Model of the toxic complex of anthrax: Responsive conformational changes in both the lethal factor and the protective antigen heptamer

Florence Tama, Gang Ren, Charles L. Brooks, III and Alok K. Mitra

References	This article cites 57 articles, 18 of which can be accessed free at: http://www.proteinscience.org/cgi/content/full/15/9/2190#References
Email alerting service	Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here**

Notes

To subscribe to *Protein Science* go to:
http://www.proteinscience.org/subscriptions/

Model of the toxic complex of anthrax: Responsive conformational changes in both the lethal factor and the protective antigen heptamer

FLORENCE TAMA,[1,4] GANG REN,[2,5] CHARLES L. BROOKS III,[1] AND ALOK K. MITRA[3]

[1]Department of Molecular Biology, The Scripps Research Institute, La Jolla, California 92037, USA
[2]Department of Cell Biology, The Scripps Research Institute, La Jolla, California 92037, USA
[3]School of Biological Sciences, University of Auckland, Auckland 1020, New Zealand

Abstract

The toxic complex of anthrax is formed when the monomeric protective antigen (PA) (83 kDa), while bound to its cell-surface receptor, is first converted to PA63 heptamers (PA63h) following N-terminal proteolytic cleavage, and then lethal (LF) (90 kDa) or edema factor (EF) binds to the heptamer. We report a "pseudoatomic" model for the complex of PA63h and full-length LF determined by applying the normal-mode flexible fitting procedure to a ~18 Å cryo-electron microscopy (EM) density map of the complex. The model describes the interacting surface that buries a total area of ~10,140 $Å^2$ comprising ~40% charged, and ~30% each of polar and hydrophobic residues. For the heptamer, the buried surface, composed of ~110 residues, involves primarily three monomers and includes for two, similar stretches of the polypeptide chain from domain 1. For LF, the interface again involves ~110 residues, mostly from the N-terminal domain I (LF_N), and the structurally homologous C-terminal domain IV. Most interestingly, bound LF displays a marked conformational change resulting from a "collapse" of domains I, III, and IV on domain II, with the largest movement of ~9 Å noted for domain I. On the other hand, primarily, rigid-body movements, larger than ~10 Å for three PA63 monomers, cause the hourglass-shaped heptamer lumen to enlarge by as much as ~50% near the middle of the molecule. Such concerted structural rearrangements in LF and the heptamer can facilitate ingress of the ligand into the heptamer lumen prior to unfolding and release through the PA63h channel formed in the acidic late endosomal membrane.

Keywords: anthrax toxin complex; lethal factor; cryo-EM; normal mode flexible fitting; conformational change

Anthrax is caused by one of the two virulence factors produced by the Gram-positive bacteria *Bacillus anthracis*. This factor is a cocktail of three monomeric proteins, lethal factor (LF) (90 kDa), edema factor (EF) (89 kDa), and the protective antigen (PA) (83 kDa). The protective antigen upon binding to a host-cell surface receptor (ATR/TEM8 or CMG2) (Bradley et al. 2001; Scobie et al. 2003) is cleaved at the N terminus, which results in spontaneous oligomerization into a heptamer (PA63h) (Milne et al. 1994). The binding of enzymic subunits lethal factor and/or edema factor to PA63h leads to the formation of the toxic entity. The complex of the PA63h–ligand and the seven bound receptor molecules is endocytosed in a clathrin-dependent process. Under the acidic condition of the endosome, PA63h inserts into the membrane bilayer to form a voltage-gated pore of minimum diameter of ~12 Å (Zhao et al. 1995), which ultimately translocates the ligand. In the case of LF, this

---

translocation has been shown to be vectorial, in an N- to C-terminal direction (Zhang et al. 2004a).

PA63h is a ring-shaped molecule that encloses a negatively charged lumen $\sim$20–35 Å in diameter (Petosa et al. 1997). Based on the study of Petosa et al. (1997), a model for membrane insertion, supported by biophysical studies (Benson et al. 1998; Nassi et al. 2002), has been proposed involving an amphipathic, 23–amino acid long, chymotrypsin-sensitive segment of domain 2 (β3–β4 loop). In this model, analogous to that in staphylococcal α-hemolysin, each amphipathic segment contributes two β-strands to form a 14-stranded, porin-like β-barrel channel (Benson et al. 1998). Low pH-induced protonation of key His residues in the β3–β4 loop is thought to trigger membrane insertion. More recently, the X-ray crystal structures of the complex of the recombinant soluble domain of the CMG2 receptor, and the monomeric PA determined at 2.5 Å resolution (Santelli et al. 2004) or the heptameric PA63h determined at 4.3 Å resolution (Lacy et al. 2004) revealed the aforementioned loop in an intimate interaction with the receptor in addition to the already known involvement of domain 4 in the receptor binding interface. This arrangement of the amphipathic loop is thought to prevent premature membrane insertion until exposure to acidic environment in the endosome.

We had demonstrated that the binding stoichiometry of full-length LF and PA63h is 1:1 by direct visualization of the complex by cryo-electron microscopy (EM) image analysis (Ren et al. 2004). Recently, Neumeyer et al. (2006) showed that the blockage of the PA63 heptamer ion channel by the full-length lethal factor is also consistent with such a 1:1 stoichiometry. On the other hand, earlier studies to characterize biochemically the complex of the lethal factor and PA63h instead utilized the recombinant N-terminal domain LF_N (Elliott et al. 2000; Mogridge et al. 2002a,b), presumably because it can also bind tightly to the heptamer and is able to translocate heterologous fusion proteins into the cytosol (Ballard et al. 1996; Goletz et al. 1997). Even though a possible maximum occupancy of 3 LF_N molecules per PA63h (Mogridge et al. 2002b) has been indicated, a recent study by Zhang et al. (2004b) showed that the efficiency of the translocation of LF_N is the same, irrespective of the amount (e.g., up to saturating levels) of ligand present. Site-directed mutagenesis of a stretch of residues in domain 1 of PA63, which, based on the 4.5 Å resolution X-ray structure of PA63h (Petosa et al. 1997), was postulated to harbor the ligand-binding site, demonstrated their role in ligand binding at the interface between adjacent PA63 monomers (Cunningham et al. 2002). Similar mutagenesis studies, directed by the X-ray crystal structure of LF, identified those cognate binding sites that are located on the N-terminal domains of lethal and edema factors (Lacy et al. 2002).

Studies with PA and the recombinant LF_N–DTA chimera, where DTA is the catalytic fragment of diphtheria toxin (DT), showed that delivery of the polypeptide into the cell was blocked when engineered DTA variants contained intramolecular disulfide links (Wesche et al. 1998). This supports the notion that, in general, the ligand in the A–B family of toxins undergoes some degree of unfolding en route to transit into the cytosol. Unfolding of the 90-kDa LF protein is necessary to facilitate translocation across the narrow 12 Å pore and, as shown by in vitro experiments on LF_N, could be affected by acid-induced unfolding in the endosome (Krantz et al. 2004).

We had earlier directly visualized the PA63h–LF complex from analysis of images of specimens suspended in vitrified buffer by cryo-EM (Ren et al. 2004) and single-particle image processing. To complement emerging information from modest-resolution cryo-EM experiment, one can use existing X-ray structural data to construct ''pseudoatomic'' models. This has been shown to be a useful approach in many instances (Volkmann and Hanein 1999; Wriggers et al. 1999; Jiang et al. 2001; Rossmann et al. 2001; Chacon and Wriggers 2002; Gao et al. 2003). In our earlier study (Ren et al. 2004), the docking of full-length LF and the seven PA63 monomers in the density map of the complex was carried out visually, using the program O. In the present study, we have arrived at an accurate model of the liganded complex using the flexible fitting procedure (Tama et al. 2004a,b) coupled with energy minimization to relieve short contacts. In the NMFF (Normal Mode Flexible Fitting) technique, a linear combination of low-frequency normal modes known to well represent intrinsic flexibilities of a molecule is used in an iterative manner to fit the modeled structure optimally into the low- to medium-resolution electron density map (Mitra et al. 2005). Using NMFF and the previously determined density map of the PA63h–LF complex (Ren et al. 2004), an optimally fitted model for the ligand-bound structure is constructed. We find that in the complex, the LF molecule displays a marked conformational change leading to a compaction, which is derived largely from a movement of the N-terminal domain I toward the C-terminal domain IV. In addition, upon complexation, rigid-body movements of the PA63 monomers cause the entrance and the volume of the hourglass-shaped heptamer lumen to enlarge. The details of the conformational change of full-length LF and the nature of the binding interface for the ligand and the heptamer are discussed vis-à-vis for their possible implication in the biological activity of anthrax toxin.

Results

In our earlier manually built model of the toxic complex (Ren et al. 2004), the orientations of the two protein

moieties (PA63 monomer and LF) defined by their respective X-ray structures were varied only as rigid bodies. Such a model (Fig. 1a), which had a cross-correlation coefficient (CC) of 0.77, was not optimal considering the resolution of the map and also that it failed to account for several volumes of density for the complex, especially those in the spatial location of LF. In the density map of the complex, the seven PA63 monomers can be visually easily positioned in their appropriate density volumes. But, because a rigid-body structure of LF does not completely account for its density, we examined whether a conformationally alternative structure of LF can be found that (1) is consistent both stereochemically as well as chemically in terms of what is known about the interaction between the ligand and the heptamer, and (2) leads to a more optimal fit with the cryo-EM density map.

Site-specific alanine substitutions at seven sites in LF_N (ASP182, ASP187, LEU188, TYR233, HIS229, LEU235, and TYR236) produce significant reduction in binding,

while alanine mutations for residues ARG178, LYS197, ARG200, PRO205, ILE207, ILE210, and LYS214 in PA63 inhibit ligand binding (Gupta et al. 2001; Chauhan and Bhatnagar 2002; Cunningham et al. 2002; Lacy et al. 2002). A constraint in building our model into the cryo-EM map (Ren et al. 2004) was that the binding interface between LF and PA63h should include or be proximal to the aforementioned residues.

Starting with the manually built model, a more optimal model that explains the cryo-EM density was constructed. For this purpose, it was necessary to first accurately isolate the density that corresponds to LF. In order to achieve this, we optimized the locations of each of the seven PA63 monomers in the model (see Materials and Methods) as follows: Using the manually fitted model for the PA63 heptamer in the complex, we generated a difference map to isolate, one at a time, the density for each of the seven monomers. We then performed rigid-body fitting of a PA63 monomer into these seven volumes of density, followed in each case by flexible fitting with
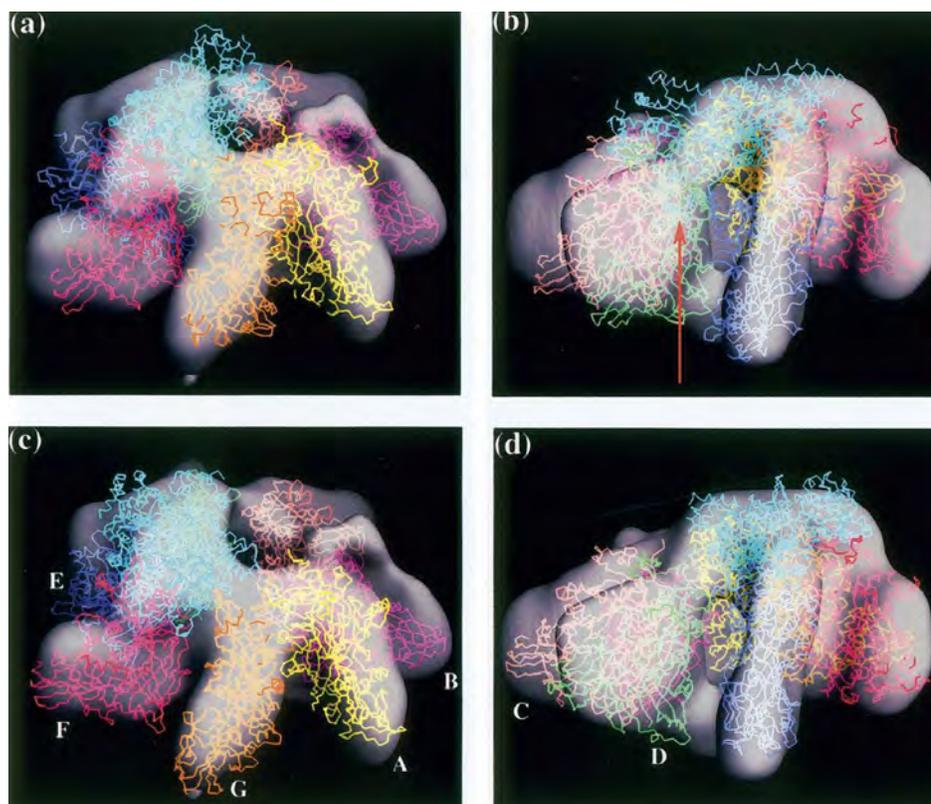

Figure 1. Pseudoatomic model of the complex of PA63 heptamer (PA63h) and LF fitted into the 18 Å resolution cryo-EM density map (Ren et al. 2004). Two views for the earlier manually built model (*a,b*) (Ren et al. 2004) and for the refined model (*c,d*) obtained from flexible fitting and consideration of mutagenesis data are shown. The PA63 monomers are labeled (Ren et al. 2004), as shown in *c* and *d*, and LF is colored in cyan. The red arrow indicates the region where steric clashes were present in the manually built model but are now absent in the flexibly fitted model. In addition, improved fitting of LF within the density envelope is apparent so that the overall cross-correlation coefficient for the final model is 0.91 compared to 0.77 for the manually built model. The figure was created using PYMOL (DeLano Scientific).

NMFF. After this initial procedure, a starting model for the heptamer was generated for the ligand-bound state. This model was then energy-minimized to improve the stereochemistry (see Materials and Methods). When compared to the crystal structure (Petosa et al. 1997), the root-mean-squared-deviation (RMSD) of a PA63 monomer ranged from ~2 Å to ~4 Å.

Next, a density map of the modeled heptamer was calculated (at ~18 Å resolution), which was then subtracted from the experimental density map of the complex (Ren et al. 2004) to isolate the density corresponding to bound LF. A rigid-body fitting of LF (Pannifer et al. 2001) was then performed (see Materials and Methods). The top five solutions generated for the orientation of LF were each subjected to the flexible fitting procedure. After this exercise, we found that the CC was rather similar for each of the models, but, in four of the five modeled structures, the locations of the domains of LF were such that the residues known to be involved in interaction with PA63 were spatially distant (>15 Å) as reflected by their closest C^α–C^α distance with residues in the PA63 monomers. Therefore, these structures were disregarded, and only the model that was consistent with residue proximity reflected by the mutagenesis data was examined for further optimization. In general, in cases where it is clear that the conformation indicated by the cryo-EM density map is similar to that in the crystal structure, rigid-body fitting will provide a set of very similar solutions with comparable CC values. However, in our particular case, where substantial conformational changes are apparent, the fitting procedure becomes more challenging so that conformationally distinct structures can result, again with very comparable values of CC. Therefore, it is necessary to exploit, if available, ancillary experimental data to adjudicate the most favored solution. In the model that survived, residues HIS229 and TYR236 in LF known to be important for binding were in close proximity to residues in some of the PA63 monomers (e.g., C^α–C^α distance was ≤7 Å). This model for the PA63h–LF complex was then energy-minimized in order to alleviate any steric clashes. We then performed a further round of flexible fitting on the whole PA63h–LF complex to improve the fitting to the cryo-EM density map. The CC for the final model is 0.91.

Model of the PA63h–LF complex

This current model shows minimal steric overlap between LF and PA63 monomers and displays several close contacts between residues of LF and PA63, defining the molecular interacting surface that is discussed further below. The binding of LF disrupts the sevenfold symmetry in the PA63 heptamer. In Figure 1a, the original, visually fitted model for the PA63h–LF complex (Ren et al. 2004) is shown. As can be seen, in this model, parts of LF are outside the experimentally derived density envelope. In addition, on closer examination, we found that this original model suffered from having several steric clashes between atoms of LF and PA63 monomers. In Figure 1b, the final "pseudoatomic" model generated using the protocol described above is also shown. As can be seen, a much better overall fit of the model, in particular the fit of the LF molecule within the density envelope, is now apparent. Overall, the PA63 monomers also show an improvement in their fit into the respective density envelopes compared to our earlier analysis where the monomers were fitted only as rigid bodies. The flexibly docked monomers have a RMSD between ~2 Å and 4 Å relative to the crystal structure. After least-squares alignment of the centers of masses of the seven chains in the modeled liganded structure of the heptamer and the unliganded crystal structure, <~5 Å displacements are seen for three of the chains (B, C, and D), whereas the rest of the chains show displacements ranging from ~8 Å to ~15 Å.

Description of the conformational change of LF

We observe that an optimal accommodation of the full-length LF in its density envelope necessarily translates to a significant conformational change. This involves global domain movements that are described in Figure 2. Compared to the 2.2 Å resolution crystal structure (Pannifer et al. 2001), the RMSD of the fitted LF after the NMFF procedure is 6.5 Å. Whereas in the crystal structure, LF is ~105 Å tall and ~70 Å wide at its "base," these values, while bound in the complex, are smaller, leading to a reduction in the accessible surface area by ~16% and a reduction of overall volume from ~24,240 A^3 to ~24,190 A^3. In particular, in two orthogonal directions, the size of the molecule is decreased by ~15 Å and by ~10 Å (see Fig. 2). The right panel of Figure 2 illustrates the details of the conformational rearrangements. Each arrow represents the direction and amplitude of motion of individual C^α atoms leading to the altered conformation. Large movements are seen for the N-terminal domain I (residues 1–245 equivalent to the recombinant LF_N). In particular, its center of mass translates by ~9 Å and the domain overall rotates by ~16° toward the central domain II (residues 263–297 and 385–550). Similar motions are observed for domain III and IV that rotate by ~23° and 11° and translate by ~8 Å and ~5 Å, respectively. On the other hand, domain II has a rotation of ~10° but a rather small translation of its center of mass (~1 Å), which indicates that upon binding of LF to PA63h, domains I, III, and IV "collapse" on domain II.

In order to get a better understanding of these domain movements, we examined correlation of the atomic displacements upon conformational rearrangements modeled for LF (see Fig. 3). The conformational change involves
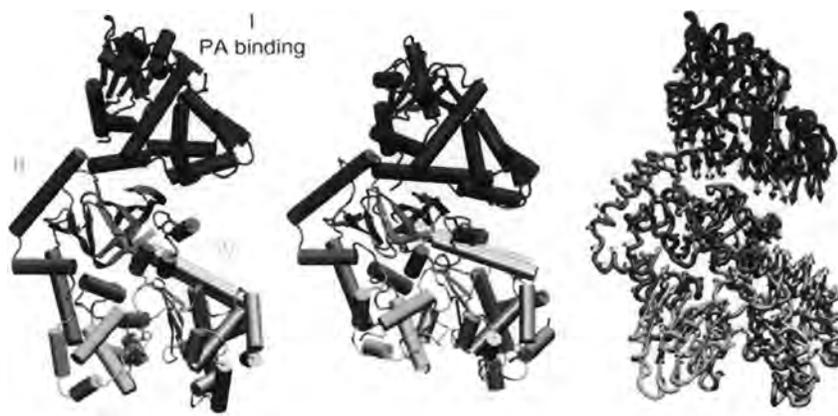
Figure 2. Conformational change in full-length LF upon binding to the PA63 heptamer. The *left* panel shows the X-ray structure (Pannifer et al. 2001) used as the starting model for flexible fitting that led to the model shown in the *center* panel, which fits the cryo-EM density optimally. In the *right* panel, the arrows define the directions and the amplitudes of displacements of all the C^{α} atoms leading to the modeled conformational change, which is characterized by the ''collapse'' of domains I, III, and IV onto domain II.

relative, correlated, and anti-correlated motions between the different domains of LF. In particular, anti-correlated motions of domain I relative to the three other domains are observed, while overall, domains II, III, and IV show correlated motions. Our analysis would suggest that upon binding with PA63 monomers at the ''top'' of the heptamer, long-range communication between the various domains of LF is triggered. This, in turn, can promote the necessary rearrangements, i.e., the observed compaction of LF.

Interactions between PA63h and LF

In the final refined model, LF interacts with several PA63 monomers, E, F, and G, and to a lesser extent, with monomer D. The monomer nomenclature follows that introduced earlier (Ren et al. 2004), as indicated in Figure 1. We observe a cluster of residues from LF that is known to be important for binding to be in close proximity of a cluster of several critical residues on PA63 monomers (Fig. 4). The full interaction surfaces of LF and the heptamer, delineated based on residues in PA63h and LF that are within 7 Å in the final model, are illustrated in Figure 5, and Table 1 details contiguous patches of amino acids constituting the buried interface. For both LF and PA63h, ~39% of the total residues in the interaction surface are charged; ~30% and 35% are hydrophobic, while ~31% and ~25% are polar, respectively. The total surface area buried for the two molecules in the complex is ~10,140A^2 calculated using CNS (Brünger et al. 1987). This surface area is expected to correspond to ~200 amino acid residues based on a value of ~50 $Å^2$ for the exposed area of a residue in a folded polypeptide (Wang and Chen 2003). Consistent with this, we find that the buried area involves ~110 residues of the ligand

interacting with an approximately equal number of residues mostly distributed on three PA63 monomers.

In the refined model, primarily, residues from domain I and structurally homologous domain IV of full-length LF (Pannifer et al. 2001) interact with residues of PA63. Thus, chain E interacts exclusively with residues in domain IV, chain F exclusively with residues in domain I, and chain G with residues in domain I and in domain II (Table 1). Overall, ensemble of residues in the same region of chain E and chain F (i.e., residues ~173–214 in domain 1) interact with domain IV and I of LF,
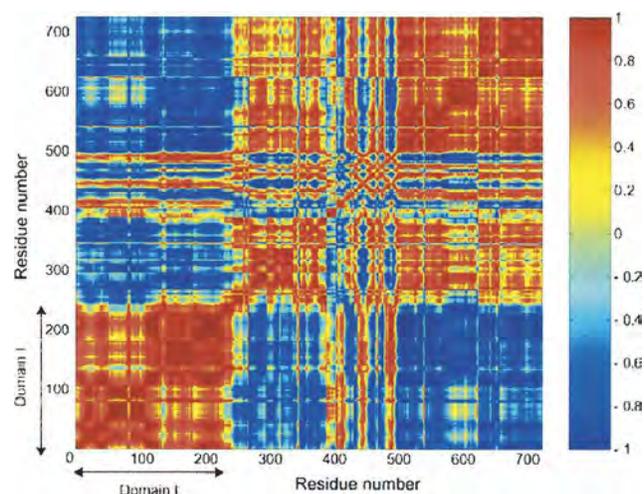

Figure 3. Correlation map for the displacement of the amino acid residues in the modeled conformational rearrangement in LF. For every pair of residues, the red (positive) and blue (negative) regions correspond to movements that are strongly correlated (i.e., displacements in the same direction) and strongly anti-correlated (i.e., displacements in the opposite directions), respectively.
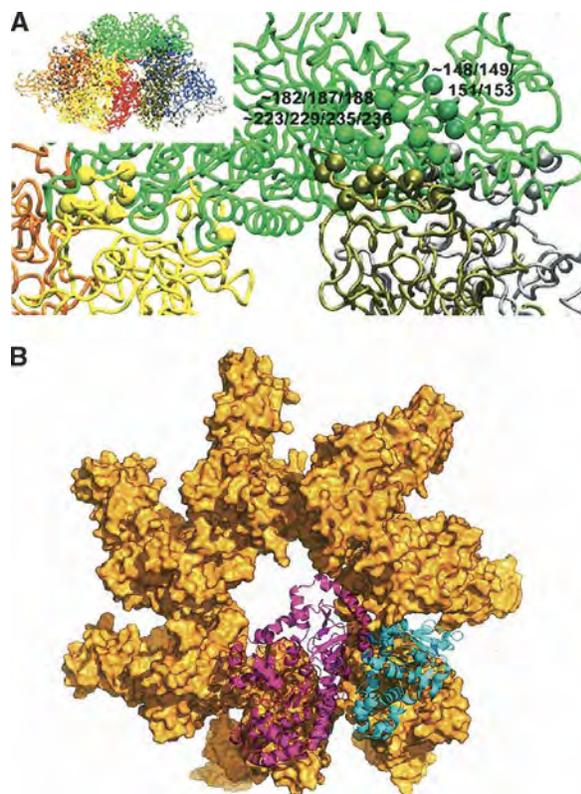
Figure 4. Interaction at the interface of full-length LF and PA63 heptamer in the model of the complex. (*A*) Residues in LF, which are suggested to be involved in binding to the heptamer from mutagenesis studies, are shown as green spheres. Also indicated are residues (colored spheres) in three of the PA63 monomers (E, F, and G as in Fig. 1) that are within 7 Å from residue(s) in LF. The *inset* shows the full complex in the given orientation. (*B*) A view of the atomic model from the "top" showing the disposition of the fitted molecule on the heptamer. Domain I of LF is shown in cyan; the rest is in purple. The figure was created using PYMOL (DeLano Scientific).

respectively. Residues in the stretch R178 to R200 of PA63, deemed important in binding to LF_N, are proximal to residues in domain I, while such residues in the stretch I207 to K214 are close to residues in domain I as well as domain IV of full-length LF. Interestingly, our model suggests that in addition to the residues in domain I of LF implicated in binding from mutagenesis studies mentioned above, residues ~38–42 and ~75–78 from domain I, as well as residues ~430–433 in domain IV, are in close proximity of several residues from the PA63 monomer (chain G) that are known to interact with LF (residues 195–201).

Discussion

In this study, we re-examined the cryo-EM density map of the complex of PA63h and full-length LF (Ren et al. 2004) using NMFF to arrive at an optimized "pseudoatomic"

model consistent with available results from site-directed mutagenesis studies and the cryo-EM density envelope. Even visually, the approximate locations of the seven PA63 monomers in the cryo-EM density map could be assigned unequivocally, as was followed in our earlier study (Ren et al. 2004). Careful analysis in this study, using rigid-body refinement followed by NMFF procedure led to a more accurate delineation of the heptamer structure in the complex. Comparison of the calculated map of the fitted heptamer in the complex and the experimental map of the complex yielded the density volume for the ligand by examining a difference map. This volume, as before (Ren et al. 2004), could account for one full-length LF and no more. A recent electrophysiological study of the anthrax protective ion channel (PA63 heptamer) in the presence of full-length LF has also indicated such a 1:1 stoichiometry (Neumeyer et al. 2006). Based on Figure 4B, one could argue, based on purely geometrical grounds, that there is enough space for another full-length LF to bind on "top" of the "unoccupied" PA63 monomers. However, the location of an additional LF straddling another set of three monomers appears to lead to short contacts between portions of the LF polypeptide chain overlaying the lumen entrance. The fact that, clearly, the density for only one LF molecule is seen from careful analysis of the cryo-EM map suggests that binding of additional LF is destabilizing, very likely because of steric clashes and/or due to transduced structural perturbations in the heptamer, possibly of the form observed in the current study.

Conformational change in LF

For many biological systems, observed conformational change can be represented by one of the low-frequency normal modes associated with that conformation (Tama
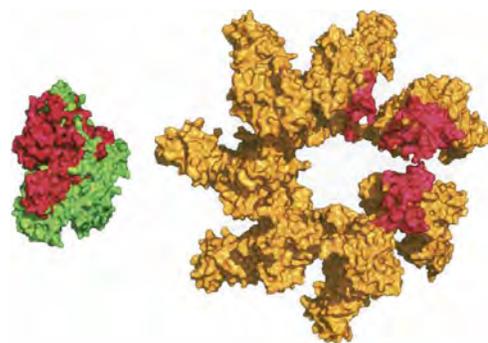
Figure 5. The interaction interface between LF and the PA63 heptamer represented by residues that are within C^α–C^α distance of 7 Å. The purple-colored region spread over three PA63 monomers (*right* panel) and that on LF (*left* panel) define the interaction surface. In order to show the cognate binding surface on LF, it is displayed after a rotation by 180° around the vertical axis from its binding site on the heptamer. The figure was created using PYMOL (DeLano Scientific).

Table 1. *Residues of PA63 and LF in the interaction surface*

PA63h	LF	
CHAIN E		
GLY173…THR174	ASN638…ILE639	
VAL175…ARG178	THR584…HIS588	
PRO184…SER186	THR584…HIS588	
GLU188…VAL189	LYS581…LEU582	
GLU190…TYR192	GLY625…ASN626	DOMAIN IV
PHE202…PRO205	GLU662….. ARG664	
ILE210…GLY215	GLY697…TYR698	
	LYS702…ASP706	
THR516…LYS518	PRO577…TYR579	
CHAIN F		
	GLU64…LYS65	
PRO176…ASN180	ASN248…GLU249	
	TYR223…ILE225	
PRO184…LEU187		
SER204…PRO205	ARG229…TYR236	DOMAIN I
VAL189…GLU190		
ASP195…LYS199	ASN179…ASP182	
ILE 207…LYS214	ASP182…LEU188	
	VAL232…PRO238	
ASN463…PHE464	GLU27…ARG28	
CHAIN G		
	GLU38…LYS41	
ASP195…THR201	LYS75…GLY78	DOMAIN I
	THR430…ASN433	DOMAIN II

Contiguous regions in the polypeptide chains of PA63 and in LF that are proximal (≤ 7 Å between C^{α}) in the "pseudoatomic" model are listed. Not every residue in a given stretch of PA63 is in close contact with all residues in the cognate stretch in LF and vice versa.

and Sanejouand 2001). In our flexible fitting procedure, displacements along some of the lowest frequency normal modes are applied to flex the structure iteratively such that a better fit with the density envelope can be obtained. By definition, such a technique examines conformational changes that can result naturally in the polypeptide chain. We find that ~60% of the overall, generated conformational change in LF was captured by just two normal modes. This observation thus points to the fact that a set of predominant directions of protein motion is closely followed during the refinement. It stands to reason that the fitting procedure that we employed does not correspond to a random search and is therefore reliable, and the results attained are not due to overfitting of the data.

Acid-induced unfolding study has quantified how the conformation of LF_N changes at low pH (Krantz et al. 2004). These, by definition, are substantial structural alterations possibly mimicking unfolding events necessary to "stretch" the ligand so that it can traverse the narrow pore in the endosomal membrane. The energetics involved in domain movements observed in our study for the bound, full-length LF is expected to be modest, being, as discussed above, derived from low-frequency normal modes, i.e., intrinsic natural low-energy deformations of

the polypeptide chain. On the other hand, we anticipate that the observed concomitant perturbations in the structure of the heptamer are derived from the interaction energy accompanying ligand binding (Table 1). Movements of site-specific labels engineered in LF and monitored by FRET (Deniz et al. 2001) can test the modeled conformational change in LF as wells as the displacements of the monomers in PA63h.

Refined model of the PA63h–LF complex

The "pseudoatomic" model presented here integrates experimental cryo-EM data and data from site-directed mutagenesis to arrive at a possible model for the toxic complex. Overall, the PA63 monomers show an RMSD between ~2 Å and 4 Å relative to the crystal structure. Proximity of residues 229–236 in LF to the binding site on a PA63 monomer (monomer designated F on PA63h) was used as a critical constraint in adjudicating the choice of the model derived from the fitting procedure. Only one out of the five initial "best" models of LF, which were subjected to fitting within the density envelope attributed to LF in the cryo-EM map, survived the analysis. In this chosen model, in addition, and in agreement with implications from mutagenesis results, we find that residues from several PA63 monomers, in the peptide patch 170–220, are in close proximity (C^{α}–C^{α} distance ≤ 7 Å) to several LF residues. LF-mutants Y148A, Y149A, I151A, and K153A, which reside in the central 1a4 helix (Pannifer et al. 2001), are unable to lyse macrophage, possibly due to impaired translocation. This has been suggested to be as a result of inefficient binding to PA63h (Gupta et al. 2001). In our model, while the C^{α}–C^{α} distance of these residues in LF with some of the residues from PA63 are ~7 Å, we do not see significant rearrangements around this central helix of LF that could have brought the aforementioned residues on the surface of LF and thereby facilitated possible intermolecular binding with PA63 monomers. Mutations at these sites are more likely to affect the local structure of this helix (Lacy et al. 2002), which could affect structural stability and impair binding.

A publication describing a model of LF_N bound to a dimer of PA63 recently appeared (Lacy et al. 2005) while this manuscript was in preparation. The work reported in that paper was a modeling exercise that satisfied proximity information implied in site-directed mutagenesis studies (as also used in our study) and site-specific disulfide cross-linking results. The investigators found a "twist" in LF_N conformation in the energy-minimized model. In both their model and our model, the N-terminal helix of the crystal structure (starting with Glu27 of LF_N) is positioned at the lumen entrance, complementing the picture that LF enters at the N-terminal

end. Also, a majority of the interaction of LF$_N$ is with one subunit (chain F in our case), and there are a number of charged residues in the buried interface, as seen by us also (Table 1). However, we do not see proximity of Y108 (LF$_N$) and N209 (PA63) and E135 (LF$_N$) and K197 (PA63) in our model.

The large conformational change in LF and the breathing motion of PA63h upon complexation appear to go hand in hand. We should emphasize here that it is not uncommon that in the event of binding two partners, conformational changes in both are observed. It is, for example, well known that during the synthesis of the polypeptide chain in the cell, the binding of the Elongation Factor G to the ribosome involves large rearrangements in both partners (Frank and Agarwal 2000; Valle et al. 2003a). Such conformational changes have also been observed for other factors binding to the ribosome, such as the elongation factor Tu bound to the transfer RNA molecule (Valle et al. 2003b) and the termination factor RF2 (Rawat et al. 2003).

The optimized model for the complex described here provides a picture of the structural rearrangement of the heptamer, which is largely brought about by rigid-body movements of the monomers. The lumen, shaped approximately as an hourglass, becomes asymmetric and is considerably enlarged in the liganded state, especially near the middle of the molecule where the largest width is increased by ~16 Å or by as much as ~50%. At this location, in the unliganded heptamer, the lumen diameter is at its narrowest (Fig. 6). In addition, as noted earlier, there is a compaction in the structure of the LF molecule. Thus, the enlarged lumen volume can now accommodate a more compact LF and can therefore explain the passage of LF from the distal binding site at the ''top'' of the heptamer to the entrance of the pore formed in the endosomal membrane. We suggest that these structural motions in both PA63h and full-length LF, acting in concert, constitute the obligatory first step in ligand translocation. However, this compaction alone is insufficient to allow passage through the pore formed in the endosomal membrane, which has been indicated to be only ~12 Å in diameter. Here, for instance, the observed ability of LF$_N$ to unfold in acidic condition could come into play (Krantz et al. 2004).

Materials and methods

Normal mode flexible fitting

Starting with the crystal structure of LF, we used the NMFF approach to fit full-length LF into the cryo-EM map by optimally flexing its structure. In this approach, we use the normal mode coordinates associated with the eigenvectors as variables that are optimized (Go et al. 1983). These coordinates, in turn, reflect collective conformational change of the molecule and modes with lower frequencies representing preferential, low-energy, large, and collective motions (Tama and Sanejouand 2001). A major problem in existing methods that use flexible fitting is that the molecule is generally deformed in a rather ad hoc fashion, without consideration of the energetic cost of deformation. In NMFF, energetically unrealistic structural alterations are avoided because such modifications in the molecule are directed by the low-frequency normal modes. In addition, we use the normal mode analysis iteratively, which results in a more accurate description of the conformational change. Such a treatment also reduces the possibility of introducing distortions that can appear if the structure is displaced too far along a given normal mode coordinate (Miyashita et al. 2003; Tama et al. 2004a). Finally, since it is only necessary to optimize a few (of the order of 10–30 normal modes) degrees of freedom, errors due to ''overfitting'' of data are minimized. We also note that it has been shown that for simulated EM data, even at 30 Å resolution, the NMFF approach is able to identify large conformational changes (Tama et al. 2004a).

Normal mode analysis (NMA)

For the NMA we used the simplified elastic network representation of the potential energy function (Tirion 1996). Such a form of potential energy has been shown to be successful in reproducing large and collective motions of macromolecules (Bahar et al. 1999; Atilgan et al. 2001; Tama and Sanejouand 2001; Miyashita et al. 2003; Tama et al. 2003). Within the framework of the elastic network model, the standard semi-empirical potential energy function is replaced by

$$E_P = \sum_{d_{ij}^0 < R_c} \frac{C}{2} (d_{ij} - d_{ij}^0)^2 \qquad (1)$$

where d_{ij} is the distance between atoms i and j, d_{ij}^0 is the distance between these two atoms in the given studied structure, and R_c is an arbitrary cutoff parameter delimiting the distance beyond which elastic bonds are not considered between centers of interacting atom pairs. R_c was set to 8 Å, and C, the strength of the potential, is assumed the same for all interacting pairs. In such formalism, C is only involved in the definition of the overall energy scale. The use of this simplified potential energy representation is optimal in our iterative normal mode analysis because, the new set of d_{ij}^0 values place the model at an energy minimum after each iterative step and no energy minimization is necessary to accompany the normal mode analysis. Because of the modest ~18 Å resolution of the experimental density map, polypeptides chains defined by only the C$^\alpha$ atoms were considered for the fitting procedure. Such a representation has been shown to yield a good approximation for the low-frequency normal modes (Bahar et al. 1997; Atilgan et al. 2001; Tama and Sanejouand 2001; Delarue and Sanejouand 2002).

Normal mode analysis requires the diagonalization of the $3N \times 3N$ (N being the number of atoms) matrix of the second derivatives of the potential energy L, also called the Hessian. We used the rotation–translation block (RTB) algorithm, which allows for fast and memory-efficient diagonalization of the Hessian (Durand et al. 1994; Tama et al. 2000). In the RTB method, the molecule is divided into n_b blocks (one each for a stretch of few consecutive amino acid residues), and, rather than considering simultaneously all degrees of freedom for all the atoms, only rotational and translational displacements of these blocks are considered. Such
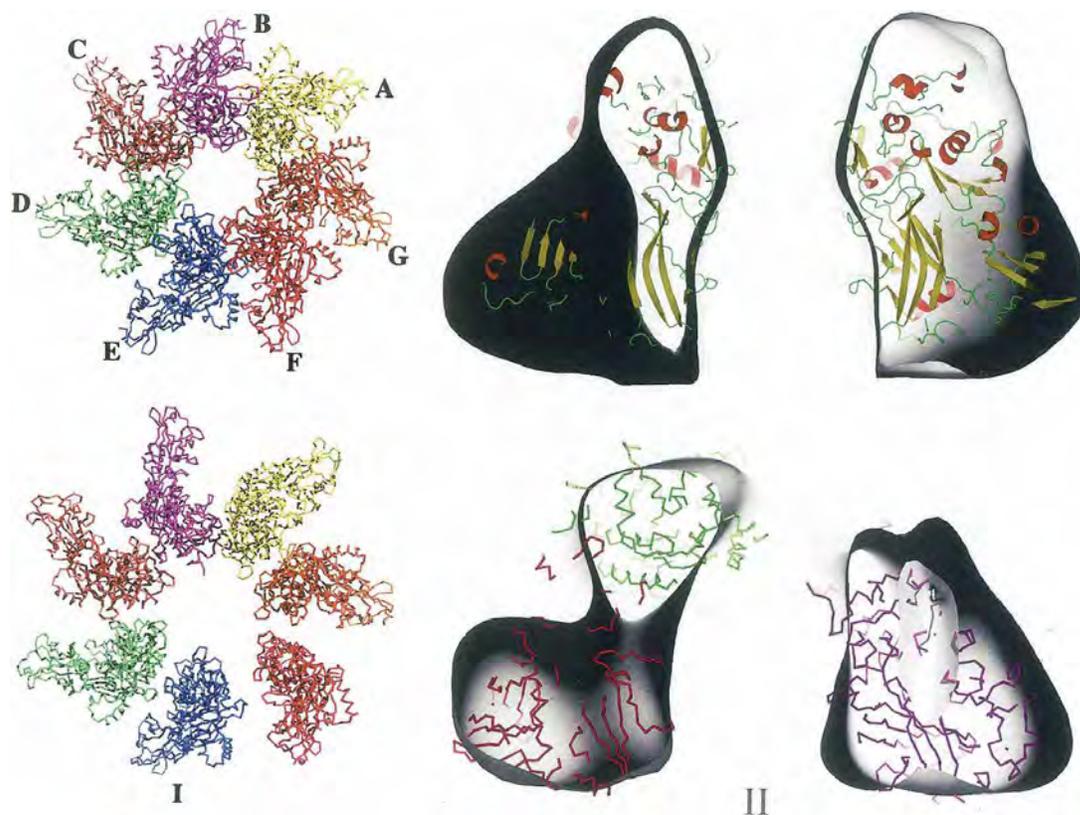
Figure 6. Structural change in the heptamer upon binding of full-length LF. The *left* panel (I) shows the enlargement of the lumen in cross-section upon ligand binding; unliganded (*top*, crystal structure), refined heptamer structure in the liganded complex (*bottom*). The *right* panel (II) shows slabs through the cryo-EM density envelopes slicing the density map approximately near the center. The constriction in the lumen in the unliganded state is relaxed in the complex (*bottom* figure of the *right* panel) where LF is shown in green. Figures in the *right* panel were created using PYMOL (DeLano Scientific).

an arrangement still provides a good approximation of the lowest frequency normal modes. In the present application, the RTB projection is carried out on blocks containing three individual amino acid residues. Diagonalization of the Hessian was carried out using the ARPACK package (Lehoucq et al. 1998).

Rigid-body fitting

In order to initiate flexible fitting, it is necessary to first place the X-ray structure approximately in the same orientation as that of the density map. This can be achieved by first performing a rigid-body fitting of the crystal structure into the given density map, which then serves as the initial model for the flexible fitting. Several packages are available for the rigid-body fitting of an X-ray structure into a low-resolution density map such as that generated by cryo-EM. In this study, we used the program package SITUS (version 1.4 Wriggers et al. 1999). We defined six codebook vectors, which are a set of control points that provide information about the shape and the density distribution within the molecule (Wriggers and Birmanns 2001). We then performed a rigid-body fitting using the "qdock" command. We used the recommended ranking procedure; i.e., the fitted structures were ranked according to the RMSD between the codebook vectors defined by the EM map and the codebook vectors defined

by the X-ray structure. The structure ranked first was then used as the starting model for the flexible fitting.

Cross-correlation coefficient was calculated based on the correlation of the cryo-EM density map with the calculated density map of the given atomic model that was truncated to a limiting resolution of 18 Å.

All-atom reconstruction

The C^{α}-based models obtained from NMFF were used to reconstruct an all-atom model using the MMTSB package (Feig et al. 2004). The reconstruction was performed using SCWRL (Canutescu et al. 2003), and energy minimizations of the models to reduce steric clashes were performed with CHARMM30 (Brooks et al. 1983).

Acknowledgments

References

Atilgan, A.R., Durell, S.R., Jernigan, R.L., Demirel, M.C., Keskin, O., and Bahar, I. 2001. Anisotropy of fluctuation dynamics of proteins with an elastic network model. *Biophys. J.* **80:** 505–515.

Bahar, I., Atilgan, A.R., and Erman, B. 1997. Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential. *Fold. Des* **2:** 173–181.

Bahar, I., Erman, B., Jernigan, R.L., Atilgan, A.R., and Covell, D.G. 1999. Collective motions in HIV-1 reverse transcriptase: Examination of flexibility and enzyme function. *J. Mol. Biol.* **285:** 1023–1037.

Ballard, J.D., Collier, R.J., and Starnbach, M.N. 1996. Anthrax-toxin mediated delivery of a cytotoxic T-cell epitope *in vivo*. *Proc. Natl. Acad. Sci.* **93:** 12531–12534.

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

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

Brooks, B., Bruccoleri, B., Olafson, D., States, D., Swaminathan, S., and Karplus, M. 1983. CHARMM: A program for macromolecular energy, minimization and dynamics calculation. *J. Comput. Chem.* **4:** 187–217.

Brünger, A.T., Kuriyan, J., and Karplus, M. 1987. Crystallographic R factor refinement by molecular dynamics. *Science* **235:** 458–460.

Canutescu, A.A., Shelenkov, A.A., and Dunbrack, R.L. 2003. A graph-theory algorithm for rapid protein side-chain prediction. *Protein Sci.* **12:** 2001–2014.

Chacon, P. and Wriggers, W. 2002. Multi-resolution contour-based fitting of macromolecular structures. *J. Mol. Biol.* **317:** 375–384.

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

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

Delarue, M. and Sanejouand, Y.H. 2002. Simplified normal mode analysis of conformational transitions in DNA-dependent polymerases: The Elastic Network Model. *J. Mol. Biol.* **320:** 1011–1024.

Deniz, A.A., Laurence, T.A., Dahan, M., Chemla, D.S., Shultz, P.G., and Weiss, S. 2001. Ratiometric single-molecule studies of freely diffusing biomolecules. *Annu. Rev. Phys. Chem.* **52:** 233–253.

Durand, P., Trinquier, G., and Sanejouand, Y.H. 1994. New approach for determining low-frequency normal-modes in macromolecules. *Biopolymers* **34:** 759–771.

Elliott, J.L., Mogridge, J., and Collier, R.J. 2000. A quantitative study of the interactions of *Bacillus anthracis* edema factor and lethal factor with activated protective antigen. *Biochemistry* **39:** 6706–6713.

Feig, M., Karanicolas, J., and Brooks III, C.L. 2004. MMTSB tool set: Enhanced sampling and multiscale modeling methods for applications in structural biology. *J. Mol. Graph.* **22:** 377–395.

Frank, J. and Agrawal, R.K. 2000. A ratchet-like inter-subunit reorganization of the ribosome during translocation. *Nature* **406:** 318–322.

Gao, H.X., Sengupta, J., Valle, M., Korostelev, A., Eswar, N., Stagg, S.M., Van Roey, P., Agrawal, R.K., Harvey, S.C., Sali, A., et al. 2003. Study of the structural dynamics of the *E-coli* 70S ribosome using real-space refinement. *Cell* **113:** 789–801.

Go, N., Noguti, T., and Nishikawa, T. 1983. Dynamics of a small globular proteins in terms of low-frequency vibrational modes. *Proc. Natl. Acad. Sci.* **80:** 3696–3700.

Goletz, T.J., Klimpel, K.R., Arora, N., Leppla, S.H., Keith, J.M., and Berzofsky, J.A. 1997. Targeting HIV proteins to the major histocompatibility complex class I processing pathway with a novel gp120-anthrax fusion protein. *Proc. Natl. Acad. Sci.* **94:** 12059–12064.

Gupta, P., Singh, A., Chauhan, V., and Bhatnagar, R. 2001. Involvement of residues 147VYYEIGK153 in binding of lethal factor to protective antigen of *Bacillus anthracis*. *Biochem. Biophys. Res. Commun.* **280:** 158–163.

Jiang, W., Baker, M.L., Ludtke, S.J., and Chiu, W. 2001. Bridging the information gap: Computational tools for intermediate resolution structure interpretation. *J. Mol. Biol.* **308:** 1033–1044.

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

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

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

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

Lehoucq, R.B., Sorensen, D.C., and Yang, C. 1998. *ARPACK user's guide: Solution of large-scale eigenvalue problems with implicitly restarted Arnoldi methods*. SIAM, Philadelphia.

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

Mitra, K., Schaffitzel, C., Shaikh, T., Tama, F., Jenni, S., Brooks III, C.L., Ban, N., and Frank, J. 2005. Structure of the *E. coli* protein-conducting channel bound to a translating ribosome. *Nature* **438:** 318–324.

Miyashita, O., Onuchic, J.N., and Wolynes, P.G. 2003. Nonlinear elasticity, proteinquakes, and the energy landscapes of functional transitions in proteins. *Proc. Natl. Acad. Sci.* **100:** 12570–12575.

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

Mogridge, J., Cunningham, K., Lacy, D.B., Mourez, M., and Collier, R.J. 2002b. The lethal and edema factors of anthrax toxin bind only to oligomeric forms of the protective antigen. *Proc. Natl. Acad. Sci.* **99:** 7045–7048.

Nassi, S., Collier, R.J., and Finkelstein, A. 2002. PA(63) channel of anthrax toxin: An extended β-barrel. *Biochemistry* **41:** 1445–1450.

Neumeyer, T., Tonello, F., Molin, F.D., Schiffler, B., Orlik, F., and Benz, R. 2006. Anthrax lethal factor (LF) mediated block of the anthrax protective antigen (PA) ion channel: Effect of ionic strength and voltage. *Biochemistry* **45:** 3060–3068.

Pannifer, A.D., Wong, T.Y., Schwarzenbacher, R., Renatus, M., Petosa, C., Bienkowska, J., Lacy, D.B., Collier, R.J., Park, S., Leppla, S.H., et al. 2001. Crystal structure of the anthrax lethal factor. *Nature* **414:** 229–233.

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

Rawat, U.B.S., Zavialov, A.V., Sengupta, J., Valle, M., Grassucci, R.A., Linde, J., Vestergaard, B., Ehrenberg, M., and Frank, J. 2003. A cryo-electron microscopic study of ribosome-bound termination factor RF2. *Nature* **421:** 87–90.

Ren, G., Quispe, J., Leppla, S.H., and Mitra, A.K. 2004. Large-scale structural changes accompany binding of lethal factor to anthrax protective antigen: A cryo-electron microscopic study. *Structure* **12:** 2059–2066.

Rossmann, M., Bernal, G.R., and Pletnev, S.V. 2001. Combining electron microscopic with X-ray crystallographic structures. *J. Struct. Biol.* **136:** 190–200.

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

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

Tama, F. and Sanejouand, Y.H. 2001. Conformational change of proteins arising from normal mode calculations. *Protein Eng.* **14:** 1–6.

Tama, F., Gadea, F.X., Marques, O., and Sanejouand, Y.H. 2000. Building-block approach for determining low-frequency normal modes of macromolecules. *Proteins* **41:** 1–7.

Tama, F., Valle, M., Frank, J., and Brooks III, C.L. 2003. Dynamic reorganization of the functionally active ribosome explored by normal mode analysis and cryo-electron microscopy. *Proc. Natl. Acad. Sci.* **100:** 9319–9323.

Tama, F., Miyashita, O., and Brooks III, C.L. 2004a. Flexible multi-scale fitting of atomic structures into low-resolution electron density maps with elastic network normal mode analysis. *J. Mol. Biol.* **337:** 985–999.

———. 2004b. NMFF: Flexible high-resolution annotation of low-resolution experimental data from cryo-EM maps using normal mode analysis. *J. Struct. Biol.* **147:** 315–326.

Tirion, M.M. 1996. Large amplitude elastic motions in proteins from a single-parameter, atomic analysis. *Phys. Rev. Lett.* **77:** 1905–1908.

Valle, M., Zavialov, A., Sengupta, J., Rawat, R., Ehrenberg, M., and Frank, J. 2003a. Locking and unlocking of ribosomal motions. *Cell* **114:** 123–134.

Valle, M., Zavialov, A., Li, W., Stagg, S.M., Sengupta, J., Nielsen, R.C., Nissen, P., Harvey, S.C., Ehrenberg, M., and Frank, J. 2003b. Incorporation

of aminoacyl-tRNA into the ribosome as seen by cryo-electron microscopy. *Nat. Struct. Biol.* **10:** 899–906.

Volkmann, N. and Hanein, D. 1999. Quantitative fitting of atomic models into observed densities derived by electron microscopy. *J. Struct. Biol.* **125:** 176–184.

Wang, J. and Chen, L. 2003. Domain motions in GroEL upon binding of an oligopeptide. *J. Mol. Biol.* **334:** 489–499.

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

Wriggers, W., Milligan, R.A., and McCammon, J.A. 1999. SITUS: A package for docking crystal structures into low-resolution maps from electron microscopy. *J. Struct. Biol.* **125:** 185–195.

Wriggers, W. and Birmanns, S. 2001. Using Situs for flexible and rigid-body fitting of multiresolution single-molecule data. *J. Struct. Biol.* **133:** 193–202.

Zhang, S., Finkelstein, A., and Collier, R.J. 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.* **101:** 16756–16761.

Zhang, S., Cunningham, K., and Collier, R.J. 2004b. Anthrax protective antigen: Efficiency of translocation is independent of the number of ligands bound to the prepore. *Biochemistry* **43:** 6339–6343.

Zhao, J.M., Milne, J.C., and Collier, R.J. 1995. Effect of anthrax toxins lethal factor on ion channels formed by the protective antigen. *J. Biol. Chem.* **270:** 18626–18630.