H., Janowiak, B. E., Brzozowski, M., Juryck, J., Falke, S., Gogol, E. P., Collier, R. J., and Fisher, M. T. (2008) GroEL as a molecular scaffold for structural analysis of the anthrax toxin pore. *Nat. Struct. Mol. Biol.* **15**, 754–760

36. Kintzer, A. F., Thoren, K. L., Sterling, H. J., Dong, K. C., Feld, G. K., Tang, I. I., Zhang, T. T., Williams, E. R., Berger, J. M., and Krantz, B. A. (2009) The protective antigen component of anthrax toxin forms functional octameric complexes. *J. Mol. Biol.* **392**, 614–629

37. Mogridge, J., Cunningham, K., and Collier, R. J. (2002) Stoichiometry of anthrax toxin complexes. *Biochemistry* **41**, 1079–1082

38. Lacy, D. B., Wigelsworth, D. J., Melnyk, R. A., Harrison, S. C., and Collier, R. J. (2004) Structure of heptameric protective antigen bound to an anthrax toxin receptor: a role for receptor in pH-dependent pore formation. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 13147–13151

39. Santelli, E., Bankston, L. A., Leppla, S. H., and Liddington, R. C. (2004) Crystal structure of a complex between anthrax toxin and its host cell receptor. *Nature* **430**, 905–908

40. Lacy, D. B., Lin, H. C., Melnyk, R. A., Schueler-Furman, O., Reither, L., Cunningham, K., Baker, D., and Collier, R. J. (2005) A model of anthrax toxin lethal factor bound to protective antigen. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 16409–16414

41. Miller, C. J., Elliott, J. L., and Collier, R. J. (1999) Anthrax protective antigen: prepore-to-pore conversion. *Biochemistry* **38**, 10432–10441

42. Blaustein, R. O., Koehler, T. M., Collier, R. J., and Finkelstein, A. (1989) Anthrax toxin: channel-forming activity of protective antigen in planar phospholipid bilayers. *Proc. Natl. Acad. Sci. U.S.A.* **86**, 2209–2213

43. Nguyen, T. L. (2004) Three-dimensional model of the pore form of anthrax protective antigen. Structure and biological implications. *J. Biomol. Struct. Dyn.* **22**, 253–265

44. Krantz, B. A., Trivedi, A. D., Cunningham, K., Christensen, K. a, and Collier, R. J. (2004) Acid-induced unfolding of the amino-terminal domains of the lethal and edema factors of anthrax toxin. *J. Mol. Biol.* **344**, 739–756

45. Thoren, K. L., Worden, E. J., Yassif, J. M., and Krantz, B. A. (2009) Lethal factor unfolding is the most force-dependent step of anthrax toxin translocation. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 21555–21560

46. Petosa, C., Collier, R. J., Klimpel, K. R., Leppla, S. H., and Liddington, R. C. (1997) Crystal structure of the anthrax toxin protective antigen. *Nature* **385**, 833–838

47. Duesbery, N. S., Webb, C. P., Leppla, S. H., Gordon, V. M., Klimpel, K. R., Copeland, T. D., Ahn, N. G., Oskarsson, M. K., Fukasawa, K., Paull, K. D., and Vande Woude, G. F. (1998) Proteolytic inactivation of MAP-kinase-kinase by anthrax lethal factor. *Science* **280**, 734–737

48. Leppla, S. H. (1982) Anthrax toxin edema factor: a bacterial adenylate cyclase that increases cyclic AMP concentrations of eukaryotic cells. *Proc. Natl. Acad. Sci. U.S.A.* **79**, 3162–3166

49. Leppla, S. H. (1984) Bacillus anthracis calmodulin-dependent adenylate cyclase: chemical and enzymatic properties and interactions with eucaryotic cells. *Adv. Cyclic Nucleotide Protein Phosphorylation Res.* **17**, 189–198

50. Drum, C. L., Yan, S. Z., Bard, J., Shen, Y. Q., Lu, D., Soelaiman, S., Grabarek, Z., Bohm, A., and Tang, W. J. (2002) Structural basis for the activation of anthrax adenylyl cyclase exotoxin by calmodulin. *Nature* **415**, 396–402

51. Molloy, S. S., Bresnahan, P. A., Leppla, S. H., Klimpel, K. R., and Thomas, G. (1992) Human furin is a calcium-dependent serine endoprotease that recognizes the sequence Arg-X-X-Arg and efficiently cleaves anthrax toxin protective antigen. *J. Biol. Chem.* **267**, 16396–16402

52. Klimpel, K. R., Molloy, S. S., Thomas, G., and Leppla, S. H. (1992) Anthrax toxin protective antigen is activated by a cell surface protease with the sequence specificity and catalytic properties of furin. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 10277–10281

53. Milne, J. C., Furlong, D., Hanna, P. C., Wall, J. S., and Collier, R. J. (1994) Anthrax protective antigen forms oligomers during intoxication of mammalian cells. *J. Biol. Chem.* **269**, 20607–20612

54. Kintzer, A. F., Sterling, H. J., Tang, I. I., Abdul-Gader, A., Miles, A. J., Wallace, B. A., Williams, E. R., and Krantz, B. A. (2010) Role of the protective antigen octamer in the molecular mechanism of anthrax lethal toxin stabilization in plasma. *J. Mol. Biol.* **399**, 741–758

55. Cunningham, K., Lacy, D. B., Mogridge, J., and Collier, R. J. (2002) Mapping the lethal factor and edema factor binding sites on oligomeric anthrax protective antigen. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 7049–7053

56. Chauhan, V., and Bhatnagar, R. (2002) Identification of amino acid residues of anthrax protective antigen involved in binding with lethal factor. *Infect. Immun.* **70**, 4477–4484

57. Lacy, D. B., Mourez, M., Fouassier, A., and Collier, R. J. (2002) Mapping the anthrax protective antigen binding site on the lethal and edema factors. *J. Biol. Chem.* **277**, 3006–3010

58. Mourez, M., Yan, M., Lacy, D. B., Dillon, L., Bentsen, L., Marpoe, A., Maurin, C., Hotze, E., Wigelsworth, D., Pimental, R. A., Ballard, J. D., Collier, R. J., and Tweten, R. K. (2003) Mapping dominant-negative mutations of anthrax protective antigen by scanning mutagenesis. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 13803–13808

59. Melnyk, R. A., Hewitt, K. M., Lacy, D. B., Lin, H. C., Gessner, C. R., Li, S., Woods, V. L., and Collier, R. J. (2006) Structural determinants for the binding of anthrax lethal factor to oligomeric protective antigen. *J. Biol. Chem.* **281**, 1630–1635

60. Schueler-Furman, O., Wang, C., and Baker, D. (2005) Progress in protein-protein docking: atomic resolution predictions in the CAPRI experiment using RosettaDock with an improved treatment of side-chain flexibility. *Proteins* **60**, 187–194

61. Bradley, K. A., Mogridge, J., Mourez, M., Collier, R. J., and Young, J. A. (2001) Identification of the cellular receptor for anthrax toxin. *Nature* **414**, 225–229

62. Scobie, H. M., Rainey, G. J. A., Bradley, K. A., and Young, J. A. T. (2003) Human capillary morphogenesis protein 2 functions as an anthrax toxin receptor. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 5170–5174

63. Koehler, T. M., and Collier, R. J. (1991) Anthrax toxin protective antigen: low-pH-induced hydrophobicity and channel formation in liposomes. *Mol. Microbiol.* **5**, 1501–1506

64. Qa'dan, M., Christensen, K. A., Zhang, L., Roberts, T. M., and Collier, R. J. (2005) Membrane insertion by anthrax protective antigen in cultured cells. *Mol. Cell. Biol.* **25**, 5492–5498

65. Zhang, S., Udho, E., Wu, Z., Collier, R. J., and Finkelstein, A. (2004) Protein translocation through anthrax toxin channels formed in planar lipid bilayers. *Biophys. J.* **87**, 3842–3849

66. Basilio, D., Kienker, P. K., Briggs, S. W., and Finkelstein, A. (2011) A kinetic analysis of protein transport through the anthrax toxin channel. *J. Gen. Physiol.* **137**, 521–531

67. Krantz, B. A., Melnyk, R. A., Zhang, S., Juris, S. J., Lacy, D. B., Wu, Z., Finkelstein, A., and Collier, R. J. (2005) A phenylalanine clamp catalyzes protein translocation through the anthrax toxin pore. *Science* **309**, 777–781

68. Pentelute, B. L., Sharma, O., and Collier, R. J. (2011) Chemical dissection of protein translocation through the anthrax toxin pore. *Angew. Chem. Int. Ed. Engl.* **50**, 2294–2296

69. Pentelute, B. L., Barker, A. P., Janowiak, B. E., Kent, S. B. H., and Collier, R. J. (2010) A semisynthesis platform for investigating structure-function relationships in the N-terminal domain of the anthrax Lethal Factor. *ACS Chem. Biol.* **5**, 359–364

70. Basilio, D., Jennings-Antipov, L. D., Jakes, K. S., and Finkelstein, A. (2011) Trapping a translocating protein within the anthrax toxin channel: implications for the secondary structure of permeating proteins. *J. Gen. Physiol.* **137**, 343–356

71. Basilio, D., Juris, S. J., Collier, R. J., and Finkelstein, A. (2009) Evidence for a proton–protein symport mechanism in the anthrax toxin channel. *J. Gen. Physiol.* **133**, 307–314

72. Janowiak, B. E., Fischer, A., and Collier, R. J. (2010) Effects of introducing a single charged residue into the phenylalanine clamp of multimeric anthrax protective antigen. *J. Biol. Chem.* **285**, 8130–8137

73. Fischer, A., Holden, M. A., Pentelute, B. L., and Collier, R. J. (2011) Ultrasensitive detection of protein translocated through toxin pores in droplet-interface bilayers. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 16577–16581

74. Blaustein, R. O., Lea, E. J., and Finkelstein, A. (1990) Voltage-dependent block of anthrax toxin channels in planar phospholipid bilayer membranes by symmetric tetraalkylammonium ions. Single-channel analysis. *J. Gen. Physiol.* **96**, 921–942

75. Jennings-Antipov, L. D., Song, L., and Collier, R. J. (2011) Interactions of anthrax lethal factor with protective antigen defined by site-directed spin labeling. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 1868–1873

76. Harsman, A., Bartsch, P., Hemmis, B., Krüger, V., and Wagner, R. (2011) Exploring protein import pores of cellular organelles at the single molecule level using the planar lipid bilayer technique. *Eur. J. Cell Biol.* **90**, 721–730

77. Grigoriev, S. M., Muro, C., Dejean, L. M., Campo, M. L., Martinez-Caballero, S., and Kinnally, K. W. (2004) Electrophysiological approaches to the study of protein translocation in mitochondria. *Int. Rev. Cytol.* **238**, 227–274

78. Henry, J. P., Juin, P., Vallette, F., and Thieffry, M. (1995) Characterization and function of the mitochondrial outer membrane peptide-sensitive channel. *J. Bioenerg. Biomembr.* **27**, 15–22

79. Movileanu, L., Schmittschmitt, J. P., Scholtz, J. M., and Bayley, H. (2005) Interactions of peptides with a protein pore. *Biophys. J.* **89**, 1030–1045

80. Mahendran, K. R., Romero-Ruiz, M., Schlösinger, A., Winterhalter, M., and Nussberger, S. (2012) Protein translocation through Tom40: kinetics of peptide release. *Biophys. J.* **102**, 39–47

81. Senzel, L., Huynh, P. D., Jakes, K. S., John Collier, R., and Finkelstein, A. (1998) The diphtheria toxin channel-forming T domain translocates its own NH2-terminal region across planar bilayers. *J. Gen. Physiol.* **112**, 317–324

82. Montal, M. (2009) Translocation of botulinum neurotoxin light chain protease by the heavy chain protein-conducting channel. *Toxicon* **54**, 565–569

83. Hoch, D. H., Romero-Mira, M., Ehrlich, B. E., Finkelstein, A., DasGupta, B. R., and Simpson, L. L. (1985) Channels formed by botulinum, tetanus, and diphtheria toxins in planar lipid bilayers: relevance to translocation of proteins across membranes. *Proc. Natl. Acad. Sci. U.S.A.* **82**, 1692–1696

84. Friedlander, A. M. (1986) Macrophages are sensitive to anthrax lethal toxin through an acid-dependent process. *J. Biol. Chem.* **261**, 7123–7126

85. Zhang, S., Finkelstein, A., and Collier, R. J. (2004) Evidence that translocation of anthrax toxin's lethal factor is initiated by entry of its N terminus into the protective antigen channel. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 16756–16761

86. Blanke, S. R., Milne, J. C., Benson, E. L., and Collier, R. J. (1996) Fused polycationic peptide mediates delivery of diphtheria toxin A chain to the cytosol in the presence of anthrax protective antigen. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 8437–8442

87. Katayama, H., Wang, J., Tama, F., Chollet, L., Gogol, E. P., Collier, R. J., and Fisher, M. T. (2010) Three-dimensional structure of the anthrax toxin pore inserted into lipid nanodiscs and lipid vesicles. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 3453–3457

88. Nassi, S., Collier, R. J., and Finkelstein, A. (2002) PA63 channel of anthrax toxin: an extended beta-barrel. *Biochemistry* **41**, 1445–1450

89. Benson, E. L., Huynh, P. D., Finkelstein, A., and Collier, R. J. (1998) Identification of residues lining the anthrax protective antigen channel. *Biochemistry* **37**, 3941–3948

90. Meador, W. E., Means, A. R., and Quiocho, F. A. (1993) Modulation of calmodulin plasticity in molecular recognition on the basis of x-ray structures. *Science* **262**, 1718–1721

91. Meador, W. E., Means, A. R., and Quiocho, F. A. (1992) Target enzyme recognition by calmodulin: 2.4 A structure of a calmodulin-peptide complex. *Science* **257**, 1251–1255

92. Finkelstein, A. (2009) Proton-coupled protein transport through the anthrax toxin channel. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* **364**, 209–215

93. Martin, A., Baker, T. A., and Sauer, R. T. (2008) Pore loops of the AAA+ ClpX machine grip substrates to drive translocation and unfolding. *Nat. Struct. Mol. Biol.* **15**, 1147–1151

94. Zimm, B. H., and Bragg, J. K. (1959) Theory of the phase transition between helix and random coil in polypeptide chains. *J. Chem. Phys.* **31**, 526–535

95. Brockwell, D. J., Paci, E., Zinober, R. C., Beddard, G. S., Olmsted, P. D., Smith, D. A., Perham, R. N., and Radford, S. E. (2003) Pulling geometry defines the mechanical resistance of a beta-sheet protein. *Nat. Struct. Biol.* **10**, 731–737

96. Cecconi, C., Shank, E. A., Bustamante, C., and Marqusee, S. (2005) Direct observation of the three-state folding of a single protein molecule. *Science* **309**, 2057–2060

97. Borgia, A., Williams, P. M., and Clarke, J. (2008) Single-molecule studies of protein folding. *Annu. Rev. Biochem.* **77**, 101–125

98. Crampton, N., and Brockwell, D. J. (2010) Unravelling the design principles for single protein mechanical strength. *Curr. Opin. Struct. Biol.* **20**, 508–517

99. Wickner, W., and Schekman, R. (2005) Protein translocation across biological membranes. *Science* **310**, 1452–1456

100. Sauer, R. T., Bolon, D. N., Burton, B. M., Burton, R. E., Flynn, J. M., Grant, R. A., Hersch, G. L., Joshi, S. A., Kenniston, J. A., Levchenko, I., Neher, S. B., Oakes, E. S. C., Siddiqui, S. M., Wah, D. A., and Baker, T. A. (2004) Sculpting the proteome with AAA(+) proteases and disassembly machines. *Cell* **119**, 9–18

101. Huang, S., Ratliff, K. S., and Matouschek, A. (2002) Protein unfolding by the mitochondrial membrane potential. *Nat. Struct. Biol.* **9**, 301–307

102. Shariff, K., Ghosal, S., and Matouschek, A. (2004) The force exerted by the membrane potential during protein import into the mitochondrial matrix. *Biophys. J.* **86**, 3647–3652

103. Cross, R. L., and Müller, V. (2004) The evolution of A-, F-, and V-type ATP synthases and ATPases: reversals in function and changes in the H+/ATP coupling ratio. *FEBS Lett.* **576**, 1–4

104. Sowa, Y., and Berry, R. M. (2008) Bacterial flagellar motor. *Q. Rev. Biophys.* **41**, 103–132

105. Berg, H. C. (2003) The rotary motor of bacterial flagella. *Annu. Rev. Biochem.* **72**, 19–54

106. Oster, G., and Wang, H. (1999) ATP synthase: two motors, two fuels. *Structure* **7**, 67–72

107. Abrahams, J. P., Leslie, A. G. W., Lutter, R., and Walker, J. E. (1994) Structure at 2.8-angstrom resolution of F1-ATPase from bovine heart-mitochondria. *Nature* **370**, 621–628

108. Gregorini, M., Wang, J., Xie, X. S., Milligan, R. A., and Engel, A. (2007) Three-dimensional reconstruction of bovine brain V-ATPase by cryo-electron microscopy and single particle analysis. *J. Struct. Biol.* **158**, 445–454

109. Nakanishi-Matsui, M., Sekiya, M., Nakamoto, R. K., and Futai, M. (2010) The mechanism of rotating proton pumping ATPases. *Biochim. Biophys. Acta* **1797**, 1343–1352

110. Finbow, M. E., and Harrison, M. A. (1997) The vacuolar H+-ATPase: a universal proton pump of eukaryotes. *Biochem. J.* **324**, 697–712

111. Boyer, P. D. (1997) The ATP synthase--a splendid molecular machine. *Annu. Rev. Biochem.* **66**, 717–749

112. Noji, H., Yasuda, R., Yoshida, M., and Kinosita Jr., K. (1997) Direct observation of the rotation of F1-ATPase. *Nature* **386**, 299–302

113. Voos, W., Martin, H., Krimmer, T., and Pfanner, N. (1999) Mechanisms of protein translocation into mitochondria. *Biochim. Biophys. Acta* **1422**, 235–254

114. Llopis, J., McCaffery, J. M., Miyawaki, A., Farquhar, M. G., and Tsien, R. Y. (1998) Measurement of cytosolic, mitochondrial, and Golgi pH in single living cells with green fluorescent proteins. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 6803–6808

115. Wang, J., Song, J., Seong, I., and Franklin, M. (2001) Nucleotide-dependent conformational changes in a protease-associated ATPase HslU. *Structure* **9**, 1107–1116

116. Lum, R., Niggemann, M., and Glover, J. R. (2008) Peptide and protein binding in the axial channel of Hsp104. Insights into the mechanism of protein unfolding. *J. Biol. Chem.* **283**, 30139–30150

117. Martin, A., Baker, T. A., and Sauer, R. T. (2008) Protein unfolding by a AAA+ protease is dependent on ATP-hydrolysis rates and substrate energy landscapes. *Nat. Struct. Mol. Biol.* **15**, 139–145

118. DeLaBarre, B., and Brunger, A. T. (2005) Nucleotide dependent motion and mechanism of action of p97/VCP. *J. Mol. Biol.* **347**, 437–452

119. Astumian, R. D., and Bier, M. (1994) Fluctuation driven ratchets: molecular motors. *Phys. Rev. Lett.* **72**, 1766–1769

120. Yamano, K., Kuroyanagi-Hasegawa, M., Esaki, M., Yokota, M., and Endo, T. (2008) Step-size analyses of the mitochondrial Hsp70 import motor reveal the Brownian ratchet in operation. *J. Biol. Chem.* **283**, 27325–27332

121. Stuart, R. A., Cyr, D. M., Craig, E. A., and Neupert, W. (1994) Mitochondrial molecular chaperones: their role in protein translocation. *Trends Biochem. Sci.* **19**, 87–92

122. Matouschek, A., Azem, A., Ratliff, K., Glick, B. S., Schmid, K., and Schatz, G. (1997) Active unfolding of precursor proteins during mitochondrial protein import. *EMBO J.* **16**, 6727–6736

123. Huang, S., Ratliff, K. S., Schwartz, M. P., Spenner, J. M., and Matouschek, A. (1999) Mitochondria unfold precursor proteins by unraveling them from their N-termini. *Nat. Struct. Biol.* **6**, 1132–1138

124. Feld, G. K., Brown, M. J., and Krantz, B. A. (2012) Ratcheting up protein translocation with anthrax toxin. *Protein Sci.* **21**, 606–624

125. Kintzer, A. F., Sterling, H. J., Tang, I. I., Williams, E. R., and Krantz, B. A. (2010) Anthrax toxin receptor drives protective antigen oligomerization and stabilizes the heptameric and octameric oligomer by a similar mechanism. *PLoS ONE* **5**, e13888

126. Song, L., Hobaugh, M. R., Shustak, C., Cheley, S., Bayley, H., and Gouaux, J. E. (1996) Structure of staphylococcal alpha-hemolysin, a heptameric transmembrane pore. *Science* **274**, 1859–1866

127. Vernier, G., Wang, J., Jennings, L. D., Sun, J., Fischer, A., Song, L., and Collier, R. J. (2009) Solubilization and characterization of the anthrax toxin pore in detergent micelles. *Protein Sci.* **18**, 1882–1895

128. Mueller, P., Rudin, D. O., Tien, H. T., and Wescott, W. C. (1963) Methods for the formation of single bimolecular lipid membranes in aqueous solution. *J. Phys. Chem.* **67**, 534–535

129. Robinson, R. A., and Stokes, R. H. (2002) *Electrolyte Solutions*, 2nd Ed., Dover, Mineola, New York

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

131. Emsley, P., and Cowtan, K. (2004) COOT: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132

132. Woodhull, A. M. (1973) Ionic blockage of sodium channels in nerve. *J. Gen. Physiol.* **61**, 687–708

133. Levinthal, C. (1968) Are there pathways for protein folding? *J. Chim. Phys.* **65**, 44–45

134. Wynia Smith, S. L., Brown, M. J., Chirichella, G., Kemalyan, G., and Krantz, B. A. (2012) Electrostatic ratchet in the protective antigen channel promotes anthrax toxin translocation. *J. Biol. Chem.* **287**, 43753–43764

135. Schwass, D. E., and Finley, J. W. (1984) Heat and alkaline damage to proteins: racemization and lysinoalanine formation. *J. Agric. Food Chem.* **32**, 1377–1382

136. Adzhubei, A. A., and Sternberg, M. J. E. (1993) Left-handed polyproline II helices commonly occur in globular proteins. *J. Mol. Biol.* **229**, 472–493

137. Finley, J. W., and Kohler, G. O. (1979) Processing conditions to inhibit lysinoalanine formation in alkaline-treated proteins. *Cereal Chem.* **56**, 130–132