Structure of exotoxin A of *Pseudomonas aeruginosa* at 3.0-Ångstrom resolution

VILOYA S. ALLURED*, R. JOHN COLLIER†, STEPHEN F. CARROLL†, AND DAVID B. MCKAY*

*Department of Chemistry, University of Colorado, Boulder, CO 80309; and †Department of Microbiology and Molecular Genetics and Shipley Institute of Medicine, Harvard Medical School, Cambridge, MA 02115

Communicated by A. M. Pappenheimer, Jr., October 15, 1985

**ABSTRACT** Exotoxin A of *Pseudomonas aeruginosa* is a secreted bacterial toxin capable of translocating a catalytic domain into mammalian cells and inhibiting protein synthesis by the ADP-ribosylation of cellular elongation factor 2. The protein is a single polypeptide chain of 613 amino acids. The x-ray crystallographic structure of exotoxin A, determined to 3.0-Å resolution, shows the following: an amino-terminal domain, composed primarily of antiparallel β-structure and comprising approximately half of the molecule; a middle domain composed of α-helices; and a carboxyl-terminal domain comprising approximately one-third of the molecule. The carboxyl-terminal domain is the ADP-ribosyltransferase of the toxin. The other two domains are presumably involved in cell receptor binding and membrane translocation.

Exotoxin A of *Pseudomonas aeruginosa* is one member of a family of secreted bacterial toxins that are capable of covalently modifying specific target proteins within mammalian cells (1). Included in this family are the exotoxins of *Corynebacterium diphtheriae* (diphtheria toxin) and *Vibrio cholerae* (cholera toxin), *Escherichia coli* heat-labile toxin (LT), and exotoxins of *Shigella dysenteriae* (shiga toxin) and *Bacillus anthracis* (anthrax toxins) as well as exotoxin A (2). Despite their diversity in size, subunit composition, cell specificity, and enzymatic activity, these toxins appear to share a similar multistep mechanism in which (i) the toxin binds to a receptor on the membrane surface of a target cell; (ii) the catalytic domain of the toxin is translocated into, or at a minimum into contact with, the cell cytoplasm; (iii) the catalytic moiety is then able to modify its target substrate. The toxins thus must have a receptor binding activity, a membrane translocation mechanism, and an enzymatic domain. It is characteristic that the receptor binding function and the enzymatic activity reside in separate structural components of the molecules, in separate subunits of an oligomer (cholera toxin, LT, shiga toxin) (3–5), in separate proteins (anthrax system) (6), or within a single monomeric polypeptide (diphtheria toxin, exotoxin A) (7, 8).

Several of the toxins (cholera toxin, LT, diphtheria toxin, and exotoxin A) catalyze transfer of the ADP-ribose moiety of oxidized nicotinamide adenine dinucleotide (NAD⁺) to target substrates (9–12). Diphtheria toxin and exotoxin A specifically ADP-ribosylate a modified histidine (diphthamide) of protein synthesis elongation factor 2, thereby inactivating the elongation factor and terminating peptide chain elongation in a target cell (13).

Several intriguing mechanistic questions arise: (i) What are the mechanisms of membrane translocation by which the toxic factors enter the target cell cytoplasm? (ii) How is the membrane translocation and enzymatic activation process controlled during intoxication? (iii) What is the mechanism of the ADP-ribosyltransferase reaction? Little structural information is available for members of this class of bacterial toxins. Crystals suitable for high resolution structural analysis have been reported for cholera toxin (14), diphtheria toxin (15), and one protein component of the anthrax system (16), as well as for exotoxin A (17). In this report we present the three-dimensional molecular structure of exotoxin A of *P. aeruginosa*. The molecule is secreted from *P. aeruginosa* in a proenzyme form; it cannot catalyze the ADP-ribosylation of elongation factor 2 prior to its activation upon entering target cells. The proenzyme form of the molecule, as isolated from bacterial cell culture supernatants, is the structure we have crystallized and solved.

**MATERIALS AND METHODS**

Exotoxin A was purified and crystallized as reported (17). Initially, data on potential heavy atom isomorphous derivatives were collected by diffractometer to 5.5-Å resolution. Data were subsequently collected to 3.0-Å resolution on a combination of derivatives that gave an interpretable low-resolution electron density map: a native dataset and one heavy atom derivative dataset were collected by diffractometer (18); background, radiation decay and transmission corrections were applied to integrated intensities. Three heavy atom derivative datasets were collected on a multiwire area detector at the University of Virginia (19); background and radiation decay corrections were applied. A rudimentary but effective local scaling method was utilized to correct for systematic differences between datasets. A native dataset composed of the strongest 20% of the intensities uniformly distributed in resolution to 3.0 Å was measured on a single crystal with a diffractometer and was used as a standard to which other data were scaled. The local scaling procedure used consisted of the following: (i) dividing reciprocal space into blocks of equal volume containing 50–100 measured reflections; (ii) computing the ratio of the sums of intensities within blocks; (iii) applying the scale factor to all intensity measurements in one of the symmetry-related blocks, leaving the other unchanged. Using this procedure, asymmetric units measured on an individual crystal were scaled to each other; then all crystals contributing to a given derivative were scaled together; and each heavy atom derivative dataset was scaled to the native dataset.

Heavy atom positions were refined and Blow–Crick "best" multiple isomorphous replacement phases were computed by standard methods (20) (Table 1). Bijvoet differences computed from intensities measured by diffractometer were used to determine the correct enantiomorph of the heavy atom solutions. The o-chloromercuriurotropophen and methylmercury chloride derivatives shared one heavy atom site that accounted for 44% and 27% of their heavy atom scattering density (occupancy), respectively; this overlap was considered small enough to allow them to be treated as independent derivatives. The mercuric iodine, platinum ethylenediamine, and platinum nitrate derivatives also had a
shared site, distinct from the sites of the o-chloromercurinitrophenol and methymercury chloride derivatives; this site contributed 100%, 45%, and 38% of their respective heavy atom scattering densities. These derivatives were also treated as independent.

Phases computed to 5.5-Å resolution by using solely diffractometer data had an overall figure-of-merit of 0.91. The resulting low-resolution electron density map displayed well-defined molecular boundaries and showed much of the overall secondary structure and connectivity of the molecule clearly.

The rapid falloff of figure-of-merit of phases with resolution above 5.0 Å made interpretation of a molecular model from the 3.0-Å resolution electron density map somewhat difficult. An initial α carbon backbone trace was determined from "minimaps"; then a polyalanine backbone was built into the map by using computer graphics systems and the model-building program FRODO. In two regions of the map, the trace was ambiguous at this point. When the amino acid sequence became available (21), it was possible to place residues 3–438 of the molecule in the map; the final 175 residues could not be traced unambiguously. To resolve the uncertainties, the phase combination procedure of Hendrickson and Lattman (22), following the strategy of Remington et al. (23), was used. In our implementation of the method, residues whose placement was impossible or appeared questionable were omitted from the model, and residues where the sequence placement appeared questionable were put in as alanines, or, in the first cycle, amino acid side chains that filled the side chain density in the map. After omissions, greater than 90% of the atoms of the molecule were still present. The approximate molecular model was then refined by using a least-square, stereochemically restrained reciprocal space refinement package developed by L. F. Ten Eyck and D. Tronrud (personal communication). A typical cycle involved approximately 11,000 measured amplitudes between 10-Å and 3.0-Å resolution, as observations, and geometric restraints including approximately 4400 bond lengths, 5800 bond angles, 600 planar groups, and 100 trigonal groups (torsion angles were not tightly restrained) for a total of 21,900 observations plus restraints. Refined variables consisted of the atomic coordinates for approximately 4400 atoms; individual temperature factors were not refined; this resulted in 13,200 variables, for an observation-plus-restraint to variable-parameter ratio of 1.7, ensuring adequate stability in the refinement procedure when the molecular geometry is adequately constrained. Statistics of the refinements are given in Table 2.

Sim-weighted model phases were then computed (36). A new map was computed, using multiple isomorphous replacement phases and F0 amplitudes exclusively below 5.5-Å resolution, combined phases (obtained by multiplying the multiple isomorphous replacement and model phase probability distributions) and 2 F0 – Fc amplitudes between 5.5 Å and an outer resolution of 3.0 Å (in cycles 1–3) or 3.5 Å (in 4 and subsequent cycles), and model phases between 3.5 Å and 3.0 Å after cycle 4 and subsequent cycles. The model was then rebuilt to the resulting phase-combined map. Efficacy of the method relies on the partial model improving the phases significantly, with a consequent improvement of the map in regions where the model has been omitted.

After the first cycle, it was possible to extend the trace uninterrupted through residue 482, and to make a tentative sequence placement, with two breaks, in the remainder of the

<table>
<thead>
<tr>
<th>Table 1. Heavy-atom derivative preparation and refinement statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Soaking conditions</td>
</tr>
<tr>
<td>--------------------</td>
</tr>
<tr>
<td>Saturated mercuric iodine, 2 days*</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Satuated</td>
</tr>
<tr>
<td>O-chloromercurin-</td>
</tr>
<tr>
<td>nitrophenol, 2 days†</td>
</tr>
<tr>
<td>Satuated</td>
</tr>
<tr>
<td>Platinum ethylenedi-</td>
</tr>
<tr>
<td>amine, 2 days‡</td>
</tr>
<tr>
<td>10 mM Platinum</td>
</tr>
<tr>
<td>nitrite, 20 h§</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Overall</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

*Collected by diffractometer: f_{m}/E_{rms} = \|f_{m}/|f_{m}|\|^{2/3}, E_{rms} = \|f_{m} – |F_0| |I_0|\|^{2/3}, R_{c} = \|f_{m} – |F_0| |I_0|\|/\|f_{m}\|, where I = measured diffraction intensity, F_0 = structure factor amplitude of native crystal = |f_{m}|; |I_0| = structure factor of native crystal = |f_{m}| \exp(i\alpha), where \alpha is computed native protein phase; |F_0| = structure factor amplitude of derivative crystal; f_{m} = calculated structure factor for heavy atoms of derivative; f_{m} = calculated structure factor amplitude for heavy atoms of derivative = \|f_{m}\|^{2/3}. Summations are overall reflections, n.

‡Low resolution collected by diffractometer; high resolution collected by multi-wire area detector.

§(m) = mean figure of merit.
molecule. In subsequent cycles, parts of the model, where continuity was clear from the multiple isomorphous replacement map but precise amino acid placement was difficult, were omitted from the model, as a check of their placement. The current molecular model has two short breaks in the final 130 residues of the polypeptide chain where it is not possible to unambiguously fit the protein sequence into a rather noisy positive electron density. Consequently, the sequence placement of the final 130 amino acids must, in a strict interpretation, be regarded as tentative, although their overall secondary structure is clear; future extension of the structure determination and refinement to at least 2.5-Å resolution should resolve current uncertainties. The conclusions presented below are not compromised by the polypeptide breaks in the carboxyl-terminal region of the molecular model.

RESULTS

Molecular Structure. The exotoxin A molecule has three distinct structural domains (Fig. 1).

Domain I. Domain I, the lower left domain in the view shown in Fig. 1, is an antiparallel β-structure. It includes residues 1–252 and residues 365–404. It has 17 β-strands; all but the final short strand run antiparallel to neighbor strands (Fig. 2a). The first 13 strands form the structural core of an elongated β-barrel, the topology of which can be considered an extension of a familiar β-roll folding pattern. β-Rolls observed in other structures to date, including those found in influenza hemagglutinin (24), catabolite gene activator protein (25), tomato bushy stunt virus (26), and southern bean mosaic virus (27), consist of, with minor variation, eight strands with the folding topology shown in Fig. 3a. In exotoxin A, if one considers strands 12 and 13 to be topologically equivalent to a single continuous strand, the β-roll consists of 12 topological antiparallel strands (Fig. 3b).

Following strand 13 of domain I, the peptide chain traverses one face of the barrel, leading into the second domain.

Domain II. Domain II (residues 253–364; top/middle of Fig. 1) is composed of six consecutive α-helices with one disulfide linking helix A and helix B (Fig. 2b). Helices B and E are approximately 30 Å in length; helices C and D are approximately 15 Å long.

Domain III. Domain III, appearing as a structurally separate domain at the right of Fig. 1, is comprised of the carboxyl-terminal third of the molecule, residues 405–613. The most notable structural feature of domain III is its extended cleft (Fig. 2c). The domain manifests a less regular secondary structure than domains I and II.

Disulfides. The disulfide arrangement in exotoxin A has not been previously reported; the structure reveals that the eight cysteines form four disulfides in sequential order:

Cys-11 forms a disulfide with Cys-15, as does Cys-197 with -214, Cys-265 with -287, and Cys-372 with -379. Since exotoxin A is secreted from P. aeruginosa into the extracellular environment during its synthesis, it would be interesting to consider to what extent the sequential formation of disulfides, as the protein is synthesized and secreted, is a major determinant in proper folding of the tertiary structure.

DISCUSSION

The Enzymatic Domain. Several pieces of evidence identify domain III as being the enzymatic domain of exotoxin A. Early experiments demonstrated that complete chemical cleavage of exotoxin A after cysteines or methionines generates enzymatically active fragments of approximate molecular weights 30,000 and 50,000, respectively (28). The cysteine and methionine most distal from the amino terminus are residues 379 and 234, respectively; thus these chemical scissions would have left domain III intact, while excising portions of the domains I and II. More recently, carboxyl-terminal fragments of the exotoxin A structural gene have been cloned and expressed; it has been demonstrated that a fragment consisting of amino acid residues 308–613 is enzymatically active, and a fragment of residues 492–613 is active, albeit unstable (21).

It is notable that domain III does not manifest the classical nucleotide binding domain structure found in the dehydrogenases and other NAD⁺/NADH-binding enzymes (29). This presumably is a consequence of the different chemical role of NAD⁺ in this system. In dehydrogenases, NAD⁺/NADH is used as a cofactor to/from which a hydride ion is transferred. Typical dissociation constants in dehydrogenases are $10^{-3}$–$10^{-4}$ M for NAD⁺ and $10^{-5}$–$10^{-6}$ M for NADH (30). In exotoxin A, NAD⁺ is used as a substrate, in which the bond between nicotinamide and ribose is broken. For an active fragment of exotoxin A, the $K_m$ for NAD⁺ in the ADP-ribosylation of elongation factor-2 is 8 μM, while NADH binds much less tightly (28).

The structure of exotoxin A discussed in this paper is the proenzyme form of the molecule; it cannot catalyze the ADP-ribosylation of its in vivo target substrate, elongation factor-2, without activation. It does appear, however, to catalyze the hydrolysis of NAD⁺ to nicotinamide and ADP-ribose, albeit with a substantially weaker $K_m$ for the NAD⁺ substrate (S.F.C. and R. J. Collier, unpublished results).

Thus the crystallographic and biochemical data suggest that enzymatic activation of exotoxin A requires only removing the steric constraints imposed on domain III by other parts of the molecule; it is unlikely that a major refolding of the domain, even in vivo where it must translocate across a membrane, is a prerequisite for full enzymatic activity.

Fig. 1. Stereo drawing of the α carbon backbone of exotoxin A. In the view shown, the amino terminus is at the lower left of the molecule, and the carboxyl terminus is on the right side of the figure.
Receptor Binding and Membrane Translocation. Since domain III is the enzymatic moiety of exotoxin A, the receptor binding and membrane translocation activities presumably reside in domains I and II. It is attractive to suggest a distinct division of function between the two domains, with domain I, whose structure is arguably suggestive of a ligand binding structure, functioning in receptor binding, and with domain II inducing membrane translocation of the catalytic domain of exotoxin A. However, biochemical evidence to corroborate such a suggestion is currently lacking.

Although the general pathway by which extracellular ligands are internalized into intracellular endosomes has been well delineated (31), the mechanisms by which some viral coat proteins induce membrane fusion and some bacterial exotoxins induce membrane translocation are poorly understood. Many of the current models for the processes rely on triggering a conformational change within the protein, with a consequent exposure of a hydrophobic surface of the protein that subsequently inserts into membrane bilayers. One conventional diagnosis for suggesting which part of a protein is likely to interact with membranes is to search for strongly hydrophobic stretches in the amino acid sequence (32).

We have computed the net hydropathy (as defined in ref. 32) of exotoxin A, averaging individual amino acid hydrop-
phobicity, therefore, (33-35). It is apparent that there is no long continuous stretch of strongly nonpolar sequence, as was observed in diphtheria toxin (33-35). Further, inspection of the structure shows that local peaks in hydropathy of exotoxin A correspond to segments of the peptide that are packed internally in the molecule; there is no patch of nonpolar amino acids on the surface of the molecule.

Arguments founded solely on amino acid sequence hydrophobicity, therefore, give no compelling evidence for indicating a specific part of the exotoxin A molecule in membrane insertion. It would be sterically feasible for the helices of domain II, protruding at the top of the molecule, to undergo a significant conformational change and to insert into and disrupt a phospholipid bilayer. This could occur in the absence of any major change in structure of domains I and III. Such a transient membrane disruption could allow relatively nonspecific membrane translocation of proteins, including exotoxin A. Whether such a scenario actually occurs is a question that requires further examination.

We are indebted to the following people: Drs. Fred Richards and Tom Steitz for use of data collection facilities supported by National Institute of Health (NIH) Award GM-22778 to Yale University; Drs. Robert Kretsinger and colleagues for data collected on the Multiwire Area X-ray Diffractometer Biotechnology Resource at the University of Virginia; Drs. Ray Slaemme and Mike Schmid for graphics facilities use at the University of Arizona; Drs. Gerald Johnson, Pat Fitzhorn, Pat Lenders, and the Colorado State University College of Engineering for access to graphic facilities; Drs. Brian Matthews, Steve Remington, and Dale Troumud for access to computer facilities and assistance in molecular refinement at the University of Oregon; Kerri Wang and Lisa Case for technical assistance. This work has been supported by New Investigator Research Award 1028 from the Cystic Fibrosis Foundation and Award AI-19762 from NIH to D.B.M., a postdoctoral fellowship from the American Cancer Society to M.S.A., and Awards AI-20201 and CA-39217 from NIH to R.J.C. In addition, this project was supported in part by Grants RR07013-17 and RR07013-18 awarded by the Biomedical Research Support Grant Program, Division of Research Resources, NIH, to the University of Colorado.