Title
Mechanisms of translocation-coupled protein unfolding using anthrax toxin as a model

Author
Thoren, Katie Lynn

Publication Date
2011

Peer reviewed|Thesis/dissertation

Mechanisms of translocation-coupled protein unfolding using anthrax toxin as a model

by

Katie Lynn Thoren

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Chemistry

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Bryan Krantz, Chair
Professor Judith Klinman
Professor John Kuriyan
Professor Susan Marqusee

Spring 2011

Mechanisms of translocation-coupled protein unfolding using anthrax toxin as a model

Katie Lynn Thoren

Abstract

Mechanisms of translocation-coupled protein unfolding using anthrax toxin as a model

by

Katie Lynn Thoren

Doctor of Philosophy in Chemistry

University of California, Berkeley

Professor Bryan Krantz, Chair

Molecular machines face a number of challenges in transporting a protein either across a membrane or into a proteolytic complex. In many cases, a substrate protein must first be unfolded before being transported through a narrow channel. Despite its importance and relevance to a variety of different processes in the cell, translocation-coupled protein unfolding is still not well understood. In effort to determine the general biophysical mechanisms of this process, anthrax toxin is used as a model system.

In order to understand how a protein unfolds on a translocase channel, planar lipid bilayer electrophysiology, site-directed mutagenesis and thermodynamic stability studies were used to first identify the barriers in the translocation pathway and determine which barrier corresponds to substrate unfolding. Working under conditions where substrate unfolding is rate-limiting, we were then able to map how LF_N actually unfolds on the surface of the PA channel.

Next, the role of the channel in substrate unfolding and translocation is discussed. In particular, a novel substrate binding site on the surface of PA was identified from the crystal structure of a PA octamer bound to four LF_N substrates. This structure, which was solved by my colleague, Geoffrey Feld, reveals that the first α helix and β strand of each LF_N unfold and dock into a deep amphipathic cleft, termed the α clamp. Through extensive mutatgenesis studies on both PA and LF_N, Geoff and I determined that this site can bind a broad array of polypeptide substrates. The role of the α clamp in substrate unfolding, channel oligomerization and translocation is investigated and discussed.

Finally, in effort to further probe the α clamp's role in translocation, binding to the site is disrupted and the effects on translocation are investigated. Preliminary hypotheses and future directions are discussed.

Dedicated to Geoff and my family.

Table of Contents

List of Figures

List of Tables

List of Abreviations

$\Delta\Psi$	membrane potential
ΔpH	proton gradient
ATP	adenosine-5'-triphosphate
Atx	anthrax toxin
cAMP	3’-5’-cyclic adenosine monophosphate
CD	circular dichroism
CD_{222}	CD signal at 222 nm
DTA	diphtheria toxin A chain
EDTA	ethylenediaminetetra acetic acid
EF	edema factor
EF_N	amino-terminal domain of edema factor
FA	fluorescence anisotropy
GdmCl	guanidinium chloride
Lethal toxin	lethal factor + PA
LF	lethal factor
LF_N	amino-terminal domain of lethal factor
MAP	mitogen-activated protein
MIL	membrane insertion loop
PA	protective antigen
$PA^{\Delta MIL}$	protective antigen Δ303-324, V303P and H304G
PA_{20}	~20 kDa fragment of PA
PA_{63}	~63 kDa fragment of PA
PA_{83}	~83 kDa fragment of PA
PDB	protein data bank
PMSF	phenylmethylsulfonyl fluoride
UBB	universal bilayer buffer

Acknowledgements

I would like to thank all of the people who donated generously of their time, experience, and resources during the course of my graduate career. I am greatly indebted to the people who have mentored me along the way. First and foremost, I would like to thank my advisor, Bryan Krantz. I was fortunate to work closely with him in the lab, and I am extremely grateful for the scientific training and guidance he has provided me. I would also like to thank my thesis committee for their mentorship and time: Judith Klinman, John Kuriyan and Susan Marqusee.

I wish to thank the former and current members of the Krantz lab for their support and camaraderie: Mike Brown, Jen Colby, Geoff Feld, Shon Greenberg, Melinda Hanes, Alex Kintzer, Adam Schawel, Sam Stephenson, Yoki Tang, Sarah Wynia Smith, and Jaime Yassif. Thank you also to Evan Worden, an undergraduate student who worked with me for two summers, and Harry Sterling, a graduate student in the Williams lab, who both significantly contributed to the work I present in this thesis.

I wish to thank members of Marqusee lab, especially Katelyn Connell, Katie Tripp and Phillip Elms for their friendship and scientific advice. I am also grateful to Albert Lang for keeping me grounded and well fed. Finally, this thesis would not have been possible without the support of my family and my boyfriend, Geoff Brunn.

This work was supported by the National Institutes of Health Research Grant R01-AI077703 (to Bryan Krantz), and National Institutes of Health Training Grant T32GM08295.

Chapter 1

Introduction to Protein Translocation

A protein's function depends not only on its structure, but also on its correct location. Since most protein synthesis occurs in the cytosol, proteins that function within a membrane, outside the cell or within a subcellular compartment must translocate into or across a membrane. In eukaryotic cells, it is estimated that almost half of all proteins have to cross a membrane in order to reach their site of function (Schatz & Dobberstein, 1996). While prokaryotes lack the complex compartmentalization of eukaryotes, a significant portion of all bacterial proteins (25-30%) function outside of the cytosol and therefore must be translocated at least across the cytoplasmic membrane (Driessen & Nouwen, 2008). Gram-negative bacteria also have an outer membrane, which presents an additional challenge for protein secretion.

The fact that transmembrane protein translocation is a fundamental process in the cell is exemplified by the number and variety of transport systems observed in both eukaryotes and prokaryotes. Translocation systems are found in virtually every membrane-encapsulated organelle in the cell. They play an essential role in many biological processes such as membrane and organelle biogenesis, protein trafficking, antigen presentation, microbial pathogenesis via Types I-VI secretion apparatuses and toxin delivery into host cells, and protein quality control through the endoplasmic reticulum associated degradation pathway.

Despite variations in location, size and specific architecture, the translocation machinery involved in these processes shares many common features. Importantly, many translocation systems are comprised of a membrane-spanning proteinaceous channel, or translocase, that serves as the passageway for proteins to cross or insert into the membrane. While some channels, like the twin arginine translocase, are wide enough to accommodate folded proteins (Sargent, 2007), many are so narrow that their substrates must be transported as unfolded chains (Lee & Schneewind, 2002, Falnes *et al.*, 1994, Eilers & Schatz, 1986). In some systems, translocation occurs co-translationally or the protein is delivered to the translocase in a chaperone-maintained, unfolded state. In other instances, however, the substrate protein folds prior to import. Thus in these cases, the translocation reaction essentially breaks down into two fundamental and interconnected steps: protein unfolding and transport through a narrow channel. Substrate unfolding is required in a number of different translocation reactions including transport of bacterial toxins into host cells (Falnes et al., 1994, Haug *et al.*, 2003, Wesche *et al.*, 1998), translocation through the Type III and Type IV secretion systems (Lee & Schneewind, 2002), and translocation into several organelles, such as mitochondria (Eilers & Schatz, 1986, Rassow *et al.*, 1990), chloroplasts (Walker *et al.*, 1996, Ruprecht *et al.*), and in some cases, the endoplasmic reticulum (ER) (Paunola *et al.*, 1998).

1.1 Protein degradation and disaggregase machines

The fundamental and interconnected steps of protein unfolding and subsequent transport through a narrow channel are also relevant to soluble molecular machines in the cell such as protein degradation complexes and disaggregases. These machines play an important role in the maintenance of proteins within the cell as they are responsible for degrading unneeded or abnormal proteins (Gottesman & Maurizi, 1992). Removing proteins when they are no longer needed is an essential regulatory mechanism that is important for cellular processes such as cell cycle control, signal transduction, and antigen presentation (Gottesman, 2003, Hershko &

Ciechanover, 1998). Degradation of misfolded or abnormal proteins is also critical in order to prevent a variety of diseases characterized by protein aggregation, such as Alzheimer's, Parkinson's and diabetes (Wickner *et al.*, 1999).

Four classes of energy-dependent proteases have been identified in *E. coli* (ClpAP/XP, Lon, FtsH and ClpYQ/HsIUV); the same proteases are also found in other prokaryotes and in eukaryotic organelles (mitochondria and chloroplasts) (Gottesman, 2003). These proteases, and the 26S proteasome, which is responsible for most protein degradation in the eukaryotic cytosol, share a common architecture. In general, these degradation machines are comprised of two major components: (i) a hollow, cylinder-shaped protease component, and (ii) a ring-shaped, ATPase component that caps both ends of the proteolytic chamber (Figure 1.1). Substrates enter the proteolytic chamber through the narrow channel formed by the ATPase ring. This narrow opening offers some advantages to the protease in that it restricts access to the proteolytic active sites and sequesters substrate proteins within the proteolytic chamber. However, the narrow aperture also means that the substrate must be unfolded before being transported into the chamber for degradation.

How does the ATPase component unfold and translocate proteins into the protolytic chamber? The ATPase domains involved in protein degradation are members of the AAA+ (ATPases associated with various cellular activities) family. As such, they contain a homologous ATP-binding motif, and undergo nucleotide-dependent conformational changes that are coupled to protein unfolding and translocation. Interestingly, some ATPases can function in the absence of the proteolytic component (Hoskins *et al.*, 2001). ClpB (Hsp104 in eukaryotes), for example, does not associate with a protease, and may only act as a chaperone to disaggregate insoluble protein aggregates (Zolkiewski, 1999, Motohashi *et al.*, 1999, Doyle & Wickner, 2009). Understanding how ATP-dependent conformational changes are coupled to protein unfolding and translocation is an active area of research.

1.2 Questions

Despite its importance and relevance to a variety of processes in the cell, translocation-coupled protein unfolding is still not well understood. The processes of protein unfolding and translocation present numerous challenges to molecular machines (Figure 1.2) and investigating how these challenges are overcome touches upon exciting questions in structural and molecular biology. For example, how do molecular machines harness a source of free energy to drive substrate unfolding and translocation? How are molecular machines able to handle such a wide variety of protein substrates? In addition, how do they process unfolded substrates, which not only present a wide array of side chain chemistries but also occupy a large configurational space? How are counterproductive diffusive forces mitigated and/or harnessed by the transporter during translocation? Insight into these mechanisms is not only relevant to protein translocation and protein degradation, but because the ATPases that regulate protein degradation are members of the AAA+ family, common mechanisms of action may also apply to molecular machines involved in a variety of cellular roles.

1.3 Anthrax toxin as a model system

Anthrax toxin, a key virulence factor secreted by *Bacillus anthracis*, is an ideal model system to study the processes of protein unfolding and translocation. The toxin is composed of three proteins: an 83-kDa translocase-forming protein, called protective antigen (PA), and two ~90 kDa enzymatic factors, called lethal factor (LF) and edema factor (EF). In order for the toxin

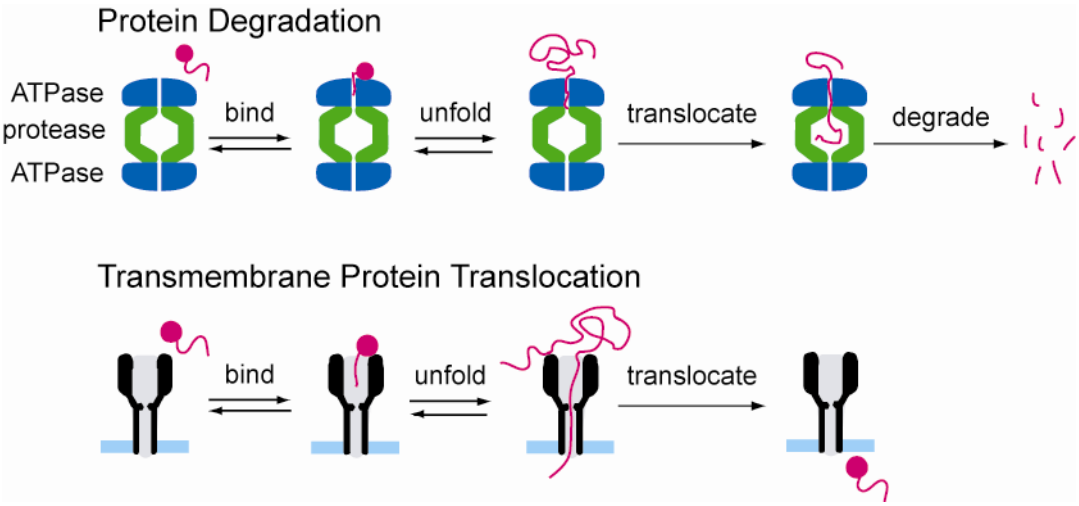

Figure 1.1. General steps of (non-lysosomal) protein degradation and transmembrane protein translocation. Protein degradation machines, which are generally comprised of a hollow, cylinder-shaped proteolytic chamber (green) capped by ATPase rings (blue), must first bind their substrates (pink). The ATPase domains then unfold and translocate the substrate through its narrow pore into the proteolytic chamber for degradation. These same steps are involved in transmembrane protein translocation.

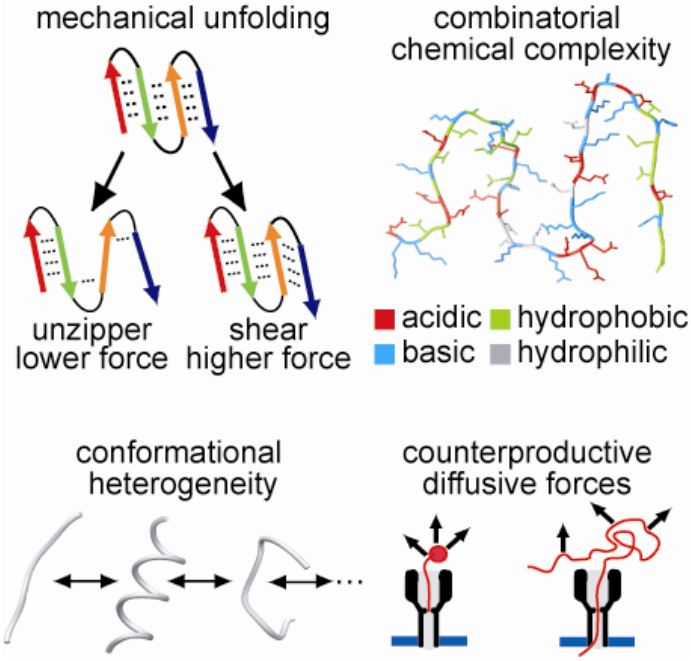

Figure 1.2. Challenges encountered during translocation. During translocation substrates are mechanically unfolded by the driving force; the mechanical resistance, however, can vary dramatically depending on the relative orientation of the substrate to the force vector. Combinatorial chemical complexity arises as the unfolded chain presents a wide array of side chain chemistries to the channel. Conformational heterogeneity is also present in the unfolded substrate polypeptides. Combinatorial chemical complexity and conformational heterogeneity present significant challenges for substrate recognition. Finally, during translocation counterproductive diffusive forces must also be overcome.

to function, it must assemble into an active holotoxin complex, which contains multiple copies of LF and EF bound to a ring-shaped PA oligomer (Figure 1.3). Proteases present either on host-cell surfaces or in blood serum (Mabry *et al.*, 2006) potentiate toxin assembly by proteolytically nicking PA (Ezzell & Abshire, 1992, Milne *et al.*, 1994, Kintzer et al., 2009). Dissociation of a 20-kDa amino-terminal fragment from nicked PA exposes substrate binding sites and permits assembly of the remaining 63-kD portion of PA into ring-shaped homoheptamers, PA_7, (Milne et al., 1994, Petosa *et al.*, 1997, Katayama *et al.*, 2008, Kintzer et al., 2009), or homooctamers, PA_8 (Kintzer et al., 2009). Once assembled, the complex is taken up by the host cell through receptor-mediated endocytosis, and transferred to an acidic compartment. Under these acidic conditions, the PA oligomer undergoes a conformational change that allows it to insert into the membrane, forming a cation-selective channel (Blaustein *et al.*, 1989, Miller *et al.*, 1999). Using the membrane potential ($\Delta\Psi$) (Zhang *et al.*, 2004) and proton gradient (ΔpH) that develops across the endosomal membrane, the PA channel is able to unfold and translocate LF and EF into the cytosol of the host cell (Krantz *et al.*, 2005, Krantz *et al.*, 2006, Thoren *et al.*, 2009). Once inside the cytosol, LF [a zinc-metaloprotease (Duesbery *et al.*, 1998)] and EF [a calcium- and calmodulin-activated adenylcyclase (Leppla, 1982, Leppla, 1984)] disrupt a variety of cell-signalling pathways, manifesting ultimately in general immune system dysfunction and potentially death.

Anthrax toxin has become a useful structure/function model system to characterize transmembrane protein translocation because the three proteins can be expressed recombinantly and purified independently. Also, translocation can be monitored directly using planar lipid bilayer electrophysiology. In this assay, PA channels are inserted into model membrane bilayers and protein translocation is monitored by measuring the restoration of ion conductance once a substrate completely traverses the channel (Krantz et al., 2006, Krantz et al., 2005, Zhang et al., 2004) (Figure 1.4). A great advantage of this assay is that the applied driving force can be externally controlled and continuously adjusted (Zhang et al., 2004, Krantz et al., 2006). This feature is critical when trying to determine the force-dependencies of energy barriers and ultimately dissect the mechanism of translocation (Feld *et al.*, 2010, Krantz et al., 2006, Krantz et al., 2005, Thoren et al., 2009).

1.4 Current state of the translocation-coupled protein unfolding field

Prior studies have investigated the processes of unfolding and translocation in a variety of different systems. General themes to describe different aspects of the unfolding and translocation reactions have emerged from the research, and will be summarized below. I will also discuss how this thesis fits into the context of the unfolding/transport field.

How is free energy harnessed to drive unfolding and translocation?

Transmembrane protein translocases and soluble molecular machines, like those involved in protein degradation and disaggregation share several common features. Notably, these systems are comprised of narrow, protein channels that require the substrate to be first unfolded before it is transported through the channel across a membrane or into a proteolytic complex. These unfolding and translocation steps require some sort of chemomechanical coupling of an energy source to the physical unwinding of the substrate polypeptide, namely through ATP hydrolysis or dissipation of a chemical gradient, such as a proton gradient.

Three general models that describe how a machine harnesses a driving force to physically unwind and transport a substrate protein have been proposed: (i) an active pushing/pulling

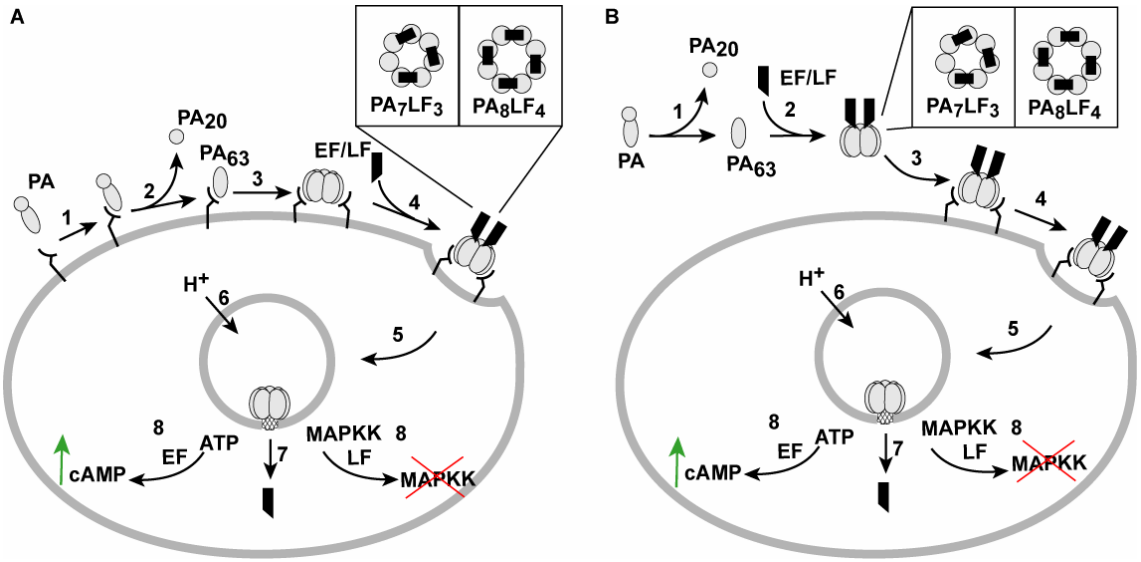

Figure 1.3. Anthrax toxin assembly and translocation. Two possible anthrax toxin assembly pathways are depicted. See the text for more details. **(A)** A receptor-dependent, cell-surface assembly pathway. **1**, PA binds an anthrax toxin receptor; **2**, proteolytic activation of PA by a furin-type protease cleaves off a 20-kD portion; **3**, the 63-kDa portion of PA can assemble into a homoheptamers or homooctamers; **4**, EF and LF bind to the prechannel oligomer; **5**, endocytosis; **6**, acification of the endosome causes a conformational change in the PA oligomer that allows it to insert into the membrane; **7**, a membrane potential and/or a proton gradient can drive the unfolding and translocation of LF and EF into the cytosol; **8**, LF (a Zn-protease that cleaves mitogen-activated protein kinase kinases (MAPKK)) and EF (a Ca^{2+}/calmodulin-activated adenylyl cyclase) disrupt normal cellular physiology by modifying cytosolic substrates. **(B)** A receptor-free assembly pathway. **1**, PA is proteolytically-activated by a serum protease in the blood; **2**, LF- and EF-binding drive the assembly of toxic complexes containing heptameric and octameric (Kintzer *et al.*, 2009) PA; **3**, these holotoxin complexes bind cells; **4-8**, the subsequent steps of the pathway occur as in (A).

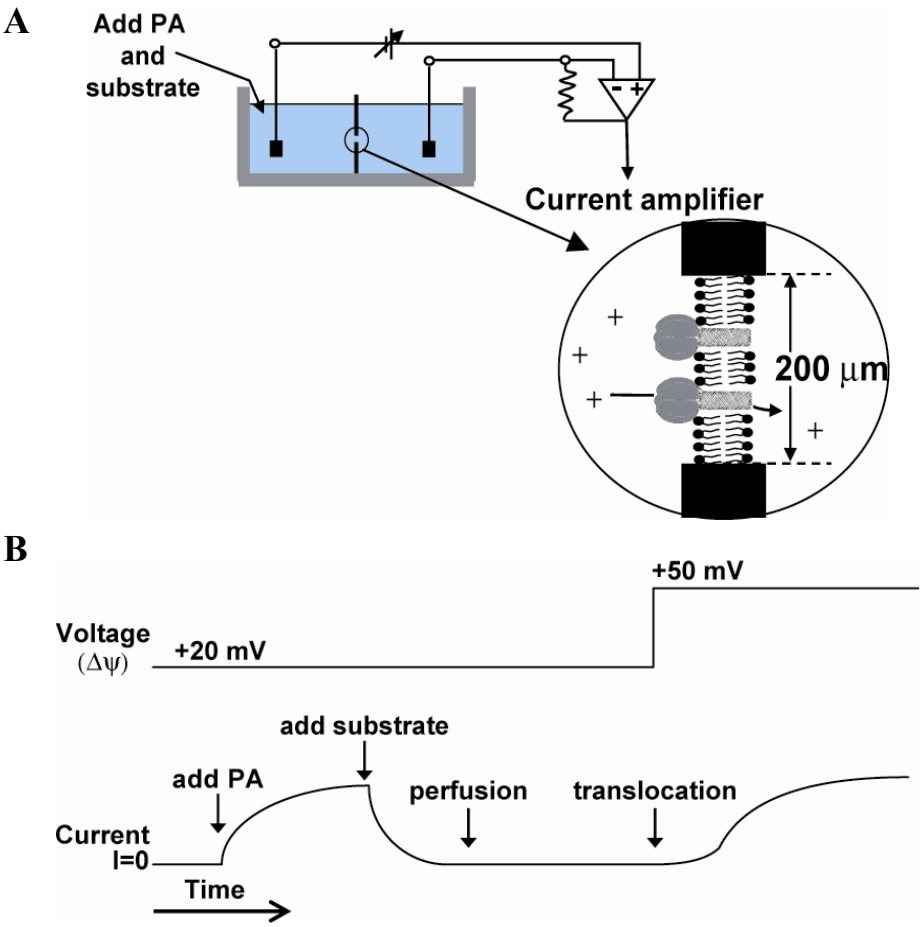

Figure 1.4. Translocation is studied using planar lipid bilayer electrophysiology. **(A)** Schematic of our instrument setup. **(B)** Typical recording of a translocation experiment. A bilayer is formed by applying a lipid solution over a small hole in a plastic partition. PA, the channel-forming protein is added to the front chamber of the apparatus and it inserts into the membrane. A substrate protein is added, blocks the channels and therefore blocks current. Translocation through PA is initiated by increasing the applied voltage or by creating a pH gradient. Changes in ionic current allow us to monitor formation of the bilayer, PA insertion and substrate translocation.

mechanism, (ii) a passive Brownian-ratchet mechanism and (iii) an entropic pulling mechanism. In the active pushing/pulling mechanism (Glick, 1995), structures within the molecular machine (usually central pore loops) or chaperones associated with the channel can contain critical substrate binding sites, which engage the substrate protein (Wang *et al.*, 2001, Hinnerwisch *et al.*, 2005, Lum *et al.*, 2008, Martin *et al.*, 2008, Glynn *et al.*, 2009). Upon cycles of ATP hydrolysis or proton binding and release, these components move like actuators to push/pull the substrate polypeptide through the central pore of the complex (Glick, 1995, Wang et al., 2001, Hinnerwisch et al., 2005, Lum et al., 2008, Martin et al., 2008, Zimmer *et al.*, 2008, Glynn et al., 2009). Thus unfolding forces in this mechanism are generated rather directly by movement of these substrate-binding components, and translocation directionality can be enforced by allowing substrate interactions to occur in only the power-stroke direction and not during the resetting of these loops.

In the Brownian-ratchet mechanism, the thermal diffusive motion of the translocating polypeptide chain is biased in a directional manner by means of an external energy gradient. In a theoretical proposal, Oster and colleagues suggested that a chemical gradient of heat shock proteins could prevent backward/retrograde diffusion by binding the substrate chain as it emerged from a translocase channel, thereby acting as a steric clamp (Simon *et al.*, 1992). Thus the key distinguishing feature between this mechanism and the active push/pull mechanism is how forces are developed on the substrate. For the Brownian-ratchet mechanism, Brownian-motion itself becomes rectified, and for the active mechanism, the ATP-dependent power stroke is directly coupled to ATP hydrolysis. The Brownian-ratchet mechanism has, however, drawn criticism principally on the issue of substrate unfolding, because it is not expected to generate enough force to denature a folded protein (Glick, 1995). Thus it has been postulated that the Brownian-ratchet mechanism would only be able to act upon pre-unfolded substrates, and that translocation would be limited by the spontaneous unfolding of the substrate protein.

The power stroke and Brownian ratchet model are not mutually exclusive. A third model, called entropic pulling, has been proposed and is a modified version of the Brownian ratchet model (De Los Rios *et al.*, 2006, Goloubinoff & Rios, 2007). This model proposes that an unfolding force can be generated based on entropy loss due to excluded volume effects. As a substrate protein emerges from a channel, a chaperone/motor protein binds to the polypeptide chain and thus reduces the number of conformations accessible to the chain. More conformations become available to the chaperone-bound chain as it moves further away from the channel's exit. Since systems tend to move toward a higher entropy state, a pulling force can be generated that is proportional to the difference in free energy. Thus, this model incorporates features of both the power stroke and Brownian ratchet model. Like the power stroke model, entropic pulling exerts a force on the substrate that accelerates unfolding and translocation, but like the Brownian ratchet model, force is not generated by the lever-arm movement of a chaperone/motor against a fulcrum.

Driving forces in anthrax toxin translocation

How are driving forces harnessed to unfold and translocate anthrax toxin? Part of the answer to this question, in the case of anthrax toxin, may be inferred by its endocytic trafficking mechanism. Anthrax toxin enters host cells through endocytosis into vesicles which acidify as they mature (Friedlander, 1986). Acidification of the endosome is believed to drive substrate unfolding and translocation in two ways. First, solution studies indicate that LF and EF are destabilized under the mildly acidic pH conditions likely encountered in the endosome (pH 5 to

6) (Krantz *et al.*, 2004). The amino-terminal domains of LF and EF (LF_N and EF_N) were examined in these experiments and they were found to populate a fairly compact molten-globule intermediate. This intermediate state contains a large amount of secondary structure, but its tertiary packing interactions are largely disrupted (Krantz et al., 2004). Obviously, this finding has implications for translocation because unfolding is a required step in the reaction.

In addition to destabilizing the substrate proteins, acidification of the endosome also results in the formation of a ΔpH across the endosomal membrane. This gradient, which is estimated to be ~ 2 units in late endosomes ($pH_{endosome}$ ~ 5.5, $pH_{cytosol}$ = 7.3), was shown to promote translocation through the PA channel (Krantz et al., 2006). In planar lipid bilayer experiments, full-length LF translocates poorly under a pure ΔΨ, and a one-to-two unit ΔpH resembling that expected naturally across the endosomal membrane is required to observe significant translocation (Krantz et al., 2006). More recently, it has been shown that a ΔpH is sufficient to drive translocation in electrophysiology assays (Brown *et al.*, 2011). While a positive membrane potential ($\Psi_{endosome} > \Psi_{cytosol}$) is believed to exist across the endosomal membrane, it is estimated to be small (only about 10-30 mV) due to chloride ion import during acidification (Van Dyke & Belcher, 1994, Rybak *et al.*, 1997, Sonawane *et al.*, 2002). Thus, a ΔpH is likely to be the principle driving force of anthrax toxin translocation in cells, and there has been much interest in how the channel harnesses a ΔpH to drive unfolding and translocation (Brown et al., 2011, Basilio *et al.*, 2009, Pentelute *et al.*).

A Brownian-ratchet model (referred to as the charge-state ratchet) has been invoked to explain how a ΔpH is harnessed by the channel during translocation (Krantz et al., 2006). In this mechanism, acidic residues in the translocating chain protonate upon entering the PA channel, because the channel is cation-selective (and therefore anion repulsive). Once these groups emerge from the PA channel into the higher pH on the cytosolic side of the membrane, they can deprotonate, thereby allowing an electrostatic repulsion to develop between the channel and the exiting polypeptide chain. This electrostatic repulsion may effectively capture Brownian-motion, further driving translocation and enforcing directionality. Some more recent tests of this Brownian-ratchet hypothesis have lent further support to the model. For example, it has been shown that negative charges from sulfonic acid groups attached artificially to the substrate protein inhibit translocation (Basilio et al., 2009). Since these groups cannot be protonated and their charge cannot be neutralized, it was proposed that the channel's cation-selective filter was rejecting these strong anions. Finally, this ΔpH mechanism appears to generate enough force to unfold substrate proteins, which will be discussed in more detail in Chapter 2.

Interactions with the channel

The mechanism by which molecular machines harness a source of free energy to drive unfolding and translocation has been the focus of debate in the field, especially for mitochondrial import (Matouschek *et al.*, 2000, Pfanner & Geissler, 2001, Neupert & Brunner, 2002, Liu *et al.*, 2003). Although important to understanding translocation, many of these prior studies are overlooking a critical issue. Namely, how does a substrate protein get to the motor? In many of these systems, the motor/chaperone component is located on the opposite side of the membrane as the folded substrate protein and thus, the substrate must first traverse the length of the channel before it contacts the motor. During import into the mitochondria, for example, the substrate must translocate across the outer and inner membrane in order to reach the ATP-dependent Hsp70 motor protein located in the mitochondrial matrix. Researchers found that the presequence (an amino-terminal extension that targets the substrate to mitochondria) must be at

least 60 amino acids long to reach the motor (Ungermann *et al.*, 1994). However, the length of presequence varies depending on the substrate, and most presequences are only about 20-35 amino acids long (Schatz & Dobberstein, 1996). Thus, some unfolding and translocation of the substrate probably takes place in the absence of the motor/chaperone, at least until the substrate emerges from the channel. How does the translocation machinery accomplish these processes?

I hypothesize, and present evidence in this thesis, that translocase channels are not passive conduits for their substrates. Instead, multiple binding sites, or clamps, within the channel are able to interact with the translocating chain to facilitate unfolding and translocation. By recognizing general features of an unfolded polypeptide, a translocase channel is able to achieve such broad substrate specificity and handle the wide array of chemistries presented by an unfolded protein. In addition, multiple clamping sites could explain how the channel overcomes counterproductive diffusive forces.

On one hand, the idea that numerous nonspecific binding sites can form nonspecific interactions with the substrate appears to be quite problematic, as extensive binding would create thermodynamic traps and impair translocation. However, the penalty of having numerous clamping sites may be offset by some of the following benefits: orienting the substrate toward the central lumen, stabilizing partially unfolded intermediates, and minimizing the diffusional mobility and backward motion during translocation.

A simple, back-of-the-envelope calculation illustrates the advantages of this model in terms of reducing the configurational entropy of an unfolded substrate protein (Figure 1.5). Consider a 100-residue, unfolded protein. Let's assume that each residue has 4 conformations in the unfolded state and all of the conformations are equal in energy. When confined in the lumen of a translocase channel, each residue only has one conformation. Based on the change in entropy, it would take 69 kJ mol^{-1} of energy to move 20 residues of this protein into the channel (at 300 K) (Figure 1.5A).

Now imagine that the unfolded protein can bind to the channel (Figure 1.5B). Binding to the channel would reduce the number of conformations accessible to the unfolded, un-translocated protein. In addition, multiple binding sites could allow translocation to proceed in small, incremental steps. Thus, for our back-of-the-envelope calculation, let's assume that the residues constrained between multiple binding sites only have 2 possible conformations (instead of 4), and that translocation occurs in 2 steps. Here, each step requires about 15 kJ mol^{-1} of energy at 300 K (Figure 1.5B). Compared to the 69 kJ mol-1 of energy required above, this amount of energy is more manageable. Although we have only considered the entropy of the substrate here, interactions with the channel would also provide enthalpic stabilization and further reduce the amount of energy required for translocation.

Clamp structures in anthrax toxin

In the PA channel, several discrete clamping sites are observed that address these challenging aspects of translocation. In an effort to understand how a molten-globular, partially-folded substrate would be linearized and fully unfolded during translocation, Krantz *et al.* sought to identify key active site residues lining the PA channel (Krantz et al., 2005). While most of the residues lining the interior of the channel's cap domain are mainly small hydrophilic residues [recapitulating an earlier set of results in the β-barrel stem (Nassi *et al.*, 2002)], one prominant hydrophobic residue was identified within the channel, Phe-427 (Krantz et al., 2005). Electron paramagnetic resonance spectroscopy studies revealed that when the channel forms, Phe residues from each subunit converge, forming a radially-symmetrical aromatic clamp site, called the ϕ

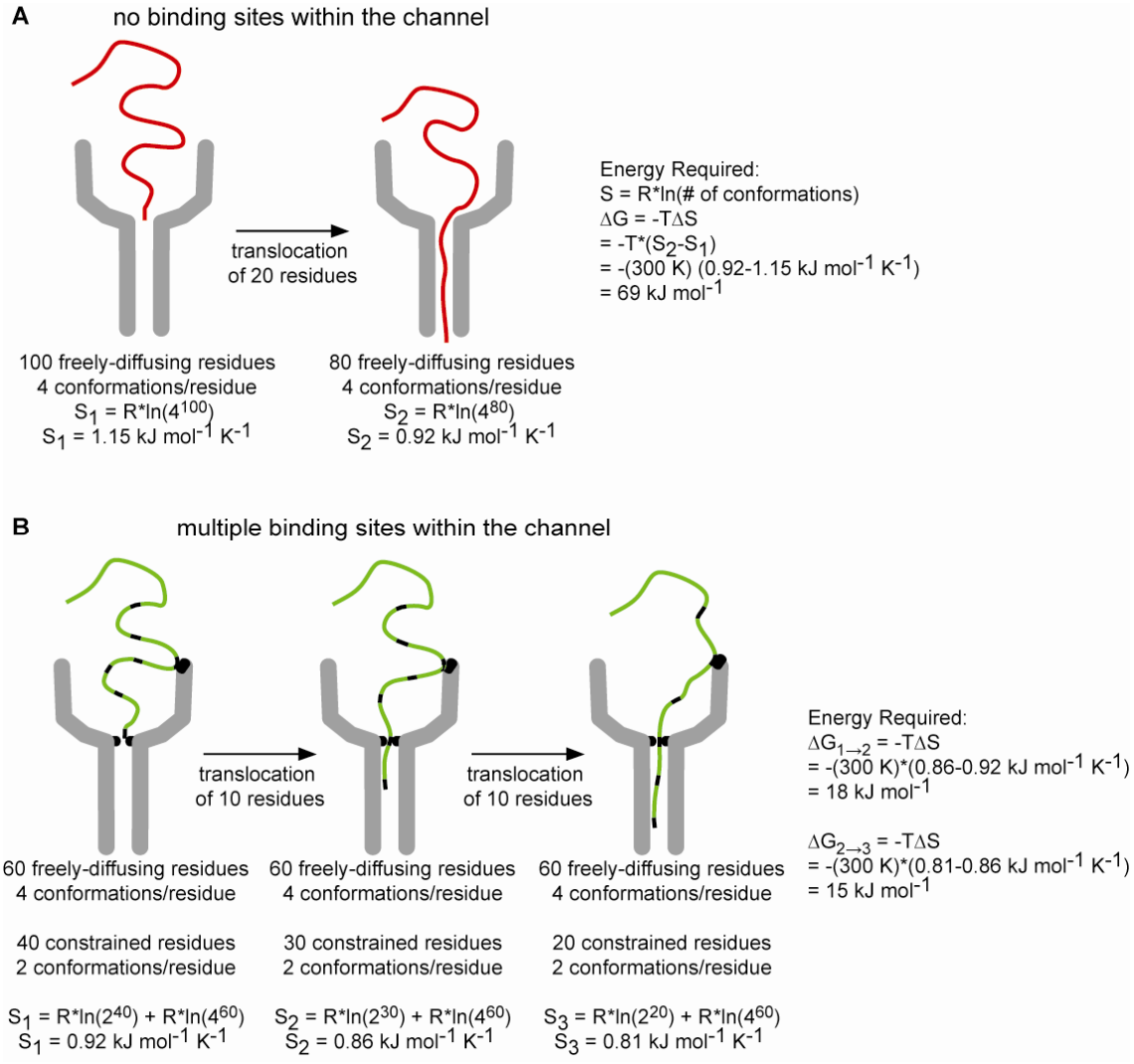

Figure 1.5. Multiple binding sites within the channel reduce the entropy of the substrate. A crude approximation of the free energy required to translocate 20 residues of a substrate into the channel based on the change in entropy. **(A)** In this situation, the substrate protein does not bind to the channel. Let's assume that each residue has 4 accessible conformations when in the unfolded, freely diffusing state, but each only has one conformation when confined to the lumen of the channel. Based on the change in entropy, 69 kJ mol^{-1} of free energy is required to translocate 20 residues of this 100-residue, unfolded protein into the channel at 300 K. **(B)** Here, the substrate can bind to multiple sites in the channel, thus reducing the entropy of the un-translocated substrate and allowing translocation to proceed in small, incremental steps. There are only 2 accessible conformations for residues that are constrained between multiple binding sites in the channel. When the translocation of 20 residues is broken down into two steps, each step requires about 15 kJ mol^{-1} of free energy at 300 K.

clamp (Krantz et al., 2005). Electrophysiology studies then showed that the ϕ-clamp structure is required for protein translocation. Furthermore, the ϕ-clamp site forms an ion-conductance bottleneck in the channel, suggesting these Phe-427 residues make a narrow approach and could form a polypeptide binding site in the channel. This model was confirmed when it was observed that mutations to the ϕ-clamp site allow the substrate to backslide or retrotranslocate in an unproductive manner, inhibiting efficient translocation. Model-compound-binding studies revealed that the ϕ-clamp site possesses broad substrate specificity, where the multifaceted aromatic surfaces of the Phe residues preferred cationic, aromatic and hydrophobic substrates, consistent with the π-cloud electrons of the Phe residues making π-π, cation-π, and π-dipole interactions (Krantz et al., 2005). The ϕ-clamp site thus serves a chaperone-like function, where it may interact with a broad spectrum of sequences presented by the protein substrate as it translocates. It is therefore hypothesized that the ϕ clamp plays an important role in translocation-coupled unfolding by binding to and stabilizing unfolded intermediates (Krantz et al., 2005).

In addition to the ϕ clamp, a second substrate binding site in PA was recently identified. The details of this site, termed the α clamp, and its role in substrate binding and translocation are discussed in Chapters 3 and 4.

How are proteins unfolded on a molecular machine?

We've already discussed ways in which a source of free energy could be coupled to drive protein unfolding, and how the channel may play a role in the process. But what features of the substrate protein determine if and how it is unfolded during translocation? Prior studies have examined the correlation between protein stability and translocation kinetics (Krantz et al., 2006, Lee *et al.*, 2001, Kenniston *et al.*, 2003, Burton *et al.*, 2001). In some cases, the translocation rate depends on the thermodynamic stability of the protein (Eilers & Schatz, 1986, Krantz et al., 2006), but in other cases, it does not (Huang *et al.*, 1999, Burton et al., 2001). In order to reconcile these conflicting pieces of data, it has been proposed that resistance to unfolding is determined by the local structure of the protein next to the signal sequence, rather than the protein's global thermodynamic stability (Huang et al., 1999, Kenniston et al., 2003). For example, it was found that mitochondrial presequences that were next to an α helix were more easily imported than when the signal sequence was next to a buried β strand (Wilcox *et al.*, 2005). This finding may be explained by the fact that α helices are generally found on the surface of a protein, whereas β strands are usually buried in the hydrophobic core of the protein (Branden & Tooze, 1991). Once this local secondary structure element is unraveled, the rest of the protein is thought to denature rapidly (Huang et al., 1999) since protein unfolding is a highly cooperative process.

Because solution-unfolding studies (Englander *et al.*, 2002) have generally shown that the rate-limiting unfolding step involves an extensive region of structure, the local stability model is somewhat surprising. The observation that small, local regions of structure are rate limiting to unfolding could mean that the unfolding transition state is shifted toward the native state. However, it is also possible that the global unfolding transition was not rate-limiting under the conditions tested. Due to the inability to continuously tune the applied driving force in these prior translocation studies, it is unclear which steps are rate-limiting and how the substrate protein unfolds during translocation.

We revisit the local structure hypothesis in Chapter 2. Using site-directed mutagenesis, thermodynamic stability studies, and planar lipid bilayer electrophysiology experiments which

allow us to control and continuously adjust the driving force, we identify the barriers in the translocation mechanism and map how a substrate protein would actually unfold on the surface of the PA channel. In addition, the role of the ϕ clamp in LF_N translocation was investigated using double mutant cycle analysis.

In Chapters 3 and 4, I discuss the α clamp, a recently-identified substrate binding site in the PA channel. Chapter 3 describes the crystal structure of a core of a lethal toxin complex solved by my colleague, Geoff Feld (Feld et al., 2010). The most interesting feature of the structure reveals that the first α helix and β strand of each LF_N unfold and dock into the α clamp, a deep amphipathic cleft on the surface of the PA. Through extensive mutagenesis of both PA and the substrate LF_N, Geoff and I determined that this site can bind a broad array of polypeptide substrates. Finally, the role of the α clamp in substrate unfolding, channel oligomerization and translocation is discussed.

In Chapter 4, I further investigate the role of the α clamp in translocation. We hypothesize that the non-specific binding activity of the site implies a means for the PA channel to unfold multidomain substrates, where the process may be repeated on each folded domain during translocation. Here, I test this hypothesis and probe the α clamp's role in translocation by disrupting binding to the site.

Chapter 2

Lethal factor unfolding is the most force-dependent step of anthrax toxin translocation

2.1 Abstract

Cellular compartmentalization requires machinery capable of translocating polypeptides across membranes. In many cases, transported proteins must first be unfolded by means of the proton motive force and/or ATP hydrolysis. Anthrax toxin, which is composed of a channel-forming protein and two substrate proteins, is an attractive model system to study translocation-coupled unfolding since the applied driving force can be externally controlled and translocation can be monitored directly using electrophysiology. By controlling the driving force and introducing destabilizing point mutations in the substrate, we identified the barriers in the transport pathway, determined which barrier corresponds to protein unfolding, and mapped how the substrate protein unfolds during translocation. In contrast to previous studies, we find that the protein's structure next to the signal tag is not rate-limiting to unfolding. Instead, a more extensive part of the structure, the amino-terminal β-sheet subdomain, must disassemble to cross the unfolding barrier. We also find that unfolding is catalyzed by the channel's phenylalanine-clamp active site. We propose a broad molecular mechanism for translocation-coupled unfolding, which is applicable to both soluble and membrane-embedded unfolding machines.

2.2 Introduction

Folded proteins are ~5-10 kcal mol^{-1} more stable than their unfolded states. Therefore, the disassembly and translocation of folded proteins often require a molecular machine and a source of free energy. These ubiquitous multi-protein complexes include soluble degradation machinery, such as the proteasome or the Clp bacterial proteases (Sauer *et al.*, 2004), which unfold and degrade proteins, and some, but not all, membrane-embedded translocase channels, which can unfold and transport proteins across membranes (Matouschek, 2003). There are general features shared between these soluble and membrane-embedded translocase machines: a narrow central pore first engages the protein substrate on its free end; the substrate is unfolded mechanically; and the unfolded chain is translocated through the narrow pore, allowing it ultimately to either cross a membrane or enter into a proteolytic complex for degradation. Protein unfolding and translocation in these systems are often driven by ATP-hydrolysis (Sauer *et al.*, 2004, Matouschek, 2003), a membrane potential (ΔΨ) (Zhang *et al.*, 2004b, Matouschek, 2003), and/or a proton gradient (ΔpH) (Krantz *et al.*, 2006). The molecular mechanism of translocation-coupled unfolding, however, is poorly understood.

Prior studies examining the correlation between substrate protein stability and translocation kinetics have produced conflicting results. Some ligand-stabilized substrates translocate inefficiently because they are too thermodynamically stable (Eilers & Schatz, 1986); however, other substrates show little change in the rate of translocation when destabilized by mutagenesis (Huang *et al.*, 1999, Burton *et al.*, 2001). To resolve these conflicting results, it was proposed that translocation-coupled unfolding (Huang *et al.*, 1999, Kenniston *et al.*, 2003) depends on the mechanical stability of the local structure adjacent to the signal tag. Once this local secondary structure element is unraveled, the rest of the protein is thought to denature rapidly (Huang *et al.*, 1999).

This local-stability model is surprising when considering solution-unfolding studies (Englander *et al.*, 2002), which generally show that the rate-limiting unfolding step involves an extensive region of structure. While local structure may unfold as a cooperative unit, as shown by native state hydrogen exchange (Bai *et al.*, 1995), these local-unfolding events are not rate-limiting to the larger-scale global unfolding (Figure 2.1). Local-unfolding intermediates are often found on the native side of the major rate-limiting unfolding barrier, and they can be similar in free energy to the native state. The observation that a small, local region of structure is rate-limiting in these translocation studies suggests one of two possible explanations. Either the unfolding transition state is shifted toward the native state, or the global unfolding transition was not rate-limiting under the conditions tested. Furthermore, the magnitude of the applied driving force may alter the translocation pathway, depending upon the force-dependencies of the underlying barriers in the mechanism. Thus due to the inability to continuously tune the applied driving force in prior translocation studies, it is unclear which steps are rate-limiting and how the substrate protein unfolds during translocation.

Anthrax toxin (Young & Collier, 2007) is well suited to study protein unfolding during translocation because the applied driving force can be externally controlled and continuously adjusted using electrophysiology. The toxin is comprised of three protein components: lethal factor (LF), edema factor (EF) and protective antigen (PA). The PA component first forms a ring-shaped oligomer that binds LF and EF. These toxin complexes are endocytosed and delivered to an acidic compartment in the cell. The PA oligomer then converts into a membrane-spanning translocase channel, which is so narrow that LF and EF must unfold to translocate (Krantz *et al.*, 2004). The acidic conditions encountered following endocytosis not only destabilize LF and EF (Krantz *et al.*, 2004), but they also generate a transmembrane ΔpH that drives translocation (Krantz *et al.*, 2006). In addition, a narrow ring of symmetric phenylalanine residues contained within the PA channel (called the ϕ clamp) provides an aromatic interaction surface that is critical for catalyzing translocation (Krantz *et al.*, 2005).

Here we develop a general framework to define the energy landscape of a translocation mechanism. Using anthrax toxin, we establish the major barriers of translocation, define their force-dependencies, and determine which barriers involve protein unfolding. Through site-directed mutagenesis, we then probe how the substrate protein unfolds during translocation and determine whether the channel's ϕ-clamp site drives translocation by catalyzing substrate unfolding.

2.3 Results

Voltage-dependence of translocation. We use planar lipid bilayer electrophysiology to study translocation. PA channels are first inserted into the bilayer at a low membrane potential (ΔΨ). LF's amino-terminal, 263-residue, PA-binding domain (LF_N) is added to the bilayer, where it binds to PA and blocks conductance by inserting into the narrow channel. Excess LF_N is perfused, and translocation is initiated by raising the ΔΨ. As LF_N translocates, the channels become unblocked, and the restoration of conductance reports on the translocation kinetics in real time. The observed translocation kinetics are multi-exponential and complex (Zhang *et al.*, 2004b, Krantz *et al.*, 2006, Krantz *et al.*, 2005) (Figure 2.2). We estimate the rate constant, k, for the translocation reaction using the half time ($t_{1/2}$), which is the time (measured in seconds) for half of the translocated protein to move through the channel. Since $t_{1/2} \propto 1/k$, the approximation is adequate for our analysis. We calculate an empirical activation energy ($\Delta G‡$) at a particular ΔΨ with $\Delta G‡ = RT \ln t_{1/2} / c$. Here, R and T are the gas constant and temperature, and c is an

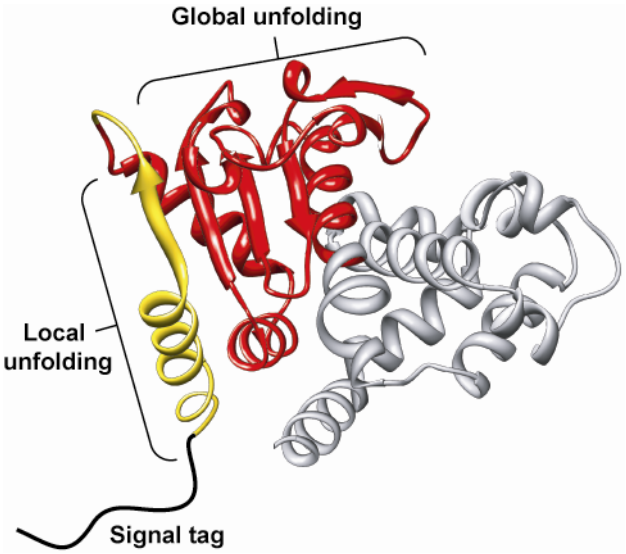

Figure 2.1. Models of translocation-coupled unfolding. A depiction of the amount of structure that unfolds when crossing the unfolding barrier during protein translocation. On one extreme, local structure (yellow) near the presequence or degradation tag (black line) may be rate-limiting to unfolding. On the other extreme, more extensive global unfolding (red), which is typically observed in solution, is rate-limiting.

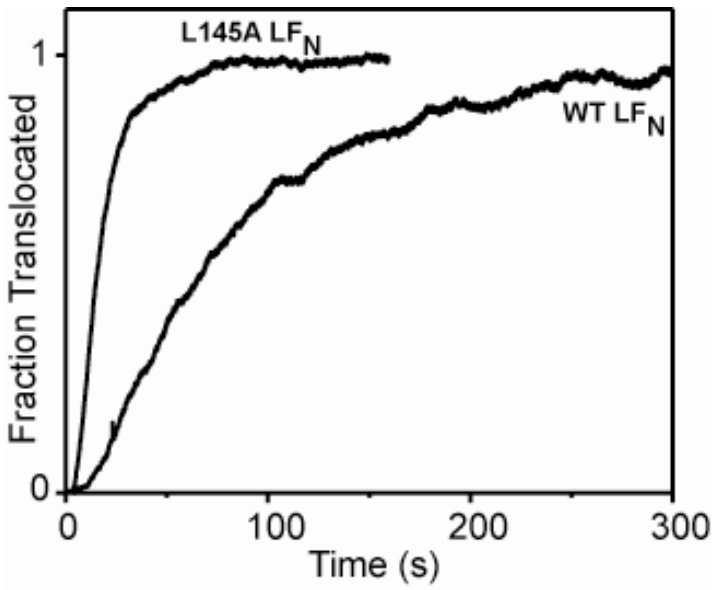

Figure 2.2. Translocation is multi-exponential and complex. Ensemble translocation records for WT LF_N and a destabilized mutant (L145A LF_N) at a $\Delta\Psi$ of 40 mV normalized to the fraction of channels that translocated, which is generally ~90%.

arbitrary reference, which we define as 1 s. We find that the translocation $\Delta G‡$s have a biphasic $\Delta\Psi$ dependence with two different linear extremes at high and low voltages (Figure 2.3A).

We considered this biphasic $\Delta\Psi$ dependence in terms of the Eyring-Woodhull model (Woodhull, 1973). In this model, the transport of a charged species across a membrane is both barrier-limited and modulated by the applied $\Delta\Psi$. For a simple, single-barrier case, the dependence of an activation energy and $\Delta\Psi$ should be linear; i.e., $\Delta G‡(\Delta\Psi) = \Delta G‡_o + zF\Delta\Psi$. Here, z represents the number of charges acted upon in the rate-limiting step, $\Delta G‡_o$ is the activation energy determined at 0 mV, and F is Faraday's constant. Our biphasic dependence, however, shows that there are two different barriers in the translocation mechanism: a $\Delta\Psi$-sensitive barrier is limiting at low $\Delta\Psi$s; and a relatively $\Delta\Psi$-insensitive barrier is limiting at high $\Delta\Psi$s. We fit the observed $\Delta\Psi$ dependencies to a two-barrier model (see Materials and Methods), and the linear extremes at low and high $\Delta\Psi$ correspond to z-value slopes of 4.2(±0.2) and 0.22(±0.05), respectively. Therefore, due to the ~20-fold difference in z-values, we may specifically probe the kinetics at either extreme in voltage to determine which structures in the substrate and channel may limit each barrier.

ΔpH-driven translocation. The ΔpH is also a potent driving force for translocation. We wanted to test whether the more ΔpH-dependent phase was analogous to the $\Delta\Psi$-sensitive phase encountered in $\Delta\Psi$-driven translocation. To produce a ΔpH, we changed the pH on the trans side of the membrane and maintained a constant pH on the cis side, which is the side that LF_N and PA are added. This procedure ensures that neither LF_N's stability nor its binding interaction with PA would be altered, as these quantities are pH dependent (Krantz *et al.*, 2004). Translocations were driven by a constant $\Delta\Psi$ (60 mV), and the ΔpH ($\Delta pH \equiv pH_{trans} - pH_{cis}$) was varied over a range of ±1 unit (Figure 2.3B).

We measured a biphasic ΔpH dependence for WT LF_N (Figure 2.3B). For a simple single-barrier case, the following chemical potential relationship, $\Delta G‡ = \Delta G‡_o - 2.3nRT\Delta pH$, applies. Here n is the number of protons required to cross the barrier. The biphasic data we obtained required a two-barrier form of the above relationship (see Materials and Methods). The two n-values that define the limiting slopes in the biphasic relationship are 2.09(±0.06) and 0.21(±0.03), and correspond to a ΔpH-sensitive and ΔpH-insensitive barrier, respectively. Since the two different types of driving forces ($\Delta\Psi$ and ΔpH) produce a similar biphasic relationship with $\Delta G‡$, we infer that these driving forces do not modulate the structure of the channel, but rather they act directly upon the substrate.

Probing how protein stability affects translocation. To determine which barrier(s) relate to unfolding, we introduced multiple point mutations into LF_N. The thermodynamic stability of each mutant was assessed by chemical denaturant titration probed by various spectroscopic methods (Figure 2.4). These data were fit to a four-state equilibrium unfolding model, $N \leftrightarrow I \leftrightarrow J \leftrightarrow U$ (Krantz *et al.*, 2004), to obtain ΔG values for each transition. The ΔGs between the N and I states (ΔG_{NI}) were then used to compute the difference between the mutant and WT protein, and these $\Delta\Delta G_{NI}$ values ranged from ~0.4 to ~3 kcal mol^{-1} (Table 2.1).

Ensemble translocation experiments were performed on the destabilized LF_N mutants. Initially, we examined the $\Delta\Psi$ dependence of the translocation activation energy for a destabilized LF_N mutant (L145A) and found that it translocated faster than WT at lower voltages; however, at higher voltages the mutant translocated like WT LF_N (Figure 2.3A). Importantly, the shape of the curve and its z-values were identical to WT, indicating that the position of the barriers did not change and only the barrier heights were affected by the destabilizing mutation. Likewise we measured the ΔpH dependence of the translocation kinetics of the L145A LF_N

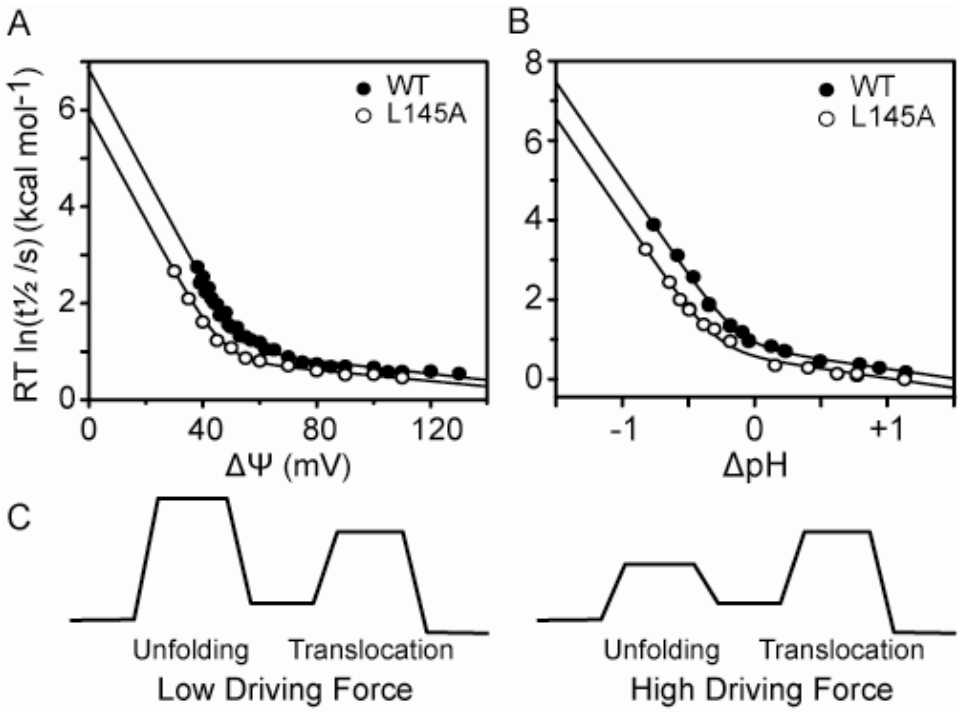

Figure 2.3. Substrate unfolding coincides with the more force-dependent step. (A) The ΔΨ dependence of the translocation activation energy ($\Delta G‡ = RT \ln t_{1/2} / c$) for WT (●) and L145A LF_N (○), where the reference, c, is 1 s. The two arms of the curves are fit to two-barrier Eyring-Woodhull relations (Eq. 1, Materials and Methods) with limiting z-value slopes of 4.2(±0.2) and 0.22(±0.05), respectively. (B) The ΔpH dependence of the translocation $\Delta G^‡$ for WT (●) and L145A LF_N (○). The two arms of the curves are fit to chemical-potential-modulated, two-barrier activation energy relations (Eq. 2, Materials and Methods) with limiting slopes, n, of 2.09(±0.06) and 0.21(±0.03) protons, respectively. (C) A two-barrier translocation mechanism is shown. The first barrier is ΔΨ- or ΔpH-dependent and limited by unfolding. The second barrier is largely driving-force independent and may be limited by translocation of the unfolded substrate.

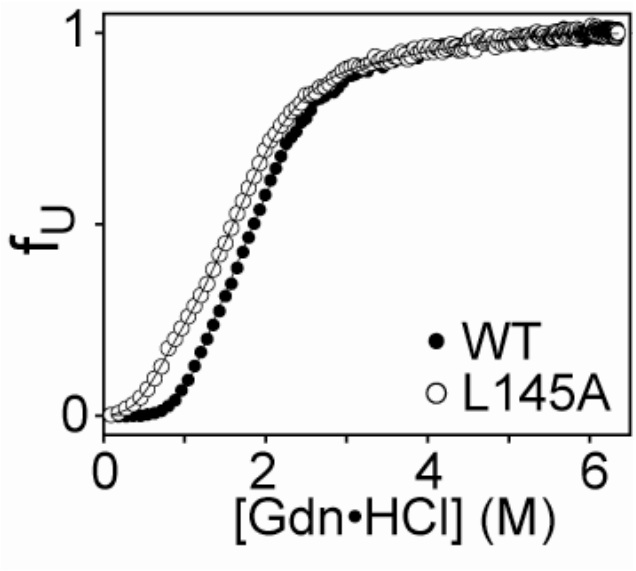

Figure 2.4. Equilibrium denaturant titrations for WT and L145A LF_N probed by CD at 222 nm. These profiles are fit to a four-state equilibrium unfolding model, $N \leftrightarrow I \leftrightarrow J \leftrightarrow U$ (Krantz *et al.*, 2004), and all thermodynamic parameters are given in Table 2.1.

Table 2.1. Thermodynamic unfolding[1] and kinetic translocation[2] parameters for LF_N.

Mutation	Secondary structure[3]	ΔG_{NI} (kcal mol^{-1})	ΔG_{IJ} (kcal mol^{-1})	ΔG_{JU} (kcal mol^{-1})	$\Delta\Delta G_{NI}$[6] (kcal mol^{-1})	$\Delta\Delta G‡$[7] (kcal mol^{-1})	ϕ [8]
WT	N/A	-5.39 ± 0.01 (-3.99 ± 0.04) [4] [-5.17 ± 0.03] [5]	-2.6 ± 0.1 (-2.58 ± 0.01) [-2.81 ± 0.05]	-4.61 ± 0.04 (-4.55 ± 0.08) [-4.60 ± 0.06]	0	0	N/A
I39A	α1	-4.54 ± 0.07	-2.0 ± 0.1	-4.61 ± 0.05	0.84 ± 0.07	0.2 ± 0.2	0.2 ± 0.1
M40A	α1	-4.77 ± 0.05	-2.1 ± 0.1	-4.71 ± 0.05	0.61 ± 0.05	-1.1 ± 0.3	-1.7 ± 0.5
L62A	α2	-4.0 ± 0.2 (-2.10 ± 0.07)	-2.07 ± 0.07 (-2.33 ± 0.06)	-4.67 ± 0.09 (-3.8 ± 0.2)	1.6 ± 0.2	0.6 ± 0.1	0.4 ± 0.1
V70A	α3	-4.18 ± 0.04 (-3.5 ± 0.1)	-2.1 ± 0.2 (-2.06 ± 0.06)	-4.5 ± 0.1 (-3.2 ± 0.3)	0.8 ± 0.1	0.59 ± 0.09	0.7 ± 0.2
H91N	Loop1	-4.53 ± 0.08	-2.0 ± 0.1	-4.61 ± 0.06	0.86 ± 0.08	0.3 ± 0.1	0.4 ± 0.2
A112G	Loop1	-4.89 ± 0.04 (-3.1 ± 0.1) [-4.44 ± 0.09]	-2.1 ± 0.1 (-2.44 ± 0.03) [-2.32 ± 0.09]	-4.75 ± 0.05 (-4.1 ± 0.2) [-4.27 ± 0.04]	0.7 ± 0.2	0.4 ± 0.1	0.6 ± 0.2
V119A	ß3	-5.07 ± 0.01	-2.13 ± 0.09	-4.28 ± 0.07	0.32 ± 0.01	0.37 ± 0.06	1.2 ± 0.2
E126G	turn, between ß3 and ß4	-4.8 ± 0.1 (-2.43 ± 0.07) [-3.5 ± 0.1]	-2.06 ± 0.09 (-2.41 ± 0.04) [-2.12 ± 0.01]	-4.72 ± 0.07 (-3.1 ± 0.9) [-4.12 ± 0.02]	1.3 ± 0.2	0.40 ± 0.06	0.31 ± 0.06
L129A	ß4	-3.2 ± 0.3	-2.46 ± 2.82	-4.6 ± 0.4	2.22 ± 0.07	0.9 ± 0.2	0.39 ± 0.08
L145A	α4	-4.01 ± 0.03 (-2.1 ± 0.1) [-3.52 ± 0.05]	-1.99 ± 0.06 (-2.22 ± 0.02) [-2.07 ± 0.02]	-4.53 ± 0.04 (-4.41 ± 0.06) [-3.95 ± 0.02]	1.6 ± 0.1	0.9 ± 0.1	0.5 ± 0.1
V147A	α4	-3.6 ± 0.2	-1.89 ± 0.07	-4.75 ± 0.06	1.8 ± 0.2	0.4 ± 0.1	0.21 ± 0.07
Y148A	α4	-3.4 ± 0.1	-2.17 ± 0.02	-4.9 ± 0.1	2.0 ± 0.1	0.9 ± 0.1	0.42 ± 0.07
I154A	α4	-4.75 ± 0.01	-2.50 ± 0.08	-4.70 ± 0.04	0.63 ± 0.02	0.2 ± 0.1	0.3 ± 0.2
L155A	α4	-4.19 ± 0.02 (-1.84 ± 0.09) [-3.4 ± 0.2]	-2.28 ± 0.06 (-2.07 ± 0.03) [-1.89 ± 0.07]	-4.53 ± 0.04 (-3.2 ± 0.5) [-3.51 ± 0.05]	1.7 ± 0.2	0.14 ± 0.09	0.08 ± 0.06
L174A[9]	α5	-3.84 ± 0.08	-0.9 ± 0.4	-5.4 ± 0.1	1.6 ± 0.1	-0.03 ± 0.08	-0.02 ± 0.05

Mutation	Secondary structure[3]	ΔG_{NI} (kcal mol^{-1})	ΔG_{IJ} (kcal mol^{-1})	ΔG_{JU} (kcal mol^{-1})	$\Delta\Delta G_{NI}$[6] (kcal mol^{-1})	$\Delta\Delta G‡$[7] (kcal mol^{-1})	ϕ [8]
V213A	α9	-4.72 ± 0.01 (-2.28 ± 0.09) [-3.77 ± 0.08]	-2.52 ± 0.06 (-2.20 ± 0.02) [-2.16 ± 0.03]	-4.65 ± 0.04 (-4.55 ± 0.06) [-3.99 ± 0.02]	1.3 ± 0.1	0.30 ± 0.08	0.24 ± 0.07
F217A	α9	-3.06 ± 0.02	-2.19 ± 0.03	-4.44 ± 0.03	2.32 ± 0.02	-0.1 ± 0.1	-0.03 ± 0.05
A220G	α9	-3.31 ± 0.02 (-1.72 ± 0.09) [-2.42 ± 0.1]	-2.20 ± 0.05 (-2.14 ± 0.03) [-1.89 ± 0.01]	-4.48 ± 0.04 (-5.4 ± 0.1) [-3.86 ± 0.02]	2.4 ± 0.2	0.3 ± 0.2	0.12 ± 0.07
F221A	α9	-2.22 ± 0.02	-2.13 ± 0.02	-4.38 ± 0.03	3.17 ± 0.03	0.9 ± 0.1	0.30 ± 0.03
D245G	α11	-4.95 ± 0.02 (-2.8 ± 0.2)	-2.29 ± 0.08 (-2.35 ± 0.04)	-4.63 ± 0.04 (-4.2 ± 0.2)	0.8 ± 0.2	-0.04 ± 0.06	-0.04 ± 0.08
F247A	α11	-3.93 ± 0.02	-2.10 ± 0.04	-4.46 ± 0.03	1.46 ± 0.03	0.36 ± 0.09	0.25 ± 0.06

[1]The four-state equilibrium free energy parameters, ΔG_{NI}, ΔG_{IJ}, ΔG_{JU}, are generally obtained from fitting equilibrium denaturation experiments probed using CD. The denaturant sensitivity for each thermodynamic transition, or *m* value, were generally fit using fixed values consistent with previously published data (Krantz *et al.*, 2004), where m_{NI}, m_{IJ}, m_{JU} are 3.59, 1.37, and 1.05, respectively (unless noted otherwise).The fit model has been described elsewhere (Krantz *et al.*, 2004).

[2]Kinetic activation free energy parameters are obtained from protein translocation experiments using an ensemble planar lipid bilayer assay using the relation, $\Delta G‡ = -RT \ln t_{1/2}$, where R, T, and $t_{1/2}$ are the gas constant, temperature, and time for half of the translocated protein to translocate.

[3]Secondary structure indicates where in structure the mutation is localized; the numbering scheme for these structures was determined from a crystal structure and has described elsewhere (Pannifer *et al.*, 2001).

[4]Those equilibrium free energy parameters in parenthesis were obtained identically to CD-probed denaturation experiments, but unfolding was probed using FRET. The ΔG_{NI} values for FRET were consistently lower than the CD ΔG_{NI} perhaps due to the labeling and/or the K14C and N242C mutations. However, most of the ddG values measured by CD and FRET are similar.

[5]The equilibrium free energy parameters in brackets were obtained using EDANS anisotropy.

[6]The equilibrium free energy differences ($\Delta\Delta G_{NI}$) are computed as $\Delta\Delta G_{NI} = \Delta G_{NI}(\text{mutant}) - \Delta G_{NI}(\text{WT})$. For the mutants where several probes were used, the $\Delta\Delta G_{NI}$'s were averaged.

[7]The kinetic activation free energy differences ($\Delta\Delta G‡$) between mutant and wild type are computed as $\Delta G‡(\text{mutant}) - \Delta G‡(\text{WT})$.

[8]The ϕ values are obtained by the relation, $\phi = \Delta\Delta G‡ / \Delta\Delta G_{NI}$.

[9]The m values were fit independently for the L174A mutant as 2.62 ± 0.04 and 0.59 ± 0.1 for m_{NI} and m_{IJ}, respectively.

mutant. We found that protein unfolding mainly corresponds to the ΔpH-sensitive phase (Figure 2.3B), mirroring what we observed for ΔΨ-driven translocation. Thus we conclude that translocation-coupled unfolding is a mechanical process that can be driven by any type of externally applied force.

We then analyzed the translocation kinetics of all 21 mutants at two different voltages, 40 and 70 mV. The lower ΔΨ of 40 mV was used to probe the steeply ΔΨ-dependent step; and the higher ΔΨ of 70 mV was used to probe the largely ΔΨ-independent step. At 40 mV, the activation energy difference between mutant and WT ($\Delta\Delta G‡$) crudely tracks with the measured $\Delta\Delta G_{NI}$ (Figure 2.5A). There are exceptions to the extent of the correlation between the $\Delta\Delta G‡$ and $\Delta\Delta G_{NI}$, which likely relate to the location of the mutation. At 70 mV, most of the LF_N mutants translocate with comparable rates to WT, and their $\Delta\Delta G‡$s at this voltage are invariant with protein stability (Figure 2.5B). Interestingly, one outlier mutant, M40A, translocated ~10-fold slower than WT at both 40 and 70 mV despite being destabilized 0.6 kcal mol^{-1}. In general, we find that the major unfolding transition coincides with the more driving-force-dependent, ΔΨ-sensitive barrier (Figure 2.3C).

Using ϕ analysis to probe unfolding during translocation. The $\Delta\Delta G‡$ values obtained at low voltage show a large degree of scatter when plotted against $\Delta\Delta G_{NI}$ (Figure 2.5A). The apparent noise in the correlation may reflect that some mutations apportion more of their equilibrium stability difference to the unfolding rate than other mutations. Thus certain sites in the structure are more critical to the unfolding pathway. To determine the location of these critical sites, we calculated the relative effects of each mutation using ϕ analysis (Matouschek *et al.*, 1989). This analysis normalizes the relative kinetic effect of a destabilizing mutation ($\Delta\Delta G‡$) to the equilibrium change in stability between the mutant and WT proteins ($\Delta\Delta G$), using $\phi = \Delta\Delta G‡ / \Delta\Delta G$. These ϕ values are often fractional, ranging between zero and unity. Here ϕ values near unity indicate that the mutation disrupts structure that must unfold in order to cross the rate-limiting barrier. However, ϕ values near zero indicate that the opposite is true, and the mutation disrupts structure that unfolds after crossing the rate-limiting barrier. Therefore, ϕ analysis allows us to map the portions of the protein that must unfold to cross the ΔΨ-sensitive, unfolding barrier during translocation.

We computed ϕ values for each of our 21 mutants from the $\Delta\Delta G‡$ values obtained at 40 mV and the $\Delta\Delta G_{NI}$ values obtained from equilibrium denaturant titration experiments. Overall, we find that LF_N unfolds via a cooperative transition during translocation, where higher ϕ values are observed in a specific region of the structure. This structure is located in or adjacent to the β sheet of LF_N (Figure 2.6). Namely, H91 and A112 are in the loop connecting β strands 2 and 3; and V119 is in β3. Other high-ϕ-value residues, L62, V70 and L145, are located in helices α2, α3 and α4, respectively, which flank the β sheet. The low-ϕ-value sites (L155, L174, F217, A220 and D245) cluster in the carboxy-terminal, α-helical subdomain (Figure 2.6), which is required for binding to PA (Lacy *et al.*, 2002, Lacy *et al.*, 2005). Thus this carboxy-terminal subdomain unfolds only after the unfolding barrier is traversed, and we conclude that the overall unfolding transition is extensive, polarized, and specifically localized in the protein's β-sheet subdomain.

Outlier LF_N mutants. The M40A LF_N mutant is unusual in that it translocates ~10-fold slower than WT despite being destabilized 0.6 kcal mol^{-1} (Figure 2.7A). The mechanism behind this mutational effect appears in conflict with our unfolding model. When M40 is substituted with other large residues (D, K, L or W), the mutants translocate like WT LF_N (Figure 2.7B). However, when M40 is substituted with small residues, such as S, T or V, the translocation rate is slow and resembles that observed for the M40A mutant (Figure 2.7B). The tolerance of the site

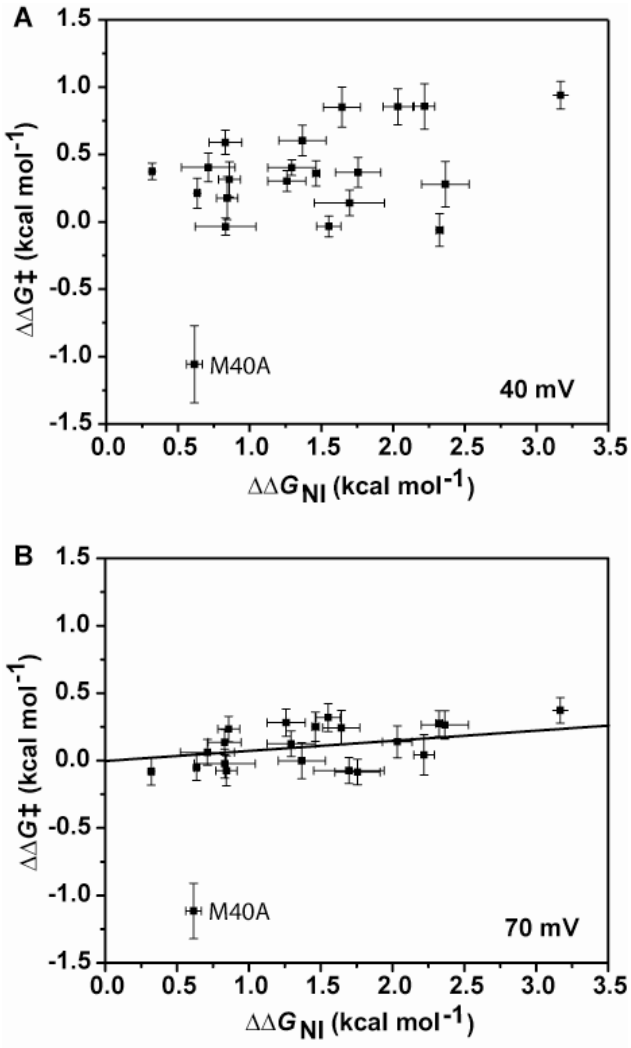

Figure 2.5. The complex relationship between protein stability and translocation activation energies. A plot of the translocation activation energy difference ($\Delta\Delta G‡$) between mutant and WT LF_N at **(A)** 40 mV and **(B)** 70 mV versus their corresponding equilibrium stability differences ($\Delta\Delta G_{NI}$) between their *N* and *I* states.

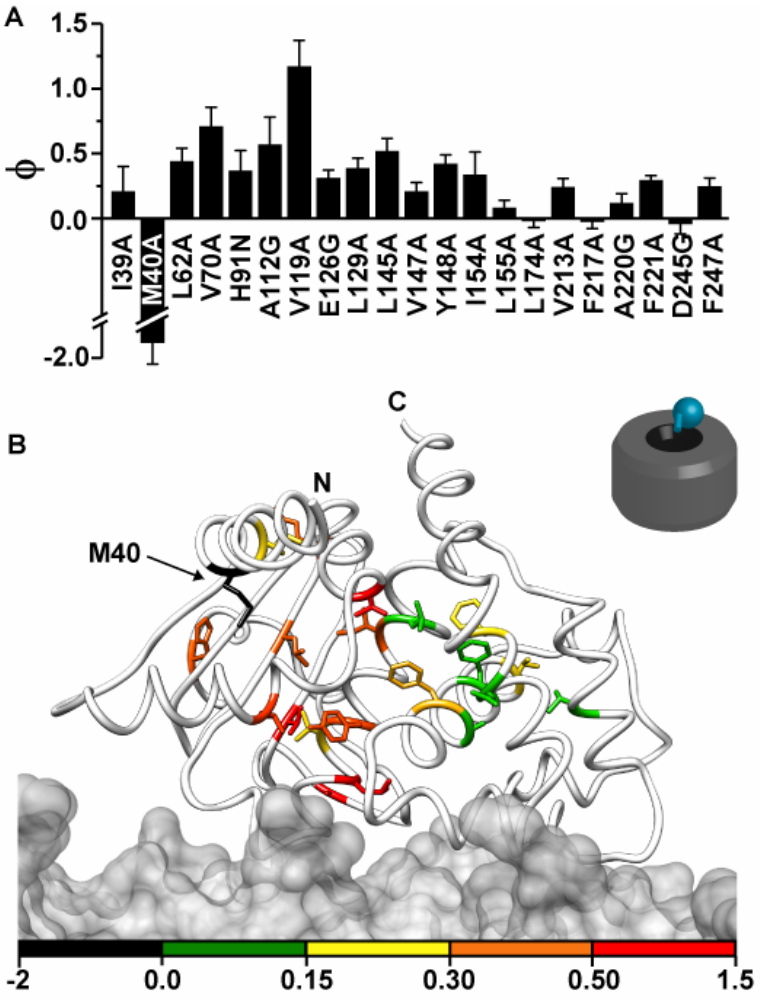

Figure 2.6. The β-sheet subdomain is rate-limiting to the unfolding step of translocation. (A) Mutational ϕ values for the translocation kinetics at 40 mV are shown for each destabilized LF_N mutant. **(B)** The effect of destabilizing mutations on the translocation kinetics is mapped onto a ribbons depiction (Pettersen *et al.*, 2004) of the three-dimensional structure of LF_N. Sites are colored according to their respective ϕ values: >0.5 (red), 0.5-0.3 (orange), 0.3-0.15 (yellow), and 0.15-0 (green). Unusual sites ($\phi < 0$) are black. LF_N is docked on the PA oligomer (gray surface) according to a computational model (Lacy *et al.*, 2005). The orientation of LF_N on the PA oligomer is given in the upper right corner, where LF_N and the PA pre-channel are blue and black, respectively.

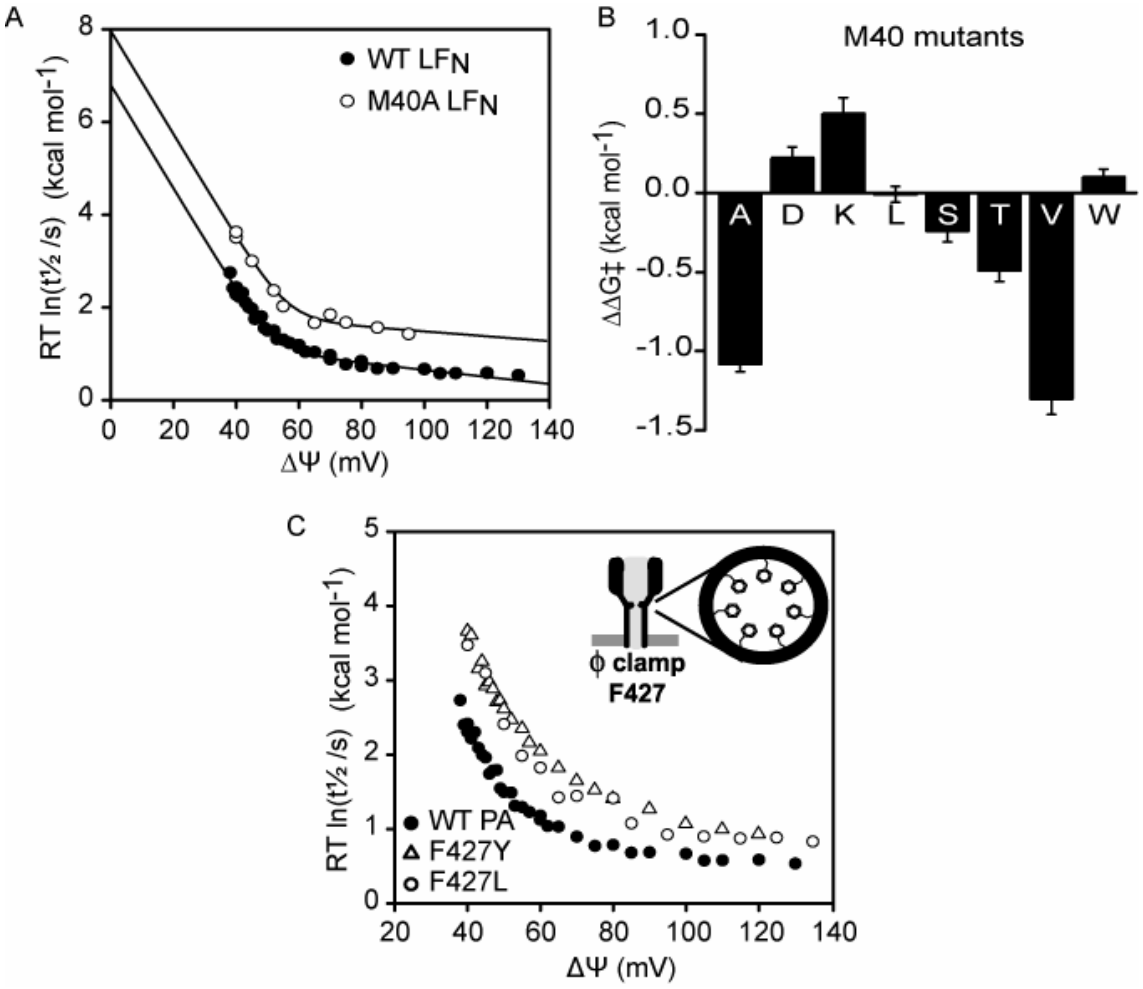

Figure 2.7. Hydrophobic interactions between the channel and substrate catalyze unfolding during translocation. (A) A comparison of the ΔΨ dependence of the ΔG‡ for WT (●) and M40A LF_N (○), where WT PA channels are used in each case. Curves are fit to two-barrier Eyring-Woodhull relations. **(B)** A plot of the translocation activation energy differences (ΔΔG‡) between M40 LF_N mutants and wild type. All activation energies were obtained at 40 mV. A negative ΔΔG‡ value indicates that the rate of translocation slowed relative to wild type. **(C)** The ΔΨ dependence of ΔG‡ for WT LF_N, using either WT PA (●), F427Y PA (Δ) and F427L PA (○) channels. (Inset) Diagram depicting the structure of the PA channel and the ϕ-clamp site, a ring of exposed phenylalanine residues in its lumen that catalyzes protein translocation (Krantz *et al.*, 2005).

to various substitutions suggests that the interaction with PA is nonspecific and may function to destabilize the substrate. Thus, we conclude that mutations at M40 may disrupt a putative interaction between LF and PA, effectively stabilizing the substrate and thereby impeding translocation.

Role of the ϕ clamp in substrate unfolding. The ϕ clamp in the PA channel is required for translocase activity (Krantz *et al.*, 2005). This site is comprised of an F427 residue from each PA subunit in the oligomer, and it forms a critical hydrophobic constriction point in the PA channel (Figure 2.7C, inset). We tested two functional ϕ-clamp mutants to determine how the mutations affected the ΔΨ dependence of translocation. We find that the ϕ-clamp mutants, F427L and F427Y, translocate slower than WT at both high and low ΔΨs (Figure 2.7C). This shift in the translocation kinetics is similar to what we observe for both M40A LF_N translocation and WT LF_N translocation under high-pH conditions, which are known to be stabilizing (Krantz *et al.*, 2006, Krantz *et al.*, 2004). Thus we propose that the substrate is effectively stabilized when translocated via a channel with a mutated ϕ clamp.

To further test whether the ϕ-clamp site has unfoldase activity, we used double-mutant cycle analysis (Horovitz, 1996), where mutations at the ϕ-clamp site could be analyzed in the context of destabilizing mutations in the substrate protein. If the ϕ clamp plays a role in substrate unfolding, then destabilized LF_N mutants should complement a defective ϕ-clamp mutant, and a negative interaction energy ($\Delta\Delta G_{int}$) should be measured. $\Delta\Delta G_{int}$ is calculated from the ΔG‡ values for all four combinations of mutant (MUT) and WT substrate and channel: $\Delta\Delta G_{int}$ = ΔG‡[PA(MUT), LF_N(MUT)] – ΔG‡[PA(WT), LF_N(MUT)] – ΔG‡[PA(MUT), LF_N(WT)] + ΔG‡[PA(WT),LF_N(WT)]. Using the F427L and F427Y PA mutants, we measured interaction energies of about -0.2 to -0.3 kcal mol^{-1} for the V70A, L145A and V119A LF_N mutants (Table 2.2). To control for the small, negative $\Delta\Delta G_{int}$s observed, we tested two carboxy-terminal sites, F217A and F221A, that strongly destabilized LF_N (2-3 kcal mol^{-1}) but only marginally increased the activation energy of translocation; these $\Delta\Delta G_{int}$s were essentially zero, +0.2 and 0.0 kcal mol^{-1}, respectively (Table 2.2). Thus the negative $\Delta\Delta G_{int}$s observed in sites rate-limiting to unfolding suggest that the ϕ clamp is functionally linked to unfolding.

2.4 Discussion

Barriers. To elucidate the molecular mechanism of translocation, not only must the barriers be defined, but the barrier heights must be manipulated in a controlled manner. The anthrax toxin system is advantageous because a constant ΔΨ or ΔpH may be applied, and these driving forces are continuously tunable. Using the Eyring-Woodhull model for electrodiffusion ($\Delta G‡ = \Delta G‡_o + zF\Delta\Psi$) and the chemical potential relationship ($\Delta G‡ = \Delta G‡_o - 2.3nRT\Delta pH$), $\Delta G‡$ should vary linearly with either ΔΨ or ΔpH. We find that the relationship between the $\Delta G‡$ of translocation and the driving forces (ΔΨ and ΔpH) is not linear. The dependencies are, however, biphasic with two different linear extremes (Figure 2.3A,B), indicating that two barriers are crossed during translocation: a more force-dependent barrier and a less force-dependent barrier (Figure 2.3C). Through our analysis of the open and closed lifetimes during single-channel translocations, we were able to distinguish a third barrier-limited process related to substrate docking (Figure 2.8; Supporting Data, p. 37). We do not see this step in ensemble translocation measurements because LF_N is already docked in the channel before translocation initiates. Thus we propose the following three-barrier model for translocation: upon binding the surface of a PA channel, the amino terminus of LF_N first docks inside the channel at the ϕ-clamp site; LF_N then unfolds; and finally the unfolded chain translocates through PA (Figure 2.9A).

Table 2.2. Thermodynamic interaction energies obtained from double mutant cycle analysis.

Destabilizing LF_N Mutant	**PA ϕ clamp Mutant**	**$\Delta\Delta G_{int}$ (kcal mol^{-1})**
Test mutants		
V70A	F427Y	-0.2 (± 0.1)
L145A	F427Y	-0.1 (± 0.1)
V119A	F427Y	-0.3 (± 0.1)
V70A	F427L	-0.2 (± 0.1)
L145A	F427L	-0.3 (± 0.1)
Control mutants		
F217A	F427Y	0.2 (± 0.2)
F221A	F427Y	0.0 (± 0.2)

Double mutant cycle analysis was conducted using mutant (MUT) and WT versions of PA and LF_N. An interaction energy, $\Delta\Delta G_{int}$, was calculated by the following equation: $\Delta\Delta G_{int}$ = $\Delta G‡[PA(MUT), LF_N(MUT)] - \Delta G‡[PA(WT), LF_N(MUT)] - \Delta G‡[PA(MUT), LF_N(WT)] + \Delta G‡[PA(WT),LF_N(WT)]$. The LF_N test mutants were chosen because they had high ϕ values and, therefore, represent critical sites that must unfold to cross the rate-limiting unfolding barrier. Residue F427 of PA forms a critical hydrophobic constriction in the PA channel called the ϕ clamp. The F427Y and F427L ϕ clamp mutants were chosen because they retain some function. Errors in $\Delta\Delta G_{int}$ are shown in parentheses and were determined from linear regression analysis of the translocation activation energies for a datasets of ~4-20 ΔG‡ measurements made between ~38-50 mV. The LF_N control mutants were chosen because they had near zero ϕ values and, therefore, represent non-critical sites that unfold after the rate-limiting unfolding barrier has been crossed.

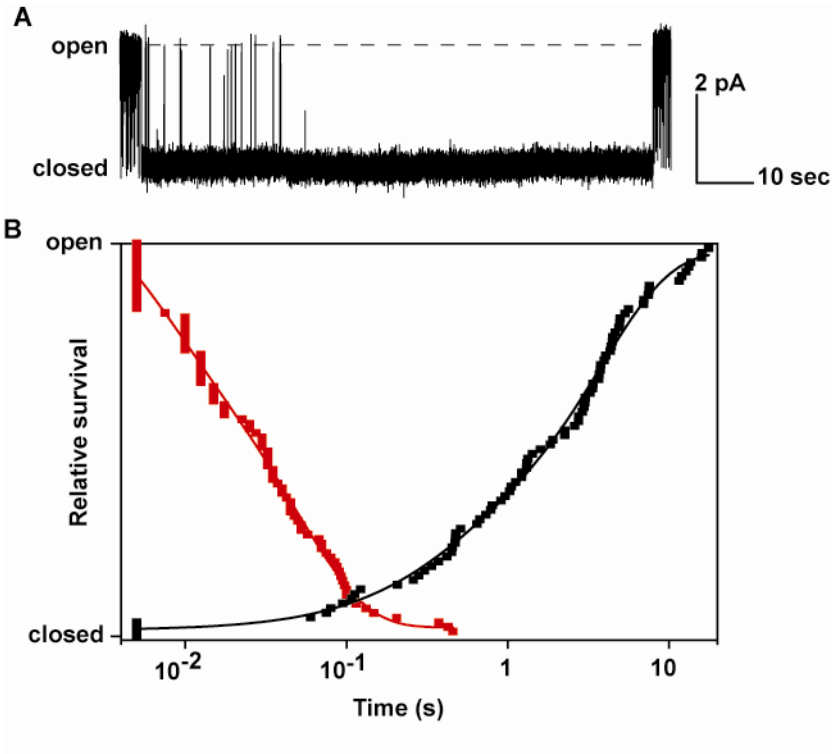

Figure 2.8. Analysis of single-channel translocation events. (A) Example of a single-channel translocation event. Translocation events are observed as extended periods of blocked conductance with occasional spikes to the open state. **(B)** The durations of open and closed events were plotted in normalized survival plots. The survival of the open state events (red trace), which represent how fast the channel re-closes, fits to a single exponential function. The survival of the closed state events (black trace), which represent how fast the channel re-opens, fits to a double exponential function. Exponential fits used, $A(t) = \sum A_i \times \exp(-t/\tau_i) + c$, where A_i is the amplitude for each phase, τ_i is the lifetime of each phase, and c is an offset.

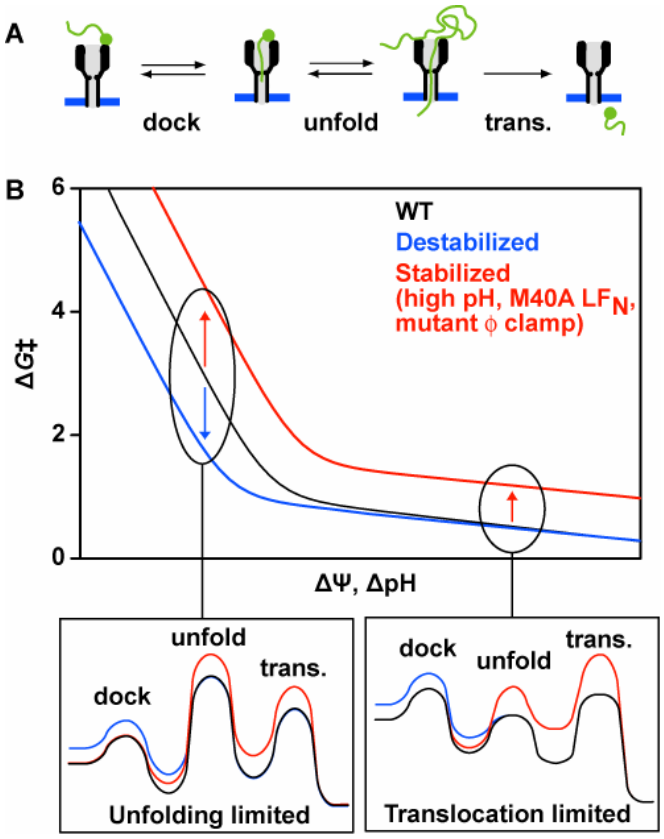

Figure 2.9. Kinetic models for translocation-coupled unfolding. (A) A diagram of the proposed kinetic mechanism for translocation, showing the docking (dock), unfolding (unfold) and translocation (trans.) steps. **(B)** The $\Delta G‡$ versus driving-force curve for WT LF_N via WT PA (black) is compared to scenarios in which the substrate is either destabilized (blue) or effectively stabilized (red), i.e., by channel ϕ-clamp mutations, M40A LF_N, or higher pH. Energy barriers that represent docking, unfolding and translocation are shown for lower ΔΨ translocations (bottom left), when unfolding is limiting, and higher ΔΨ translocations (bottom right), when translocation of the unfolded chain is limiting.

Changes in the observed translocation $\Delta G‡$ versus driving-force plots (Figure 2.9B) indicate how these barriers are affected. Destabilized LF_Ns affect the translocation mechanism in the following manner. Given that the folded state of a destabilized mutant is higher in free energy than that of WT, destabilization lowers the force-sensitive, unfolding barrier and results in faster translocation. At high ΔΨs, the height of the translocation barrier should not change relative to the unfolded well (between the unfolding and translocation barriers) since the mutants translocate at the same rate as WT at high ΔΨs (Figure 2.9B).

Stabilizing the substrate protein should have a different effect on the translocation mechanism. For example, under higher symmetrical pHs, which are thermodynamically stabilizing (Krantz *et al.*, 2004), LF_N translocates more slowly at both high and low ΔΨ (Krantz *et al.*, 2006). These $\Delta G‡$ versus ΔΨ plots are also biphasic; however, they are shifted upward compared to WT translocations at lower pH conditions (Figure 2.9B). Similarly, PA ϕ-clamp mutants and mutations at M40 in LF_N translocate slower than WT LF_N via WT PA (Figure 2.7). For these situations, we propose that the substrate proteins are effectively more stable. Although different mechanisms may be involved, substrate stabilization increases the energy of the unfolded well relative to the folded, docked well, which thermodynamically limits the translocation step (Figure 2.9B). The effective translocation rate decreases proportionately and is consistent with what we observe at high ΔΨs.

We hypothesize that the ϕ clamp and other hydrophobic sites in the channel favor substrate destabilization by providing hydrophobic interaction sites for the unfolded substrate to partition (Krantz *et al.*, 2005). This feature, in addition to the mechanical unfolding force, lowers the unfolding barrier and may explain how proteins can unfold more rapidly during translocation than in solution. Mutating the ϕ clamp may disrupt the hydrophobic interaction with the unfolded substrate, thereby leading to the stabilization of the substrate. In a similar way, the M40A mutant is effectively stabilized in the context of the PA translocase. We propose that a putative M40-PA interaction maintains the substrate in a more unfolded state. Therefore, disrupting hydrophobic interactions between the channel and substrate may effectively stabilize the substrate.

Driving forces. Within the cell, unfolding is most likely accomplished by mechanical denaturation, although acid denaturation is also relevant for toxins that translocate from the endosomal compartment (Krantz *et al.*, 2004). Unfolding forces generally range from 10 to 300 pN (Oberhauser & Carrion-Vazquez, 2008). As previously shown in mitochondrial import (Huang *et al.*, 2002, Shariff *et al.*, 2004), the ΔΨ can generate a mechanical-unfolding force called the Lorentz force, $F = -1.6\Delta\Psi z/d$, where F is the force (in pN) and d is the distance (in Å). In this system, the applied electric field acts upon positive charge in the translocating chain: acidic pH conditions combined with the PA channel's cation selectivity (Blaustein *et al.*, 1989) may induce a net-positive protonation state in the translocating chain, which can be driven productively by the positive membrane potential.

We were able to elaborate on the prior (Huang *et al.*, 2002, Shariff *et al.*, 2004) Lorentz-force translocation model in several ways. By measuring the ΔΨ dependence of the kinetics, we obtained a z-value of 4.2 charges for the unfolding transition. It is likely that these charges are in the amino-terminal leader sequence, which contains a high density of basic residues, and is required to initiate translocation (Zhang *et al.*, 2004a). Considering our z-value of 4.2 charges, we estimate that LF_N experiences forces of 2.7 - 27 pN at 40 mV (for d-values ranging from 100 - 10 Å). The value for d is difficult to predict and complicated by features in the channel, such as the ϕ-clamp site. The ϕ-clamp constriction may concentrate the electric field lines and, therefore, result in a higher force applied over a shorter distance. Finally, we find that the second barrier is

also slightly ΔΨ dependent (z = 0.22). Since this barrier is not related to unfolding, this dependence may indicate that positive charge downstream of the amino-terminal leader sequence is acted upon when crossing the translocation barrier.

We also tested how a proton gradient (ΔpH) would affect the two barriers. This type of driving force is especially relevant to toxin translocation since LF and EF are transported from an acidic endosomal compartment to the cytosol. A ΔpH is sufficient to drive LF_N translocation and is required for the full-length LF and EF substrates (Krantz *et al.*, 2006). The ΔpH may generate a force through a charge-state Brownian ratchet (Krantz *et al.*, 2006). In this mechanism, the protonation state of anionic residues in the translocating chain and/or the channel can change during successive cycles of translocation. As the substrate exits the channel under Brownian motion, anionic residues are more likely to deprotonate as they enter the higher pH side of the membrane, and they may develop an electrostatic repulsion with the anionic channel. Thus Brownian motion is rectified, favoring productive translocation.

Analogous to ΔΨ-driven translocation, the activation energy of ΔpH-driven translocation is biphasic with two linear extremes. Once again, this biphasic dependence indicates that two barriers are crossed during translocation: one barrier is strongly ΔpH dependent, and the other is ~10-fold less ΔpH dependent. We found that destabilizing mutations affect the ΔpH-sensitive barrier most (Figure 2.3B), and we can generalize that unfolding is the most force-dependent step in the translocation mechanism. Compared to the unfolding step, the translocation step is about 10-20 fold less-force-dependent, which makes physical sense. While a large amount of force is required to disassemble the hydrophobic core of a protein, a smaller force may be sufficient to overcome unproductive diffusive motion and guide the unfolded chain through a channel.

Structures. Using mutagenesis, we sought to determine how LF_N unfolds during translocation. We avoided mutating regions of LF_N that are expected to bind the channel (Lacy *et al.*, 2005, Lacy *et al.*, 2002). The Y236A LF_N mutant, which is defective in binding to PA at neutral pH (Lacy *et al.*, 2002), is also defective in binding to PA at low pH. We conclude that PA's LF binding sites undergo little remodeling once the oligomer converts to its membrane-inserted channel state. We also presume that LF_N begins translocation in a largely native topology on the surface of PA, and we use the $\Delta\Delta G_{NI}$ values accordingly in our ϕ analysis. Finally, it should be noted that this structural picture is consistent with biochemical data on the LF_N-PA complex (Young & Collier, 2007).

The fact that LF_N's carboxy terminal subdomain is bound to PA in a native conformation prior to translocation may be important because it acts as the point of resistance (or fulcrum) for the force that is applied on the amino-terminal end (Figure 2.10). If the carboxy-terminal end is bound tightly to PA, then a force will be applied somewhere in the structure of LF_N. We adapted mutational ϕ analysis (Matouschek *et al.*, 1989) to determine the location of the structure acted upon by the applied force. Our analysis identified a core of high ϕ-value residues that must unfold to cross the unfolding barrier; these sites are located towards the amino terminus in and adjacent to the β1-β4 sheet and make up about a third of the protein (Figure 2.6B). This core is consistent with mechanical unfolding studies, which show that β sheets often represent the *mechanical breakpoint*, or rate-limiting structure in the unfolding pathway. However, the direction in which the force is applied can greatly affect the protein's mechanical stability (Carrion-Vazquez *et al.*, 2003). For example, prior studies suggest that it takes more force to pull apart β sheets in a *shear* topology (i.e., when the hydrogen bonds are perpendicular to the force vector) than it does when β sheets are in a *zipper* topology (i.e., when the hydrogen bonds are parallel to the force vector) (Brockwell *et al.*, 2003). Considering LF_N's orientation when bound

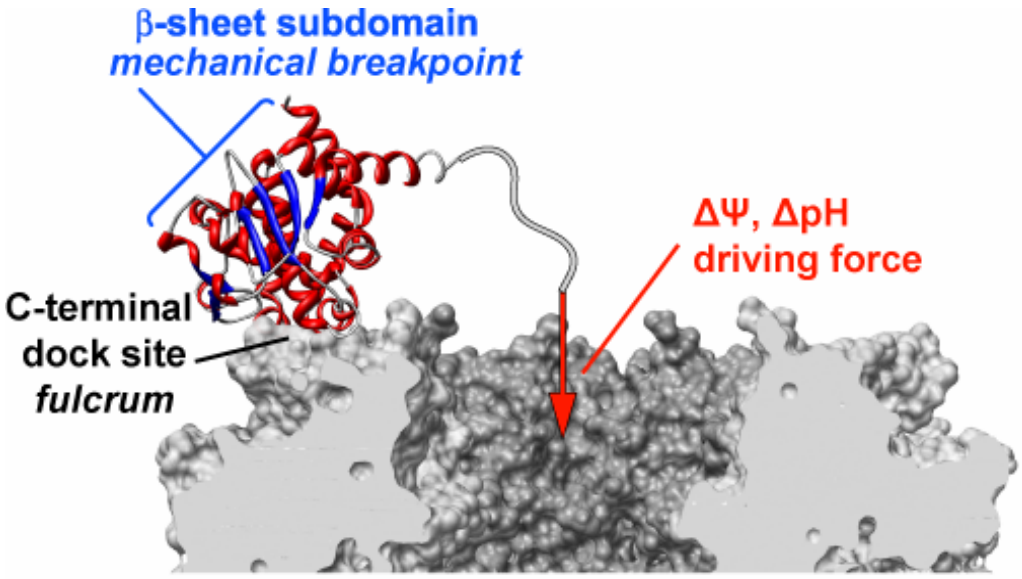

Figure 2.10. Structural model for translocation-coupled unfolding. In the proposed mechanical unfolding model, leverage is produced according to the following scheme. The carboxy-terminal end of LF_N binds on the PA channel, acting as the fulcrum or point of resistance to the force applied on its amino-terminal end. The unfolding force may then be focused upon LF_N's β-sheet subdomain—the mechanical breakpoint in the unfolding pathway. LF_N's α helices (red) and β strands (blue) are shown in the native state while docked on PA (Lacy *et al.*, 2005).

to PA (Lacy *et al.*, 2005), we believe the β sheet would be pulled apart in a zipper topology (Figure 2.10). Thus the geometry of the LF_N-PA complex may be such that it best exploits a force-dependent denaturation mechanism. The force may be concentrated on the β-sheet subdomain, and the orientation of the structure of LF_N may favor the zipper topology.

Does local structure control translocation-coupled unfolding? Prior studies indicate that translocation is limited by the local stability near the signal tag (Kenniston *et al.*, 2003, Kenniston *et al.*, 2004, Matouschek, 2003). Once this local secondary structure near the signal tag is unfolded, the rest of the protein is thought to denature rapidly (Huang *et al.*, 1999). This model does not appear to apply to anthrax toxin translocation, since a larger portion of LF_N must unfold to cross the unfolding barrier. To observe this unfolding step, we relied on our ability to control the driving force and measured the driving force dependencies of the translocation activation energy. Without this control, we may have not observed the unfolding step, because protein unfolding is not rate-limiting under large driving force conditions.

Moreover, local unfolding events that result in complete collapse of the protein are generally not observed in solution-based protein unfolding studies. Native state hydrogen exchange studies on barnase (Vu *et al.*, 2004) (the substrate used in mitochondrial import studies), show that local structure may unfold as a cooperative unit, but these local-unfolding or fraying events are not rate-limiting to unfolding in solution. While the mechanical unfolding energy landscape may differ from that encountered in solution, it seems unlikely that the unfolding of a single α helix or loop on the surface of a protein represents its rate-limiting structure, or mechanical breakpoint.

We suggest instead that local unfolding, or fraying, may simply lengthen the signal tag, allowing the protein to better engage with the channel or an ATP-dependent motor. Better engagement with the channel or motor could increase the driving force applied to the substrate such that unfolding is no longer rate-limiting. In fact, when the 12-residue ClpXP degradation tag (Kenniston *et al.*, 2004) and 65-residue mitochondrial import presequence (Matouschek *et al.*, 1997) are lengthened ~10 residues, the rate of translocation is accelerated ~10- and ~3-fold. Therefore, local unfolding near the signal tag may result in partially unfolded intermediates, but the driving force may have been too high, such that the unfolding of the remaining structure in the protein was not rate-limiting.

General model for translocation. We propose a broad model for translocation-coupled unfolding, where unfolding is limited by the protein's mechanical breakpoint. This breakpoint corresponds to a significant portion of the protein's structure, and its location can vary depending on the structure of the protein. Sometimes the breakpoint is located immediately following the signal tag; other times, it is located deeper in the structure. Finally, mechanical breakpoints may only be observed under low driving force conditions since the unfolding barrier is the most force-dependent.

Translocation may follow a three-barrier model. The first barrier corresponds to docking or engagement with the channel or motor. For the latter two barriers, one is strongly force-dependent; and the other barrier is ~10-fold less force dependent. The strongly force-dependent step corresponds to a cooperative unfolding transition and is limited by the protein's mechanical breakpoint. We presume that the less force-dependent step likely involves the translocation of the unfolded chain, since it is not limited by unfolding. Remarkably, this putative translocation barrier imposes an overall speed limit on the translocation of unfolded protein substrates (on the order of ~10 seconds for ~100-700 residue proteins) for our system (Krantz *et al.*, 2006) and others (Burton *et al.*, 2001, Huang *et al.*, 1999, Kenniston *et al.*, 2003, Kenniston *et al.*, 2004);

and this limit on the translocation step is invariant with the type of driving force applied (whether ATP, ΔΨ, or ΔpH). Understanding the common features shared among these and other translocases will advance our understanding of cellular protein unfolding.

2.5 Materials and Methods

Proteins. PA, LF_N (residues 1-263 of LF), and their mutants were produced as described (Krantz *et al.*, 2005, Krantz *et al.*, 2004). His_6 tags were removed by bovine α-thrombin. Oligomeric PA was prepared as described (Krantz *et al.*, 2005).

Fluorescence labeling. LF_N labeled for FRET [called LF_N*, LF_N K14C N242C fluorescently labeled with Cys-reactive Alexa Fluor 488 and 546 C_5 maleimides (Invitrogen)] and LF_N labeled for fluorescence anisotropy [called $LF_N^{\#}$, LF_N K14C N242C fluorescently labeled with Cys-reactive dye 5-(2-iodoacetylaminoethylamino) naphthalene-1-sulfonic acid (Invitrogen)] were prepared as described (Krantz *et al.*, 2004).

Equilibrium chemical denaturation experiments. Guanidine hydrochloride (GdmCl) (MP Biomedicals, Solon, OH) denaturation profiles were obtained as described (Krantz *et al.*, 2004). An AVIV Model 400 (AVIV Biomedical) or a Jasco Model 810 spectropolarimeter (Jasco Inc.) was used for circular dichroism (CD) studies on unlabeled LF_N at 222(±2) nm. A FluoroMax-3 spectrofluorimeter (HORIBA Jobin Yvon) was used for fluorescence anisotropy (FA) studies on $LF_N^{\#}$ at 350(±16) nm excitation and fluorescence resonance energy transfer studies on LF_N* at an excitation at 488 nm and an emission ratio of 520 to 570(±16) nm. For fluorescence anisotropy studies, the emission was > 420 nm; parallel ($F_{\|}$) and perpendicularly ($F\perp$) polarized emission intensities were recorded to calculate anisotropy, a, by: $a = F_{\|} - F\perp / (F_{\|} + 2F\perp)$. The buffer was 10 mM sodium phosphate, 100 mM sodium chloride, 0.75 M trimethylamine *N*-oxide at pH 7.5 ($[LF_N]$ = 10 nM to 1 μM). Increments of titrant containing this same buffer and protein concentration plus ~7 M GdmCl were added via a Hamilton titrator until equilibrium was reached prior to measurement. Unfolding curves were fit to a four-state model [$N \rightleftarrows I \rightleftarrows J \rightleftarrows U$ (Krantz *et al.*, 2004)]. These probes generally produced similar free energy parameters when comparing each mutant to their respective WT (either LF_N, $LF_N^{\#}$ or LF_N*).

Translocation kinetics. To form planar lipid bilayers, membrane-forming solution [3% 1,2-diphytanoyl-*sn*-glycerol-3-phosphocholine (Avanti Polar Lipids) in *n*-decane] was applied across a 100-μm aperture in a 1-mL, white-Delrin cup, using the brush technique and capacitance test as described (Krantz *et al.*, 2006). The membrane separated two chambers (cis and trans) containing 1 mL of pH 5.6 universal bilayer buffer [UBB: 100 mM KCl, 1 mM EDTA, and 10 mM each of oxalic acid, 2-(*N*-morpholino)ethanesulfonic acid, and phosphoric acid]. KCl-agar salt bridges linked the Ag/AgCl electrodes to each side. Current responses were amplified by an Axoclamp 200B and were recorded using CLAMPEX10 (Axon Instruments, Union City, CA). $\Delta\Psi \equiv \Psi_{cis} - \Psi_{trans}$ ($\Psi_{trans} \equiv 0$ V). All experiments were conducted at room temperature.

ΔΨ-dependent translocation experiments were performed as described in the text (Krantz *et al.*, 2006). After blocking channel conductance with substrate, translocation was driven by increasing the ΔΨ. For most proton gradient-driven translocation experiments (where ΔpH = ±0.5), the cis and trans chambers contained UBB differing only in pH, and the cis compartment was held at a constant pH of 5.6. After perfusion of the cis chamber, translocation was initiated by stepping to a higher +ΔΨ. For larger +ΔpHs > +0.5, LF_N was added at a ΔΨ of +1 mV to prevent translocation during perfusion due to the higher ΔpH.

Two-barrier analysis of empirical activation-energies. The $t_{1/2}$ values are analogous to a mean transit time, τ, for the two sequential first-order reactions, such that $\tau = \tau_1 + \tau_2$. The ΔΨ

and ΔpH dependencies of the translocation half-times (measured in seconds) were fit to the following two-barrier models, where $\Delta G‡(\Delta\Psi, \Delta pH) = RT \ln t_{1/2} / c$, where c is a 1-s reference factor:

$$\Delta G‡(\Delta\Psi) = RT\ln[\exp((\Delta G‡_{o1} + z_1F\Delta\Psi)/RT) + \exp((\Delta G‡_{o2} + z_2F\Delta\Psi)/RT)] \quad \text{(Eq. 1)}$$

$$\Delta G‡(\Delta pH) = RT\ln[\exp((\Delta G‡_{o1} - 2.3n_1\Delta pH)/RT) + \exp((\Delta G‡_{o2} - 2.3n_2\Delta pH)/RT)] \quad \text{(Eq. 2)}$$

Single-channel translocation. PA oligomer was applied directly to membranes at ~10^{-15} M in UBB. Single channel insertion is observed by a discrete step in current under an applied voltage. Once a single channel has inserted into the membrane, LF_N was added at ~10^{-11} M and the voltage was increased to +50 mV. Data were acquired at 400 Hz. The data were analyzed as described in the text.

2.6 Supporting Data

Single-channel translocation. Single channel translocation events were recorded using WT LF_N. Here, a single channel is formed, and LF_N is added to the membrane under a constant voltage. Translocation events are observed as extended periods of blocked conductance with occasional spikes to the open state (Figure 2.8). These are productive translocation events (as opposed to unproductive events, where the substrate dissociates from the cis face of the channel), because the overall lifetimes of the conductance blocks are consistent with our ensemble translocation measurements.

We analyzed the channel's open and closed times in survival plots to determine the lifetimes, τ, where $\tau = 1/k$, of the underlying kinetic processes. A survival plot of the duration of the spikes in the open state fits to a single exponential (τ = 49 ms) (Figure 2.8). A survival plot of the time between spikes fits to a double exponential with longer lifetimes (τ_1 = 330 ms; τ_2 = 3.8 s) (Figure 2.8). This analysis shows that the underlying kinetic mechanism contains multiple barriers. Secondly, the analysis indicates that there are at least two forward kinetic steps prior to translocation. Our model includes barrier(s) not typically observed in ensemble measurements. We propose that the shorter ~50-ms step is related to docking, since the channel opens and then closes again; and here we imagine that the protein undocks from the channel and then tries to translocate again. This docking barrier is small, since the fast undocking τ is ~300 ms. The longer backward τ of ~3.8 s may be related to a second undocking process and/or protein folding; however, this possibility must be tested with destabilized LF_Ns.

Chapter 3

Structural basis for the unfolding of anthrax lethal factor by protective antigen oligomers

3.1 Abstract

The protein transporter, anthrax lethal toxin, is comprised of protective antigen (PA), a transmembrane translocase, and lethal factor (LF), a cytotoxic enzyme. Following assembly into holotoxin complexes, PA forms an oligomeric channel that unfolds LF and translocates it into the host cell. We report the crystal structure of the core of a lethal toxin complex to 3.1-Å resolution; the structure contains a PA octamer bound to four LF PA-binding domains (LF_N). The first α helix and β strand of each LF_N unfold and dock into a deep amphipathic cleft on the surface of the PA octamer, which we call the α clamp. The α clamp possesses nonspecific polypeptide binding activity and is functionally relevant to efficient holotoxin assembly, PA octamer formation, and LF unfolding and translocation. This structure provides insight on the mechanism of translocation-coupled protein unfolding.

3.2 Introduction

Protein secretion and degradation are essential cellular processes that allow for protein trafficking, organelle biogenesis, protein quality control, and cell-cycle regulation (Wickner & Schekman, 2005, Navon & Ciechanover, 2009, Sauer *et al.*, 2004). Since folded proteins are thermodynamically stable under typical conditions, these processes often require complex, energy-consuming molecular machines (Sauer et al., 2004, Cheng, 2009, Young & Collier, 2007, Matouschek, 2003), which catalyze a series of unfolding and translocation reactions (Krantz *et al.*, 2006, Krantz *et al.*, 2005, Thoren *et al.*, 2009, Kenniston *et al.*, 2003, Martin *et al.*, 2008, Huang *et al.*, 2002, Huang *et al.*, 1999). Anthrax toxin (Young & Collier, 2007, Smith & Keppie, 1954), a three-protein virulence factor secreted by *Bacillus anthracis*, is an example of such a transmembrane protein delivery system. This bacterial toxin follows the classical two-component *AB* paradigm, where the *A* component is an active enzyme that localizes to and enters cells by forming complexes with the cell-binding, or *B* component. Anthrax toxin is composed of two *A* components, LF (91 kDa) and edema factor (EF, 89 kDa), and one *B* component, PA (83 kDa). Therefore, two different toxic complexes can form: lethal toxin (LT, PA plus LF) and edema toxin (ET, PA plus EF). LT (which we focus on herein) causes macrophage lysis (Friedlander, 1986), immune system suppression (Agrawal & Pulendran, 2004), and death (Smith & Keppie, 1954).

For LT to inflict its cytotoxic effects, PA and LF must assemble into active holotoxin complexes, which can translocate LF into host cells. Proteases present either on host-cell surfaces or in blood serum potentiate LT assembly by proteolytically nicking PA, yielding $_nPA$ (Ezzell & Abshire, 1992, Milne et al., 1994, Kintzer et al., 2009). Dissociation of a 20-kDa amino-terminal fragment from $_nPA$ exposes LF-binding sites, permitting assembly. The resulting LT complex contains multiple copies of LF bound to either a ring-shaped PA homoheptamer, PA_7 (Milne et al., 1994, Petosa *et al.*, 1997, Katayama *et al.*, 2008, Kintzer et al., 2009), or homooctamer, PA_8 (Kintzer et al., 2009). Octameric PA forms more robust LT complexes than heptameric PA under physiological conditions (Kintzer *et al.*, 2010). The crystal structures of the individual PA and LF monomers (Pannifer et al., 2001, Petosa et al., 1997) and the assembled

PA heptamer (Lacy *et al.*, 2004) and octamer (Kintzer et al., 2009) are known. However, an atomic-resolution X-ray crystal structure of a lethal toxin co-complex has not been described.

After the LT complex is endocytosed, the PA oligomer transforms into a transmembrane, β-barrel channel (Benson *et al.*, 1998) through which LF translocates to enter the cytosol. Due to the narrowness of the channel, LF unfolds during translocation. The acidic endosomal pH conditions required for toxin action (Friedlander, 1986) not only aid in the destabilization of LF (Krantz et al., 2004) but also drive further LF unfolding (Thoren et al., 2009) and translocation by means of a proton motive driving force (Krantz et al., 2006). This driving force is comprised of a proton gradient (ΔpH) and membrane potential (ΔΨ). Efficient coupling of the ΔpH requires a catalytic active site in the channel, called the ϕ clamp, composed of a narrow ring of phenylalanine residues (Krantz et al., 2006, Krantz et al., 2005). The ϕ clamp forms a narrowly apposed substrate clamping site in the central lumen of the PA channel (Krantz et al., 2005), and it allows the channel to catalyze unfolding (Thoren et al., 2009) and translocation (Krantz et al., 2005) presumably by forming transient interactions with the unfolded translocating chain (Krantz et al., 2005).

Many, but not all, protein processing machines that translocate, unfold and/or refold proteins utilize analogous polypeptide clamping features to denature a protein and engage with its unfolded structure. The features that bind to unstructured or unfolded polypeptides include hydrophobic/aromatic pore loops (Krantz et al., 2005, Van den Berg *et al.*, 2004, Martin et al., 2008, Lum *et al.*, 2008, Wang *et al.*, 2001), polypeptide clamping sites (Krantz et al., 2005, Zimmer *et al.*, 2008), and other substrate binding clefts or adapters (Hinnerwisch *et al.*, 2005, Levchenko *et al.*, 2005, Levchenko *et al.*, 2003). Some of these machines utilize tandem polypeptide binding sites (Hinnerwisch et al., 2005, Krantz et al., 2005, Thoren et al., 2009): one site is a substrate docking site, which feeds into a second hydrophobic site found deeper within the pore. Questions surround the mechanism of these clamping sites and their interactions with unfolded substrates. How do these sites unfold proteins? How do they process the wide chemical complexity and configurational flexibility contained in an unfolding substrate? These questions have remained unanswered, in part because atomic resolution structures of unfolding intermediates in complex with these clamps have not been described. Here we report a structure of a partially unfolded substrate, the PA-binding domain of LF, in complex with its unfolding machine, the PA oligomer.

3.3 Results

Crystal structure of the $PA_8(LF_N)_4$ complex. For these crystallographic studies, we focused on the PA_8 oligomer, considering its enhanced thermostability as well as its advantageous fourfold, square-planar symmetry (Kintzer et al., 2009). All of the crystallography work was done by Geoff Feld, and the mass spectrometry data was collected by Harry Sterling in the Williams lab. By mass spectrometry, we find that the PA_8LF_4 complex is physiologically relevant, as it assembles from the full-length, wild-type (WT) PA and LF subunits (Figure 3.1a). Our best diffracting crystals contain LF_N (LF residues 1-263) and a PA construct lacking its membrane-insertion loop (Kintzer et al., 2009), which is superfluous to the known PA-LF_N interaction (Cunningham *et al.*, 2002). LF_N, the minimal portion of LF that specifically binds PA (Arora & Leppla, 1993), can translocate heterologous domains as amino- or carboxy-terminal fusions into cells (Arora & Leppla, 1994, Milne *et al.*, 1995). EF contains a homologous PA-binding domain, and the PA-LF_N interaction is likely general to LT and ET complexes (Lacy *et al.*, 2002). Homogenous $PA_8(LF_N)_4$ complexes (Figure 3.1b) form crystals in the $P42_12$ space

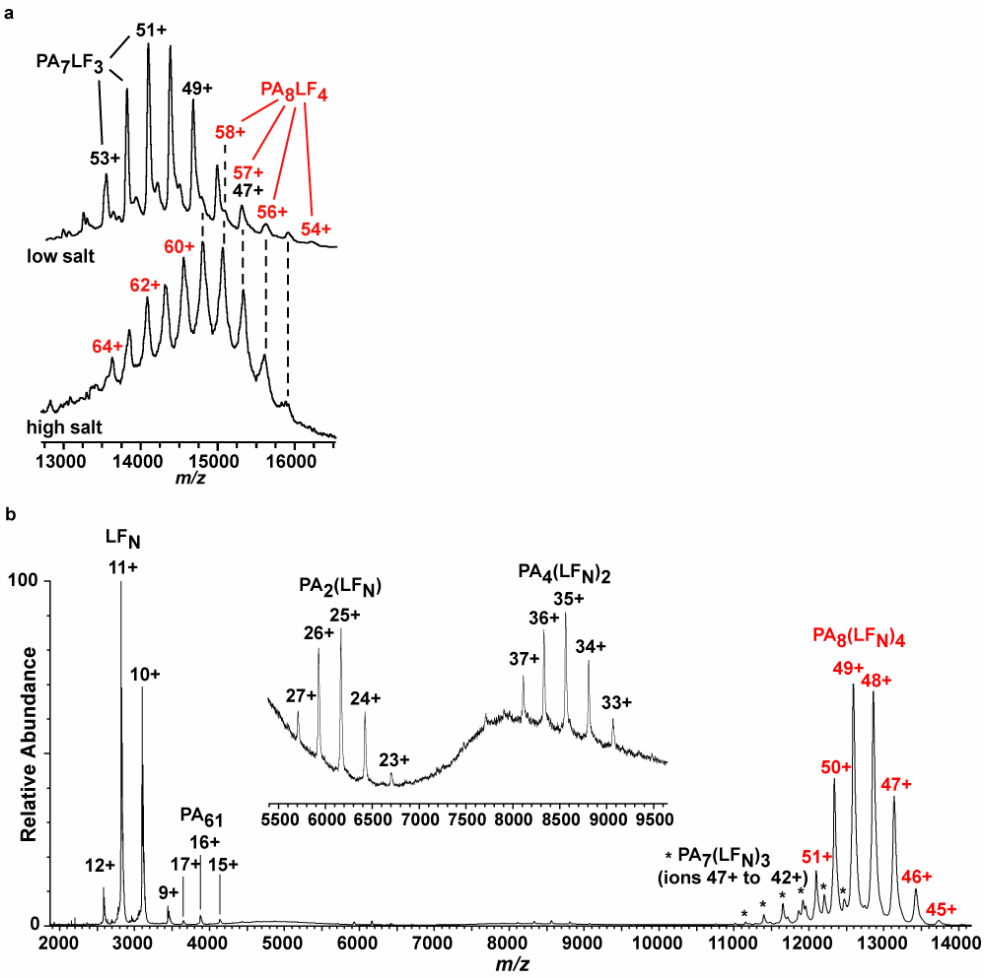

Figure 3.1. Mass spectrometry analysis of the PA_8LF_4 and $PA_8(LF_N)_4$ complexes. (a) LT complexes were assembled from WT $_n$PA and WT LF for 1 hour at room temperature at pH 7.8 in either low salt (150 mM ammonium acetate, upper panel) or high salt (300 mM ammonium acetate, lower panel). NanoESI-MS spectra of the LT assembly reactions were obtained. Molecular masses of 718,726 (±491) and 873,881 (±288) Da are obtained from the charge-state distributions (or *m/z* peaks), which correspond to the PA_7LF_3 (black) and PA_8LF_4 (red) complexes, respectively. **(b)** The PA construct used in the crystallographic experiments, $PA^{\Delta MIL}$, is a membrane-insertion-loop-deleted construct. This $PA^{\Delta MIL}$ construct was assembled, and the octameric assembly state was purified (Kintzer et al., 2009). PA octamers were liganded with LF_N at a 1:1 stoichiometric ratio (LF_N:PA monomer) to make $PA_8(LF_N)_4$ complexes. The $PA_8(LF_N)_4$ complexes were analyzed by mass spectrometry. The mass spectrum revealed the largest relative abundances for free LF_N and the octameric complexes, $PA_8(LF_N)_4$. The octameric complex's *m/z* peaks are numbered in red. Low relative abundances were observed for PA_{20}-dissociated $PA^{\Delta MIL}$ monomer (PA_{61}) and $PA_7(LF_N)_3$ as well as $PA_2(LF_N)$ and $PA_4(LF_N)_2$ (magnified inset). The latter two species are likely intermediates in the assembly of the $PA_8(LF_N)_4$ complex (Kintzer et al., 2009).

group that diffract X-rays to 3.1 Å (Table 3.1). Molecular replacement solutions indentify two PA_2 complexes and significant (2.7σ) unassigned electron density (F_o-F_c) for α helices located proximal to the domain 1′ (D1′) surface of each PA_2 complex. Rounds of polyalanine-helix modeling and refinement reveal that the novel helical density aligns well with α2, α4, α9, and α10 of LF_N. The two occurrences of the PA_2LF_N ternary complex (Figure 3.2a) in the asymmetric unit are structurally identical; its PA subunits are structurally similar to the full-length PA monomer (Petosa et al., 1997) and the PA subunits observed in the PA_7 and PA_8 prechannel oligomers (Lacy et al., 2004, Kintzer et al., 2009). Thus the biological unit—the $PA_8(LF_N)_4$ prechannel complex (Figure 3.2b)—is comprised of four PA_2LF_N ternary complexes (Figure 3.2c).

Interestingly, LF_N α1/β1 (residues 29-50) unfolds and adopts a novel conformation relative to free LF (1J7N (Pannifer et al., 2001)). LF_N α1/β1 docks in the cleft formed between adjacent PA subunits and aligns well with the experimental electron density (Figure 3.3a,b). We can assign this unique conformation of α1/β1 since it extends from LF_N α2 as a contiguous stretch of electron density contoured at σ=1 (Figure 3.4a). LF_N's carboxy terminus also reveals well-defined electron density (Figure 3.3c). Overall, LF_N excludes 1900 $Å^2$ of solvent accessible surface area (SASA) on the PA dimer. This surface is comprised of two discontinuous LF_N-binding subsites (Figure 3.2a) formed by adjacent PA subunits, termed PA_N and PA_C (to reflect whether the PA subunit interacts primarily with the amino terminus or carboxy terminus of LF_N, respectively). The details of these respective subsites, called the α-clamp binding subsite and the carboxy-terminal binding subsite, are depicted in Figure 3.5. Thus upon binding the PA oligomer, LF_N partially unfolds, whereby its first α helix and β strand (i) separate from the main body of the protein, (ii) dock into the cleft between two adjacent PA subunits (Figure 3.2a), and (iii) orient toward the center of the PA oligomer lumen (Figure 3.2b).

The carboxy-terminal binding subsite. At the carboxy-terminal subsite, LF_N's carboxy-terminal subdomain excludes ~900 $Å^2$ on PA_C (Figure 3.5b). The structure reveals a hydrophobic interface, involving PA_C Phe202, Pro205, Ile207, and Ile210 and LF Val232, Leu235, His229, Tyr223, Leu188, and Tyr236. In particular, LF Tyr236 is well packed against PA_C Ile210 (Figure 3.3c) and its phenol hydroxyl forms a hydrogen-bonding network with PA_C His211 and Asp195 near the center of the hydrophobic interface (Figure 3.5b). Additional electrostatic interactions surround this hydrophobic core. The carboxyl side chain of PA_C Glu190 forms a pair of hydrogen bonds with both the γ hydroxyl and amide nitrogen of LF Thr141; PA_C Lys197, Lys213, Lys214 and Lys218 form salt bridges with LF Asp182, Asp187, Asp184, and Glu142, respectively; and PA_N Arg200 forms a salt bridge with LF Glu139. PA and LF residues localized in this binding subsite are corroborated by mutagenesis studies, probing binding (Figure 3.6a,b), assembly/binding (Figure 3.7a), (Lacy *et al.*, 2005, Lacy et al., 2002, Melnyk *et al.*, 2006, Cunningham et al., 2002, Chauhan & Bhatnagar, 2002) and cytotoxicity (Chauhan & Bhatnagar, 2002) (Figure 3.7b,c).

The α-clamp binding subsite. At the α-clamp subsite, PA_N and PA_C interact with LF_N's unfolded α1 and β1 structures (Figure 3.5a). Remarkably, hydrogen bonds lost upon LF_N unfolding are reformed on the surface of PA: LF_N α1 maintains a similar helical conformation; and LF_N β1 (Ile43 and Lys45) forms parallel β-sheet hydrogen bonds with Leu203 in PA_N β13 (Figure 3.3b). PA_N Pro205, which is positioned at the end of PA_N β13, terminates the parallel-sheet interactions with LF_N β1. Overall, LF_N α1/β1 excludes 1000 $Å^2$ of SASA on PA. LF_N α1 is docked deep into the α-clamp cleft at the interface of adjacent PA subunits (Figures 3.2a and 3.5a). Reminiscent of calmodulin complexes with peptide helices

Table 3.1: **Data collection and refinement statistics**

	$PA_8(LF_N)_4$ [a]
Data collection	
Space group	$P42_12$
Cell dimensions	
a, *b*, *c* (Å)	178.38, 178.38, 240.36
α, β, γ (°)	90, 90, 90
Resolution (Å)	49.8-3.1(3.2-3.1)[b]
$R_{p.i.m.}$	6.9(46.0)
$I / \sigma I$	11.4(2.2)
Completeness (%)	92.0(78.0)
Redundancy	7.9(8.0)
Refinement	
Resolution (Å)	49.8-3.1
No. reflections	65,165
R_{work} / R_{free}	24.9/ 28.1
No. atoms	
Protein	20,397
Ligand/ion	8
Water	4
B-factors	
Protein	100.7
Ligand/ion	53.3
Water	56.7
R.m.s. deviations	
Bond lengths (Å)	0.005
Bond angles (°)	0.610

[a]Data for this complex were collected from a single crystal.
[b]Values in parentheses are for the highest-resolution shell.

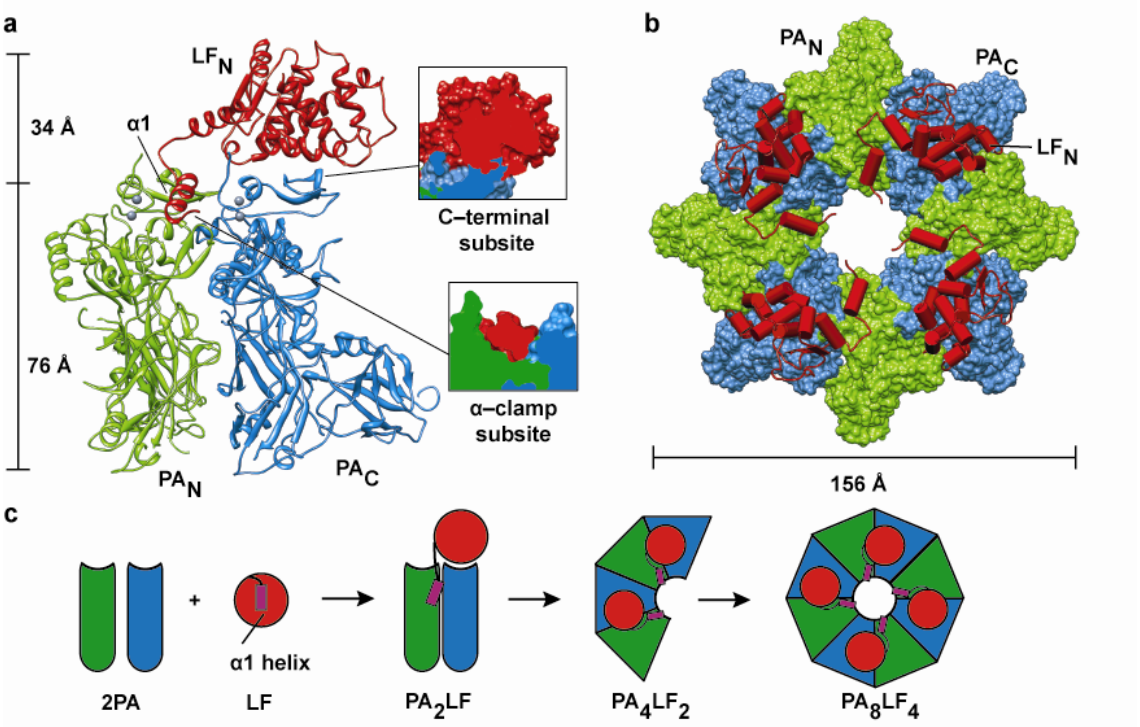

Figure 3.2. Structure of LF's PA-binding domain in complex with the PA octamer. **(a)** (Left) Ribbon depiction of the PA_2LF_N ternary complex. LF_N (red), PA_N (green), PA_C (blue), and calcium ions (gray spheres). (Right) Slices through a surface rendering of the two LF_N-binding subsites, with the carboxy-terminal binding subsite at the top and the α-clamp subsite at bottom. **(b)** Axial rendering of the biological unit, the $PA_8(LF_N)_4$ complex, colored as in (b). The PA octamer is shown as a molecular surface, and LF_N's helices and strands are cylinders and planks, respectively. **(c)** LF_N α1-β1 binds the α-clamp subsite formed at the interface of two PA subunits, driving the assembly of dimeric and tetrameric PA intermediates (Kintzer et al., 2009), which in turn form PA_8 complexes.

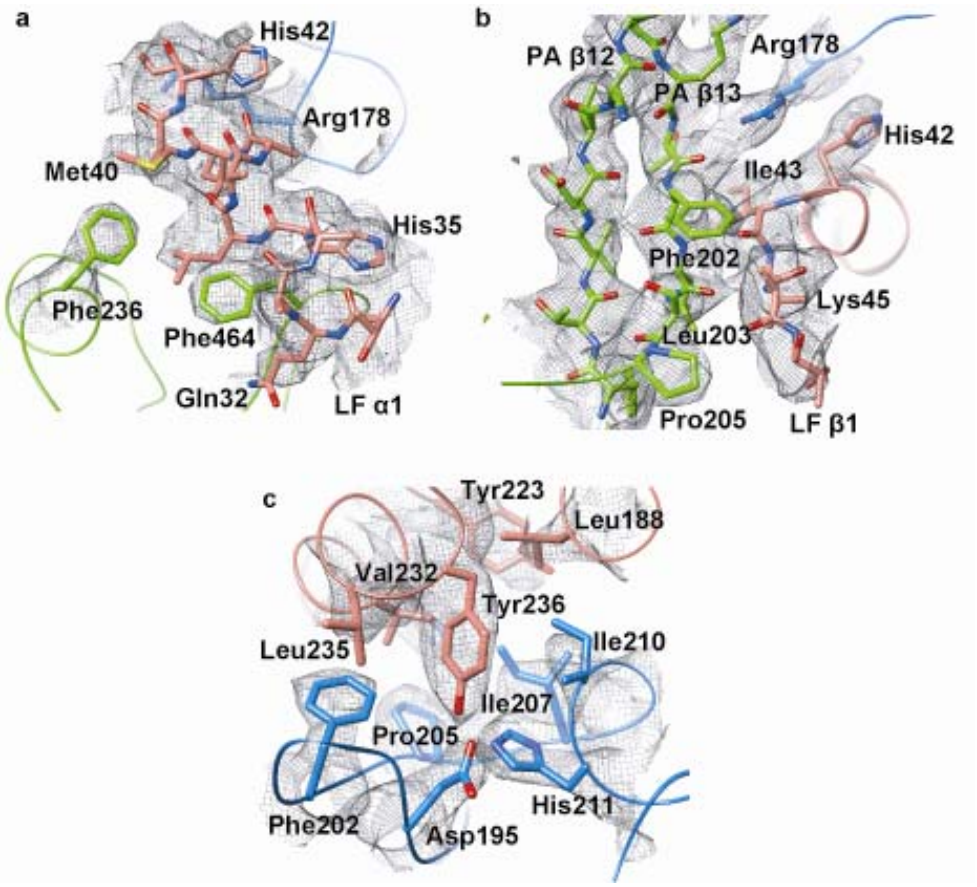

Figure 3.3. LF_N electron density in the $PA_8(LF_N)_4$ complex. A composite simulated-annealing (SA) omit map calculated in PHENIX (Adams *et al.*, 2004) to 3.1 Å contoured at σ=1 (gray mesh). The models of PA_N, PA_C and LF_N are rendered in green, blue, and red, respectively. Secondary structure elements and individual residues are labeled. Nitrogen, oxygen and sulfur atoms are colored blue, red and yellow, respectively. **(a)** LF_N α1 (residues 31-42) in complex with PA_N. Lysine and glutamate residues are truncated to Cβ for clarity. **(b)** LF_N β1 in complex with PA_N β12-13. LF_N Lys45 is truncated to Cβ for clarity. **(c)** LF_N's carboxy-terminal binding subsite interaction with PA_C. Additional stereo-pair images of LF_N omit maps following SA refinement are depicted in Figure 3.4.

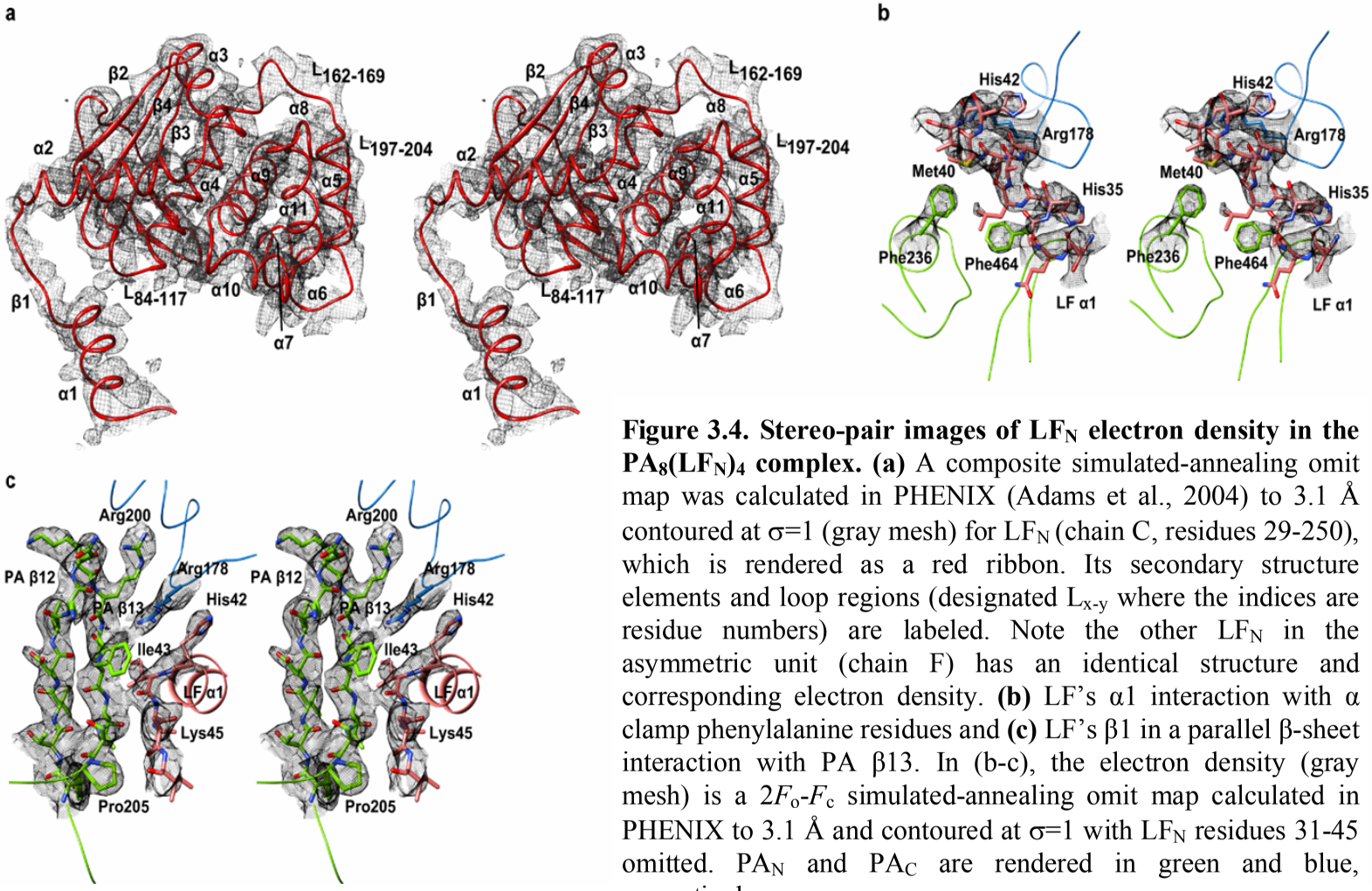

Figure 3.4. Stereo-pair images of LF_N electron density in the $PA_8(LF_N)_4$ complex. (a) A composite simulated-annealing omit map was calculated in PHENIX (Adams et al., 2004) to 3.1 Å contoured at σ=1 (gray mesh) for LF_N (chain C, residues 29-250), which is rendered as a red ribbon. Its secondary structure elements and loop regions (designated $L_{x\text{-}y}$ where the indices are residue numbers) are labeled. Note the other LF_N in the asymmetric unit (chain F) has an identical structure and corresponding electron density. **(b)** LF's α1 interaction with α clamp phenylalanine residues and **(c)** LF's β1 in a parallel β-sheet interaction with PA β13. In (b-c), the electron density (gray mesh) is a $2F_o$-F_c simulated-annealing omit map calculated in PHENIX to 3.1 Å and contoured at σ=1 with LF_N residues 31-45 omitted. PA_N and PA_C are rendered in green and blue, respectively.

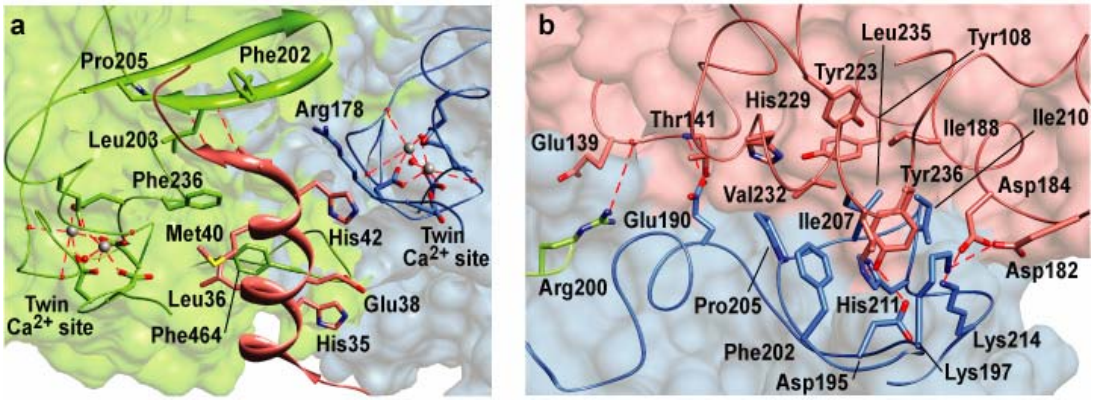

Figure 3.5. The PA octamer binds LF_N in two distinct subsites. Detailed views of **(a)** the α-clamp binding subsite and **(b)** the carboxy-terminal binding subsite are depicted. Highlighted non-covalent interactions are indicated with red dashed lines. Chains and Ca^{2+} ions are colored as in Figure 3.2a.

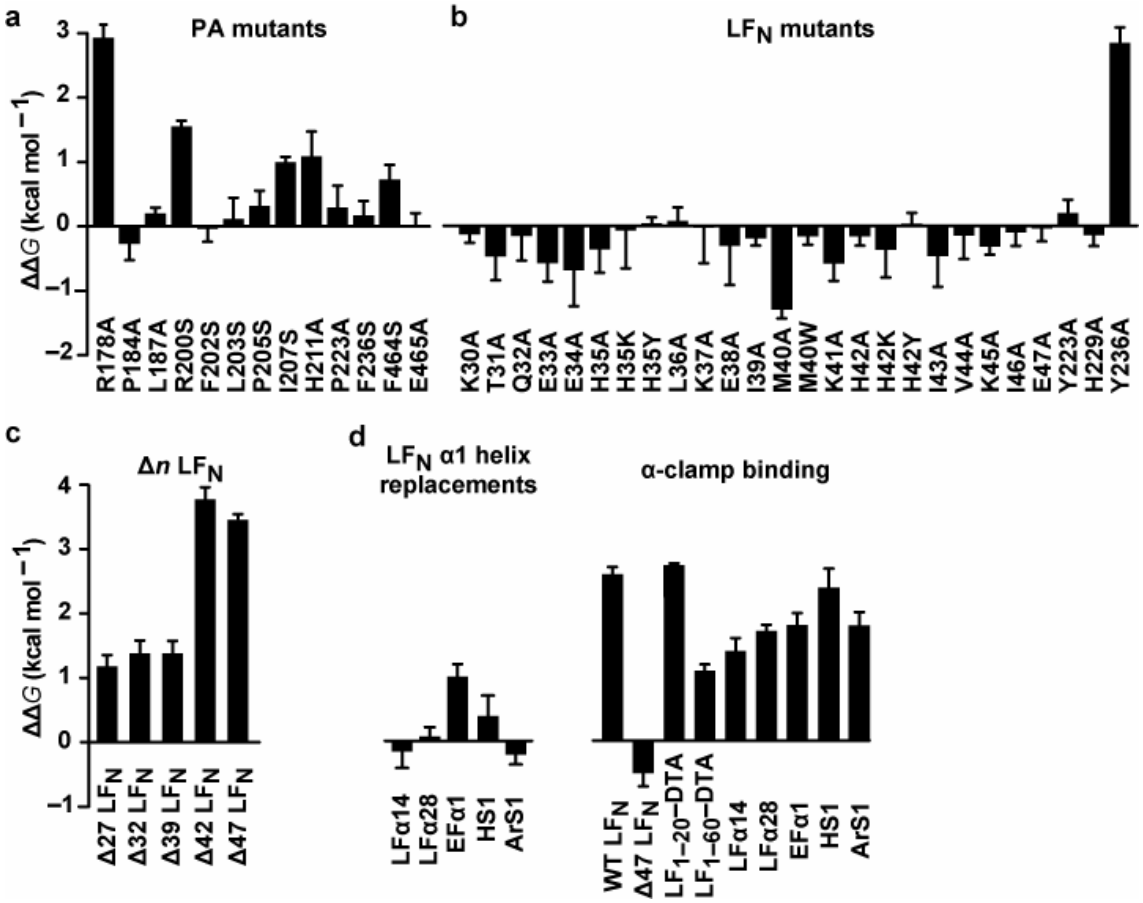

Figure 3.6. Changes in equilibrium binding free energy (ΔΔ*G*) for PA channel complexes, comparing **(a)** site-directed mutants of PA, **(b)** site-directed mutants of LF_N, and **(c)** Δ*n* LF_N amino-terminal truncation mutants. In (a-c), the reference state is WT LF_N:WT PA. **(d)** (Left) LF_N α1/β1-replacement mutant binding to WT PA; ΔΔ*G* values are referenced to WT LF_N. (Right) $LF_{1\text{-}20}$-DTA, $LF_{1\text{-}60}$-DTA, Δ47 LF_N and LF_N α1/β1-replacement mutant binding to PA R178A; ΔΔ*G* values are referenced to WT PA. LF_N α1/β1-replacement mutants either include multiple point mutations in the α1/β1 sequence (^{32}QEEHLKEIMKHIVK^{46}I) or replacements of the α1/β1 sequence with other sequences from LF or EF. The name, replacement sequence, and sequence identity (%) are listed for each: LFα14, SEEGRGLLKKLQI (23%); LFα28, NSKKFIDIFKEEG (23%); EFα1, EKEKFKDSINNLV (31%); hydrophilic sequence 1 (HS1), QEEHSKEISKHSVKS (73%); aromatic sequence 1 (ArS1), QEEHFKEIFKHFVKF (73%). See Figure 3.10 for alignments and helical-wheel depictions of the α1/β1-replacement sequences. In (a-d), $\Delta\Delta G = RT \ln K_d^{MUT} / K_d^{WT}$, where the equilibrium dissociation constants (K_d) were measured for the mutant (MUT) and WT proteins at pH 7.4, ΔΨ = 0 mV (Figure 3.9); *R* is the gas constant; and *T* is the temperature. The error bars are the mean ±s.d. (*n* = 2-6).

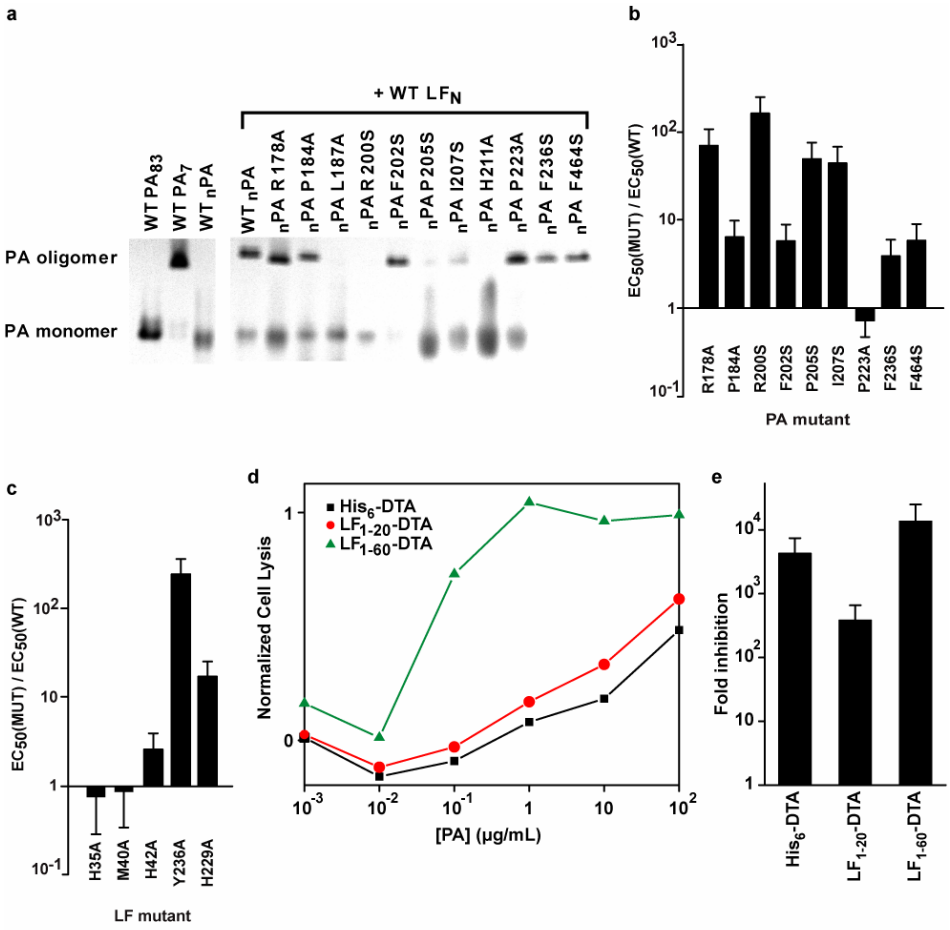

Figure 3.7. Probing the role of the PA-LF$_N$ interaction in toxin assembly and cytotoxicity. (a) Native PAGE analysis of LF$_N$-driven PA assembly. (Right) WT $_n$PA and site-directed $_n$PA mutants at 1 mg/ml were mixed with a stoichiometric equivalent of WT LF$_N$ and incubated for 3 hours at 25 °C, pH 8.0. Protein complexes were separated on a 4-20% gradient gel under native conditions and stained with coomassie blue. Individual lanes show the assembly reactions for either nicked WT PA or nicked PA ($_n$PA) point mutants and WT LF$_N$. The bands corresponding to assembled PA oligomer and unassembled full-length PA monomer (PA$_{83}$) are indicated based upon controls (shown on the Left). WT PA$_7$ refers to $_n$PA assembled over a Q-sepharose column (Blaustein *et al.*, 1989); the resulting oligomers are predominantly heptameric (Milne *et al.*, 1994, Kintzer et al., 2009). **(b)** LT cytotoxicity was monitored by an enzyme-coupled LDH release assay (Decker & Lohmann-Matthes, 1988) using immortalized bone marrow macrophage cells from 129 mice. Various PA and LF mutants were assayed in two types of experiments. Relative cytotoxicity of site-directed mutants of PA. Each PA mutant and WT PA were held at a constant concentration, and the concentration of WT LF was varied. Each bar represents the ratio of the effective-concentration-for-50% lysis (EC$_{50}$) when comparing the mutant PA to WT PA (MUT:WT). **(c)** Relative cytotoxicity of site-directed mutants of LF. WT PA was held at a constant concentration, and the concentration of WT LF or a LF mutant was varied. Each bar represents the ratio in EC$_{50}$ (MUT:WT) for the LF mutants. EC$_{50}$ values used to calculate the ratios in each panel were determined by curve fits to the normalized-cell-lysis curves as a function of the LF concentration. Errors in (a) and (b) are the mean ±s.d. (n = 3). **(d)** The substrate, His$_6$-DTA, LF$_{1-20}$-DTA, or LF$_{1-60}$-DTA, was added to the macrophage cell suspensions at 10 μg/ml, and the WT PA concentration was varied. Cell lysis was monitored by LDH release after 20-24 hours of exposure to the mixture of PA and the various DTA substrates. Normalized cell lysis curves are plotted for His$_6$-DTA (black ■), LF$_{1-20}$-DTA (red ●) and LF$_{1-60}$-DTA (green ▲). Plotted values are the mean of three trials. **(e)** Fold inhibition of DTA cytotoxicty was determined by measuring the individual EC$_{50}$s for each DTA construct in the presence of PA R178A or WT PA. Fold inhibition = EC$_{50}$(PA R178A) / EC$_{50}$(WT PA). The error bars are the mean ±s.d. (n = 3).

(Meador *et al.*, 1992, Meador *et al.*, 1993), PA's twin Ca^{2+}-binding sites scaffold the cleft and define its distinct shape and chemical character, including: (i) a delocalized anionic potential created by the excess of negatively-charged PA residues chelating the two Ca^{2+} ions and (ii) a large proportion of SASA contributed by PA backbone atoms. LF_N's side chains are not well-packed with side chains in the α-clamp cleft, in contrast to the carboxy-terminal binding subsite (Figure 3.5). Interestingly, PA contacts the side chains of LF Met40 and His35 through backbone interactions. PA_C Arg178 contacts the hydrophilic face of α1 at LF His42 while maintaining a hydrogen bond with the backbone carbonyl of PA_N Thr201. Aromatic residues, PA_N Phe236 and Phe464, and aliphatic residues, PA_N Leu187 and Leu203, line the cleft face opposite of PA_C Arg178. Upon binding LF_N, PA_N Phe202 repositions its phenyl group toward LF_N β1, shielding β1's backbone hydrogen bonds with PA_N Leu203. The chemical nature of the α-clamp cleft suggests that it is well-suited to bind an unfolded β strand and an amphipathic helix with a positively-charged face.

Both LF-binding subsites are critical for cytotoxicity activity. We initially characterized the PA-LF binding interaction using cytotoxicity assays. Site-directed mutagenesis studies on PA and LF residues involved in either binding subsite reveal defects in LT-induced macrophage cytolysis (Figure 3.7b,c). To further address the interaction between LF_N's α1/β1 sequence and the α clamp, we created fusions of the first 20 or 60 residues of LF and the A fragment from diphtheria toxin (DTA), called LF_{1-20}-DTA and LF_{1-60}-DTA, respectively. When co-administered with PA, we find LF_{1-60}-DTA is 100-fold more cytotoxic than LF_{1-20}-DTA or hexahistadine-tagged DTA (His_6-DTA, DTA with an amino-terminal, 18-residue leader containing the hexahistidine sequence, Figure 3.7d). Interestingly, despite lacking the α1/β1 sequence, His_6-DTA (Blanke *et al.*, 1996) and LF_{1-20}-DTA are cytotoxic when co-treated with WT PA (Figure 3.7d); however, all of these DTA constructs are ~1000-fold less cytotoxic when co-treated with the α-clamp mutant, PA R178A (Figure 3.7e). Thus the α clamp has broad substrate specificity. However, the role of the α1/β1-α-clamp interaction in toxin function is difficult to surmise from cytotoxicity assays alone, since toxin uptake involves multiple steps (e.g., PA assembly, LF binding, unfolding and translocation).

The role of the α clamp in LT assembly. To determine the role of the α clamp in LT assembly, we performed multiple in vitro PA-LF_N assembly assays. By native PAGE, we find that PA mutations introduced into the LF_N-PA-binding interface disrupt PA co-assembly with LF_N (Figure 3.7a). To focus on the role of LF_N α1/β1 in PA assembly, we labeled PA K563C with two different fluorescent probes. A 1:1 ratiometric mixture of these labeled $_n$PA K563C constructs ($_n$PA*) produces an increase in fluorescence resonance energy transfer (FRET) upon assembly with LF_N (Christensen *et al.*, 2006). Using this FRET assay, we find that 5-fold more $_n$PA* assembles with WT LF_N than with the Δ47 LF_N amino-terminal truncation (which lacks both α1 and β1, Figure 3.8a). The circular dichroism (CD) spectra of Δ47 and WT LF_N are comparable, demonstrating that the assembly defect is not due to the misfolding of Δ47 LF_N (Figure 3.8b). Using electron microscopy (EM), native PAGE, and mass spectrometry, we find that the percentage of octameric PA oligomers is greatly reduced for Δ47 LF_N relative to WT LF_N (Figure 3.8c-e). By EM, we estimate that ~3% of the PA oligomers produced with Δ47 LF_N are octameric (10-fold less than that observed with WT LF_N, Figure 3.8d). Thus LF_N's α1 and β1 structures not only drive PA oligomerization, but also they are critical to the mechanism of PA octamer formation (Figure 3.2c).

Mapping the LF_N-binding interaction with the PA channel. Using electrophysiology, we measure LF_N binding by observing kinetic and equilibrium changes in channel conductance

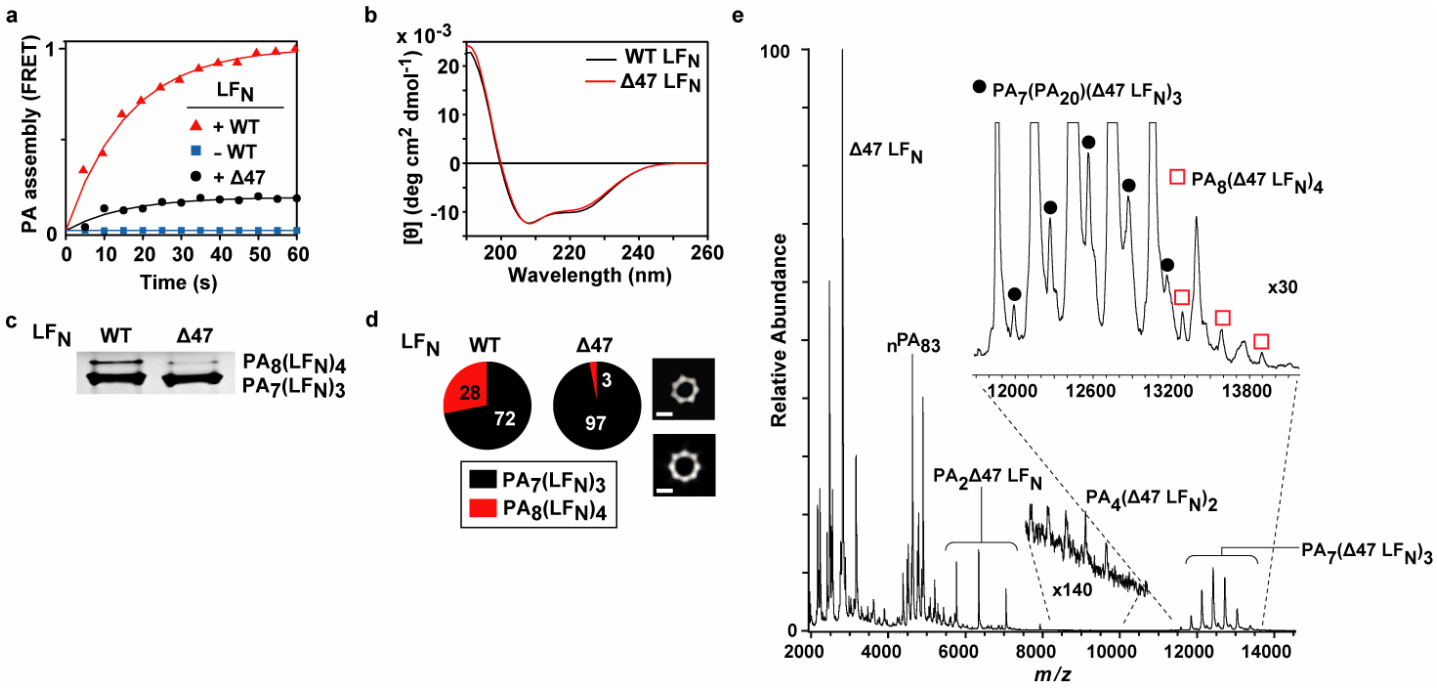

Figure 3.8. Role of LF_N α1 and β1 in PA assembly. (a) PA assembly was monitored by fluorescence resonance energy transfer (FRET) using a mixture of fluorescently-labeled $_nPA$ ($_nPA$*). Shown are FRET-probed assembly records for $_nPA$* in the presence of WT LF_N (red ▲) and Δ47 LF_N (black ●) as well as a control record for assembly in the absence of LF_N (blue ■) at room temperature, pH 7.4. **(b)** Circular dichroism spectra (average of six scans, 2 nm bandwidth, 0.1 cm pathlength) of WT (black) and Δ47 LF_N (red) in 20 mM sodium phosphate, pH 7.5. **(c)** Native PAGE and **(d)** EM analysis of the oligomeric distribution of PA_7 and PA_8 comzplexes produced after one hour of $_nPA$ co-assembly with either WT LF_N or Δ47 LF_N at room temperature, pH 8. In the assembly reaction, 1 mg/ml WT $_nPA$ was incubated with either WT LF_N or Δ47 LF_N added at 1:1 molar stoichiometry (with respect to $_nPA$). (Left) Pie charts report the distribution for WT-LF_N-assembled and Δ47-LF_N-assembled PA observed by EM. (Right) Representative class-average EM images of the PA_7 and PA_8 complexes are shown with a 5-nm scale bar (white). **(e)** WT $_nPA$ (1 mg/ml) was mixed with Δ47 LF_N at a 1:1 stoichiometric ratio (Δ47 LF_N:PA monomer) in ammonium acetate buffer at room temperature, pH 7.8. The products of the assembly reaction were analyzed using nanoESI-MS. The mass spectrum revealed the largest relative abundances for Δ47 LF_N and unassembled WT $_nPA$. The PA_7(Δ47 LF_N)$_3$ complex represented >90% of the observed oligomers. Low relative abundances were observed for PA_8(Δ47 LF_N)$_4$ complexes (red □) and other dimeric and tetrameric PA intermediates. Of note, an aberrant oligomeric state, containing $PA_7(PA_{20})$(Δ47 LF_N)$_3$ (black ●), was also observed at low relative abundance, albeit at higher levels than octameric PA_8(Δ47 LF_N)$_4$ complexes. The aberrant complex contains the PA_{20} moiety, and it can be modeled free of steric hindrance, allowing 3 LF_N and 1 PA_{20} to simultaneously bind to the PA heptamer.

(Krantz et al., 2005) (Figure 3.9a-c); i.e., when LF_N binds to the PA channel, it inserts its amino-terminal end into the channel and blocks conductance. We monitor binding in the absence of an applied $\Delta\Psi$ to eliminate its influence on the channel-substrate interaction. Since PA_7 and PA_8 have similar translocation (Kintzer et al., 2009) and cell cytotoxicity (Kintzer et al., 2010) activities, we use the PA_7 oligomer to maintain consistency with prior reports. (Krantz et al., 2006, Krantz et al., 2005, Thoren et al., 2009) To determine the overall thermodynamic contribution of LF_N α1/β1, we made a series of additional Δ*n* LF_N amino-terminal truncations (where *n* is the number of deleted residues). These Δ*n* LF_N do not block PA channel conductance, as they lack sufficient unfolded/unstructured sequence on their amino termini. We use a competition assay to measure Δ*n* LF_N binding: first we block PA channel conductance with WT LF_N (~100 pM); then we add the competitor Δ*n* LF_N and monitor the restoration of the conductance (Figure 3.9d,e). We find that Δ42 and Δ47 LF_N reduce WT PA-channel-binding affinity by 3.6-3.8 kcal mol^{-1} relative to WT LF_N (Figure 3.6c). However, since Δ27, Δ32, and Δ39 LF_N destabilize the complex by about 1.2-1.4 kcal mol^{-1}, the α1/β1 interaction is worth ~2.5 kcal mol^{-1}. We assume that downstream interactions within the channel provide the additional ~1 kcal mol^{-1} of stabilization. We conclude that LF_N α1/β1 binds to the PA channel and provides substantial stabilization of the PA-LF_N complex.

To investigate the details of the interaction between the PA channel and LF_N, we engineered point mutations into residues localized in either LF_N binding subsite and estimated their relative energetic contribution to channel binding (Figure 3.6a,b). Several mutations localized in the carboxy-terminal binding subsite, PA R200S, I207S, and H211A, disrupt LF_N binding by 1-1.5 kcal mol^{-1}. These residues form two binding "hotspots", i.e., locations where point mutations disrupt binding most severely (Clackson & Wells, 1995). By contrast, the mutations, F202S and P205S, located between these two carboxy-terminal-site hotspots have minimal effects on LF_N binding, reflecting that LF_N's carboxy terminus does not make substantial contact with these residues (Figure 3.5b). The LF_N Y236A mutant most appreciably perturbs PA-channel binding and represents the LF_N hotspot in the carboxy-terminal subsite interaction. Other adjacent LF_N residues in the carboxy-terminal subsite interaction have minimal effects on PA channel binding.

We then investigated the relative energetic contribution of residues localized in the α-clamp binding subsite (Figure 3.6a,b). We find that PA Arg178 comprises the major hotspot site in PA's α clamp, where the R178A mutation destabilizes the complex by 2.9 kcal mol^{-1}. While the aromatic PA mutant, F464S, destabilizes LF_N binding at the α-clamp site by 0.7 kcal mol^{-1}, the PA F236S mutant does not. Additionally, we find that none of 23 point mutations introduced into LF_N α1 and β1 destabilizes the LF_N-PA channel complex. Interestingly, the mutation, LF_N M40A, stabilizes the complex 1.3 kcal mol^{-1} (Figure 3.6b). These results indicate contrasting binding energetic behaviors for the two different LF_N-binding subsites. At the carboxy-terminal subsite, a classical interface is observed, where specific LF_N and PA side chains comprise the respective hotspots on either interface. At the α-clamp subsite, while we identify PA Arg178 as a major hotspot residue, no clear hotspot can be identified on LF_N α1/β1. These observations suggest that the stabilizing interactions in the α-clamp subsite do not involve specific LF_N side chains, but rather the ~2.5 kcal mol^{-1} of binding stabilization is due to the formation of nonspecific contacts and the more general exclusion of SASA.

The PA α clamp possesses nonspecific binding activity. The robustness of the binding interaction is intriguing given the paucity of specific α-clamp interactions. To test the specificity of the α-clamp interaction, we either replaced the entire LF_N α1/β1 sequence with other non-

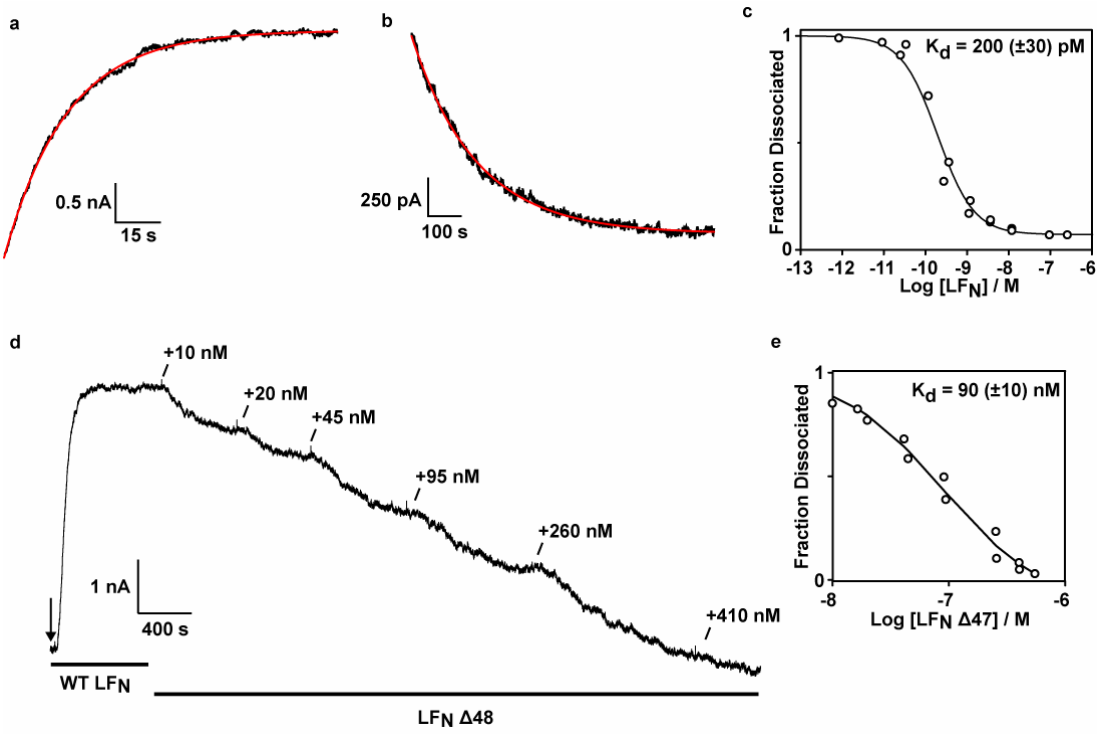

Figure 3.9. Equilibrium and kinetic measurements of LF_N binding to the PA channel. Channel binding was measured by recording equilibrium and kinetic changes in current, I, in the absence of an applied potential. Representative **(a)** association and **(b)** dissociation kinetics for WT LF_N binding to WT PA channels at pH 7.4. Association kinetics were measured at 3 nM WT LF_N. Dissociation was initiated by perfusing the cis chamber and/or adding 1 μM Δ47 LF_N competitor. Each current versus time, t, record (black line) was fit (red line) to a single exponential, $I = I_o\exp(-kt) + c$, to obtain the rate constant, k. I_o is the amplitude, and c is an offset constant. In the depicted example in (a), the association rate constant, k_a, is determined by $k_a = k/[L]$, and the k_a is 1.63×10^7 ($\pm3\times10^5$) M^{-1} s^{-1}. From (b), the dissociation rate constant, k_d, is 0.00518 ($\pm1\times10^{-5}$) s^{-1}. **(c)** Equilibrium binding currents as a function of WT LF_N ligand concentration, $[L]$, fit to a single-site binding model, $I = I_o/(1 + K_d/[L])$. The equilibrium dissociation constant values, K_d, determined in kinetic (318 pM) and equilibrium experiments (200 pM) are in good agreement. **(d)** Depicted is an example of equilibrium-binding-competition experiment used to determine the apparent K_d of ligands that bind PA but cannot block its conductance. After PA-channel insertion stabilized, the pH was adjusted to 7.4. Then 0.8 nM WT LF_N was added (indicated by the arrow). Once the WT-LF_N-conductance block stabilized, small increments of Δ47 LF_N were added. As Δ47 LF_N competes for WT LF_N sites, the channels reopen. Increments of added Δ47 LF_N are indicated. **(e)** The degree of reopening as a function of total Δ47 LF_N concentration was used to measure Δ47 LF_N's K_d, which is 90 (±10) nM.

homologous sequences from LF and EF or introduced multiple mutations into α1/β1 (Figure 3.10). Interestingly, we find that these LF_N α1/β1 replacements bind with similar affinities as WT LF_N (differing by 0.2 to 1.0 kcal mol^{-1}, Figure 3.6d). Furthermore, multisite LF_N mutants in which the buried hydrophobic face of α1/β1 is replaced with either four Ser residues (LF_N HS1) or four Phe residues (LF_N Ar1) bind PA with similar affinity as WT LF_N (Figure 3.6d), indicating that the α clamp also binds non-amphipathic helices. Finally, we find that these LF_N α1/β1-replacement constructs bind 1.3-2.4 kcal mol^{-1} less tightly to PA R178A relative to WT PA (Figure 3.6d), thereby confirming that this nonspecific-binding activity is localized to the α-clamp subsite. Thus the α clamp binds a broad array of sequences, providing 1.5-4 kcal mol^{-1} of stabilization (depending upon the identity of the α1/β1 sequence).

LF_N must unfold to bind the α-clamp subsite. Our crystal structure and thermodynamic binding data indicate that the α-clamp subsite binds nonspecifically to unfolded protein substrates. This model is well supported by several additional lines of evidence. First, the thermodynamic comparison of WT LF_N and the truncated Δ*n* LF_N mutants is appropriate because these mutants have similar folded secondary structure content as WT LF_N (Figure 3.8b). Moreover, the Δ47 LF_N construct binds similarly to PA R178A as WT PA (Figure 3.6d), confirming that the Δ47 LF_N truncation does not bind at the α-clamp site, as implied by the structure (Figure 3.2a). Second, fusions of LF's amino terminus and DTA (LF_{1-60}-DTA and LF_{1-20}-DTA) are sufficient to bind to the α-clamp site, since their affinity for the PA channel is disrupted by the PA R178A mutation (Figure 3.6d and Figure 3.11). This result indicates that the α clamp is an independent binding site capable of binding to unstructured sequences at the amino-terminus of a substrate. Third, knowing that LF_N α1/β1 unfolds upon binding PA (Figure 3.12a), we engineered the double mutant, LF_N I39C E72C ($LF_N^{C39\text{-}C72}$, which forms a disulfide bond that prevents α1/β1 unfolding). Interestingly, $LF_N^{C39\text{-}C72}$ has 10^4-fold reduced affinity for PA channels under non-reducing conditions (Figure 3.12b); however, under reducing conditions (in the presence of dithiothreitol, DTT), $LF_N^{C39\text{-}C72}$ binds with the same affinity as WT LF_N (Figure 3.12b). We also kinetically observe a DTT-dependent $LF_N^{C39\text{-}C72}$ blockade of PA channels (Figure 3.13a). Therefore, LF_N must unfold α1 and β1 to properly bind the α clamp and interact stably with PA oligomers.

Binding to PA induces strain and disorder into LF_N. We then asked how the unfolding of LF_N α1/β1 on the surface of PA affects the remaining folded structure of LF_N. First, we measured the stability of the Δ*n* LF_N mutants using chemical denaturant titrations probed by CD at 222 nm (CD_{222}). The Δ*n* mutants' stabilities are estimated by fitting the CD_{222}-probed titration data to a four-state equilibrium unfolding model ($N \leftrightarrow I \leftrightarrow J \leftrightarrow U$) (Krantz et al., 2004) (Figure 3.13b and Table 3.2). We find the truncation mutants possess native (*N*), intermediate (*I* and *J*), and unfolded (*U*) states. The truncations, however, destabilize the *N* state by ~1.2 kcal mol^{-1}, where the deletion of the α1 helix is more destabilizing than the deletion of the β1 strand (Figure 3.12c). Second, we compared the crystallographic atomic displacement parameters (*B* factors) of bound LF_N with free LF_N (1J7N (Pannifer et al., 2001)). In this analysis, we calculate the relative change in normalized *B* factor (ΔB_{norm}) for each LF_N residue upon binding PA (Figure 3.12d). The β2-β4 sheet and surrounding helices increase in B_{norm} upon binding PA, whereas α1/β1 decrease in B_{norm} (Figure 3.12d). To corroborate these ΔB_{norm} values, we measure changes in backbone and side chain mobility using fluorescence anisotropy (FA). LF_N mutants with unique Cys substitutions were labeled with thiol-reactive fluorescent probes. Upon binding WT PA_7 oligomers, the fluorescent probes attached to LF_N's α1/β1 structures show gains in normalized relative FA (FA_{norm}), and conversely, probes in the β2-β4 sheet show losses in FA_{norm}

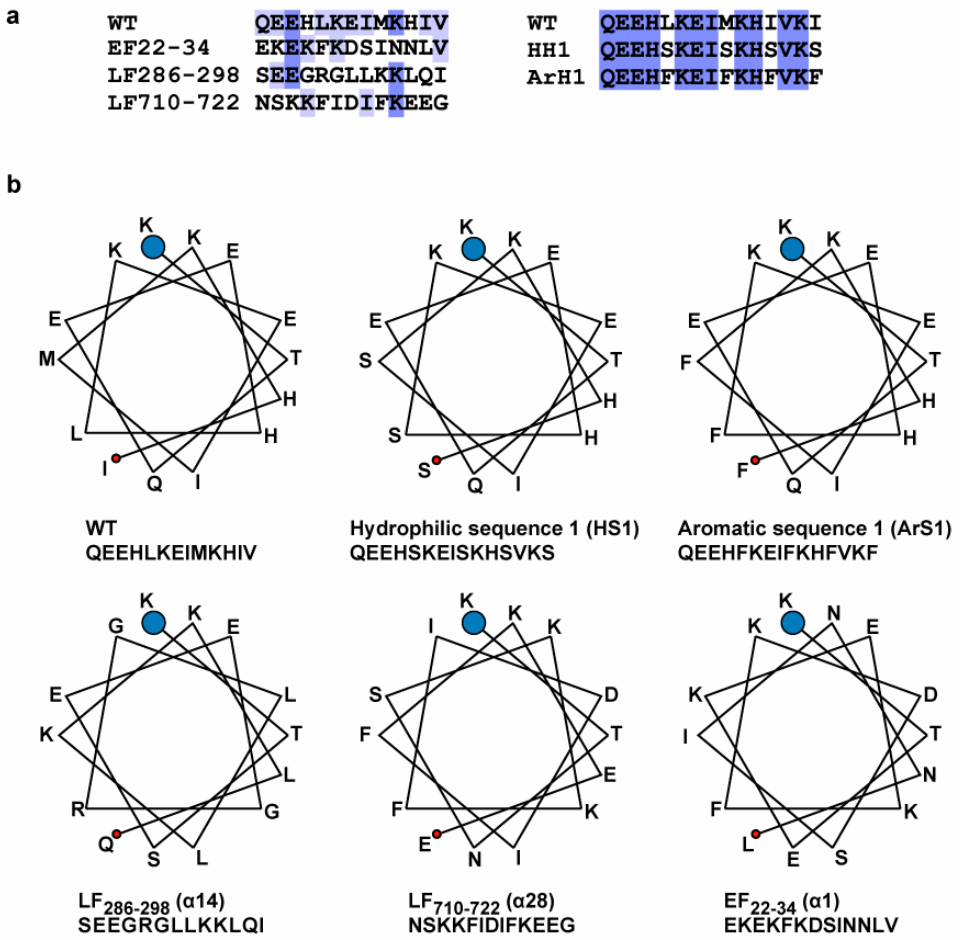

Figure 3.10. Sequence-alignment and helical-wheel analysis of α1/β1-replacement sequences. (a) A multiple sequence alignment of the 13- or 15-residue sequences engineered into the α1 position of LF_N (WT). Residues are colored according to the BELVU convention (Sonnhammer & Hollich, 2005): highly-conserved (dark blue), conserved (light blue), and non-conserved (white). The CLUSTALW (Thompson *et al.*, 1994) pairwise-percent-identity scores (relative to WT) are: HS1 (73%), ArS1 (73%), $LF_{286-298}$ α14 (23%), $LF_{710-722}$ α28 (23%), and EF_{22-34} α1 (31%). **(b)** Helical wheels are depicted, where the amino-terminal end (large blue circle at front of page) and carboxy-terminal end (small red circle at the back of the page) are indicated.

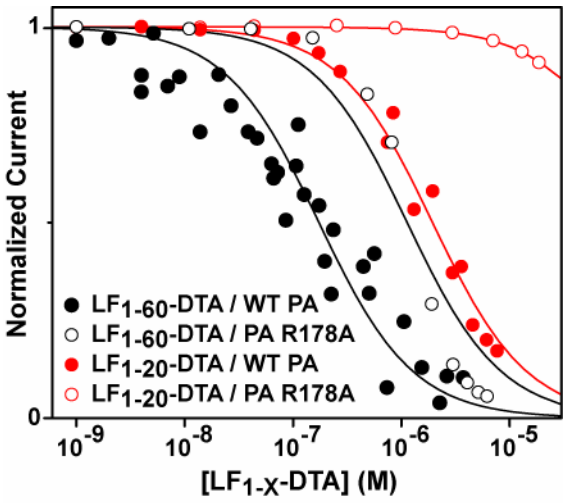

Figure 3.11. The first 20 or 60 amino-terminal residues of LF_N are sufficient to bind the PA channel at the α-clamp site. Binding curves for $LF_{1\text{-}60}$-DTA (black) and $LF_{1\text{-}20}$-DTA (red) using either WT PA channels (●) or PA R178A channels (○). In these experiments, PA channels are inserted into a planar lipid bilayer bathed in asymmetric KCl solutions buffered in 10 mM potassium phosphate, pH 7.4. An additional 0.1 M equivalent of potassium chloride was added only to the cis solution. Aliquots of $LF_{1\text{-}60}$-DTA or $LF_{1\text{-}20}$-DTA were added and the final equilibrium currents were recorded at a $\Delta\Psi$ of 0 mV. Equilibrium curves were fit to a one-binding-site model to calculate the K_d: $LF_{1\text{-}60}$-DTA with WT PA, K_d = 0.16 (±0.02) μM; $LF_{1\text{-}20}$-DTA with WT PA, K_d = 1.9 (±0.1) μM; $LF_{1\text{-}60}$-DTA with PA R178A, K_d = 1.1 (±0.2) μM; and $LF_{1\text{-}20}$-DTA with PA R178A, K_d = 190 (±4) μM.

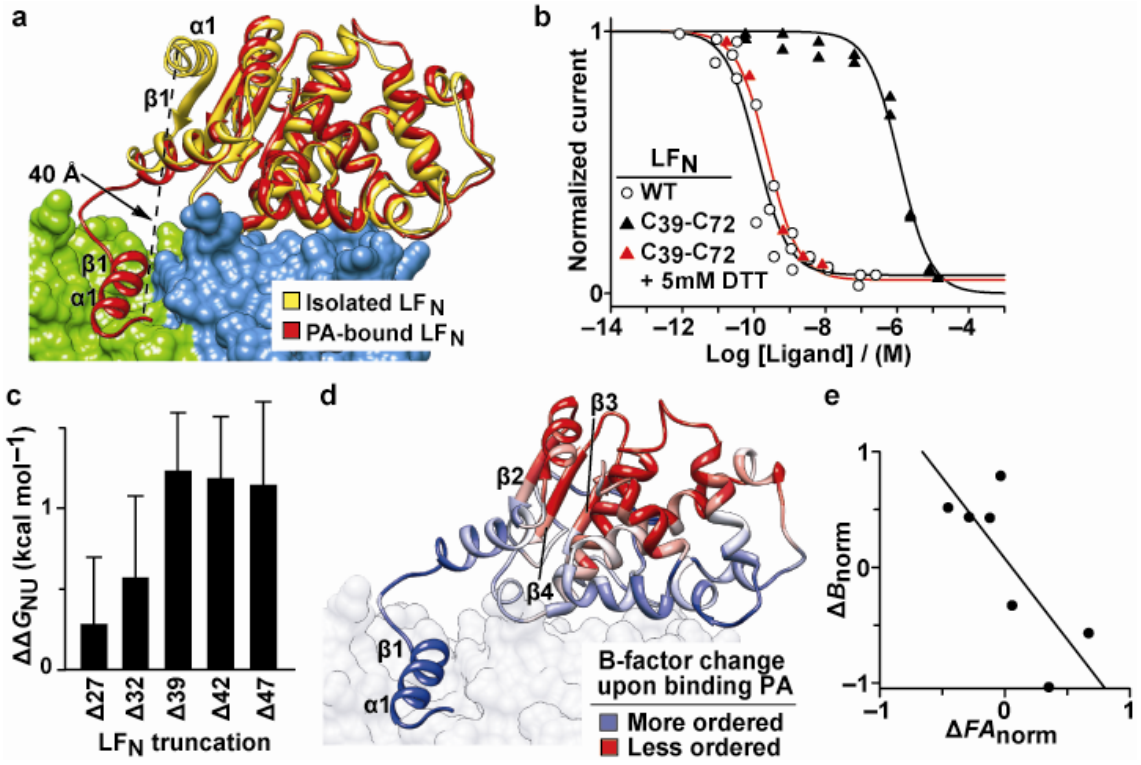

Figure 3.12. Dynamics and thermodynamics of the pre-translocation unfolding of LF_N. (a) Rendering of LF_N's unfolding transition on the surface of the PA_NPA_C dimer (green and blue, respectively). Free LF_N (gold) (PDB 1J7N (Pannifer *et al.*, 2001)) is Cα-aligned to the LF_N in the $PA_8(LF_N)_4$ complex (red). **(b)** $LF_N^{C39\text{-}C72}$ binding to WT PA channels (pH 7.4, 0 mV) in the presence of 5 mM DTT (red ▲) and in the absence of DTT (black ▲). A WT LF_N binding curve (○) is also shown. Normalized equilibrium currents were fit to single-site binding model to obtain K_d values: WT LF_N, K_d = 120 (±30) pM; $LF_N^{C39\text{-}C72}$, K_d = 1.2 (±0.1) μM; and $LF_N^{C39\text{-}C72}$ + 5 mM DTT, K_d = 240 (±60) pM. **(c)** Equilibrium stability measurements (pH 7.5, 20 °C) of amino-terminal deletions of LF_N (Δ*n* LF_N). Equilibrium free energy differences ($\Delta\Delta G_{NU}$) were obtained from denaturant titration data fit to a four-state equilibrium unfolding model (Krantz et al., 2004) (Figure 3.13b), where $\Delta\Delta G_{NU} = \Delta G_{NU}(\Delta n) - \Delta G_{NU}(WT)$. Error bars are the mean ±s.d. (*n* = 3-4). Fit parameters are listed in Table 3.2. **(d)** Residues in LF_N are colored by their differences in normalized *B* factor (ΔB_{norm}), which is obtained by comparing the model of free LF_N (1J7N, structure 1) and LF_N in complex with PA (structure 2) using $\Delta B_{norm} = B_{1,i} / \langle B_1 \rangle - B_{2,i} / \langle B_2 \rangle$. The $\langle B \rangle$ is the average *B* factor for the entire chain. **(e)** ΔB_{norm} is plotted against the normalized fluorescence anisotropy (FA) change (ΔFA_{norm}) for 7 different site-specifically-labeled residues (37, 48, 72, 126, 164, 199, and 242) in LF_N. $\Delta FA_{norm} = FA_{1,i} / \langle FA_1 \rangle - FA_{2,i} / \langle FA_2 \rangle$, where free LF_N and the LF_N-PA oligomer complex are state 1 and state 2, respectively. The linear fit is significant (p = 0.04). Raw anisotropy changes upon binding the PA oligomer for these labeled LF_N are shown in Figure 3.14.

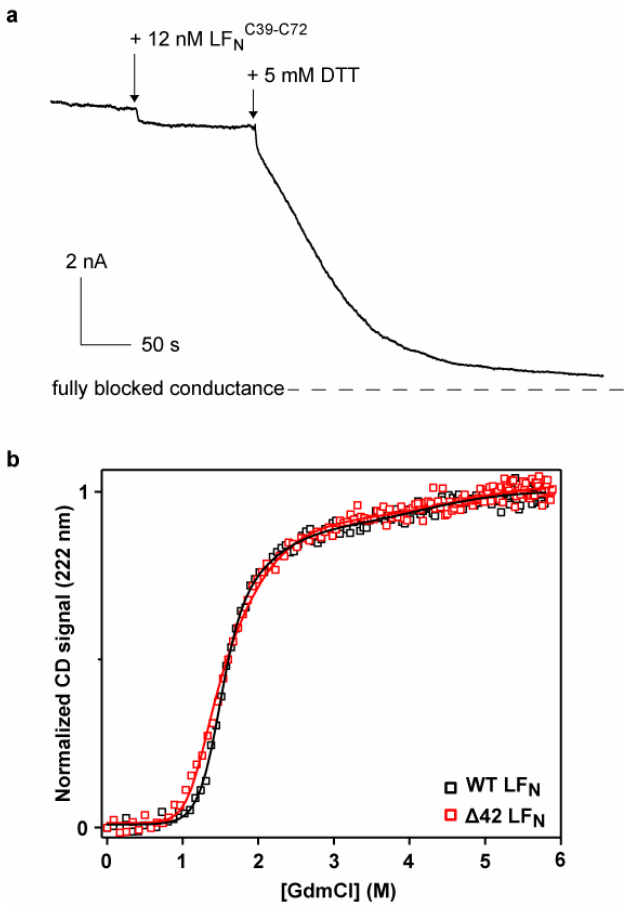

Figure 3.13. LF_N unfolding increases its affinity for PA but reduces its stability. (a) A planar bilayer recording at a ΔΨ of 0 mV of PA channels bathed in 10 mM potassium phosphate at pH 7.4. To generate a current, the salt concentration was asymmetric: 100 mM potassium chloride was added to the cis side, and 0 mM potassium chloride was added to the trans side of the bilayer. The recording begins after the PA current stabilized. Then 12 nM $LF_N^{C39-C72}$ was added to the cis side of the membrane (the side to which PA channels were inserted), a small decrease in PA current was established. Once this minor $LF_N^{C39-C72}$ block stabilized, 5 mM DTT was added to the cis side of the membrane. A rapid and nearly complete block of the PA current was subsequently observed. **(b)** Equilibrium denaturant titration profiles of WT (black squares) and Δ42 LF_N (red squares) probed by CD at 222 nm. The buffer is 20 mM sodium phosphate, 0.75 M trimethylamine *N*-oxide, pH 7.5. The normalized titration data are fit (solid line) to a four-state model ($N \leftrightarrow I \leftrightarrow J \leftrightarrow U$), which has been described elsewhere (Krantz et al., 2004). See Table 3.2 for a listing of all curve fit parameters.

Table 3.2. Thermodynamic stability free-energy parameters for Δ*n* LF_N truncations

LF_N	ΔG_{NU}	ΔG_{IU}	ΔG_{JU}	$\Delta\Delta G_{NU}$
WT	-12.6 (±0.4)	-7.2 (±0.4)	-4.7 (±0.3)	0
Δ27	-12.3 (±0.2)	-6.7 (±0.3)	-4.4 (±0.3)	0.3 (±0.4)
Δ32	-12.0 (±0.4)	-6.7 (±0.2)	-4.9 (±0.5)	0.6 (±0.5)
Δ39	-11.33 (±0.08)	-6.6 (±0.3)	-4.6 (±0.2)	1.2 (±0.4)
Δ42	-11.4 (±0.2)	-6.81 (±0.04)	-4.80 (±0.02)	1.2 (±0.4)
Δ47	-11.4 (±0.4)	-6.7 (±0.3)	-4.6 (±0.1)	1.2 (±0.5)

The four-state equilibrium free energy parameters, ΔG_{NU}, ΔG_{IU}, ΔG_{JU}, in kcal mol^{-1} are obtained from fitting equilibrium denaturant titration profiles probed by CD. The buffer and conditions are 20 mM sodium phosphate, 0.75 M trimethylamine *N*-oxide, pH 7.5, and 20 °C. The guanidinium chloride denaturant sensitivities (*m* values) defining each thermodynamic transition between states are fixed to constant values (where m_{NI}, m_{IJ}, m_{JU} are 3.59, 1.37, and 1.05 kcal mol^{-1} M^{-1}, respectively) consistent with previous values (Krantz *et al.*, 2004). The four-state fit model ($N \leftrightarrow I \leftrightarrow J \leftrightarrow U$) has been described elsewhere (Krantz et al., 2004). Example curve fits are shown in Figure 3.13b. The equilibrium free energy differences ($\Delta\Delta G_{NU}$) are computed as $\Delta\Delta G_{NU} = \Delta G_{NU}(\text{MUT}) - \Delta G_{NU}(\text{WT})$.

(Figure 3.14a). Overall, these ΔFA_{norm} values inversely correlate with ΔB_{norm} values (p value of 0.04, Figure 3.12e), confirming that the more dynamic regions in the crystal are also dynamic in solution. Therefore, we conclude that the ~2.5 kcal mol^{-1} of stabilization gained when α1/β1 binds to the α-clamp site not only offsets the ~1.2 kcal mol^{-1} of thermodynamic destabilization imparted by the unfolding of α1/β1 but also accounts for the observed entropic increases in strain and disorder throughout LF_N's remaining folded structure.

The role of the α clamp in protein translocation. To determine the role of the α clamp during protein translocation, we use planar lipid bilayer electrophysiology, which records changes in PA conductance as substrate-blocked channels translocate their substrates and reopen (Thoren et al., 2009, Krantz et al., 2005, Krantz et al., 2006). We examined 37 point mutations in PA and LF_N. Of the 13 PA mutants tested, we find that the α-clamp mutant, PA F202S, slows LF_N translocation 20-fold, or 1.7 kcal mol^{-1} (Figure 3.15a). A subset of the LF_N point mutations (H35A, M40A, and H42A), which point toward either face of the α-clamp cleft (Figure 3.5a), inhibit translocation 0.8-1.7 kcal mol^{-1} (Figure 3.15a). These translocation defects are observed for both PA_7 and PA_8 channels (Figure 3.16a). Conversely, other buried α1 sites (LF_N Leu36, Ile39, and Ile43) are tolerant to substitution and do not affect protein translocation (Figure 3.15a). Interestingly, we find that the observed positional translocation defects are restored when a bulky group is placed at position 40 (Thoren et al., 2009) and a positively-charged residue is placed at positions 35 and 42 (Figure 3.15a). All of the LF_N α1/β1 replacements translocate similarly to WT LF_N (Figure 3.15a). We conclude, therefore, that efficient LF_N unfolding and translocation are catalyzed by the aromatic α-clamp residue (PA Phe202); however, the LF_N α1/β1 sequence itself has rather minimal charge and steric requirements.

The broad substrate specificity of the α clamp led us to ask which PA residues facilitate translocation of full-length LF, a more complex, multidomain substrate. LF has a different rate-limiting step than LF_N and requires a greater driving force (Krantz et al., 2006); therefore, we measure its translocation kinetics under a ΔpH and ΔΨ. We find the PA α-clamp mutants, F202S and P205S, reduce LF translocation efficiency, ε, by ~60% (where $\varepsilon = A_{obs}/A_{exp}$, A_{exp} and A_{obs} are the expected and observed amplitudes, respectively, Figure 3.17). The PA mutants F236S and F202S inhibit the rate of LF translocation (Figure 3.15b). Interestingly, these PA mutants do not appreciably affect LF_N binding (Figure 3.6a), and only PA F202S inhibits LF_N translocation (Figure 3.15a). Finally, we find PA R178A is defective in LF_N binding but not defective in translocation. We conclude that hydrophobic and aromatic residues surrounding the α clamp (Figure 3.15c) catalyze the translocation of LF.

3.4 Discussion

Some models (Krantz et al., 2005, Zimmer et al., 2008) propose that nonspecific clamping sites are critical features of unfolding machines. In general, unfoldases are thought to denature proteins by applying mechanical forces (Thoren et al., 2009) and transiently trapping partially unfolded conformations in nonspecific binding sites (Krantz et al., 2005). Unfolded protein, however, is inherently more complex than folded protein, especially in terms of its configurational flexibility and combinatorial chemical complexity. Therefore, a translocase channel would have to accommodate an ever-changing array of possible chemistries and configurations as the unfolded chain is translocated. An elegant solution to this problem may be that unfolded sequences adopt a more rigid and uniform α-helical or β-strand conformation upon binding to an unfoldase, as we observe in the PA-LF_N complex (Figure 3.5a). Indeed we find that PA's α clamp can bind to a broad array of amino acid sequences (Figure 3.6d). This nonspecific

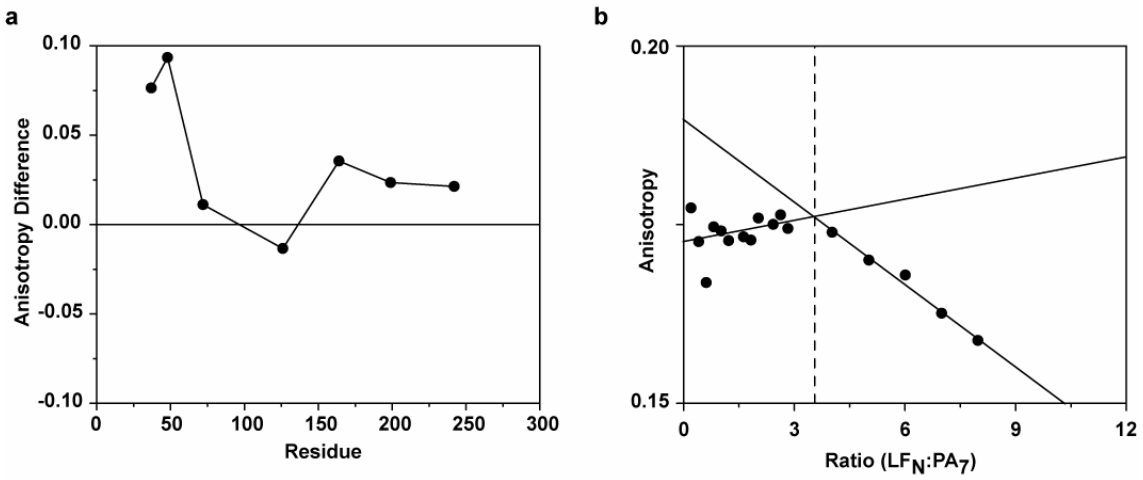

Figure 3.14. Fluorescence anisotropy changes upon PA-LF$_N$ complex formation. (a) Seven site-directed Cys mutants in LF$_N$ were labeled with EDANS (designated LF$_N$*.) The FA of each free LF$_N$* (30 nM) and each LF$_N$* in complex with PA$_7$ oligomer (30 nM each) at pH 7.5 were recorded. The plotted difference in anisotropy (ΔFA) was calculated according to $\Delta FA = FA$(LF$_N$*:PA$_7$ complex) – FA(LF$_N$*). **(b)** The FA signal for PA-LF$_N$* complex formation is saturable at the appropriate stoichiometry. PA$_7$ oligomers (30 nM) were adjusted to pH 5.5 in 0.1% (w/v) DBM, and the binding partner, LF$_N$ V48C*EDANS, was titrated. Upon reaching equilibrium, each sample's FA signal was recorded. The saturation in the FA signal occurs at a LF$_N$:PA$_7$ ratio of 3.4 (±0.2), which is consistent with the 3:1 stoichiometry of the complex reported using other methods (Kintzer *et al.*, 2009, Mogridge *et al.*, 2002).

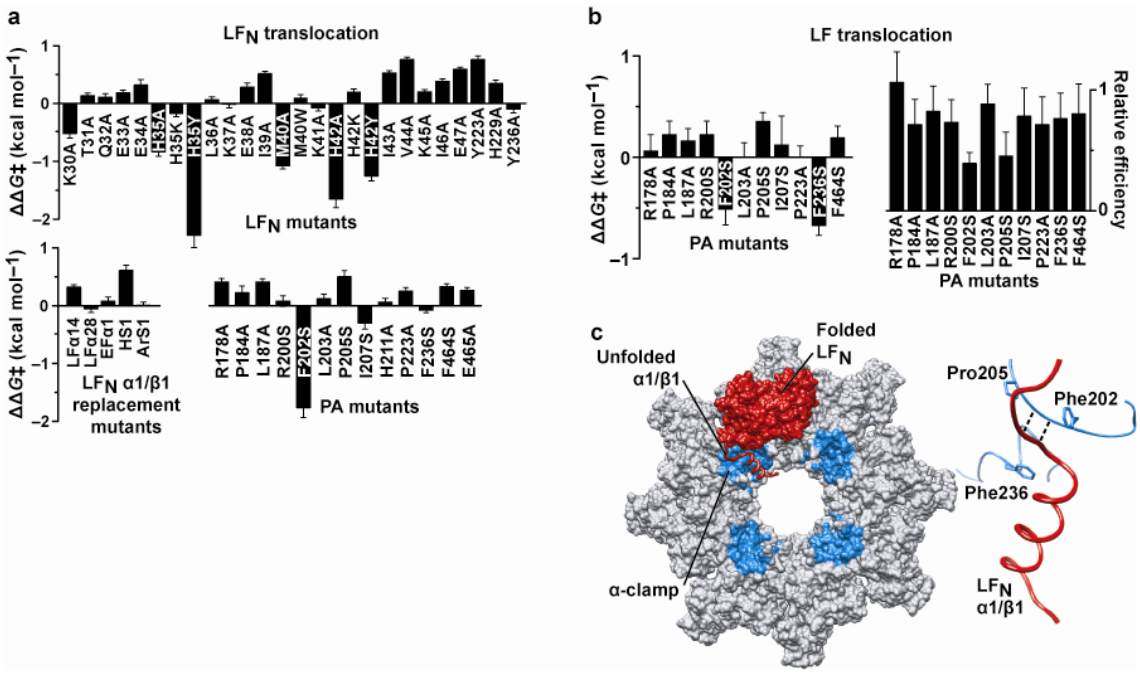

Figure 3.15. The role of the α clamp in LF_N and LF translocation. Planar lipid bilayer translocation results for various mutant channels and substrates. **(a)** Differences in translocation activation energy ($\Delta\Delta G^{\ddagger}$) for (top) LF_N mutants, (bottom left) LF_N α1/β1-replacement mutants, and (bottom right) PA mutants are shown. The reference state is WT LF_N:WT PA. $\Delta\Delta G^{\ddagger} = \Delta G^{\ddagger}(\text{WT}) - \Delta G^{\ddagger}(\text{MUT})$, and $\Delta G^{\ddagger} = RT \ln t_{1/2} / c$. The $t_{1/2}$ value is the time for half of the protein to translocate, and c is a 1-sec reference constant. All LF_N translocation rates were measured at symmetrical pH 5.6, $\Delta\Psi$ = 40 mV. A negative value indicates the rate of translocation slowed upon mutation. The relative translocation efficiencies for these LF_N translocations are given in Figure 3.16b. **(b)** Full-length LF translocation at pH_{cis} = 6.1, pH_{trans} = 7.4, ΔpH = 1.3, $\Delta\Psi$ = 20 mV. (left) $\Delta\Delta G^{\ddagger}$ values and (right) relative translocation efficiencies ($\varepsilon_{MUT}/\varepsilon_{WT}$) for mutant PA channels. Individual LF translocation records are shown in Figure 3.17. Error bars in (a-b) are the mean ±s.d. (n = 2-12). **(c)** (left) LF_N α1/β1 (red ribbon) unfolds from the structured carboxy-terminal subdomain (red surface) by binding into the α-clamp site (cyan surface) on the PA oligomer (gray surface). The interaction is comprised of nonspecific interactions. The α-clamp sites orient the unfolded structure toward the central pore, where the protein is translocated. (right) Residues in PA's α-clamp site (cyan) that affect LF_N and/or LF translocation are rendered as sticks. LF_N α1/β1 (red ribbon) and parallel β-sheet hydrogen bonds (black dotted lines) between LF_N β1 and PA β13 are shown.

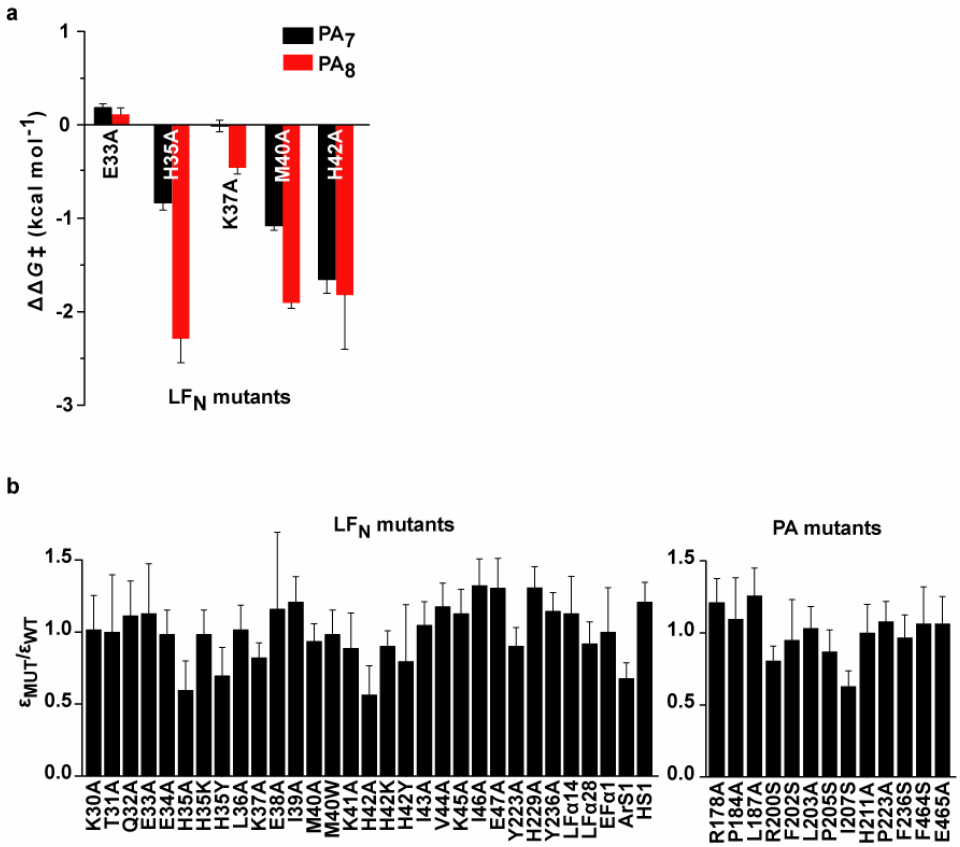

Figure 3.16. LF_N translocates similarly via PA_7 and PA_8 channels. (a) WT LF_N and the indicated LF_N mutants (MUT) were translocated at symmetrical pH 5.6, $\Delta\Psi$ = 40 mV using either heptameric (PA_7, black) or octameric (PA_8, red) channels. The translocation activation energy difference ($\Delta\Delta G^{\ddagger}$) was calculated according to $\Delta\Delta G^{\ddagger} = \Delta G^{\ddagger}(\text{WT}) - \Delta G^{\ddagger}(\text{MUT})$, where $\Delta G^{\ddagger} = RT \ln t_{1/2} / c$. The $t_{1/2}$ is the time (in seconds) for half of the protein to translocate; c is a 1-sec reference constant; R is the gas constant; and T is the temperature. A negative $\Delta\Delta G^{\ddagger}$ indicates the rate of translocation is slower for the LF_N mutant. Error bars are the mean ±s.d. (n = 2-6). **(b)** LF_N translocation was measured at symmetrical pH 5.6 and a $\Delta\Psi$ of 40 mV. The translocation efficiency (ε) is given as the fraction of channels that successfully translocate, where WT LF_N translocation efficiency is ~60% under these conditions. The relative translocation efficiency ($\varepsilon_{MUT}/\varepsilon_{WT}$) is the ratio of the mutant and WT translocation efficiencies. (Left) $\varepsilon_{MUT}/\varepsilon_{WT}$ for LF_N point mutants and α1/β1-replacement mutants translocated via WT PA; and (Right) $\varepsilon_{MUT}/\varepsilon_{WT}$ for WT LF_N translocated via mutant PAs. Errors are the mean ±s.d. (n = 2-6).

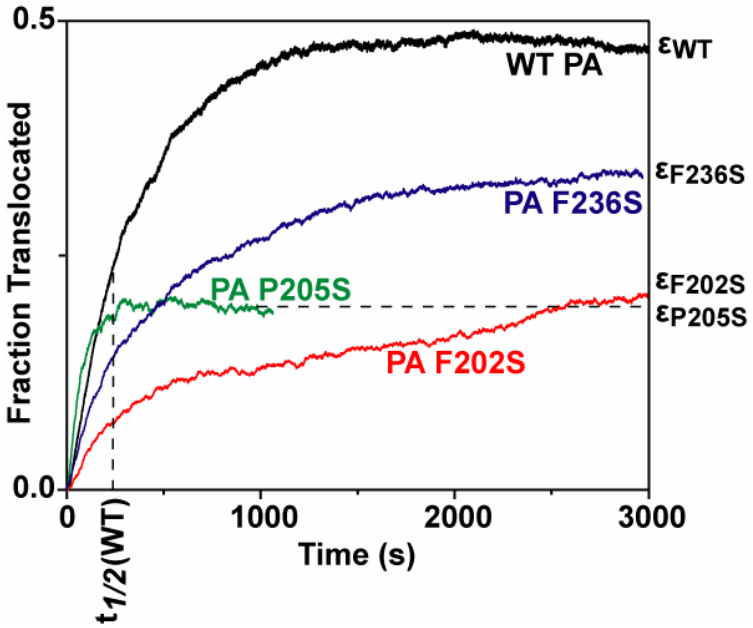

Figure 3.17. LF translocation records. LF translocation was measured using planar lipid bilayer electrophysiology under a 1.3-unit ΔpH (pH_{cis} = 6.1, pH_{trans} = 7.4) and a 20-mV ΔΨ. WT LF was translocated via WT PA (black) and the PA mutants, PA F202S (red), PA P205S (green), and PA F236S (blue). The time for half of the translocated LF to translocate ($t_{1/2}$) is indicated for WT PA. The efficiency, ε, (shown for each PA) is the ratio of the final amplitude and expected amplitude. For WT LF translocation via WT PA, ε is ~50%.

binding activity likely reflects the general helical shape complementarity of the α-clamp site, which excludes ~1000 $Å^2$ on PA without making specific side-chain-side-chain interactions. Additionally, backbone hydrogen bonds, which are ubiquitous features of polypeptides, can provide nonspecific contact points between the translocase and substrate, as we observe between LF_N β1 and PA_N β13 (Figures 3.5a and 3.15c).

Broad peptide-binding specificity has been observed in other systems, including calmodulin (Meador et al., 1992, Meador et al., 1993); the ClpXP adapter, SspB (Levchenko et al., 2005, Levchenko et al., 2003); the chaperone, GroEL/ES (Landry & Gierasch, 1991, Li *et al.*, 2009, Wang *et al.*, 1999); and the unfoldase, ClpA/Hsp100 (Hinnerwisch et al., 2005). For calmodulin, which is analogous structurally to the PA oligomer's α-clamp cleft, multiple peptide helices are recognized by the cleft formed by its twin Ca^{2+}-ion binding sites. The ClpXP adapter, SspB, binds multiple unstructured carboxy-terminal degradation signal tags in various conformations in a cleft. The chaperone complex, GroEL/ES, can bind to various amphipathic helices and strands. A substrate binding site, identified in the unfolding machine ClpA/Hsp100, is located above the ϕ-clamp-type site and may be analogous to the α-clamp site on the PA oligomer.

Our structure provides new insight on how a nonspecific polypeptide clamp can unfold its substrate. By binding to LF_N in multiple locations using nonspecific interactions [i.e., in the α clamp (Figure 3.5a) and ϕ clamp (Krantz et al., 2005)], LF_N can be partially unfolded (Figure 3.12a) and maintained in a more strained (Figure 3.12d,e) and less stable conformation (Figure 3.12c-e). The region of LF_N that is most destabilized upon binding PA (Figure 3.12d,e) coincides with LF_N's β2-β4 sheet, which was previously reported as the mechanical breakpoint, or structure that is rate-limiting to the unfolding step of translocation (Thoren et al., 2009). Therefore, we infer the α-clamp site stabilizes unfolding intermediates, introduces strain into the mechanical breakpoint, and feeds unfolded structure into the central ϕ-clamp site.

We estimate that the costs associated with binding to the α-clamp site (Figure 3.6a-c) may be offset by orienting the substrate toward the central lumen (Figure 3.15c), reducing the stability of the substrate (Figure 3.12c), and minimizing the diffusional mobility of unstructured regions before (Figure 3.12d,e) or during translocation (Krantz et al., 2005). We expect that nonspecific-clamping sites should lessen the counterproductive diffusive motions expected for large sections of unfolded polypeptide chain by maintaining contact with the unfolded chain and further reducing backbone conformational entropy, thus allowing the ΔΨ/ΔpH driving force to efficiently unfold (Thoren et al., 2009) and translocate proteins (Krantz et al., 2006) (Figure 3.15a,b). Although the α clamp forms a stable complex with unfolded structure, this intermediate does not represent a thermodynamic trap. Rather populating partially unfolded translocation intermediates would lower a much greater overall rate-limiting barrier expected in the absence of such intermediates, thereby allowing translocation to proceed on a biologically reasonable timescale.

3.5 Materials and Methods

Plasmids and proteins. Site-directed mutagenesis was performed using the commercial Quikchange procedure (Agilent Technologies) (Zheng *et al.*, 2004). WT PA and PA mutants, including the construct used in the crystallization experiments, $PA^{\Delta MIL}$ [in which the membrane insertion loop (residues 303-324) was deleted and replaced with a type II turn sequence (Kintzer et al., 2009)], were expressed and purified as described (Krantz et al., 2005). Heptameric and octameric PA oligomers were produced as described (Kintzer et al., 2009).

LF, LF_N, His_6-DTA and mutants thereof were purified from overexpressing bacteria using standard Ni^{2+}-nitrilotriacetic-acid-(NTA)-affinity chromatography and Q-sepharose, anion-exchange chromatography (GE Healthcare, USA) (Krantz et al., 2006). Their six-histidine (His_6) tags were removed with bovine α-thrombin treatment (0.5 units/mg of protein) for 30 minutes at room temperature in 20 mM Tris (pH 8), 150 mM NaCl, 2 mM $CaCl_2$ and 1 M dextrose.

Amino-terminal truncation mutants of LF_N were made by PCR amplifying the truncated sequence and cloning the sequence into pET15b-LF_N (Lacy et al., 2002) via the 5' Nde I and 3' BamH I sites. These constructs are named Δ*n* LF_N, where *n* designates the amino acids that were deleted from the amino terminus. Note that due to the design of the thrombin cleavage site in the pET15b vector, thrombin cleavage leaves an additional GSHM sequence at the amino terminus of all pET15b LF_N constructs.

Fusions of LF and DTA were produced by introducing an in-frame Sac I restriction site into the pET15b-DTA vector (Blanke et al., 1996) prior to the DTA reading frame. The first 20 or 60 residues of LF [including the His_6 tag encoded in the pET15b-LF_N vector (Lacy et al., 2002)] were subcloned into the Sac I engineered DTA vector at the vector-encoded 5' Nco I site and the silent 3' Sac I site. These His_6-tagged LF_{1-20}-DTA and LF_{1-60}-DTA fusions and His_6-tagged DTA were purified from overexpressing bacteria using Ni^{2+}-NTA-affinity chromatography, Blue-sepharose chromatography (GE Healthcare), and Q-sepharose, anion-exchange chromatography. The His_6 tag was not removed from DTA, but the His_6 tag was removed from the His_6-tagged LF_{1-20}-DTA and LF_{1-60}-DTA fusions as described above. Note the unstructured His_6 tag on the amino terminus of His_6-tagged DTA is MGSSHHHHHHSSGLVPRG.

All LF and LF_N α1/β1-replacement constructs were made using a three-step, gene-synthesis procedure, according to the following scheme:

HM ^{1}AGGHGDVGMHVKEKEKNKDENKRKDEERNKT ^{32}QEEHLKEIMKHIV
^{45}KIEVKGEEAVKKEAAEKLLEKVPSDVLEMYKAIGGKIYI ^{84}**VD**

The bold face 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 are used for cloning. Superscripted numbers indicate the numbering convention of LF residues in 1J7N (Pannifer et al., 2001). The underlined sequence (residues 32-44) is the guest site, which is replaced with the following peptides:

$LF_{286-298}$ (α14)	SEEGRGLLKKLQI
$LF_{710-722}$ (α28)	NSKKFIDIFKEEG
EF_{22-34} (α1)	EKEKFKDSINNLV

In the case of Aromatic Sequence 1 and Hydrophilic Sequence 1, the guest site is residues 32-46 ^{32}QEEHLKEIMKHIVK^{46}I, which is replaced by:

Aromatic Sequence (ArS1)	QEEHFKEIFKHFVKF
Hydrophilic Sequence 1 (HS1)	QEEHSKEISKHSVKS

Overlapping oligonucleotides encoding the desired sequences with the α1/β1 replacement were synthesized (Elim Biopharmaceuticals, Inc., Hayward, CA) and amplified by two rounds of polymerase chain reaction (PCR). In Round I, 20 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 α1/β1-replacement synthetic DNA fragments were ligated via a 5' Nde I site and 3' Sal I site into either the pET15b-LF(Sal I) or the pET15b-LF_N(Sal I) construct, which contain an in-frame, silent Sal I restriction site in LF or LF_N at V84 and D85. The synthetic LF and LF_N constructs were purified and their His_6 tags were subsequently removed as described above.

Synthesis and purification of the $PA_8(LF_N)_4$ complex used in structural studies. $PA^{\Delta MIL}$ monomer was treated with trypsin at a ratio of 1:1000 (wt/wt) for 15 minutes at room temperature and then inhibited with soybean trypsin inhibitor at 1:100 (wt/wt) and phenylmethylsulphonyl fluoride (PMSF) at 0.1 mM. Crude mixtures of oligomeric $PA^{\Delta MIL}$ were produced by anion-exchange chromatography (Kintzer et al., 2009). Homogeneous $PA^{\Delta MIL}$ octamer was made by incubating this oligomeric mixture in 74 mM sodium acetate, 7 mM Tris, 0.62 M NaCl, 37 mM tetrabutylammonium bromide, 7% ethanol, 0.07% *n*-dodecyl-β-D-maltopyranoside, pH 5.7 and then microcentrifuging for 20 minutes (14k RPM) (Kintzer et al., 2009). Soluble $PA^{\Delta MIL}$ octamer (judged pure by electron microscopy) was complexed with LF_N at a one-to-one molar ratio with respect to PA monomer, purified over S200 gel filtration in 20 mM Tris, 150 mM NaCl, pH 8.0, and tested for homogeneity by nanoESI-MS (Figure 3.1b).

Crystallization and X-ray diffraction. The $PA_8(LF_N)_4$ complex was crystallized by the hanging-drop, vapor-diffusion method (McPherson, 1976). Prior to crystallization, the protein complex was incubated with 20 mM ATP on ice for 10 minutes and then mixed one-to-one with well solution (13-17% (w/v) polyethylene glycol with average molecular weight 3000 Da, 100 mM cacodylic acid, 200 mM $MgCl_2$, pH 6.7-7.3.) Rectangular prisms grew overnight at 19 °C, maturing to dimensions of 100-300 μm. Crystals were harvested in a one-to-one mixture of well solution and cryoprotectant (50% (v/v) glycerol, 20 mM Tris-Cl, 150 mM NaCl, pH 8) and plunged into liquid N_2. X-ray diffraction data were collected at a wavelength of 1.1159 Å at 100 K on a Quantum 315r CCD detector at beamline 8.3.1 at the Lawrence Berkeley National Laboratory Advanced Light Source (MacDowell *et al.*, 2004). A single crystal, belonging to the $P42_12$ space group, diffracted X-rays to 3.1 Å and had the unit cell dimensions, 178.4, 178.4, and 240.4 Å for *a*, *b*, and *c*, respectively (Table 3.1). The diffraction data (99.8% complete) were indexed and scaled in HKL2000 (Otwinowski & Minor, 1997).

Model refinement. The $PA_8(LF_N)_4$ complex structure was solved by molecular replacement (MR) using PHASER (Storoni *et al.*, 2004). The MR search model was a loop-stripped PA dimer from 3HVD (Kintzer et al., 2009). Two PA dimers were found in the asymmetric unit. Rigid-body and TLS refinement using PHENIX (Adams et al., 2004) produced F_o–F_c electron density consistent with a helical bundle that aligned to LF_N α2, α4, α9, and α10. Rounds of polyalanine-model building in COOT (Emsley & Cowtan, 2004) and refinement in PHENIX revealed that the identified polyalanine secondary structure elements aligned well with a model of LF_N [LF residues 51-250 (PDB 1J7N (Pannifer et al., 2001))]. All of LF_N's secondary-structure elements, except the amino-terminus (LF_{1-28}) and the carboxy-terminal helix (α12), were identified and independently refined as rigid bodies to produce the initial model of the PA_2LF_N ternary complex. LF_N α1, β1, α2, α3, α4, and α5 and the loop regions, LF_{84-117}, $LF_{162-169}$ and $LF_{197-204}$, required either partial or extensive modeling to properly align them with the observed electron density. LF_{29-50} (α1/β1) was manually built extending from α2 (residue 51). Rounds of model building in COOT were followed by coordinate and *B*-factor refinement with non-crystallographic symmetry restraints in PHENIX. Backbone torsion angles were refined using the Torsion Optimization Procedure (TOP) provided by H. Gong, E. Haddadian, T.

Sosnick, and K. Freed at the University of Chicago. Molprobity analysis (Davis *et al.*, 2007) of the structure shows that 91% of residues are in the favored Ramachandran regions, yielding an overall Molprobity score of 2.88 (87th percentile for a 3.10 (±0.25) Å resolution structure). Surface burial calculations and molecular graphics were computed in CHIMERA (Pettersen *et al.*, 2004). The final model and refinement statistics are shown in Table 3.1.

Protein Data Bank accession code. The structure factors and coordinates for the $PA_8(LF_N)_4$ complex have been deposited in the PDB with the accession code 3KWV.

Planar lipid bilayer electrophysiology. Planar lipid bilayer currents were recorded using an Axopatch 200B amplifier (Molecular Devices Corp., Sunnyvale, CA) (Kintzer et al., 2009, Thoren et al., 2009). Membranes were painted on a 100-μm aperture of a 1-mL, white-Delrin cup with 3% (w/v) 1,2-diphytanoyl-*sn*-glycerol-3-phosphocholine (Avanti Polar Lipids, Alabaster, AL) in *n*-decane. *Cis* (side to which the PA oligomer is added) and *trans* chambers were bathed in various buffers as required. By convention, $\Delta\Psi \equiv \Psi_{cis} - \Psi_{trans}$ ($\Psi_{trans} \equiv 0$ V), and $\Delta pH \equiv pH_{trans} - pH_{cis}$.

To monitor LF_N, $LF_{1\text{-}20}$-DTA or $LF_{1\text{-}60}$-DTA binding to the PA channel, we first inserted PA channels into a planar lipid bilayer bathed in asymmetric KCl solutions buffered in 10 mM potassium phosphate ($[\text{added KCl salt}]_{cis}$ = 100 mM, $[\text{added KCl salt}]_{trans}$ = 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, $LF_{1\text{-}20}$-DTA or $LF_{1\text{-}60}$-DTA was then added to the cis side of the membrane at small 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, $I = I_o/(1 + K_d/[L]) + c$, to obtain the equilibrium dissociation constant, K_d, where I_o is the current amplitude and c is an offset.

In kinetic binding experiments, the rate of ligand binding and dissociation were recorded and fit to a single-exponential function, $I = I_o\exp(-kt) + c$, to obtain the observed rate constant, k, where I_o is the amplitude and c is an offset constant. The kinetic association rate constant, k_a, was computed using $k_a = k/[L]$. Dissociation of the ligand from the current-blocked complexes was initiated by perfusing the cis compartment with 5-10 volumes of buffer or by adding a 1-μM excess of Δ47 LF_N to compete with ligand binding. (The truncated form does not block the current when it binds to the PA channel.) The K_d could then be calculated from the kinetic rate constants using $K_d = k_d/k_a$. Kinetically and thermodynamically determined K_ds were self-consistent. Refer to Figure 3.9 for specific examples of these analyses.

The K_d for Δ27, Δ32, Δ39, Δ42, and Δ47 LF_N were deduced in equilibrium competition experiments with WT LF_N-PA channel complexes. PA channels were first inserted and then 0.1-0.8 nM of WT LF_N was added to the cis compartment. Once equilibrium was established the Δn LF_N competitor was added in increments. The degree of channel reopening established upon equilibration as a function of the competitor concentration was used to assess each competitor's K_d.

All LF_N translocation experiments were carried out as described previously using a universal pH bilayer buffer system (UBB: 10 mM oxalic acid, 10 mM phosphoric acid, 10 mM MES, 1 mM EDTA, and 100 mM KCl) at a symmetrical pH 5.6 (Thoren et al., 2009). Two to six replicate experiments were conducted for each mutant to establish the time (in seconds) for half of the substrate to translocate ($t_{1/2}$). The individual kinetic effects of LF_N mutations (MUT) were assessed by comparing the activation energy of translocation ($\Delta G‡$) at 40 mV for the mutant and WT LF_N, where $\Delta G‡ = RT \ln t_{1/2} / c$. R is the gas constant, T is the temperature, and c is 1

second. The change in $\Delta G‡$ ($\Delta\Delta G‡$) is reported at $\Delta\Delta G‡ = \Delta G‡(WT) - \Delta G‡(MUT)$. Efficiency, ε, was also obtained from each translocation record by the relation, $\varepsilon = A_{obs}/A_{exp}$, where A_{obs} is the observed amplitude of channels that reopened (or translocated), and A_{exp} is the expected amplitude if all of the channels reopened (or translocated).

LF translocation experiments were carried out similarly except that a 1.3-unit ΔpH was also applied during translocation. The pH of the UBB in the cis and trans chambers was adjusted to apply the proton gradient, where pH_{cis} = 6.1 and pH_{trans} = 7.4. The ΔΨ was 20 mV. LF translocation was assessed by $t_{1/2}$ and ε as described for LF_N. Relative translocation efficiency for each mutant was calculated as $\varepsilon_{MUT}/\varepsilon_{WT}$.

A special protocol was devised to analyze the PA R178A mutant due to LF's rapid dissociation from the channel. When LF was added to the channel, a 1.3-unit ΔpH was applied at a ΔΨ of 0 mV during the perfusion step. Translocation was then initiated by stepping the ΔΨ to 20 mV. This method, however, can only clamp one LF in the channel, presumably by engaging the substrate with the ΔpH-dependent mechanism involving the ϕ clamp (Krantz et al., 2006, Krantz et al., 2005). Thus the other substrates can dissociate during perfusion.

Equilibrium unfolding titrations. Guanidinium chloride titrations were performed on the Δ*n* LF_N truncations in 10 mM sodium phosphate, 0.75 M trimethylamine *N*-oxide, pH 7.5, 20 °C as described (Thoren et al., 2009, Krantz et al., 2004). Each titration point was probed by circular dichroism (CD) spectroscopy at 222 (±2) nm using a Jasco J-810 spectropolarimeter (Easton, MD). The CD-probed curves fit to a four-state thermodynamic model ($N \leftrightarrow I \leftrightarrow J \leftrightarrow U$) (Krantz et al., 2004).

Fluorophore labeling of LF_N and PA. Prior to all dye-modification reactions, excess DTT was removed from the Cys-substituted PA or LF_N proteins by buffer exchange on a G25 desalting column, equilibrated in 20 mM Tris-Cl pH 7.5, 150 mM NaCl. DTT-free PA_{83} K563C or a Cys-substituted LF_N was labeled with 10 molar equivalents of either 5-[2-[(2-Iodo-1-oxoethyl)amino]ethylamino]-1-naphthalenesulfonic acid (IEDANS), Alexa fluor 555 C_5 maleimide (AF_{555}), or Alexa fluor 647 C_5 maleimide (AF_{647}) (Invitrogen, USA) in the presence of 100 μM tris(2-carboxyethyl)phosphine (TCEP) (Sigma-Aldrich) at room temperature for 3 hours. The reaction was quenched with 5 mM DTT, and the labeled proteins were purified on a G25 desalting column to remove free-dye molecules. The labeling efficiency for each dye was determined by comparing the free-dye absorbance maximum and protein absorbance at 280 nm. The labeling efficiency was judged to be >90% in each case. For the IEDANS labeling of various Cys-substituted LF_Ns, we found the modification was >95% complete also by MALDI mass spectrometry.

Fluorescence anisotropy (FA). Fluorescence anisotropy (FA) was used to report on the changes in fluorophore mobility for LF_N labeled with IAEDANS. LF_N with single Cys residues introduced at specific sites (residues 37, 48, 72, 126, 164, 199, and 242) were modified with IEDANS dye and purified as described above. FA was measured using a FluoroMax-3 spectrofluorometer equipped with moveable linear polarizers. The excitation wavelength was 360 (±10) nm, and the emission wavelength range was 460-560 nm. The emission intensity value was the average intensity over this range. Each FA value, *a*, is calculated from the emission intensities from the parallel ($F_{\|}$) and perpendicular ($F\perp$) arrangement of the excitation and emission polarizers by $a = F_{\|} - F\perp / (F_{\|} + 2F\perp)$; however, a *G*-factor correction was applied to account for differences in sensitivities for the two different optical paths in the instrument. The FA signal change upon binding PA was not due to non-specific protein-protein associations, since the LF_N V48C*EDANS signal change is saturable at a 3:1 stoichiometry (LF_N:PA heptamer) (Figure

3.14b), which is consistent with the number of LF_N molecules that bind to PA_7 (Kintzer et al., 2009, Mogridge et al., 2002).

Normalized *B*-factor change and normalized FA-signal change calculations. Normalized differences in FA change (ΔFA_{norm}) or *B*-factor change (ΔB_{norm}) in LF_N upon binding PA were calculated by determining the mean change in the FA signal or *B* factor for all probed sites and then calculating the ratio of each individual signal change to this mean value. For the ΔB_{norm} calculation, $\Delta B_{norm} = B_{1,i} / \langle B_1 \rangle - B_{2,i} / \langle B_2 \rangle$. The $B_{1,i}$ and $B_{2,i}$ values are the average *B* factor for all atoms in the i^{th} residue in the free LF_N (structure 1) and the LF_N in complex with PA (structure 2), respectively. $\langle B_1 \rangle$ and $\langle B_2 \rangle$ are the average *B* factors for all residues in the entire LF_N chain taken from each respective structure. Likewise, for the ΔFA_{norm} calculation, $\Delta FA_{norm} = FA_{1,i} / \langle FA_1 \rangle - FA_{2,i} / \langle FA_2 \rangle$. This calculation is treated in an analogous manner, except only the 7 Cys-substituted, fluorescently-probed LF_N residues were considered.

FRET-probed, PA-assembly assay. A 1:1 mixture of dye-labeled $_n$PA ($_n$PA K563C*AF_{555} and $_n$PA K563C*AF_{647}) was diluted to 10 nM in buffer (10 mM sodium cacodylate, 100 mM potassium chloride, pH 7.4 at room temperature) either in the presence or absence of 10 nM LF_N or Δ47 LF_N, following a prior method (Wigelsworth *et al.*, 2004). Assembly was reported by the increase in the 668-nm and 566-nm emission intensity ratio (F_{668}/F_{566}) upon excitation at 555 nm, which reached a steady state in about one hour. The excitation and emission bandwidths were 2 nm. Emission intensity ratios were collected throughout the record at 5-minute intervals on a FluoroMax-3 spectrofluorometer (Horiba Jobin Yvon, Edison, NJ) using a 1×1-cm cuvette. Curves were fit with a second-order kinetic model described previously (Wigelsworth et al., 2004).

NanoESI-MS. Mass spectra of the protein complexes were acquired as described previously (Kintzer et al., 2009) using a quadrupole time-of-flight (Q-Tof) mass spectrometer with a Z-spray ion source (Q-Tof Premier, Waters, Milford, MA). Ions were formed using nanoelectrospray (nano-ESI) emitters prepared by pulling borosilicate capillaries (1.0 mm O.D./0.78 mm I.D., Sutter Instruments, Novato CA) to a tip I.D. of ~1 μm with a Flaming/Brown micropipette puller (Model P-87, Sutter). The instrument was calibrated with CsI clusters. The protein solution for the stoichiometry determinations was concentrated to 10 μM followed by dialysis into 10 mM ammonium bicarbonate, pH 7.8. Immediately prior to mass analysis, the solution was diluted 1:1 with 150-300 mM ammonium acetate, pH 7.8. A 0.127-mm-diameter platinum wire was inserted through the capillary into the solution, and electrospray was maintained by applying a 1-1.3 kV potential relative to instrument ground. Raw data were smoothed three times using the Waters MassLynx software mean smoothing algorithm with a window of 50 *m*/*z* (mass-charge ratio).

Electron microscopy. Each PA oligomer was diluted into an EM buffer (20 mM Tris, 150 mM NaCl, pH 8), making a final concentration of 20-30 nM with respect to PA monomer. 400 mesh copper grids were successively covered by a holey carbon film and a continuous carbon film. 4 μl of the diluted PA oligomer sample were applied to a freshly glow-discharged support grid for 30 s and then stained in 5 successive drops (75μl) of 2% uranyl acetate (Sigma-Aldrich). Negative-stain EM images were recorded with a Tecnai 12 (FEI Company, Hillsboro, OR) operated at 120 kV at 49,000× magnification. Images were taken using a CCD camera (2.13 Å/pixel specimen scale). Particle images were selected for each data set using automatic or manual particle picking using boxer in EMAN (Ludtke *et al.*, 1999). Reference-free processing was done using the software package, SPIDER (Frank *et al.*, 1996). Images were subjected to three successive cycles of multi-reference alignment, multivariate statistical analysis, and

classification (Stark *et al.*, 1995, van Heel *et al.*, 1996). The last classification was done using only the lowest order eigenvectors, as described elsewhere (White *et al.*, 2004), to separate the data by size and by the heptameric and octameric oligomerization states. A second method of image processing was used whereby crystal-structure-reference images were made from two-dimensional projections of low resolution density maps generated from the crystal structures of the PA heptamer (Lacy et al., 2004) and octamer (Kintzer et al., 2009) using SPIDER (Frank et al., 1996). Crystal-structure-referenced images were aligned and classified using the lowest order eigenvectors as stated above. All final class-average images were manually inspected for their oligomer number. The number of particles per classification was used to determine the percentages of heptamers and octamers in each sample. About 2000 to 10000 particles were analyzed per sample. Each method of classification, reference-free or crystal-structure-referenced, produced similar results (±2%) (Kintzer et al., 2009). The reported percentages of heptameric and octameric PA are given as the mean of the referenced and reference-free analysis.

Macrophage cytotoxicity. LT cytotoxicity was monitored by an enzyme-coupled lactose dehydrogenase (LDH) release assay (Decker & Lohmann-Matthes, 1988). Immortalized bone marrow macrophages from 129 mice (a gift from the Vance Lab at UC Berkeley) were grown to confluence in RPMI 1640 medium (Invitrogen) supplemented with 10% fetal bovine serum (Invitrogen), 100 units/mL penicillin (Sigma-Aldrich), and 100 μg/mL streptomycin (Sigma-Aldrich) in a humid, 5%-CO_2 atmosphere at 37 °C. One day prior to conducting assays, cells were trypsinized and re-plated at 10^5 cells/well. Cells were treated in triplicate with varying concentrations of LF and constant concentrations of PA. Dilutions of LF were prepared in ice-cold phosphate-buffered saline (PBS). Toxin-treated cells were incubated for 4 hours at 37 °C. Plates were then centrifuged at 1400 RPM, and 50 μL of supernatant were removed and added to a new 96-well plate. The supernatant was incubated with 20 μL of lactate solution (36 mg/mL in PBS) and 20 μL of *p*-iodonitrotetrazolium chloride (2 mg/mL in PBS with 10% dimethyl sulfoxide). The enzymatic reaction was started with the addition of 20 μL of nicotinamide adenine dinucleotide (NAD^+)/diaphorase solution (13.5 units/ml diaphorase and 3 mg/ml NAD^+). After 15 minutes, the products were analyzed on a spectrophotometric microplate reader (Bio-Rad Laboratories, Richmond, CA) at 490 nm. The change in the absorbance signal is proportional to the number of lysed cells, where the amount of LDH released was normalized to the value obtained in wells treated with 1% Triton X-100 detergent. Effective-concentration-for-50%-lysis values (EC_{50}) were determined by fitting normalized cell lysis versus PA concentration data in ORIGIN (OriginLab Corp., Northampton, MA).

PA-DTA-fusion cytotoxicity assays were carried out using J774 mouse macrophage cells grown in Dulbecco's Modified Eagle's Medium with 5% fetal bovine serum (FBS), 200 μM glutamine, and 10 μg/ml gentamycin. A 96-well plate was seeded with $5x10^4$ cells per well (in 100μl media) 16-20 hours before toxin application and incubated at 37 °C and 5% CO_2. LF-DTA fusions at varying concentrations and PA at a final concentration of 1 μg/ml were added to the cell culture (diluted in PBS supplemented with 1% bovine serum albumin). Plates were incubated at 37 °C and 5% CO_2 for 20-24 hours. Cell lysis was assayed using the LDH release assay described above.

Chapter 4

Role of the α clamp in anthrax toxin translocation

4.1 Introduction

Molecular machines involved in protein translocation and degradation must handle a wide variety of protein substrates and thus possess broad substrate specificity. In order achieve this broad specificity, it is thought that the machines recognize common features of an unfolded protein, such as hydrophobic groups or secondary structure elements (Chou & Gierasch, 2005, Wang *et al.*, 1999, Gelis *et al.*, 2007). Still, many questions remain about how channels interact with their substrates. More importantly, the implications that these interactions may have for substrate unfolding and translocation have largely been unexplored.

Recently, we reported the crystal structure of anthrax toxin's protective antigen bound to a partially unfolded substrate, the PA-binding domain of LF (LF_N) (Feld *et al.*, 2010). The first α helix and β strand of LF_N unfold and dock into a hydrophobic groove on the top of PA. Based on extensive mutational analysis, we concluded that this groove, termed the α clamp, binds to LF_N's α1/β1 region nonspecifically (Feld *et al.*, 2010). The site was shown to play an important role in substrate unfolding and PA oligomer assembly. Initial results indicate that it also plays a role in substrate translocation. We hypothesize that the nonspecific-binding capability of the α clamp may facilitate translocation by stabilizing partially-unfolded intermediates and minimizing diffusive motion.

Here we further test the α clamp's role in translocation by disrupting binding to the site. First, we show that LF_N-PA binding can be disrupted by drastically changing the shape of LF_N's α1/β1 region, or by sterically occluding the α clamp. We find that these mutants also impair translocation. While these results suggest a role for the α clamp in translocation, it is possible that some of the mutants are disrupting downstream translocation steps. I suggest some preliminary hypotheses and additional experiments that can be done in order to make more solid conclusions about the role of α clamp in translocation.

4.2 Results

Altering the shape of LF_N's α1/β1 region. Based on extensive mutational analysis we conclude that the α clamp is a non-specific binding site and may recognize its substrate based on shape complementarity (Feld *et al.*, 2010). Therefore, in order to disrupt binding to the α clamp, we wanted to drastically alter the shape of LF_N's α1 helix and β1 strand. To do this, we mutated residues in this region of LF_N (residues 30-47) to proline. Proline residues tend to disrupt α helices and β sheets because they lack an amide hydrogen and therefore cannot form the requisite hydrogen bonds. In addition, the proline side chain is detrimental to α helices because it sterically interferes with the backbone of the preceding turn. Thus, we expect proline residues in LF_N's α1/β1 region to disrupt the α1 helix and prevent the formation of β-sheet hydrogen bonds between LF_N and residue Leu203 in PA's α clamp (Feld *et al.*, 2010). Consecutive proline residues could also alter the structure of this region because they have the potential to form a polyproline II helix, which is longer and narrower than a regular α helix comprised of the same number of residues (Adzhubei & Sternberg, 1993) (Figure 4.1).

Three LF_N mutants were made with an increasing number of consecutive proline residues in their α1/β1 region: LF_N Pro43-47, LF_N Pro36-47, and LF_N Pro30-47. Here, the numbers indicate the range of residues that were mutated to proline (Figure 4.2). As a control, a fourth

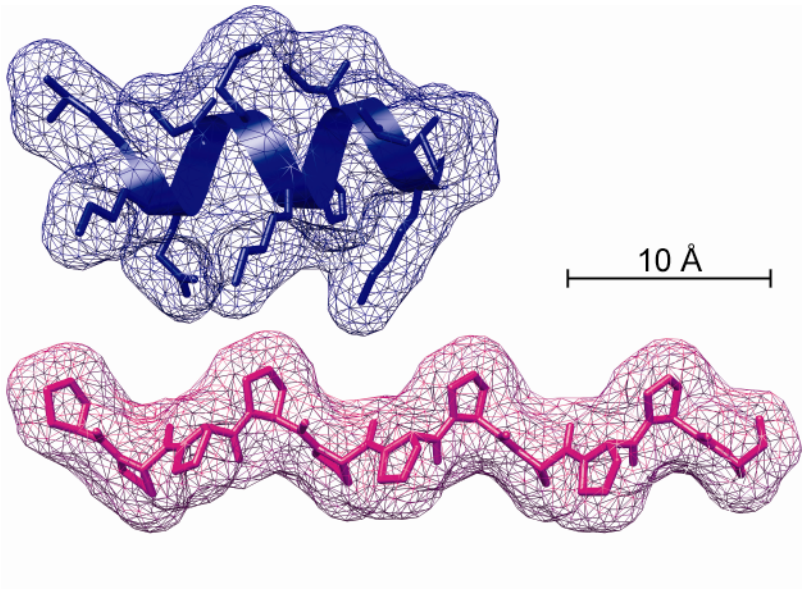

Figure 4.1. Comparison of the α1 helix of LF_N and a polyproline II helix. Residues 36-46 of LF_N's α1 helix are shown in blue. A polyproline II helix is shown in pink. Although the polyproline II helix contains the same number of residues as LF_N's α1, the structure is longer and narrower.

A

```
                    30         40         50  250          263
WT LFN...KTQEEHLKEIMKHIVKIEVKG...QEINLSLEELKDQR
LFN Pro43-47...KTQEEHLKEIMKHPPPPPVKG...QEINLSLEELKDQR
LFN Pro36-47...KTQEEHPPPPPPPPPPPPVKG...QEINLSLEELKDQR
LFN Pro30-47...PPPPPPPPPPPPPPPPPPVKG...QEINLSLEELKDQR
LFN Pro252-263...KTQEEHLKEIMKHIVKIEVKG...QEPPPPPPPPPPP
```

B

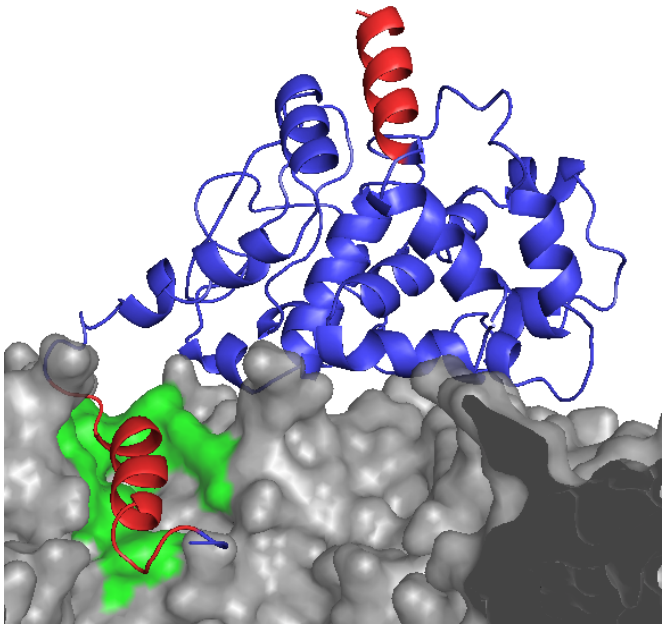

Figure 4.2. LF_N Proline mutants. (A) The sequences for WT LF_N and the four LF_N proline mutants are shown. Residue numbers are indicated at the top. **(B)** The structure of LF_N (blue) bound to the PA octamer (gray surface, which has been cut away for clarity) (PDB 3KWV) (Feld *et al.*, 2010). The α1 helix and β1 strand of LF_N unfold and dock into the α clamp (green) on the top of PA. Residues 30-47 and residues 252-263 of LF_N are highlighted in red to indicate the location of the proline mutations. Although the C-terminal helix of LF_N (residues 252-263) is not modeled in 3KWV, it is shown here based on the alignment of free LF_N (PDB 1J7N) to the LF_N molecule in 3KWV.

LF_N mutant was made, LF_N Pro252-263, where residues at LF_N's carboxy-terminus were mutated to proline (Figure 4.2). This mutant has the same number of consecutive proline residues as LF_N Pro36-47 (12 prolines).

LF_N proline mutants disrupt binding and translocation. Using electrophysiology, we can measure LF_N binding to PA by monitoring kinetic and equilibrium changes in channel conductance. When LF_N binds to the PA channel, it inserts its amino-terminal end into the channel and blocks conductance. These binding measurements are made in the absence of an applied voltage ($\Delta\Psi$) to eliminate its influence on the channel-substrate interaction. We find that the LF_N Pro43-47 mutant binds PA slightly weaker than WT LF_N (Figure 4.3). The more heavily mutated proteins, LF_N Pro36-47 and LF_N Pro30-47, bind PA ~100 times weaker than WT LF_N (Figure 4.3). In addition, these mutants only block about 80% of the bulk channel conductance, whereas WT LF_N blocks ~95-98%. Finally, we found that LF_N Pro252-263 binds PA channels like WT LF_N (Figure 4.3). Thus, we conclude that the presence of proline residues specifically in the α1/β1 region of LF_N disrupt binding to PA. In order to confirm that these binding defects are a result of the LF_N proline mutants being unable to bind the α clamp (and not another part of PA), we will measure binding to the α-clamp mutant, PA R178A (Feld *et al.*, 2010). If the LF_N proline mutants do not bind the α clamp, we would expect them to bind PA R178A with the same affinity as WT PA.

In addition to binding, we measured the translocation kinetics of these LF_N proline mutants using planar lipid bilayer electrophysiology. In these translocation experiments, PA channels are inserted into an artificial bilayer membrane. LF_N is then added to the bilayer, where it binds to PA and blocks conductance. Excess LF_N is perfused, and translocation is initiated by raising the $\Delta\Psi$ or by creating a pH gradient (ΔpH). As LF_N translocates, the channels become unblocked, and the restoration of conductance reports on the translocation kinetics in real time. From the translocation half time ($t_{1/2}$), which is the time (measured in seconds) for half of the translocated protein to move through the channel, we can calculate an empirical activation energy ($\Delta G‡$) at a particular $\Delta\Psi$, where $\Delta G‡ = RT \ln t_{1/2} / c$. Here, R and T are the gas constant and temperature, and c is an arbitrary reference, which we define as 1 s.

We find that mutants with the most proline residues in the α1/β1 region (LF_N Pro30-47 and LF_N Pro36-47) translocate significantly slower than WT LF_N at all $\Delta\Psi$s (Figure 4.4A). In fact, the LF_N Pro30-47 mutant did not translocate under a $\Delta\Psi$ driving force alone; an additional ΔpH driving force of 0.76 units was needed in order to observe any significant translocation. Even under this stronger driving force, translocation of LF_N Pro30-47 was still much slower than WT LF_N. In addition to affecting the rate of translocation, LF_N Pro30-47 and LF_N Pro36-47 translocate less efficiently than WT LF_N (Figure 4.4B). Translocation efficiency (ε) is defined as the fraction of channels that successfully transports their substrates and is calculated by $\varepsilon = A_{obs}/A_{exp}$, where A_{exp} and A_{obs} are the expected and observed amplitudes of current. Finally, the less mutated construct, LF_N Pro43-47, and the control, LF_N Pro252-263, translocate like WT LF_N in terms of both rate and efficiency. Thus, we conclude that the presence of proline residues specifically in the α1/β1 region impairs translocation.

Sterically occluding the α clamp by extending PA_{63}'s amino terminus. Our second approach to disrupt binding to the α clamp was to sterically occlude the site. We tried several intermolecular cross-linking reactions to attach various dyes or peptides to the α clamp. However, all of these attempts were unsuccessful; we still observed WT-like binding and translocation activity from these dye- or peptide-modified PA oligomers. This finding is likely due to the fact the each oligomer has seven or eight α clamps (depending on whether PA is a

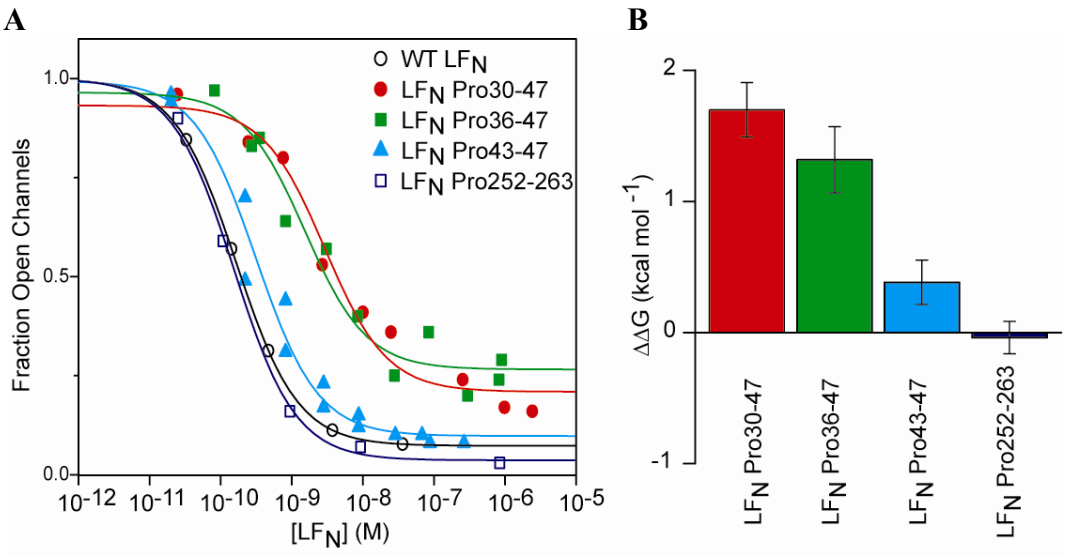

Figure 4.3. Proline residues in the α1/β1 region of LF_N disrupt binding to PA. (A) Equilibrium binding curves for WT LF_N and LF_N proline mutants using WT PA channels. Curves were fit to a simple single-binding site model, $I = I_o/(1 + K_d/[L]) + c$, to obtain K_d values, where I is the current amplitude, $[L]$ is the LF_N concentration, and c is an offset. (B) Changes in binding free energy (ΔΔG) relative to WT LF_N:WT PA, where $\Delta\Delta G = \Delta G_{MUT} - \Delta G_{WT}$. Positive ΔΔG's indicate that the mutant binds WT PA with less affinity than WT LF_N.

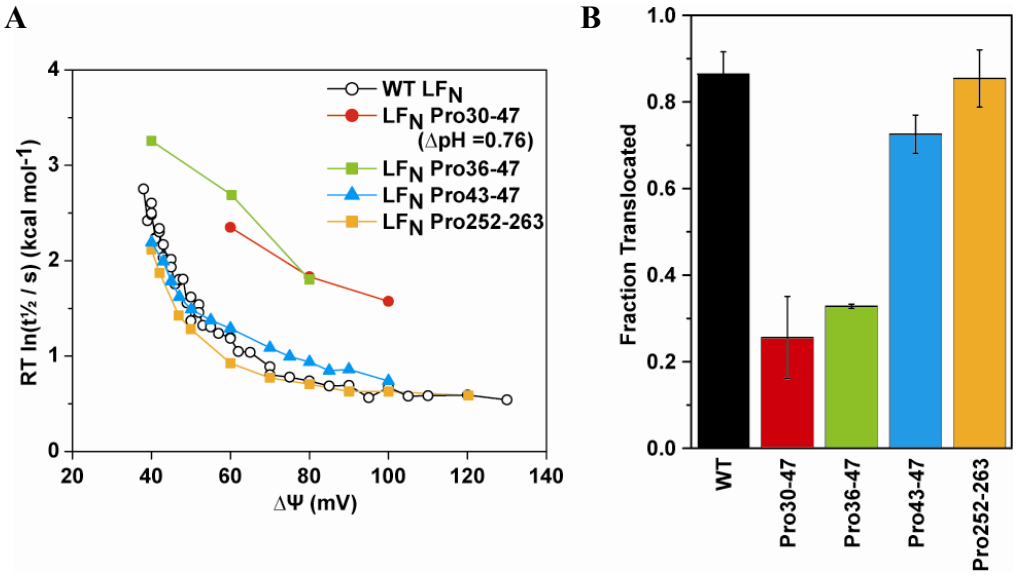

Figure 4.4. Proline residues in the α1/β1 region of LF$_N$ disrupt translocation. (A) A comparison of the ΔΨ dependence of the translocation activation energy (ΔG‡) for WT LF$_N$ and the LF$_N$ proline mutants, where WT PA channels are used in each case. $\Delta G^{\ddagger} = RT \ln t_{1/2} / c$. The $t_{1/2}$ value is the time for half of the protein to translocate, R is the gas constant, T is temperature, and c is a 1-sec reference constant. All LF$_N$ translocation rates were measured at symmetrical pH 5.6, except for LF$_N$ Pro30-47 which was translocated under a ΔpH = 0.76. (B) The translocation efficiencies for these LF$_N$ substrates at 60 mV. Efficiency (ε) is defined by $\varepsilon = A_{obs}/A_{exp}$, where A_{obs} is the observed amplitude of channels that reopened (or translocated), and A_{exp} is the expected amplitude if all of the channels reopened (or translocated).

heptamer or octamer), and even with 95-98% modification of the protein, at least one site per oligomer is left unmodified. In order to block of all α clamp sites, we attempted an intramolecular modification. Conveniently, the amino-terminus of each PA_{63} subunit in the oligomer is located in the vicinity of the α clamp (Figure 4.5). By extending this terminus and cross-linking it to the α clamp, we hypothesized that each PA subunit could "plug" its own α clamp.

In order to achieve the greatest likelihood of binding to the α clamp, we decided to append LF's α1/β1 sequence (residues 29-46) to the amino terminus of PA_{63}. The sequence had to be slightly modified due to the fact that PA_{83} monomers must be proteolytically activated in order for the oligomer to form; cleavage and dissociation of an amino-terminal, 20-kDa fragment allows the remaining PA_{63} to assemble into ring-shaped oligomers. In order to prevent proteolytic digestion of this appended sequence when PA_{83} is treated with trypsin, all four lysine residues LF's α1/β1 sequence were mutated to alanine. Thus, the final sequence grafted onto PA_{63}'s amino terminus was N**A**TQEEHL**A**EIM**A**HIV**A**I. We estimate that this appended sequence, in addition to the six unstructured residues of PA_{63}'s amino terminus, should provide enough length to fold back on itself and bind in the α clamp. To increase the flexibility of this modified amino terminus, I also mutated Pro173 of PA to a glycine residue (Figure 4.5). This mutant is called $PA_{\alpha\text{-plug}}$. In addition to this mutant, three double-cysteine versions were made in order to crosslink the appended sequence to the α clamp: $PA_{\alpha\text{-plug}}$ $M40_{LF}C$, F236C ($PA_{\alpha\text{-plug}}^{C40\text{-}C236}$); $PA_{\alpha\text{-plug}}$ $A30_{LF}C$, E465C ($PA_{\alpha\text{-plug}}^{C30\text{-}C465}$); and $PA_{\alpha\text{-plug}}$ $Q32_{LF}C$, F464C ($PA_{\alpha\text{-plug}}^{C32\text{-}C464}$). The LF numbering convention is maintained for cysteine mutations in the appended sequence, as indicated by the subscript, LF.

Initial characterization of the $PA_{\alpha\text{-plug}}$ mutants. Before measuring binding and translocation with the $PA_{\alpha\text{-plug}}$ mutants, we first wanted to verify that they oligomerize, and form ion-conducting channels in membrane bilayers. Using native PAGE, we determined that the $PA_{\alpha\text{-plug}}$ mutant does assemble into oligomers, and interestingly, it self-assembles better than WT PA. Self-assembly of WT PA is fairly slow and occurs only under dilute conditions because PA_{20} must first dissociate from PA_{63} [Kd ~ 190 nM (Christensen *et al.*, 2005)]. However, WT PA assembly is promoted by the presence of its protein substrates (EF, LF, EF_N or LF_N). In particular, the interaction between LF's α1/β1 region and the α clamp has been shown to drive assembly (Feld *et al.*, 2010). Thus, it is likely that the presence of the α1/β1 sequence in $PA_{\alpha\text{-plug}}$ allows the protein to self-assemble. Not only is this finding an interesting result in and of itself, but it also provides evidence that the appended α1/β1 sequence is actually binding to the α clamp in $PA_{\alpha\text{-plug}}$.

In addition to native PAGE, we used electron microscopy (EM) to analyze the population of heptamers and octamers formed by the $PA_{\alpha\text{-plug}}$ mutant. Because LF's α1/β1 sequence was shown to be important in the mechanism of octamer formation (Feld *et al.*, 2010), we expected $PA_{\alpha\text{-plug}}$ to form a higher percentage of octamers than WT PA. However, this is not the case; the $PA_{\alpha\text{-plug}}$ mutant forms ~9% octamer, which is only slightly higher than WT PA (<5% octamer (Kintzer *et al.*, 2009)). It is unclear why the presence of the α1/β1 sequence does not increase the octamer population in $PA_{\alpha\text{-plug}}$. Although we decided not to pursue this result here, further investigation may provide insight on the oligomerization mechanism of PA.

Lastly, we tested the channel-forming ability of these $PA_{\alpha\text{-plug}}$ mutants in membrane bilayers. In bulk experiments, all of the $PA_{\alpha\text{-plug}}$ mutants inserted into membranes and properly conducted current.

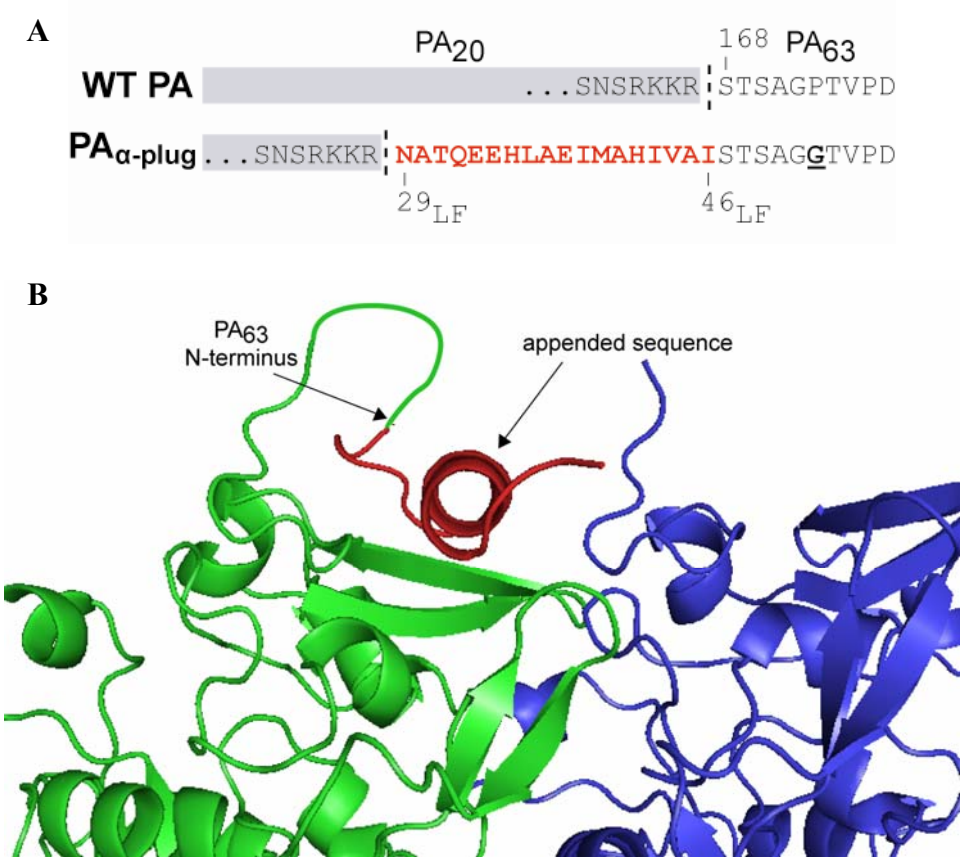

Figure 4.5. "Plugging" the α clamp by extending PA_{63}'s amino terminus. (A) Sequence of WT PA and $PA_{\alpha\text{-plug}}$. LF's α1/β1 sequence (residues 29-46), shown in red, was grafted onto the amino terminus of PA_{63}. Residue Pro173 of PA is mutated to glycine in $PA_{\alpha\text{-plug}}$, which is underlined in the sequence. PA_{20} (gray highlight) is cleaved off during proteolytic activation of PA_{83} monomers, allowing PA_{63} to assemble into oligomers. **(B)** Model of $PA_{\alpha\text{-plug}}$. Two PA_{63} subunits (green and blue ribbons) and residues 29-46 of LF_N (red ribbon) from the $PA_8(LF_N)_4$ crystal structure (PDB 3KWV) are shown (Feld *et al.*, 2010). The six unstructured residues of PA_{63}'s amino terminus have been drawn in and connected to LF residues 29-46 to illustrate how the extended terminus may look when bound to it's own α clamp.

$PA_{\alpha\text{-plug}}$ mutants disrupt LF_N binding and translocation. We measured WT LF_N binding to the $PA_{\alpha\text{-plug}}$ mutants found that they all bind WT LF_N with less affinity than WT PA (Figure 4.6). $PA_{\alpha\text{-plug}}^{C30\text{-}C465}$ and $PA_{\alpha\text{-plug}}^{C32\text{-}C464}$ are the most defective as they bind LF_N 2-2.5 kcal mol^{-1} weaker than WT PA (Figure 4.6). In addition, LF_N only blocks about 75-80% of the bulk channel conductance of these mutants compared to ~ 95% for WT PA and ~ 90% for $PA_{\alpha\text{-plug}}$ and $PA_{\alpha\text{-plug}}^{C40\text{-}C236}$.

Most of these $PA_{\alpha\text{-plug}}$ mutants are also defective in translocating WT LF_N. While $PA_{\alpha\text{-plug}}^{C40\text{-}C236}$ translocates WT LF_N with similar rates as WT PA, reaction with Ellman's reagent indicates that this mutant does not cross-link very well. Perhaps this mutant can translocate LF_N like WT PA because the un-crosslinked appendage is somehow able to move out of the way of the α clamp. The other three $PA_{\alpha\text{-plug}}$ mutants, however, translocate LF_N much slower than WT PA, especially at the higher voltages (Figure 4.7A). Again, $PA_{\alpha\text{-plug}}^{C30\text{-}C465}$ is one of the most defective; at 80 mV, it translocates LF_N about 5 times slower than WT PA. At low voltages, the mutants translocate WT LF_N at similar rates as WT PA, however, translocation is less efficient (Figure 4.7B).

Full-length LF binding and translocation. Although LF_N has been very useful as a model protein to study translocation, we ultimately would like to know how molecular machines unfold and translocate more complex, multidomain substrates. In particular, how does PA translocate its physiological substrate, LF, and what role does the α clamp play? Based on the fact that some α clamp mutants affect LF translocation more than LF_N translocation (Feld *et al.*, 2010), we hypothesize that the clamp plays an important role in unfolding later domains of LF, which lack a specific binding site for PA. In this model, the α clamp may be used repeatedly during translocation. Thus, if the α clamp is blocked, we would expect to see severe defects in translocation of LF.

In fact, this is exactly what we found with the $PA_{\alpha\text{-plug}}^{C30\text{-}C465}$ mutant; no significant LF translocation was observed under a 1.3-unit ΔpH and a 20-mV ΔΨ (Figure 4.8A). Under these conditions, WT PA translocates LF with a $t_{1/2}$ of about 80 seconds, and an efficiency of ~50% (Figure 4.8A). The fact that the $PA_{\alpha\text{-plug}}$ mutant does not translocate LF is a major finding in support of our hypothesis.

We also found that LF binding to $PA_{\alpha\text{-plug}}^{C30\text{-}C465}$ is defective. LF binds $PA_{\alpha\text{-plug}}^{C30\text{-}C465}$ about 10 times weaker than WT PA, and only blocks about 50% of the total conductance (Figure 4.8B). The partial conductance block could result from one of several reasons. First, heterogeneity in the channel or substrate population could allow some channels to be completely blocked by LF, while others remain open. Alternatively, the partial conductance block could result from a transient blockade of the channels. In this case, LF can completely block the channel, but does so only part of the time. Finally, (although unlikely) it is also possible that each LF only partially blocks the conductance of an individual channel. Single channel experiments could be used to distinguish between these possibilities and provide information about how LF initially interacts with the PA channel.

4.3 Discussion

PA's α clamp has been shown to be a non-specific binding site that plays an important role in substrate unfolding and PA oligomer assembly. Initial results indicate that it also plays a role in substrate translocation. Results presented here further support the hypothesis that there is a relationship between binding to the α clamp and translocation. We show that LF_N-PA binding can be disrupted by inserting proline residues into LF's α1/β1 region, or by sterically-occluding

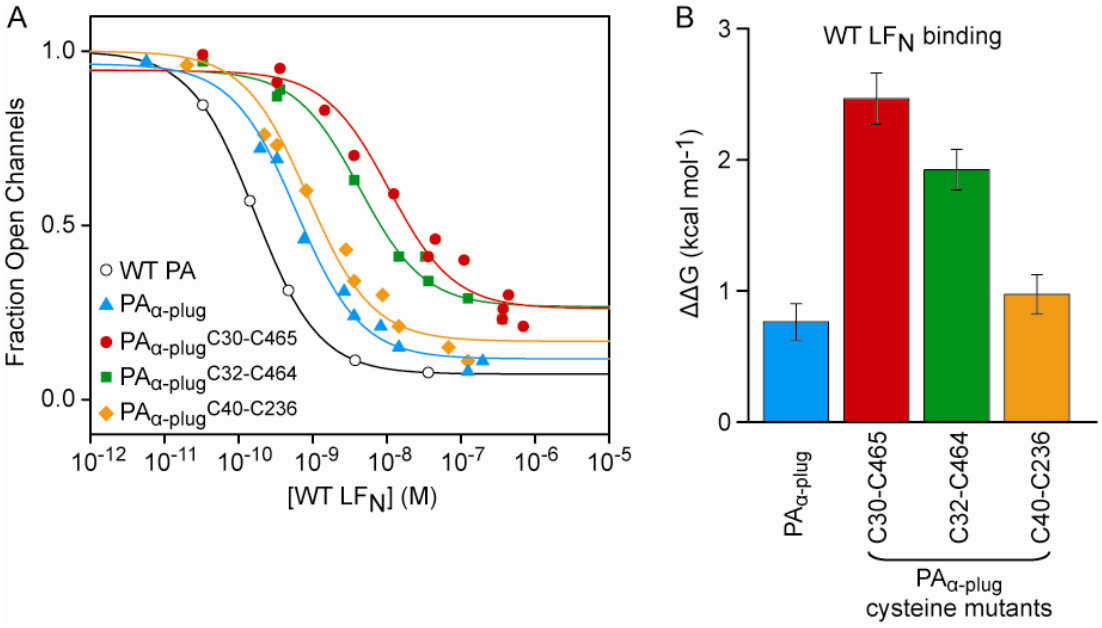

Figure 4.6. $PA_{\alpha\text{-plug}}$ mutants bind WT LF_N with less affinity than WT PA. (A) Equilibrium binding curves for WT PA and $PA_{\alpha\text{-plug}}$ mutants using WT LF_N. Curves were fit to a simple single-binding site model, $I = I_o/(1 + K_d/[L]) + c$, to obtain K_d values, where I is the current amplitude, $[L]$ is the LF_N concentration, and c is an offset. **(B)** Changes in binding free energy ($\Delta\Delta G$) relative to WT LF_N:WT PA, where $\Delta\Delta G = \Delta G_{MUT} - \Delta G_{WT}$. Positive $\Delta\Delta G$'s indicate that the mutant binds WT LF_N with less affinity than WT PA.

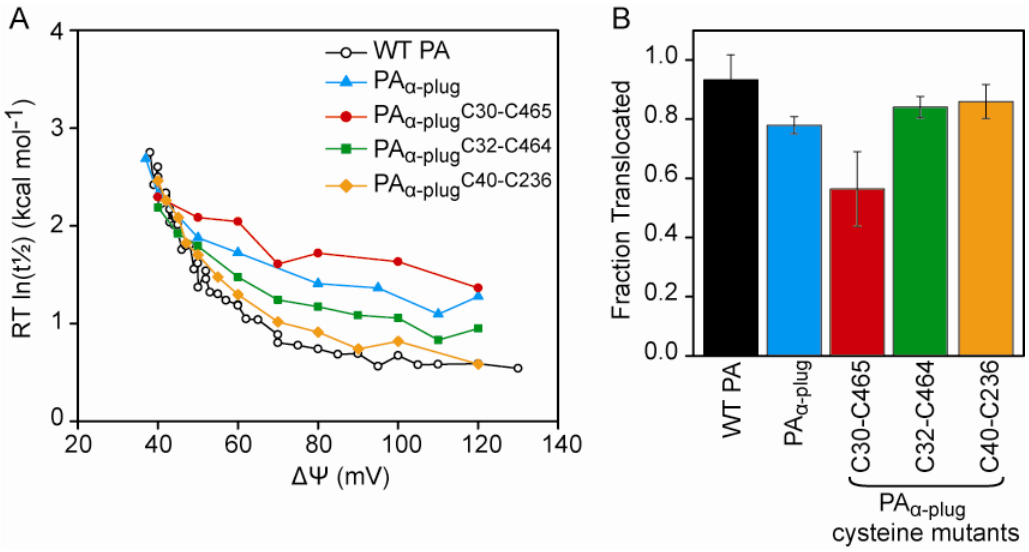

Figure 4.7. $PA_{\alpha\text{-plug}}$ mutants are defective in translocating LF_N. (A) A comparison of the ΔΨ dependence of the translocation activation energy ($\Delta G‡$) for WT PA and the PA-Nterm-LF mutants, where WT LF_N are used in each case. $\Delta G^{‡} = RT \ln t_{1/2} / c$. The $t_{1/2}$ value is the time for half of the protein to translocate, R is the gas constant, T is temperature, and c is a 1-sec reference constant. All LF_N translocation rates were measured at symmetrical pH 5.6. **(B)** The translocation efficiencies for these translocations at 60 mV. Efficiency (ε) is defined by $\varepsilon = A_{obs}/A_{exp}$, where A_{obs} is the observed amplitude of channels that reopened (or translocated), and A_{exp} is the expected amplitude if all of the channels reopened (or translocated).

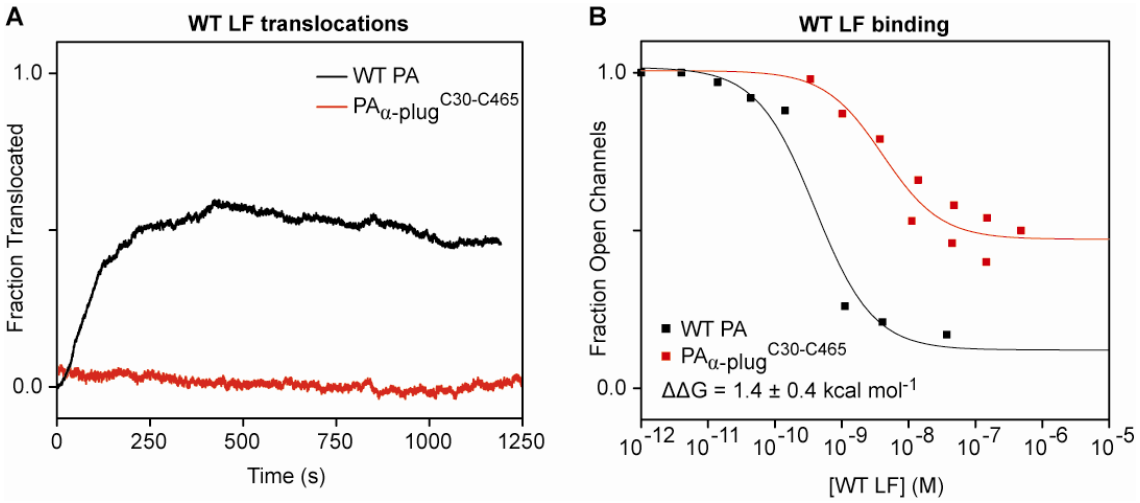

Figure 4.8. PA$_{\alpha\text{-plug}}^{C30\text{-}C465}$ is defective in LF binding and translocation. (A) LF translocation recordings using WT PA (black) and PA$_{\alpha\text{-plug}}^{C30\text{-}C465}$ (red). Translocation was measured using planar lipid bilayer electrophysiology under a 1.3-unit ΔpH (pH$_{cis}$ = 6.1, pH$_{trans}$ = 7.4) and a 20-mV ΔΨ. WT PA has a translocation efficiency of about 50%. In other words, about 50% of all the channels blocked by WT LF, are able to translocate their substrate. **(B)** Equilibrium binding curves for WT PA and the PA$_{\alpha\text{-plug}}^{C30\text{-}C465}$ mutant using WT LF. Curves were fit to a simple single-binding site model, $I = I_o/(1 + K_d/[L]) + c$, to obtain K_d values, where I is the current amplitude, $[L]$ is the LF$_N$ concentration, and c is an offset. The difference in binding free energy (ΔΔG) between PA$_{\alpha\text{-plug}}^{C30\text{-}C465}$ and WT PA is 1.4 (± 0.4) kcal mol^{-1}, where $\Delta\Delta G = \Delta G_{MUT} - \Delta G_{WT}$.

the α clamp, and these mutants also impair translocation. In addition, severe translocation defects were observed when LF was translocated using the $PA_{\alpha\text{-plug}}^{C30\text{-}C465}$ mutant. However, because translocation is a multi-step process, the mechanism by which these mutants disrupt translocation is still unclear. I suggest some preliminary hypotheses, but more experiments are needed in order to make solid conclusions about the role of the α clamp in translocation.

The nature of the α clamp/LF_N interaction.

Shape complementarity. The fact that the LF_N proline mutants disrupt binding to PA further supports the non-specific binding activity of the α clamp in which the shape of the α1/β1 region seems to be more important than specific functional groups in the sequence. When residues in LF_N's α1/β1 region were mutated to proline, the binding affinity for PA was about 10-100 times weaker than that of WT LF_N (Figure 4.3). Mutating residues in a different region of LF_N to proline had no effect on binding; LF_N Pro252-263 binds PA like WT LF_N even though it contains the same number of proline residues as LF_N Pro36-47 (Figure 4.3). In addition, the larger binding defects observed for LF_N Pro36-47 and LF_N Pro30-47 compared to LF_N Pro43-47 could indicate that the general exclusion of solvent accessible surface area of the α1 helix contributes more to the binding interaction than the β-sheet hydrogen bonds between Ile43 and Lys45 of LF and Leu203 of PA. While some additional control experiments are needed to localize the binding defect to the α clamp, evidence presented here and in Chapter 3 (Feld *et al.*, 2010) suggests that the α-helical shape of this region is important for binding to the α clamp.

Robustness of the α clamp interaction. Using a series of amino-terminal truncations of LF_N, we previously showed that the interaction between the α1/β1 region and the α clamp is worth about 2.5 kcal mol^{-1} (Feld *et al.*, 2010). Interestingly the $PA_{\alpha\text{-plug}}^{C30\text{-}C465}$ mutant binds WT LF_N about 2.5 kcal mol^{-1} weaker than WT PA. Thus, we believe that the $PA_{\alpha\text{-plug}}^{C30\text{-}C465}$ totally ablates LF_N binding at the α-clamp subsite.

On the other hand, the proline mutants do not disrupt binding by 2.5 kcal mol^{-1}. The worst LF_N proline mutant, LF_N Pro30-47, binds PA about 1.8 kcal mol^{-1} weaker than WT LF_N. These results suggest that the proline mutants are able to at least partially interact with the α-clamp site. They may be able to form some non-specific interactions with the site and gain some energy from burying solvent accessible surface area. In addition, there may be some flexibility to the α clamp/substrate interaction. In other words, these LF_N proline mutants may be able to reposition themselves in the α clamp in order to form the most favorable interactions with PA and thus compensate for the presence of the proline residues. The fact that these proline mutants are able to gain some binding energy illustrates the robustness of the α clamp's non-specific binding activity.

Relationship between α-clamp binding and translocation

Although the LF_N proline mutants and the $PA_{\alpha\text{-plug}}$ mutants both disrupt binding to the α clamp, they have very different effects on translocation. The LF_N proline mutants (LF_N Pro36-47 and LF_N Pro30-47) translocate much slower than WT LF_N at both high and low ΔΨ. In comparison, the $PA_{\alpha\text{-plug}}$ mutants only moderately impair translocation, and translocation is defective only at high voltages (Figure 4.7A). The inability of the LF_N proline mutants to bind to the clamp does not fully account for their translocation defects. Otherwise, we would expect to see the same translocation defects with the $PA_{\alpha\text{-plug}}$ mutants which also disrupt binding to the α clamp. So, why do the proline mutants impair translocation so much more than the $PA_{\alpha\text{-plug}}$ mutants? And why do the $PA_{\alpha\text{-plug}}$ mutants only affect translocation at high voltages?

LF_N proline mutants may disrupt downstream translocation steps. We hypothesize that the LF_N proline mutants are so defective at translocating because they disrupt later translocation steps. Thus, the observed translocation defects could result from a combination of disrupting downstream translocation events in addition to disrupting binding to the α clamp. It is also possible that the defects observed for these LF_N proline mutants are unrelated to their inability to bind the α clamp. Certainly, more work is needed in order to tease out the α clamp's role in translocation, and various experiments are discussed below.

In order to determine if the translocation defects observed for the LF_N proline mutants are related to their inability to bind to the α clamp, we plan to mutate a different region of LF_N to proline besides the α1/β1 region. If these new proline mutants do not disrupt binding to the α clamp, but still impair translocation, we can assume that the proline residues are disrupting a downstream translocation step, and that the translocation defects are not related to the α clamp. However, if these mutants translocate like WT LF_N, we can infer that the translocation defects observed with LF_N Pro36-47 and LF_N Pro30-47 are due to disrupting binding at the α clamp. Although we found that LF_N Pro252-263 binds and translocates like WT LF_N, the proline residues in this mutant are located at the very end of the LF_N and may not have much of an effect on translocation once the most of the protein has been translocated. Therefore, we would like to test additional proline mutants, especially ones that contain proline residues upstream of the α1/β1 sequence in LF_N's unstructured amino terminus (residues 1-29).

Even if downstream steps are disrupted by the LF_N proline mutants, understanding their effect on translocation may still help us tease out the α clamp's role in translocation, as well as provide information about other translocation steps. There are several hypotheses for how the proline mutants may disrupt translocation. First, it is possible that the LF_N proline mutants are unable to interact with the ϕ clamp, a ring of phenylalanine residues that forms a narrow constriction in the PA channel. This site has been shown to play a role in substrate unfolding (Thoren *et al.*, 2009) and translocation (Krantz *et al.*, 2005), perhaps by binding to hydrophobic and aromatic regions of a substrate protein and thus stabilizing unfolded intermediates. Proline residues, which are considered to be moderately polar (Kyte & Doolittle, 1982), may not be able to interact well with the ϕ clamp. As discussed in Chapter 2, disrupting the interaction with this site could effectively stabilize the substrate. In fact, when ϕ-clamp mutants are used to translocate LF_N, or when LF_N is translocated under higher symmetrical pHs, which are thermodynamically stabilizing (Krantz *et al.*, 2004), translocation is impaired at both high and low ΔΨ (Krantz *et al.*, 2006). Because the LF_N proline mutants affect translocation in the same way, we hypothesize that they are effectively stabilized. The inability to bind the ϕ clamp, or a combination of both the α clamp and the ϕ clamp, could lead to this stabilization and the observed translocation defects.

To test if the LF_N proline mutants disrupt the interaction with the ϕ clamp, we could measure the translocation of these mutants using PA ϕ-clamp mutants. Double mutant cycle analysis (Horovitz, 1996), as described in Chapter 2, could reveal an interaction between the ϕ clamp and the α1/β1 region. If the LF_N proline mutants don't interact well with the ϕ clamp, mutations in the ϕ clamp would not impair LF_N proline mutant translocation as much as they would impair WT LF_N translocation.

The LF_N proline mutants may also be defective in translocation because several charged residues, which are important for translocation (Brown *et al.*, 2011), have been mutated in these constructs. To test this possibility, I will be measuring the binding and translocation of LF_N Ser36-47, where residues 36-47 of LF_N have been mutated to serine. Like LF_N Pro36-47, this

mutant lacks the charged residues that could be important for translocation. However, unlike LF_N Pro36-47, this mutant is expected to maintain a helical shape in the α1/β1 region. If LF_N Ser36-47 binds and translocates like WT LF_N, we can infer that the mutating these charged residues is not the source of the translocation defects. Another way to test this hypothesis is restore the charged residues into the LF_N Pro36-37 or LF_N Pro30-47 backgrounds and see if translocation improves.

Finally, we hypothesize that the LF_N proline mutants disrupt translocation because the proline residues disrupt α helices, and thus disrupt the helical dipole. As discussed in Chapter 2, the positive membrane potential acts upon positive charge in the translocating substrate to drive unfolding and translocation. We assumed this charge is provided by basic side chains, but α helices also have a dipole moment that is positive at the amino terminus. Thus, it is possible that substrates translocate as α helices because the helical dipole provides charge for the positive voltage to act on. Because proline residues disrupt α helices, the LF_N proline mutants would have a reduced dipole moment, which could lead to impaired translocation.

In order to test this hypothesis, we will first use peptides to see if a positive dipole is important for translocation. By using peptides that contain a single cysteine residue either at their N- or C-terminus, we can create a series of dipeptides that have a variety of dipole orientations. For example, when crosslinked, the peptide that contains a cysteine at its C-terminus will produce a dipeptide where the first peptide is oriented N to C (positive dipole), and the second peptide is oriented C to N. Thus, the dipole of the second peptide is inverted with respect to the first. By comparing how these orientations affect translocation of the peptides, we hope to determine whether the backbone dipole is used in the translocation mechanism.

The role of the α clamp in reducing diffusive motion. The translocation defects for the LF_N proline mutants may result from a combination of effects and more work is need to tease out the role of the α clamp. On the other hand, the $PA_{\alpha\text{-plug}}$ mutants may provide a more direct way of determining the clamp's role in translocation. In addition, these mutants may eventually prove to be more valuable in understanding the role of the α clamp because they permanently block binding to site, whereas the LF_N proline mutants only disrupt the initial, one-time binding event.

If we look at the results from the $PA_{\alpha\text{-plug}}$ mutants, we notice that WT LF_N translocation is only defective at high voltages; at low voltages, these mutants translocate WT LF_N at about the same rate as WT PA (Figure 4.7A). We previously showed that under low driving force, substrate unfolding is rate limiting; at high driving forces, a second barrier, which is believed to be associated with translocation of an unfolded polypeptide chain, becomes rate limiting (Thoren *et al.*, 2009). Because the $PA_{\alpha\text{-plug}}$ mutants only impair LF_N translocation at high driving force, we hypothesize that the α clamp reduces the second translocation barrier by binding to the substrate after it has unfolded, thus reducing its diffusive motion and allowing for efficient translocation.

The role of the α clamp in substrate unfolding. The fact that the $PA_{\alpha\text{-plug}}$ mutants only impair translocation at high driving force (and translocate like WT PA at low driving force) may also mean that the α clamp does not play a significant role in LF_N unfolding. This interpretation seems to be in conflict with prior results which show that binding to the α clamp helps unfold and destabilize LF_N (Feld *et al.*, 2010). However, the crystal structure, thermodynamic stability measurements, and binding measurements were obtained at pH 7.4-7.5. Our translocation experiments are conducted at pH 5.6, a condition that has been shown to destabilize LF_N (Krantz *et al.*, 2004). The different conditions could explain the seemingly conflicting results about the α clamp's role in LF_N unfolding. Under the low pH conditions used for translocation, it is possible

that the barrier to unfolding LF_N's α1/β1 sequence is reduced, and the α clamp may not be needed to stabilize this unfolded intermediate. Thus, the α clamp may not be very critical in the unfolding of LF_N, at least under these low pH conditions. To test this hypothesis, we will translocate LF_N with these $PA_{\alpha\text{-plug}}$ mutants under higher pH conditions, where the substrate would be more thermodynamically stable. If the α clamp does play a role in LF_N unfolding, we should see larger defects with the $PA_{\alpha\text{-plug}}$ mutants relative to WT PA under higher pH conditions.

Although the α clamp may not play a role in unfolding LF_N, the site may be critical for unfolding and translocating the remaining domains of LF, which lack a specific binding site for PA. Based on the findings that some α clamp mutants affect LF translocation more than LF_N translocation (Feld *et al.*, 2010), and that no appreciable LF translocation occurred with the $PA_{\alpha\text{-plug}}^{C30\text{-}C465}$ mutant (Figure 4.8A), we conclude that the α clamp catalyzes LF translocation. We hypothesize that the α clamp facilitates LF unfolding and translocation by non-specifically binding its later domains. In our model, the α clamp may be used repeatedly during translocation, and binding to the site would allow the driving force to be efficiently applied to the substrate, stabilize unfolded intermediates, and minimize diffusive motion.

General model for protein translocation

Non-specific binding sites within protein translocases and degradation machines may play a critical role in substrate unfolding and translocation. Here we show that when binding to the α clamp is disrupted, major translocation defects are observed. We hypothesize that the α clamp's primary role in LF_N translocation is to reduce the configurational entropy of the substrate after it unfolds. However, we believe that the clamp plays a more critical role in unfolding the remaining domains of LF.

Future work should address how LF unfolds during translocation and continue to examine the role of the α clamp. What is the rate-limiting structure to LF unfolding? Do later parts of LF interact with the α clamp? If so, is this interaction is important for unfolding? Does the α clamp work in conjunction with another binding site in the channel (such as the ϕ clamp) to catalyze substrate unfolding and translocation? We expect that general features of the anthrax toxin system will be applicable to other protein translocases and degradation machines. By addressing these questions, we ultimately aim to understand the underlying mechanisms by which molecular machines unfold and translocate multidomain substrates.

4.4 Materials and Methods

Proteins. LF_N (residues 1-263 of LF) and mutants thereof were purified from overexpressing bacteria using standard Ni^{2+}-nitrilotriacetic-acid-(NTA)-affinity chromatography and Q-sepharose, anion-exchange chromatography (GE Healthcare, USA) (Krantz *et al.*, 2006). Their six-histidine (His_6) tags were removed with bovine α-thrombin treatment (0.5 units/mg of protein) for 30 minutes at room temperature in 20 mM Tris (pH 8), 150 mM NaCl, 2 mM $CaCl_2$ and 1 M dextrose.

The LF_N proline 43-47 mutant was made using a three-step, gene-synthesis procedure. Overlapping oligonucleotides encoding the desired sequences were synthesized (Elim Biopharmaceuticals, Inc., Hayward, CA) and amplified by two rounds of polymerase chain reaction (PCR). In Round I, 20 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 synthetic DNA fragments were ligated via a 5' Nde I site and 3' Sal I site into the pET15b-LF(Sal I) vector which contains an in-frame, silent Sal I restriction site in LF_N at V84 and D85. The synthetic LF_N constructs were purified and their His_6 tags were subsequently removed as described above. The other LF_N proline mutants were made by using multiple rounds of the Quikchange procedure (Zheng *et al.*, 2004).

WT PA and PA mutants were expressed and purified as described (Krantz *et al.*, 2005). The $PA_{\alpha\text{-plug}}$ mutants were made using the three-step, gene-synthesis procedure, as described above. Synthetic DNA fragments were ligated via a 5' Hind III site and 3' Kpn I site into the pET22b-PA(KpnI) vector which contains a silent Kpn I restriction site in PA at V175.

Planar lipid bilayer electrophysiology. Planar lipid bilayer currents were recorded using an Axopatch 200B amplifier (Molecular Devices Corp., Sunnyvale, CA) (Kintzer *et al.*, 2009, Thoren *et al.*, 2009). Membranes were painted on a 100 μm aperture of a 1-mL, white-Delrin cup with 3% 1,2-diphytanoyl-*sn*-glycerol-3-phosphocholine (Avanti Polar Lipids, Alabaster, AL) in *n*-decane. Cis (side to which the PA oligomer is added) and trans chambers were bathed in various buffers as required. By convention, $\Delta\Psi \equiv \Psi_{cis} - \Psi_{trans}$ ($\Psi_{trans} \equiv 0$ V), and $\Delta pH \equiv pH_{trans} - pH_{cis}$.

To monitor LF_N binding to the PA channel, we first inserted PA channels into a planar lipid bilayer bathed in asymmetric KCl solutions buffered in 10 mM potassium phosphate ($[\text{added KCl salt}]_{cis}$ = 100 mM, $[\text{added KCl salt}]_{trans}$ = 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 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, $I = I_o/(1 + K_d/[L]) + c$, to obtain the equilibrium dissociation constant, K_d, where I_o is the current amplitude and c is an offset.

All LF_N translocation experiments were carried out as described previously using a universal pH bilayer buffer system (UBB: 10 mM oxalic acid, 10 mM phosphoric acid, 10 mM MES, 1 mM EDTA, and 100 mM KCl) at a symmetrical pH 5.6 (Thoren *et al.*, 2009). Two to six replicate experiments were conducted for each mutant to establish the time (in seconds) for half of the substrate to translocate ($t_{1/2}$). The individual kinetic effects of LF_N mutations (MUT) were assessed by comparing the activation energy of translocation ($\Delta G‡$) at a specific $\Delta\Psi$ for the mutant and WT LF_N, where $\Delta G‡ = RT \ln t_{1/2} + c$. R is the gas constant, T is the temperature, and c is the natural log of 1 second. The change in $\Delta G‡$ ($\Delta\Delta G‡$) is reported at $\Delta\Delta G‡ = \Delta G‡(\text{WT}) - \Delta G‡(\text{MUT})$. Efficiency, ε, was also obtained from each translocation record by the relation, $\varepsilon = A_{obs}/A_{exp}$, where A_{obs} is the observed amplitude of channels that reopened (or translocated), and A_{exp} is the expected amplitude if all of the channels reopened (or translocated).

LF translocation experiments were carried out similarly except that a 1.3-unit ΔpH was also applied during translocation. The pH of the UBB in the cis and trans chambers was adjusted to apply the proton gradient, where pHcis = 6.1 and pHtrans = 7.4. The $\Delta\Psi$ was 20 mV. LF translocation was assessed by $t_{1/2}$ and ε as described for LF_N.

PA assembly. WT PA and $PA_{\alpha\text{-plug}}$ monomers at 2 mg/mL were treated with trypsin at a ratio of 1:1000 (wt/wt) for 15 minutes at room temperature and then inhibited with soybean trypsin inhibitor at 1:100 (wt/wt) and phenylmethylsulphonyl fluoride (PMSF) at 0.1 mM.

Protein products were separated on a 4-20% gradient gel under native conditions and stained with coomassie blue.

Electron microscopy. EM sample preparation and data analysis were done as previously described (Feld *et al.*, 2010). A total of 438 particles were used to calculate the oligomeric distribution produced by the $PA_{\alpha\text{-plug}}$ mutant.

References

Adams, P. D., K. Gopal, R. W. Grosse-Kunstleve, L. W. Hung, T. R. Ioerger, A. J. McCoy, N. W. Moriarty, R. K. Pai, R. J. Read, T. D. Romo, J. C. Sacchettini, N. K. Sauter, L. C. Storoni & T. C. Terwilliger, (2004) Recent developments in the PHENIX software for automated crystallographic structure determination. *J. Synchrotron Rad.* **11**: 53-55.

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

Agrawal, A. & B. Pulendran, (2004) Anthrax lethal toxin: a weapon of multisystem destruction. *Cell Mol. Life Sci.* **61**: 2859-2865.

Arora, N. & S. H. Leppla, (1993) Residues 1-254 of anthrax toxin lethal factor are sufficient to cause cellular uptake of fused polypeptides. *J. Biol. Chem.* **268**: 3334-3341.

Arora, N. & S. H. Leppla, (1994) Fusions of anthrax toxin lethal factor with shiga toxin and diphtheria toxin enzymatic domains are toxic to mammalian cells. *Infect. Immun.* **62**: 4955-4961.

Bai, Y., T. R. Sosnick, L. Mayne & S. W. Englander, (1995) Protein folding intermediates studied by native state hydrogen exchange. *Science* **269**: 192-197.

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

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

Blanke, S. R., J. C. Milne, E. L. Benson & R. J. Collier, (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.

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

Branden, C. & J. Tooze, (1991) *Introduction to Protein Structure*. Garland Publishing, Inc., New York, NY.

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

Brown, M., K. L. Thoren & B. A. Krantz, (2011) Charge requirements for proton gradient-driven transmembrane protein translocation. *J. Biol. Chem. in press*.

Burton, R. E., S. M. Siddiqui, Y. I. Kim, T. A. Baker & R. T. Sauer, (2001) Effects of protein stability and structure on substrate processing by the ClpXP unfolding and degradation machine. *EMBO J.* **20**: 3092-3100.

Carrion-Vazquez, M., H. Li, H. Lu, P. E. Marszalek, A. F. Oberhauser & J. M. Fernandez, (2003) The mechanical stability of ubiquitin is linkage dependent. *Nat. Struct. Biol.* **10**: 738-743.

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

Cheng, Y., (2009) Toward an atomic model of the 26S proteasome. *Curr. Opin. Struct. Biol.* **19**: 203-208.

Chou, Y.-T. & L. M. Gierasch, (2005) The conformation of a signal peptide bound by *Escherichia coli* preprotein translocase SecA. *J. Biol. Chem.* **280**: 32753-32760.

Christensen, K. A., B. A. Krantz & R. J. Collier, (2006) Assembly and disassembly kinetics of anthrax toxin complexes. *Biochemistry* **45**: 2380-2386.

Christensen, K. A., B. A. Krantz, R. A. Melnyk & R. J. Collier, (2005) Interaction of the 20 kDa and 63 kDa fragments of anthrax protective antigen: kinetics and thermodynamics. *Biochemistry* **44**: 1047-1053.

Clackson, T. & J. A. Wells, (1995) A hot spot of binding energy in a hormone-receptor interface. *Science* **267**: 383-386.

Cunningham, K., D. B. Lacy, J. Mogridge & R. J. Collier, (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.

Davis, I. W., A. Leaver-Fay, V. B. Chen, J. N. Block, G. J. Kapral, X. Wang, L. W. Murray, W. B. Arendall, 3rd, J. Snoeyink, J. S. Richardson & D. C. Richardson, (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* **35**: W375-383.

Decker, T. & M. L. Lohmann-Matthes, (1988) A quick and simple method for the quantitation of lactate dehydrogenase release in measurements of cellular cytotoxicity and tumor necrosis factor (TNF) activity. *J. Immunol. Methods* **115**: 61-69.

De Los Rios, P., A. Ben-Zvi, O. Slutsky, A. Azem & P. Goloubinoff, (2006) Hsp70 chaperones accelerate protein translocation and the unfolding of stable protein aggregates by entropic pulling. *Proc. Natl Acad. Sci. U.S.A.* **103**: 6166-6171.

Doyle, S. M. & S. Wickner, (2009) Hsp104 and ClpB: protein disaggregating machines. *Trends Biochem. Sci.* **34**: 40-48.

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

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

Eilers, M. & G. Schatz, (1986) Binding of a specific ligand inhibits import of a purified precursor protein into mitochondria. *Nature* **322**: 228-232.

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

Englander, S. W., L. Mayne & J. N. Rumbley, (2002) Submolecular cooperativity produces multi-state protein unfolding and refolding. *Biophys. Chem.* **101-102**: 57-65.

Ezzell, J. W. & T. G. Abshire, (1992) Serum protease cleavage of *Bacillus anthracis* protective antigen. *J. Gen. Microbiol.* **138**: 543-549.

Falnes, P. O., S. Choe, I. H. Madshus, B. A. Wilson & S. Olsnes, (1994) Inhibition of membrane translocation of diphtheria toxin A-fragment by internal disulfide bridges. *J. Biol. Chem.* **269**: 8402-8407.

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

Frank, J., M. Radermacher, P. Penczek, J. Zhu, Y. Li, M. Ladjadj & A. Leith, (1996) SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. *J. Struct. Biol.* **116**: 190-199.

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

Gelis, I., A. M. J. J. Bonvin, D. Keramisanou, M. Koukaki, G. Gouridis, S. Karamanou, A. Economou & C. G. Kalodimos, (2007) Structural basis for signal-sequence recognition by the translocase motor SecA as determined by NMR. *Cell* **131**: 756-769.

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

Glynn, S. E., A. Martin, A. R. Nager, T. A. Baker & R. T. Sauer, (2009) Structures of asymmetric ClpX hexamers reveal nucleotide-dependent motions in a AAA+ protein-unfolding machine. *Cell* **139**: 744-756.

Goloubinoff, P. & P. D. L. Rios, (2007) The mechanism of Hsp70 chaperones: (entropic) pulling the models together. *Trends Biochem. Sci.* **32**: 372-380.

Gottesman, S., (2003) Proteolysis in bacterial regulatory circuits. *Annu. Rev. Cell Dev. Biol.* **19**: 565-587.

Gottesman, S. & M. R. Maurizi, (1992) Regulation by proteolysis: energy-dependent proteases and their targets. *Microbiol. Mol. Biol. Rev.* **56**: 592-621.

Haug, G., C. Wilde, J. Leemhuis, D. K. Meyer, K. Aktories & H. Barth, (2003) Cellular uptake of Clostridium botulinum C2 toxin: membrane translocation of a fusion toxin requires unfolding of its dihydrofolate reductase domain. *Biochemistry* **42**: 15284-15291.

Hershko, A. & A. Ciechanover, (1998) The ubiquitin system. *Annu. Rev. Biochem.* **67**: 425-479.

Hinnerwisch, J., W. A. Fenton, K. J. Furtak, G. W. Farr & A. L. Horwich, (2005) Loops in the central channel of ClpA chaperone mediate protein binding, unfolding, and translocation. *Cell* **121**: 1029-1041.

Horovitz, A., (1996) Double mutant cycles; a powerful tool for analyzing protein structure and function. *Folding Des.* **1**: R121-126.

Hoskins, J. R., S. Sharma, B. K. Sathyanarayana, S. Wickner & H. Arthur, (2001) Clp ATPases and their role in protein unfolding and degradation. In: Advances in Protein Chemistry. Academic Press, pp. 413-420, IN419, 421-429.

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

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

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

Kenniston, J. A., T. A. Baker, J. M. Fernandez & R. T. Sauer, (2003) Linkage between ATP consumption and mechanical unfolding during the protein processing reactions of an AAA+ degradation machine. *Cell* **114**: 511-520.

Kenniston, J. A., R. E. Burton, S. M. Siddiqui, T. A. Baker & R. T. Sauer, (2004) Effects of local protein stability and the geometric position of the substrate degradation tag on the efficiency of ClpXP denaturation and degradation. *J. Struct. Biol.* **146**: 130-140.

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

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

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

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

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

Kyte, J. & R. F. Doolittle, (1982) A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**: 105-132.

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

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

Lacy, D. B., D. J. Wigelsworth, R. A. Melnyk, S. C. Harrison & R. J. Collier, (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.

Landry, S. J. & L. M. Gierasch, (1991) The chaperonin GroEL binds a polypeptide in an alpha-helical conformation. *Biochemistry* **30**: 7359-7362.

Lee, C., M. P. Schwartz, S. Prakash, M. Iwakura & A. Matouschek, (2001) ATP-dependent proteases degrade their substrates by processively unraveling them from the degradation signal. *Mol. Cell* **7**: 627-637.

Lee, V. T. & O. Schneewind, (2002) Yop fusions to tightly folded protein domains and their effects on Yersinia enterocolitica type III secretion. *J. Bacteriol.* **184**: 3740-3745.

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.

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.

Levchenko, I., R. A. Grant, J. M. Flynn, R. T. Sauer & T. A. Baker, (2005) Versatile modes of peptide recognition by the AAA+ adaptor protein SspB. *Nat. Struct. Mol. Biol.* **12**: 520-525.

Levchenko, I., R. A. Grant, D. A. Wah, R. T. Sauer & T. A. Baker, (2003) Structure of a delivery protein for an AAA+ protease in complex with a peptide degradation tag. *Mol. Cell* **12**: 365-372.

Li, Y., X. Gao & L. Chen, (2009) GroEL Recognizes an Amphipathic Helix and Binds to the Hydrophobic Side. *J. Biol. Chem.* **284**: 4324-4331.

Liu, Q., P. D'Silva, W. Walter, J. Marszalek & E. A. Craig, (2003) Regulated Cycling of Mitochondrial Hsp70 at the Protein Import Channel. *Science* **300**: 139-141.

Ludtke, S. J., P. R. Baldwin & W. Chiu, (1999) EMAN: semiautomated software for high-resolution single-particle reconstructions. *J. Struct. Biol.* **128**: 82-97.

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

Mabry, R., K. Brasky, R. Geiger, R. Carrion, Jr., G. B. Hubbard, S. Leppla, J. L. Patterson, G. Georgiou & B. L. Iverson, (2006) Detection of anthrax toxin in the serum of animals infected with *Bacillus anthracis* by using engineered immunoassays. *Clin. Vaccine Immunol.* **13**: 671-677.

MacDowell, A. A., R. S. Celestre, M. Howells, W. McKinney, J. Krupnick, D. Cambie, E. E. Domning, R. M. Duarte, N. Kelez, D. W. Plate, C. W. Cork, T. N. Earnest, J. Dickert, G. Meigs, C. Ralston, J. M. Holton, T. Alber, J. M. Berger, D. A. Agard & H. A. Padmore, (2004) Suite of three protein crystallography beamlines with single superconducting bend magnet as the source. *J. Synchrotron Rad.* **11**: 447-455.

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

Matouschek, A., (2003) Protein unfolding--an important process in vivo? *Curr. Opin. Struct. Biol.* **13**: 98-109.

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

Matouschek, A., J. T. Kellis, L. Serrano & A. R. Fersht, (1989) Mapping the transition state and pathway of protein folding by protein engineering. *Nature* **340**: 122-126.

Matouschek, A., N. Pfanner & W. Voos, (2000) Protein unfolding by mitochondria. The Hsp70 import motor. *EMBO Rep.* **1**: 404-410.

McPherson, A., Jr., (1976) The growth and preliminary investigation of protein and nucleic acid crystals for X-ray diffraction analysis. *Methods Biochem. Anal.* **23**: 249-345.

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

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

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

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

Milne, J. C., S. R. Blanke, P. C. Hanna & R. J. Collier, (1995) Protective antigen-binding domain of anthrax lethal factor mediates translocation of a heterologous protein fused to its amino- or carboxy-terminus. *Mol. Microbiol.* **15**: 661-666.

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

Mogridge, J., K. Cunningham & R. J. Collier, (2002) Stoichiometry of anthrax toxin complexes. *Biochemistry* **41**: 1079-1082.

Motohashi, K., Y. Watanabe, M. Yohda & M. Yoshida, (1999) Heat-inactivated proteins are rescued by the DnaK·J-GrpE set and ClpB chaperones. *Proc. Natl Acad. Sci. U.S.A.* **96**: 7184-7189.

Nassi, S., R. J. Collier & A. Finkelstein, (2002) PA_{63} channel of anthrax toxin: an extended β-barrel. *Biochemistry* **41**: 1445-1450.

Navon, A. & A. Ciechanover, (2009) The 26 S proteasome: from basic mechanisms to drug targeting. *J. Biol. Chem.* **284**: 33713-33718.

Neupert, W. & M. Brunner, (2002) The protein import motor of mitochondria. *Nat. Rev. Mol. Cell Biol.* **3**: 555-565.

Oberhauser, A. F. & M. Carrion-Vazquez, (2008) Mechanical biochemistry of proteins one molecule at a time. *J. Biol. Chem.* **283**: 6617-6621.

Otwinowski, Z. & W. Minor, (1997) Processing of X-ray diffraction data collected in oscillation mode. In: Methods in Enzymology. C. W. Carter Jr. & R. M. Sweet (eds). New York: Academic Press, Inc., pp. 307-326.

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

Paunola, E., T. Suntio, E. Jamsa & M. Makarow, (1998) Folding of active beta-lactamase in the yeast cytoplasm before translocation into the endoplasmic reticulum. *Mol. Biol. Cell* **9**: 817-827.

Pentelute, B. L., O. Sharma & R. J. Collier, Chemical dissection of protein translocation through the anthrax toxin pore. *Angew. Chem. Int. Ed.* **50**: 2294-2296.

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

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

Pfanner, N. & A. Geissler, (2001) Versatility of the mitochondrial protein import machinery. *Nat. Rev. Mol. Cell Biol.* **2**: 339-349.

Rassow, J., F.-U. Hartl, B. Guiard, N. Pfanner & W. Neupert, (1990) Polypeptides traverse the mitochondrial envelope in an extended state. *FEBS Letters* **275**: 190-194.

Ruprecht, M., T. Bionda, T. Sato, M. S. Sommer, T. Endo & E. Schleiff, On the impact of precursor unfolding during protein import into chloroplasts. *Mol. Plant* **3**: 499-508.

Rybak, S. L., F. Lanni & R. F. Murphy, (1997) Theoretical considerations on the role of membrane potential in the regulation of endosomal pH. *Biophys. J.* **73**: 674-687.

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

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

Schatz, G. & B. Dobberstein, (1996) Common principles of protein translocation across membranes. *Science* **271**: 1519-1526.

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

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

Smith, H. & J. Keppie, (1954) Observations on experimental anthrax: demonstration of a specific lethal factor produced in vivo by *Bacillus anthracis*. *Nature* **173**: 689.

Sonawane, N. D., J. R. Thiagarajah & A. S. Verkman, (2002) Chloride concentration in endosomes measured using a ratioable fluorescent Cl- indicator: evidence for chloride accumulation during acidification. *J. Biol. Chem.* **277**: 5506-5513.

Sonnhammer, E. L. & V. Hollich, (2005) Scoredist: a simple and robust protein sequence distance estimator. *BMC Bioinf.* **6**: 108.

Stark, H., F. Mueller, E. V. Orlova, M. Schatz, P. Dube, T. Erdemir, F. Zemlin, R. Brimacombe & M. van Heel, (1995) The 70S *Escherichia coli* ribosome at 23 A resolution: fitting the ribosomal RNA. *Structure* **3**: 815-821.

Storoni, L. C., A. J. McCoy & R. J. Read, (2004) Likelihood-enhanced fast rotation functions. *Acta Crystallogr. D Biol. Crystallogr.* **60**: 432-438.

Thompson, J. D., D. G. Higgins & T. J. Gibson, (1994) CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**: 4673-4680.

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

Ungermann, C., W. Neupert & D. M. Cyr, (1994) The role of Hsp70 in conferring unidirectionality on protein translocation into mitochondria. *Science* **266**: 1250-1253.

Van den Berg, B., W. M. Clemons, Jr., I. Collinson, Y. Modis, E. Hartmann, S. C. Harrison & T. A. Rapoport, (2004) X-ray structure of a protein-conducting channel. *Nature* **427**: 36-44.

Van Dyke, R. W. & J. D. Belcher, (1994) Acidification of three types of liver endocytic vesicles: similarities and differences. *Am. J. Physiol.* **266**: C81-94.

van Heel, M., G. Harauz, E. V. Orlova, R. Schmidt & M. Schatz, (1996) A new generation of the IMAGIC image processing system. *J. Struct. Biol.* **116**: 17-24.

Vu, N. D., H. Feng & Y. Bai, (2004) The folding pathway of barnase: the rate-limiting transition state and a hidden intermediate under native conditions. *Biochemistry* **43**: 3346-3356.

Walker, D., A. M. Chaddock, J. A. Chaddock, L. M. Roberts, J. M. Lord & C. Robinson, (1996) Ricin A chain fused to a chloroplast-targeting signal is unfolded on the chloroplast surface prior to import across the envelope membranes. *J. Biol. Chem.* **271**: 4082-4085.

Wang, J., J. J. Song, M. C. Franklin, S. Kamtekar, Y. J. Im, S. H. Rho, I. S. Seong, C. S. Lee, C. H. Chung & S. H. Eom, (2001) Crystal structures of the HslVU peptidase-ATPase complex reveal an ATP-dependent proteolysis mechanism. *Structure* **9**: 177-184.

Wang, Z., H. Feng, S. J. Landry, J. Maxwell & L. M. Gierasch, (1999) Basis of substrate binding by the chaperonin GroEL. *Biochemistry* **38**: 12537-12546.

Wesche, J., J. L. Elliott, P. O. Falnes, S. Olsnes & R. J. Collier, (1998) Characterization of membrane translocation by anthrax protective antigen. *Biochemistry* **37**: 15737-15746.

White, H. E., H. R. Saibil, A. Ignatiou & E. V. Orlova, (2004) Recognition and separation of single particles with size variation by statistical analysis of their images. *J. Mol. Biol.* **336**: 453-460.

Wickner, S., M. R. Maurizi & S. Gottesman, (1999) Posttranslational Quality Control: Folding, Refolding, and Degrading Proteins. *Science* **286**: 1888-1893.

Wickner, W. & R. Schekman, (2005) Protein translocation across biological membranes. *Science* **310**: 1452-1456.

Wigelsworth, D. J., B. A. Krantz, K. A. Christensen, D. B. Lacy, S. J. Juris & R. J. Collier, (2004) Binding stoichiometry and kinetics of the interaction of a human anthrax toxin receptor, CMG2, with protective antigen. *J. Biol. Chem.* **279**: 23349-23356.

Wilcox, A. J., J. Choy, C. Bustamante & A. Matouschek, (2005) Effect of protein structure on mitochondrial import. *Proc. Natl Acad. Sci. U.S.A.* **102**: 15435-15440.

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

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

Zhang, S., A. Finkelstein & R. J. Collier, (2004a) 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.

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

Zheng, L., U. Baumann & J. L. Reymond, (2004) An efficient one-step site-directed and site-saturation mutagenesis protocol. *Nucleic Acids Res.* **32**: e115.

Zimmer, J., Y. Nam & T. A. Rapoport, (2008) Structure of a complex of the ATPase SecA and the protein-translocation channel. *Nature* **455**: 936-943.

Zolkiewski, M., (1999) ClpB cooperates with DnaK, DnaJ, and GrpE in suppressing protein aggregation. *J. Biol. Chem.* **274**: 28083-28086.