Structure of the Isolated Catalytic Domain of Diphtheria Toxin†

Manfred S. Weiss,‡ Steven R. Blanke,§ R. John Collier,§ and David Eisenberg*,‡

Molecular Biology Institute and UCLA DOE Laboratory of Structural Biology and Molecular Medicine, University of California, Los Angeles, Los Angeles, California 90024-1570, and Department of Microbiology and Molecular Genetics, Harvard University, Boston, Massachusetts 02115

Received September 13, 1994; Revised Manuscript Received October 27, 1994⊗

ABSTRACT: The structure of the isolated catalytic domain of diphtheria toxin at pH 5.0 was determined by X-ray crystallography at 2.5 Å resolution and refined to an *R*-factor of 19.7%. The domain is bound to its endogenous inhibitor adenylyl(3′→5′)uridine 3′-monophosphate (ApUp). The structure of this 190-residue domain, which was expressed in and isolated from *Escherichia coli*, is essentially identical to the structure of the catalytic domain within whole diphtheria toxin determined at pH 7.5. However, there are two adjacent surface loops (loop 66−78 and loop 169−176) that exhibit clear differences when compared to the structure of the catalytic domain in whole diphtheria toxin. Although both loops are at the surface of the protein and are relatively flexible, the chain trace is well-defined in the electron density. The main structural difference is the closer approach of loops 66−78 and 169−176. We ascribe this structural change mainly to the absence of the neighboring transmembrane domain in the isolated catalytic domain as compared to whole diphtheria toxin. We suggest that this change represents the first step of the structural transition from the catalytic domain in whole diphtheria toxin to the translocated form of the domain. The changes are described in detail, and their implications for membrane translocation are discussed.

Diphtheria toxin is a 535-residue protein secreted by *Corynebacterium diphtheriae*, which causes diphtheria. Its structure consists of three domains, the catalytic domain (C-domain),[1] the transmembrane domain (T-domain), and the receptor binding domain (R-domain) (Choe *et al.*, 1992; Bennett *et al.*, 1994). The C-domain is the part of the protein that enters the cytoplasm of the target cell and ADP-ribosylates the side chain of the residue diphthamide, a posttranslationally modified histidine of EF-2 (Van Ness *et al.*, 1980), thus shutting down protein synthesis and killing the cell (Collier, 1975).

Bennett and Eisenberg (1994) proposed a stepwise structure-based mechanism for intoxication by DT, by combining structural information with results from many earlier biochemical studies. After DT is secreted from *C. diphtheriae*, a flexible loop between the C-domain and the T-domain is proteolyzed ("nicked") (Sandvig & Olsnes, 1981; Tsuneoka *et al.*, 1993). Then a prominent β-hairpin loop in the R-domain binds to the DT receptor at the surface of the target cell (Shen *et al.*, 1994), leading to receptor-mediated endocytosis (Morris *et al.*, 1985). In the endosome, the pH drops, possibly leading to the formation of an "open" monomer by protonation of acidic residues at the interface between the C- and the R-domain. In the open monomer, a structure found within dimeric DT (Bennett *et al.*, 1994), the R-domain is displaced from its position in the fully folded "closed" monomer. The T-domain then undergoes a conformational change and inserts into the membrane, facilitating the translocation of the C-domain across the membrane. However, the C-domain translocates only after the whole molecule has been nicked (Sandvig & Olsnes, 1981). Since it appears that low pH and temperature induce partial unfolding of DT (Blewitt *et al.*, 1985; Dumont & Richards, 1988; Ramsay *et al.*, 1989; Ramsay & Freire, 1990), it has been speculated that the C-domain translocates in at least a partially unfolded state. This is in accord with studies on double-cysteine mutants of the C-domain (Falnes *et al.*, 1994), where translocation was found to be slowed appreciably in four out of five cases. Sometime during or after translocation, the native disulfide bond between residues Cys186 of the C-domain and Cys201 of the T-domain is reduced (Madshus *et al.*, 1994), and the C-domain is released into the cytoplasm. This cell-mediated reduction of the disulfide bond has recently been shown to be the rate-determining step of toxin entry into the cell (Papini *et al.*, 1993).

The structures of the whole DT molecule in both monomeric and dimeric forms have been determined previously at pH 7.5 (Choe *et al.*, 1992; Bennett *et al.*, 1994; Bennett & Eisenberg, 1994). The core of the structure of the C-domain in whole DT is formed by eight β-strands forming two sheets of three and five strands, respectively. These two sheets are surrounded by six short α-helices (see Figure 1). These secondary structure elements are arranged into two subdomains, between which the presumed active site is located. The overall folding of the C-domain resembles the one found for the catalytic domain of *Pseudomonas aeruginosa* exotoxin, despite only weak sequence similarity between the two domains (Allured *et al.*, 1986). No structural

† This work was supported by an EMBO long-term fellowship (M.S.W.), by Postdoctoral Fellowship Award NIH80469 (S.R.B.), and by NIH Grant GM31299.

‡ University of California, Los Angeles.

§ Harvard University.

⊗ Abstract published in *Advance ACS Abstracts,* December 15, 1994.

[1] Abbreviations: DT, diphtheria toxin; DTA, isolated catalytic domain of DT (residues 1−190); C-domain, catalytic domain (residues 1−190) in DT; T-domain, transmembrane domain (residues 191−378) in DT; R-domain, receptor binding domain (residues 379−535) in DT; ApUp, adenylyl(3′→5′)uridine 3′-monophosphate; EF-2, elongation factor-2; σ, standard deviation; rms, root mean square; fragment A, proteolytic fragment of DT, approximately DTA; BME, β-mercaptoethanol; DTT, dithiothreitol.

0006-2960/95/0434-773$09.00/0

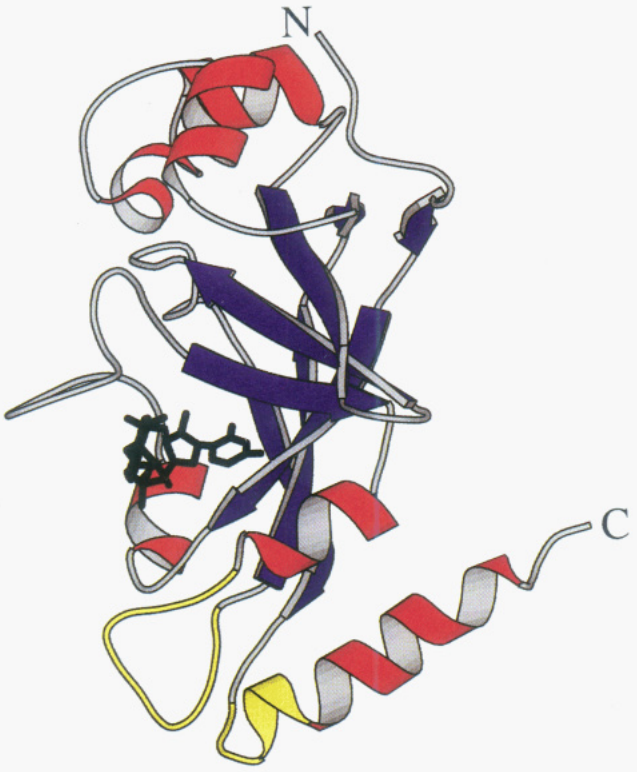

FIGURE 1: Ribbon representation of the structure of DTA at pH 5.0. α-Helices are in red, β-strands are in blue, and the inhibitor ApUp is shown in thick black lines. N- and C-termini are labeled. The two loops, 66−78 and 169−176, that are structurally different from the C-domain in DT are highlighted in yellow. The figure has been prepared using the program MOLSCRIPT (Kraulis, 1991).

information regarding the C-domain in its catalytically competent form has been available to date. In whole DT, the C-domain is essentially a proenzyme and is not able to catalyze the ADP-ribosylation of EF-2 at an appreciable rate. Fragment A, which is a proteolytic fragment of DT nearly equivalent to recombinant DTA, has also been crystallized previously in three different crystal forms (Kantardjeff *et al.*, 1989), but for none of them has a structure been yet reported. In this study, we wanted to address the following questions: (a) Is the conformation of isolated DTA the same as the conformation of the C-domain in context to the whole DT, (b) how does the conformation of DTA change with pH, (c) does the structure of DTA reveal anything about the transition from the proenzymatic form of the C-domain in DT to the enzymatically competent form of the isolated domain, and (d) does the DTA conformation at pH 5.0 give any hints about how the molecule is translocated across the membrane?

MATERIALS AND METHODS

Expression and Purification

Recombinant DNA Procedures. A recently constructed synthetic gene encoding the C-domain of diphtheria toxin was used for expression of protein (Blanke and Collier, unpublished results). The synthetic gene was designed by altering the codon usage of the corynephage β-gene to reflect the bias exhibited by highly expressed proteins in *Escherichia coli*. The gene was divided into smaller, evenly spaced fragments by engineering unique restriction sites throughout the open reading frame. The DTA gene was cloned into pET-15b (Novagen Inc.), replacing the *Nco*I−*Bam*HI fragment, and transformed into the *E. coli* strain BL21 (DE3a) for expression of proteins under transcriptional control of the T-7 promoter.

Fermentation and Harvesting of E. coli. A single 50-mL culture of L-broth (100 μg/mL ampicillin) was inoculated with a single colony from a LB-amp plate freshly streaked with BL21 (DE3a) transformed with pET-15b-DTA. This inoculum was grown to an $OD_{600} = 0.5$ and then placed at 4 °C overnight. On the second day, two 1-L baffled flasks, each containing 500 mL of L-broth (100 μg/mL ampicillin) prewarmed to 37 °C, were inoculated with 5 mL of the overnight culture. The flasks were aerated on a rotary shaker at 37 °C. The cultures were induced at an $OD_{600} = 1.0$ with 1−2 mM isopropyl β-D-galactopyranoside (IPTG). The cells were harvested after 1−2 h, immediately chilled to 4 °C, and centrifuged at 5000*g* for 15 min at 4 °C. The pellets were resuspended in 5 mL of ice-cold sonication buffer, composed of 50 mM Na_2HPO_4, pH 8.0, 100 mM KCl, 0.1% Tween-20, 1.0 mM phenylmethanesulfonyl fluoride, and 20 mM BME. The resuspended pellets were either frozen at −70 °C or sonicated immediately.

Purification of Mutant Toxins. The combined pellets were sonicated three times for 30−45 s using a Sonifier Cell Disruptor 350 (Branson Sonic Power Co.), with the power control at 5 and the duty cycle at 50%. The extracts were chilled on ice for at least 1 min between each sonication cycle. Cellular debris was pelleted by centrifugation at 3500*g* for 20 min at 4 °C. The supernatants were collected and chilled, while the pellets were resuspended in 5−10 mL of sonication buffer and sonicated again as described above.

The synthetic gene encoding DTA was constructed with an N-terminal fusion peptide of six histidines, in order to facilitate purification by nickel chelate affinity chromatography. The crude extracts were clarified immediately before chromatography by centrifugation at 20000*g* for 15 min. The extracts were loaded directly onto a 3−5-mL nickel chelate affinity column (Qiagen), preequilibrated with 5 bed volumes of sonication buffer. The column was washed with 3 bed volumes of sonication buffer, followed by 3−5 bed volumes of 15 mM imidazole in sonication buffer. It was further washed with 3−5 bed volumes of 20 mM imidazole and then 3−5 bed volumes of 25 mM imidazole, both in 25 mM Tris-HCl, pH 8.0, and 20 mM BME. The purified protein was eluted with 5 bed volumes of 50 mM imidazole in 25 mM Tris-HCl, pH 8.0, and 20 mM BME. The column fractions were analyzed by SDS−PAGE, and the purified protein was pooled.

Proteolytic Removal of the Polyhistidine Fusion Peptide. The synthetic gene was generated with a modified thrombin recognition site (Leu-Val-Pro-Arg-Gly-Ala), linking the carboxy terminus of the polyhistidine fusion peptide to the amino terminus of DTA. Cleavage of the Arg−Gly peptide bond with thrombin results in production of DTA with an authentic amino-terminal sequence (Gly-Ala). Typically, 1 mg of DTA was incubated with 2−5 units of human thrombin (Novagen) for 1−5 days at 20 °C in 20 mM Tris-HCl, pH 8.4, 150 mM NaCl, 2.5 mM $CaCl_2$, and 2 mM DTT. The intact DTA was purified using a single anion-exchange FPLC chromatography fractionation step (FPLC: Mono-Q, Pharmacia).

Alternatively, the polyhistidine fusion peptide was removed by limited trypsin digestion, also yielding an authentic amino terminus. Generally, 1.5 mg of DTA was incubated with 12 μg of bovine pancreatic trypsin (TPCK treated, Sigma) for 1 h on ice in 50 mM Tris-HCl, pH 7.4, and 10 mM BME. The reaction was stopped by the addition of 48 μg of soybean trypsin inhibitor (Sigma). The entire incubation mixture was run over a 1-mL nickel chelate affinity column to remove uncut DTA, as well as the cleaved polyhistidine peptide. The flowthrough was collected, and the column was washed twice with 1 mL of FPLC buffer A (20 mM Tris-HCl, pH 7.5, 5% glycerol, and 10 mM BME). The flowthrough and wash steps were collected and pooled.

Trypsin-digested DTA (3−6 mg) was fractionated using anion-exchange chromatography (FPLC: Mono-Q, Pharmacia). After the sample was loaded, the column was washed with 5 mL of FPLC buffer A, and the DTA was eluted using a 60-mL linear gradient of 0−125 mM NaCl in buffer A. The purified DTA generally eluted as a peak at 60−80 mM NaCl. The fractions were analyzed by SDS−PAGE. Those fractions purified to greater than 99% homogeneity were pooled and concentrated for further use. Fractions of lesser purity were pooled, desalted, and refractionated.

Crystallization and Data Collection

Crystals of the complex of DTA with its inhibitor ApUp were obtained using the hanging drop method at pH 4.7−5.3 at a protein concentration of about 9 mg/mL in the presence of 25−75 mM $MgCl_2$, 0−5 mM $CaCl_2$, 5 mM DTT, 100 mM NaOAc, and >28% PEG-3350 as a precipitant. The inhibitor concentration was 1−5 mM which is 2−10 times the protein concentration on a molar basis. Crystals appeared within 3 days and grew to their maximum size in less than 1 week. They usually grew in clumps of many crystals, with single crystals a rarity. The space group of the crystals is $C2$ with cell dimensions $a = 55.7$ Å, $b = 43.9$ Å, $c = 73.9$ Å, and $\beta = 101.5°$. With one molecule per asymmetric unit, the Matthews parameter is 2.1 Å^3/Da. The diffraction of the crystals extends to further than 2.5 Å but is very anisotropic beyond 3.0 Å (see section about refinement). There is no relationship between this crystal form and the three crystal forms of fragment A previously described by Kantardjeff *et al.* (1989).

Data were collected from a single crystal on an RAXIS imaging plate detector at 5 °C. Processing was carried out using the standard RAXIS program package. Sixty-one frames covering a total of 122° around one axis gave a data set, which is 83.7% complete to 2.5 Å resolution with an average I/σ of 1.7 in the outermost shell (Table 1). The overall R_{sym} is 6.2%, and the total number of unique reflections is 5315. Despite the relatively low R_{sym} the quality of the data decreases dramatically beyond 2.8 Å resolution because of anisotropic scattering. An analysis of mean structure factors in different areas of the reciprocal space using data from 2.8 to 2.5 Å showed that the mean structure factors in the c^* direction are larger by a factor of 2−3 than the ones in the a^* and b^* directions.

Structure Solution and Refinement

The structure was solved by molecular replacement using as a search model fragment 1−187 of dimeric DT, which is the part of the C-domain that was visible in the electron density (Bennett *et al.*, 1994). The bound inhibitor ApUp was omitted from the search model. The radius of gyration of the search model was calculated to be 17.2 Å; therefore, the maximum Patterson vector length was chosen to be 25 Å. The programs X-PLOR (Brünger, 1990; Brünger *et al.*, 1990) and ALMN (SERC Daresbury Laboratory, 1986) gave unambiguous rotation function solutions. For 1155 reflections between 8 and 4 Å using a 2σ cutoff on F, the highest peak using X-PLOR was at 9.4σ; the next highest completely unrelated peak was at 4.5σ. The Eulerian angles of the highest peak were (170.0, 82.0, 98.0). Using Patterson correlation refinement in the resolution range 8−4 Å, the peak refined to (169.59, 81.147, 98.70), giving a correlation coefficient on E^2 of 0.31 with all other refined unrelated peaks having a correlation coefficient of 0.13 or below. The translation function was calculated using data between 10 and 3 Å. Using X-PLOR, the highest peak was at fractional coordinates (0.375, 0.0, 0.063), 7.9 standard deviations above the mean. The next highest peak was at 3.1σ. The same solution was obtained with the program TSEARCH (SERC Daresbury Laboratory, 1986).

The entire refinement was carried out using the program X-PLOR (Brünger *et al.*, 1990). Electron density maps were calculated according to Read (1986), and the program FRODO (Jones, 1978) was used for displaying the model and the maps and for rebuilding the model.

The model was rotated and translated according to the molecular replacement solution, and 50 cycles of rigid body refinement at 3 Å resolution decreased the R-factor to 45.7%. After a complete refinement round including simulated annealing refinement and restrained temperature factor refinement at 2.5 Å, the R-factor was 29.0%. High R-factors in the high-resolution shells and the corresponding map, which looked much like a 3 Å map, suggested that the isotropic model did not fit the anisotropic data well. After the observed data were scaled anisotropically (see Table 1), using the isotropic model as a reference, the R-factor was decreased to 26.1% overall, and the corresponding map revealed clear difference density for the bound inhibitor ApUp. ApUp was built in and the refinement carried on with the scaled data. After the second round the R-factor was decreased to 22.9%. It was evident from the density that two loops had to be rebuilt, although the density was not of good enough quality to allow one to unambiguously trace the chain on the first try. Three more rounds of refinement and careful manual rebuilding were carried out. After each round, the now improved model was used to recalculate the anisotropic correction for F_o. The final R-factor is 19.7% for all data between 10 and 2.5 Å (5228 reflections); the final model with a total of 1505 non-hydrogen atoms consists of all 190 amino acids, with the side chains of Lys172 and Asn189 missing, 1 molecule of ApUp, and 8 well-defined water molecules. A summary of the data collection and refinement statistics is given in Table 1.

The coordinates of the atomic model and the anisotropically scaled structure factors have been deposited in the Brookhaven Protein Data Bank (identification code 1DTP).

RESULTS

Quality of the Structure. The R-factor of the final model compared to the anisotropically scaled data is 19.7% for all

Table 1: Statistics of X-ray Data Collection and Refinement of DTA

data parameters	
resolution (Å)	2.5
R_{sym}[a] (%)	6.2 (18.9)[b]
unique reflections	5315
completeness (%)	83.7 (80.4)[b]
no. of atoms	
non-hydrogen	1505
protein[c]	1454
ApUp	43
water	8
refinement parameters	
resolution range (Å)	10−2.5
R-factor[d] (%)	19.7 (35.0)[b]
no. of reflections	5228
intensity cutoff ($F/\sigma F$)	0.0
av atomic B-factors (Å^2)	
protein	38.4
ApUp	39.6
water	43.2
overall anisotropic B-factors[e] (Å^2)	
B_{11}	−9.7
B_{22}	−26.7
B_{33}	+13.9
B_{13}	+14.2
rms deviation from ideality	
bonds (Å)	0.023
angles (deg)	4.4
dihedrals (deg)	26.4
impropers (deg)	1.7

[a] $R_{sym} = 100\sum_h|I_h - \langle I\rangle|/\sum_h I_h$, where $\langle I\rangle$ is the mean intensity of all symmetry-related reflections I_h. [b] The numbers in parentheses refer to the resolution range 2.8−2.5 Å. [c] The side chains of Lys172 and Asn189 are truncated to alanines. [d] $R = 100\sum_h||F_o| - |F_c||/\sum_h|F_o|$. [e] B_{ii} are applied to F_o; $F_{o,corr} = F_o \exp[(s_{11}^2B_{11} + s_{22}^2B_{22} + s_{33}^2B_{33} + 2s_1s_2B_{12} + 2s_1s_3B_{13} + 2s_2s_3B_{23})/4]$. In space group $C2$ the mixed terms B_{12} and B_{23} are both 0.0.

reflections between 10 and 2.5 Å resolution. However, there is a steep increase in R-factor beyond 2.8 Å, showing that the quality of the higher resolution data is quite poor and/or that the anisotropic correction is not sufficient to fit the model to the data. In a Luzzati plot (Luzzati, 1952) using data to only 3 Å resolution the R-factor vs resolution curve matches the theoretical one for a coordinate error of 0.25 Å quite well (data not shown).

The final $2F_o - F_c$ electron density map shows continuous density at the 1σ level for all main-chain atoms except for two small breaks at the peptide bond of Asn16−Phe17 in the surface loop between the first and second β-strand and at Gly79 at the beginning of the fourth β-strand. An illustration of the electron density is given in Figure 2. Depicted is the inhibitor ApUp and the final SIGMAA weighted map (Read, 1986) at the 1σ level. The weaker density at the adenine ring may correspond to a weaker binding of the inhibitor at this site as compared to the uracil ring.

The stereochemical parameters commonly used to evaluate the quality of a structure are the rms deviations of bond lengths, bond angles, dihedral angles, and improper angles from ideality. For the protein part of the structure the values as determined using X-PLOR and the param19x parameter set are 0.023 Å, 4.4°, 26.4°, and 1.7°, respectively.

The program package PROCHECK (Laskowski *et al.*, 1993) assesses a number of different main- and side-chain parameters. In all categories the parameter values are found

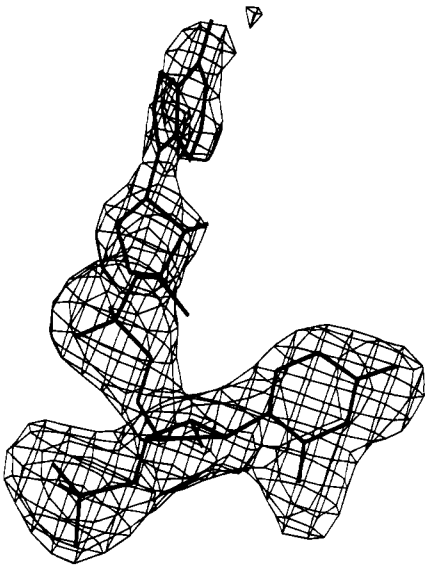

FIGURE 2: Final SIGMAA weighted $2F_o - F_c$ electron density map (Read, 1986) at the 1σ level of the active site inhibitor ApUp. The inhibitor is shown in thick lines. The orientation is roughly 90° about an axis horizontal in the plane of the picture from the ones depicted in Figures 1 and 3. Protein density around ApUp has been omitted for clarity.

to be very close to the given comparison values. For instance, 131 out of 163 non-glyine and non-proline residues (80.4%) are in the most favored regions of the Ramachandran plot, with the comparison value typical for a 2.5 Å structure being 76.6%.

The 3D−1D profile score (Bowie *et al.*, 1991; Lüthy *et al.*, 1992) of DTA averaged over a window of 21 residues shows no drop below 0.1 (data not shown), a criterion that has been established to hold true for most correct, well-refined structures.

A method to check for incorrectly determined regions of a protein structure was published by Colovos and Yeates (1993). Their procedure exploits the fact that different atom types are distributed nonrandomly with respect to each other in regions where the structure is determined correctly, whereas in incorrect regions a more random distribution should be expected. As determined with their program ERRAT, 95.7% of the structure of DTA scored below the empirically determined 95% confidence limit, which is better than the average score in the PDB for structures determined at 2.5 Å resolution or better.

Comparison of the Structure of DTA at pH 5.0 to the Structure of the C-Domain of Whole DT at pH 7.5. The structure of DTA at pH 5.0 is very similar to the structure of the C-domain of dimeric DT determined at pH 7.5 [see Figure 1 and Bennett *et al.* (1994)]. A superposition is shown in Figure 3. Using C_α atoms and the algorithm of Kabsch (1978), the structures superimpose well, yielding an overall rms deviation of 1.3 Å for all 187 C_α atom pairs; if one considers only the 179 pairs of atoms closer than 3 Å, the rms deviation is 0.8 Å. A plot of the rms difference vs the residue number is shown in Figure 4. The largest differences are observed in the two adjacent loops, loop 66−78 and loop 169−176. In dimeric DT, the two loops are spread apart by a network of polar interactions between residues in the loops and in the adjacent transmembrane domain (Table 2). The only protein−protein contacts between these two loops are formed by Gln175-NE2 on the

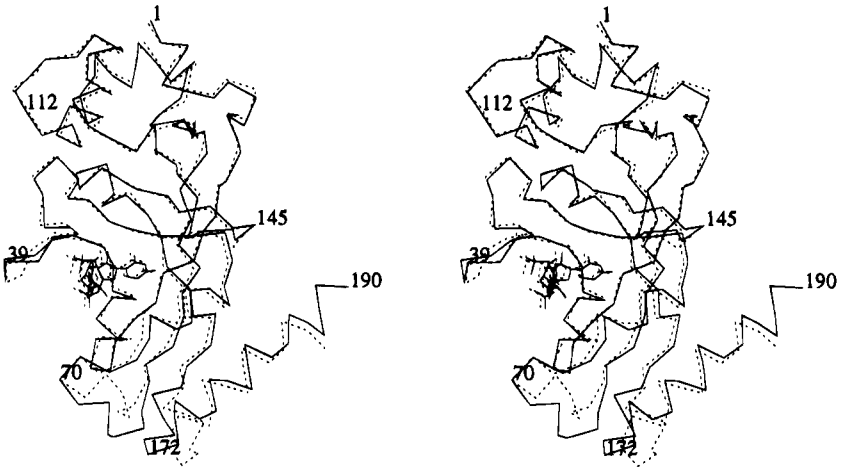

FIGURE 3: Stereo C_α superposition of the ApUp complex of the isolated catalytic domain of DT with the ApUp complex of the C-domain in dimeric DT. Some residues are labeled. Key: solid line, DTA; broken line, C-domain of dimeric DT (residues 1−187). Notable differences are seen only in the two loops at the bottom of the picture.

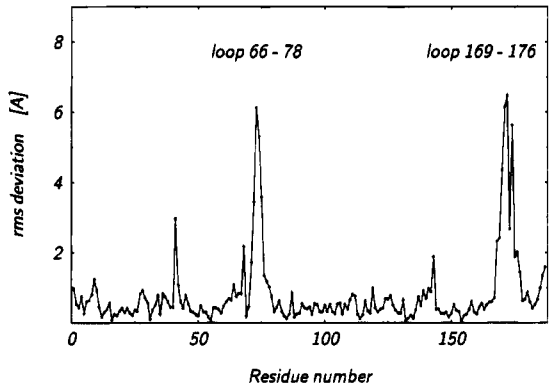

FIGURE 4: rms deviation between the structures of DTA and the C-domain of dimeric DT vs residue number. The overall value is 0.8 Å for 179 C_α atom pairs closer than 3 Å and 1.3 Å for all 187 C_α atom pairs. Notice that the only significant differences occur in the two adjacent loops, 66−78 and 169−176.

one side and Leu73-O and Gly75-O (Table 3). Upon removal of the T-domain, the two loops move closer together.

Main-chain−main-chain hydrogen bonds are formed between Gly75-O and Gln175-N, Gly75-O and Gly174-N, Lys76-O and Gln175-N, and Lys76-N and Gly171-O (Table 3). There are also polar contacts gained and lost between residues within the loops (Table 4) and between the residues of the loops and other parts of the protein (Table 5). The number of gained and lost contacts in Tables 4 and 5 is approximately balanced, whereas Table 2 lists lost contacts only and Table 3 lists mainly gained contacts.

Compared to the C-domain in DT, loops 66−78 and 169−176 shift about 2 Å closer together in DTA as determined by comparing the average coordinates of all the non-hydrogen atoms (the numbers in parentheses are the corresponding ones for C_α atoms only) of the two loops. In the C-domain in DT they are 12.8 Å (12.0 Å) apart, whereas in DTA they are only 10.4 Å (8.6 Å) apart. On the basis of the superposition of C_α atoms, the average coordinate for loop 66−78 shifts by 1.1 Å (1.3 Å) and the average coordinate for loop 169−176 by 1.9 Å (2.4 Å). Thus, the separation of the loops is reduced by about 2.4 Å (3.4 Å). The structure of the two loops will be examined in further detail in the next two sections.

The mobility of the main chain is shown in Figure 5, as a plot of mean main-chain atom temperature factor vs residue

Table 2: Polar Contacts[a] between Loop 66−78 and Loop 169−176 in the C-Domain of Dimeric DT and the T-Domain[b,c]

atom 1	atom 2	distance (Å)
Pro72-N	Glu298-OE2	3.48
Leu73-N	Glu298-OE1	3.41
Leu73-N	Glu298-OE2	3.05
Leu73-O	Asn366-OD1	3.50
Ser74-N	Glu298-OE1	2.85
Ser74-OG	Glu298-N	3.47
Ser74-OG	Glu298-OE1	2.85
Arg173-NH1	Asp207-OD1	2.62
Arg173-NH2	Asp207-OD2	2.49
Gln175-OE1	Asn366-OD1	3.16
Asp176-OD1	Arg210-NE	3.00
Asp176-OD1	Arg210-NH2	2.75
Asp176-OD1	Gln369-OE1	3.10
Asp176-OD2	Arg210-NE	2.81

[a] The criterion for assigning a polar contact is that the distance between polar atoms has to be ≤3.5 Å. The program used for calculating the list of contacts was CONTACT (SERC Daresbury Laboratory, 1986). [b] In Bennett *et al.* (1994) there are only two other polar interactions obeying the distance limit between residues in the C- and the T-domains listed. These are Asp61-OD2 to Arg377-NH1 (3.0 Å) and Ala183-O to Asn203-ND2 (3.1 Å). [c] These interactions are lost when the C-domain is removed from whole DT. There are also interactions between residues in the C-domain and the R-domain which are lost upon removal of the C-domain. Since these affect other parts of the C-domain structure, which have not been found to change, they are not considered here.

number. The overall temperature factor is 38.0 Å^2 for DTA as compared to 24.4 Å^2 for the C-domain of dimeric DT (Bennett, 1993). The largest increase in chain flexibility is found in the segments around residues 41, 69, 100, and 144. All of these segments are surface loops; therefore, the increase in temperature factor is not surprising.

The effect of the conformational changes on different parts of the structure can be assessed by using the 3D−1D profile method (Bowie *et al.*, 1991; Lüthy *et al.*, 1992). When the C-domain is computationally removed from the structure of DT, the overall profile score drops from 93.5 to 87.6. Going from the C-domain as it is in whole DT to the DTA structure, as described in this paper, the overall score stays virtually the same, with only a slight drop to 86.0. The numbers are in agreement with the fact that the isolated catalytic domain is less stable with respect to thermal denaturation than the C-domain within the context of whole DT (Ramsay *et al.*,

Table 3: Polar Contacts between Loop 66−78 and Loop 169−176 in DTA and in the C-Domain in Dimeric DT[a]

		distance (Å)	
atom 1	atom 2	DTA	C-domain
Pro72-O	Gln175-OE1	3.40	
Pro72-O	Gln175-NE2	3.05	
Leu73-O	Gln175-NE2		3.50
Ser74-O	Gln175-NE2	2.66	
Gly75-O	Gly174-N	2.69	
Gly75-O	Gln175-N	2.78	
Gly75-O	Gln175-NE2		2.87
Lys76-N	Gly171-O	3.23	
Lys76-O	Gln175-N	2.85	

[a] See Bennett *et al.* (1994). Interactions are listed when the distance criterion of ≤3.5 Å between polar atoms was met either in DTA or in the C-domain of dimeric DT. Notice that 7 contacts are gained when DTA is removed from DT and two contacts are lost.

Table 4: Polar Contacts within Loop 66−78 and Loop 169−176 in DTA and in the C-Domain in Dimeric DT[a]

		distance (Å)	
atom 1	atom 2	DTA	C-domain
	Loop 66−78		
Ser66-O	Gly78-N	3.05	2.85
Val67-O	Asn69-N		3.07
Asp68-N	Lys76-O		2.90
Asp68-O	Asn71-N		3.18
Asp68-O	Pro72-N		3.35
Asp68-OD1	Asn69-N	2.67	2.85
Asp68-OD1	Glu70-N		2.81
Asp68-OD2	Asn71-ND2		2.86
Glu70-O	Asn71-OD1		3.34
Asn71-O	Leu-73-N	3.13	3.29
Asn71-O	Ser74-N	3.27	
Asn71-O	Ser74-OG	2.91	
Asn71-O	Gly75-N		2.88
Leu73-O	Gly75-N	3.33	
Ser74-O	Lys76-N	2.85	
	Loop 169−176		
Thr169-O	Gly171-N	2.81	3.35
Arg170-O	Lys172-N[b]	2.76	3.37
Arg170-O	Arg173-N	2.52	
Gly171-O	Arg173-N	3.21	
Lys172-NZ[b]	Asp176-OD2		3.00
Arg173-O	Gln175-N		3.27
Arg173-O	Asp176-N	3.42	3.18
Arg173-O	Asp176-OD1		3.35
Arg173-O	Ala177-N	2.96	
Gly174-O	Asp176-N	3.23	
Gly174-O	Ala177-N	3.28	

[a] See Bennett *et al.* (1994). Interactions are listed when the distance criterion of ≤3.5 Å between polar atoms was met either in DTA or in the C-domain of dimeric DT. Notice that 10 contacts are gained when DTA is removed from DT and 11 contacts are lost. Of the 5 remaining contacts 3 are strengthened when DTA is removed from DT and 2 are weakened. [b] In DTA the side chains of Lys172 and Asn189 are truncated to Ala.

1989). Both the N-terminal segment from 1 to 85 and the C-terminal segment from 140 to 190 have lower scores than the C-domain in DT, but the segment in between has a higher score. Interestingly, in loop 66−78, which shifts position, the profile score drops dramatically when the C-domain is computationally removed from DT, but then recovers partly in DTA, due to the structural changes (data not shown). These conformational changes could be interpreted as a way to create more favorable residue environments when DTA is separated from the rest of DT. However, the same does not hold true for loop 169−176.

Table 5: Polar Contacts between Loop 66−78 and Loop 169−176 with Other Polar Atoms in DTA and in the C-Domain in Dimeric DT[a]

		distance (Å)	
atom 1	atom 2	DTA	C-domain
Ser66-N	Ala63-O	2.85	3.30
Ser66-OG	Ala63-O	2.63	2.80
Val67-N	Tyr65-O	3.06	
Val67-O	Lys24-NZ	3.43	
Glu70-OE1	Glu70′-OE1[b]	2.86	
Asn71-ND2	Glu142″-OE2[b]	3.19	
Ala77-O	Glu168-N	2.75	
Gly78-N	Thr23-O	3.42	
Gly78-O	Asn166-OD1	2.59	3.34
Gly78-O	Phe167-N	2.61	2.71
Gly78-O	Glu168-N	2.61	
Thr169-N	Asn166-O		3.28
Thr169-N	Asn166-OD1		3.35
Thr169-N	Phe167-O	3.24	3.34
Arg170-N	Phe167-O	3.18	3.13
Arg170-N	Glu168-O		3.20
Gly171-N	Glu168-O		3.29
Gly174-O	Met178-N	2.85	
Gln175-O	Met178-N		3.30
Gln175-O	Tyr179-N		2.98
Asp176-O	Met178-N	2.69	3.50
Asp176-O	Tyr179-N	2.69	
Asp176-O	Glu180-N		2.89

[a] See Bennett *et al.* (1994). Interactions are listed when the distance criterion of ≤3.5 Å between polar atoms was met either in DTA or in the C-domain of dimeric DT. Also listed are crystal contacts involving polar atoms of loops 66−78 and 169−176 in DTA. [b] The symmetry operator relating primed residues to the original one is $(-x, y, -z)$; for double-primed residues it is $(-x + 0.5, y + 0.5, -z)$.

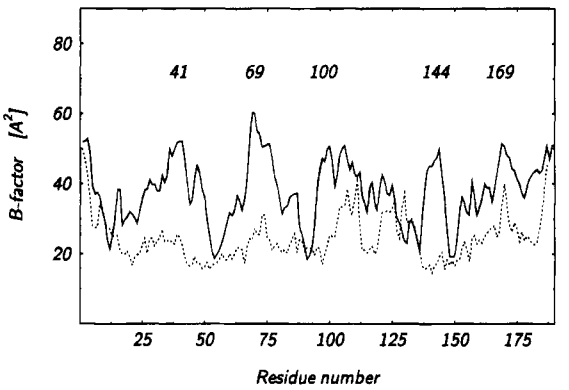

FIGURE 5: Flexibility of the main chain vs residue number. The average main-chain temperature factor is plotted for DTA as a solid line and the C-domain of dimeric DT (Bennett *et al.*, 1994) as a broken line. The overall values are 38.0 Å² for DTA and 24.4 Å² for the C-domain of dimeric DT. The respective maxima at positions 41, 69, 100, 144, and 169 are labeled. All of them occur in loops at the surface of the molecule.

Structure of Loop 66−78. Loop 66−78 is located in the sequence between the second α-helix of whole DT (Bennett *et al.*, 1994), which ends at Tyr65, and the fourth β-strand, which starts at Gly79. In the C-domain of dimeric DT this loop is anchored by a β-bridge formed by hydrogen bonds between Ser66-O and Gly78-N and between Asp68−N and Lys76-O (Table 4). Two more hydrogen-bonded reverse turns are found at positions Asn69-Glu70 and Pro72-Leu73 (Table 6). In DTA the anchoring β-bridge as well as the turn at Asn69-Glu70 is lost, whereas the other turn at Pro72-Leu73 is strengthened. Overall interactions between atoms within this loop are weaker, thus leaving it more flexible

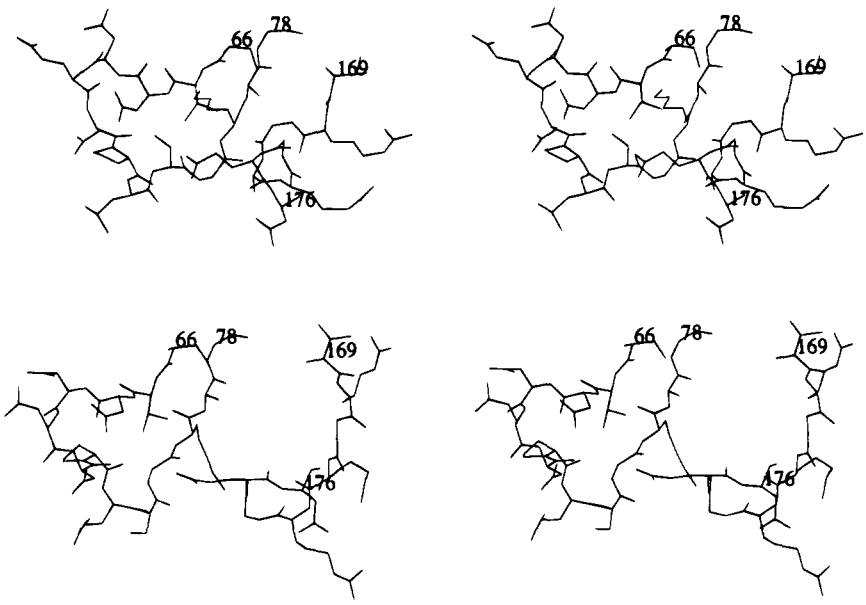

FIGURE 6: Stereo figures of the structure of chain segments 66−78 and 169−176 in DTA (a, top) and the C-domain of dimeric DT (b, bottom). The orientation is approximately the same in both figures and is also very similar to the one depicted in Figures 1 and 3. Notice that in (a) the side chain of Lys172 is truncated beyond C_β.

Table 6: Dihedral Angles of Loop 66−78 and Loop 169−176 Residues for DTA and the C-Domain of Dimeric DT[a] Showing the Major Changes in Structure between DTA at pH 5.0 and the C-Domain in DT at pH 7.5

	DTA				dimeric DT		
residue	ϕ (deg)	ψ (deg)	secondary structure[b]	distance[c] (deg)	ϕ (deg)	ψ (deg)	secondary structure[d]
Ser66	−90	+24		22	−103	+6	
Val67	−123	+165		46	−89	+134	E
Asp68	+45	−140		127	−38	+124	E
Asn69	+86	+88		190	−71	−19	T
Glu70	+56	+89		175	−76	−26	T
Asn71	+94	+87		130	−137	+75	
Pro72	−71	−9	T	14	−69	−23	T
Leu73	−78	−27	T	14	−78	−41	T
Ser74	−82	+117		120	−133	+8	
Gly75	−69	+47		131	−77	+178	
Lys76	−84	−162		46	−73	+153	E
Ala77	+178	+149		76	−106	+144	E
Gly78	−85	−67	E	9	−89	−75	E
Thr169	−78	−33	T	16	−68	−21	T
Arg170	−76	+19		45	−65	−25	
Gly171	+24	−94	T	132	+65	+29	
Lys172	−65	−36	T	164	−98	+163	
Arg173	−89	+7	T	155	−76	+161	
Gly174	−56	−61	H	120	+44	−128	T
Gln175	−58	−29	H	81	−117	+27	T
Asp176	−77	−74	H	32	−53	−53	H

[a] See Bennett *et al.* (1994). [b] Secondary structure elements were assigned manually. Nomenclature: E, β-sheet, T, hydrogen-bonded reverse turn; H, α-helix. [c] The distance was calculated as $(\Delta\phi^2 + \Delta\psi^2)^{1/2}$ with $\Delta\phi$ and $\Delta\psi$ being the smallest distance in the angular space. [d] Taken from Bennett (1993). The type I turn at residues 72−73 is listed here, despite the fact that the length of the hydrogen bond is 3.6 Å and therefore is not within the limits which were used for assigning secondary structure elements.

(see also Figure 5). The movement of this whole loop can be described as a rotation and translocation of the more or less unchanged type I turn at Pro72-Leu73. As can be seen from Table 6, the main-chain dihedral angles of Pro72 and Leu73 are preserved, whereas big movements occur at residues before and after this turn. The removal of the T-domain causes the side chain of Leu73 to become almost completely solvent exposed. The rotation and translation of this turn Pro72-Leu73 to the very tip of this loop keep it solvent exposed rather than packing it against some other side chain as might have been expected. Notable are also the positive ϕ angles observed for residues 68−71 (Table 6). Although this segment has a high temperature factor (Figure 5), it fits the electron density well. Furthermore, this segment consists of two Asn, one Asp, and one Glu, residues for which positive ϕ angles are not uncommon.

Structure of Loop 169−176. Loop 169−176 starts two residues after the eighth β-strand and ends with the first residue of the final α-helix in the C-domain. In DTA as compared to dimeric DT there are two more hydrogen-bonded reverse turns formed at positions Gly171-Lys172 and Lys172-Arg173. The first one is of type II′ and the second one (interlocked with the first one) of type I. The main-chain dihedral angles of both residues of the type II′ turn at Gly174-Gln175 in dimeric DT have become α-helical, thus extending the following α-helix by two residues (Gly174 and Gln175) in DTA. With the newly formed secondary structure elements in this loop, the increase in flexibility seen here is less than the average increase of 13.6 Å^2 for the whole chain (see Figure 5).

DISCUSSION

What Causes the Structural Transition To Occur? There are at least three possible causes for the structural transition between the C-domain in DT and DTA, which is mainly the approach of loops 66−78 and 169−176. One possibility is the different crystalline environment, the second is the difference in pH, and the third is the removal of the T-domain from the vicinity of the two loops.

First, we consider changes in crystal contacts. Of all the residues in the two loops, only Glu70 and Asn71 are engaged in polar interactions (distance between polar atoms ≤3.5 Å) to symmetry-related molecules. The side chain of Glu70 contacts its own symmetry mate, whereas Asn71-ND2 forms a hydrogen bond with Glu142-OE2 of a symmetry-related molecule. If one relaxes the criterion for crystal contacts to

distances ≤4.5 Å, then Leu73 (contacts Glu142), Ser74 (contacts Glu142), and Asp176 (contacts Ala187) have to be added to the list. Altogether, however, with only two strong, polar contacts it seems very unlikely that crystal packing forces can be the dominant driving force for the two loops to move together.

Next we consider the pH change as the primary cause for the structural transition. From the sequence of the two loops, it is not evident how a change in pH from 7.5 to 5.0 could cause the structural transition. The only histidine in the structure is located in the cleft between the two subdomains and is therefore more than 10 Å away (the distance between His21-CA and Ser66-CA is 8.7 Å). Also there is no cluster of negatively charged residues where the p*K* of one of the carboxylic acid side chains could be elevated so that protonation would occur at pH 5.0. Furthermore, the N-terminal α-NH_3 group is on the other side of the molecule, and the SH group of Cys186 in the C-terminal α-helix is also more than 15 Å away (the distance between Cys186-SG and Asp176-CA is 15.5 Å). Thus there are no acidic or basic groups near the two loops that would be expected to change charge appreciably between pH 5.0 and pH 7.5. Of course, in the absence of detailed electrostatic calculations on the structure, we cannot rule out the possibility that global effects on the electrostatic potential of the structure could cause the transition.

The remaining possibility is that the removal of the rest of the DT structure causes the loops to move closer together. In fact, there are 14 polar interactions that are lost when the T-domain is removed from DTA (Table 2). Thus the absence of the neighboring T-domain would be expected to release constraints on the neighboring loops. Once the T-domain has been removed, some seven polar interactions between residues of the two loops are gained as they move closer together, whereas only two are lost. This is shown in Table 3. The movement of the two loops is therefore energetically favored once the T-domain is removed. For loop 66−78 this interpretation is corroborated by the profile score, which drops when the C-domain is extracted from whole DT and recovers partly after the conformational change has taken place. However, no such effect could be observed for loop 169−176. In this segment the average profile score is considerably lower than in the C-domain in whole DT and also the C-domain by itself.

In summary, although we cannot completely rule out the possible influence of changes in crystal contacts and of the pH change on the approach of the loops, the current evidence favors the removal of the T-domain as the primary cause for the conformational change.

Implications for Enzymatic Function of DTA. The major function of DTA is to ADP-ribosylate diphthamide, a posttranslationally modified histidine residue of EF-2 (Van Ness *et al.*, 1980). This occurs in the cytosol of the target cell at an approximately neutral pH. Whole DT is inactive, and a possible reason for that was revealed by the X-ray structures of monomeric and dimeric DT (Bennett *et al.*, 1994; Bennett & Eisenberg, 1994). In both structures, the active site is partially blocked by the adjacent R-domain. A detailed description of the interactions between the C-domain and the inhibitor ApUp can be found in Figures 12−14 in Bennett *et al.* (1994). Neither the active site geometry nor the conformation of the bound inhibitor ApUp is different in DTA compared to what was found in the structures of dimeric and monomeric DT. Therefore, it seems unlikely that the movement of the loops has a large influence on the catalytic activity of the protein. It could be, however, that the loops are involved in recognition and binding of EF-2. Because of the large size of EF-2 (100 kDa) the recognition site can be quite a distance away from the active site. Therefore, although neither the residues in the active site nor the surrounding residues change in going from the structure of the C-domain in whole DT to the structure of DTA, it is possible that the movement of the loops, which are more than 10 Å away from the active site, forms the intact recognition and binding site for EF-2, thus rendering DTA in its catalytically competent form.

Implications for Membrane Translocation. Membrane translocation of the C-domain of DT occurs in cells following receptor-mediated endocytosis (Morris *et al.*, 1985) and in a cell-free system using purified endosomes (Beaumelle *et al.*, 1992). The common feature of these translocation systems is that they require a pH gradient. The current hypothesis for membrane translocation of DTA assumes that at low pH the T-domain inserts into the membrane and facilitates the translocation of the C-domain (Bennett & Eisenberg, 1994). The C-domain is believed to be at least partially unfolded during the translocation process (Blewitt *et al.*, 1985; Dumont & Richards, 1988; Ramsay *et al.*, 1989; Ramsay & Freire, 1990; Falnes *et al.*, 1994). This unfolding is triggered either by low pH or by the conformational change of the T-domain. Once translocated and freed from the T- and the R-domain, the C-domain refolds. The fact that the structure of DTA is virtually identical to the structure of the C-domain in whole DT shows that the isolated C-domain does not need the rest of the DT molecule to fold. That is, the translocated C-domain can refold into its active conformation. This observation is in accord with calorimetric studies, which showed that the thermal denaturation of only isolated fragment A is reversible (Ramsay *et al.*, 1989), whereas that of whole DT is not.

As mentioned above, both translocation systems require a pH gradient for translocation to occur. Not much is known though about the events that follow after DT is exposed to low pH in the acidified endosomal compartment. Does low pH trigger the change of conformation in the C-domain, then leading to the unfolding of the domain? Or is it that low pH triggers the insertion of the T-domain into the membrane and the resulting conformational change of the T-domain or a further decrease in pH which induces the C-domain to unfold? Above we argued that the change in pH down to pH 5.0 is not the primary cause for the structural transition of the C-domain. We therefore suggest that it is the T-domain which changes conformation upon a decrease in pH and that this event initiates the translocation of the C-domain. This interpretation is in accord with mutational studies that showed the first α-helix of the T-domain (residues 205−222), which is amphipathic, is the key to the initiation of translocation (Madshus, 1994; vanderSpek *et al.*, 1994).

Since in whole DT the C-domain is linked to the T-domain by a disulfide bridge, involving Cys186 proximal to the carboxy terminus of the C-domain and Cys201 close to the amino terminus of the T-domain, it is likely that the C-terminus inserts first into the membrane and the C-terminus therefore has to unfold first. This need to unfold first could be reflected in the lower profile score of the

C-terminal segment as compared to the score of the same segment of the C-domain in whole DT. The observed DTA structure could therefore be an intermediate structure, between the more stable C-domain in DT and the translocated form, ready to unfold and to insert into the membrane.

REFERENCES

Allured, V. S., Collier, R. J., Carroll, S. F., & McKay, D. B. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 1320−1324.

Beaumelle, B., Bensammar, L., & Bienvenüe, A. (1992) *J. Biol. Chem. 267*, 11525−11531.

Bennett, M. J. (1993) Ph.D. Thesis, University of California, Los Angeles.

Bennett, M. J., & Eisenberg, D. (1994) *Protein Sci. 3*, 1464−1475.

Bennett, M. J., Choe, S., & Eisenberg, D. (1994) *Protein Sci. 3*, 1444−1463.

Blewitt, M. G., Chung, L. A., & London, E. (1985) *Biochemistry 24*, 5458−5464.

Bowie, J. U., Lüthy, R., & Eisenberg, D. (1991) *Science 253*, 164−170.

Brünger, A. T. (1990) *Acta Crystallogr. A46*, 46−57.

Brünger, A. T., Krukowski, A., & Erickson, J. (1990) *Acta Crystallogr. A46*, 585−593.

Choe, S., Bennett, M. J., Fujii, G., Curmi, P. M. G., Kantardjeff, K. A., Collier, R. J., & Eisenberg, D. (1992) *Nature 357*, 216−222.

Collier, R. J. (1975) *Bacteriol. Rev. 39*, 54−85.

Colovos, C., & Yeates, T. O. (1993) *Protein Sci. 2*, 1511−1519.

Dumont, M. E., & Richards, F. M. (1988) *J. Biol. Chem. 263*, 2087−2097.

Falnes, P. Ø., Choe, S., Madshus, I. H., Wilson, B. A., & Olsnes, S. (1994) *J. Biol. Chem. 269*, 8402−8407.

Jones, T. A. (1978) *J. Appl. Crystallogr. 11*, 268−272.

Kabsch, W. (1978) *Acta Crystallogr. A34*, 827−828.

Kantardjeff, K., Collier, R. J., & Eisenberg, D. (1989) *J. Biol. Chem. 264*, 10402−10404.

Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.

Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.

Lüthy, R., Bowie, J. U., & Eisenberg, D. (1992) *Nature 356*, 83−85.

Luzzati, V. (1952) *Acta Crystallogr. 5*, 802−810.

Madshus, I. H., Wiedlocha, A., & Sandvig, K. (1994) *J. Biol. Chem. 269*, 4648−4652.

Morris, R. E., Gerstein, A. S., Bonventre, P. F., & Saelinger, C. B. (1985) *Infect. Immun. 50*, 721−727.

Papini, E., Rappuoli, R., Murgia, M., & Montecucco, C. (1993) *J. Biol. Chem. 268*, 1567−1574.

Ramsay, G., & Freire, E. (1990) *Biochemistry 29*, 8677−8683.

Ramsay, G., Montgomery, D., Berger, D., & Freire, E. (1989) *Biochemistry 28*, 529−533.

Read, R. J. (1986) *Acta Crystallogr. A42*, 140−149.

Sandvig, K., & Olsnes, S. (1981) *J. Biol. Chem. 256*, 9068−9076.

SERC Daresbury Laboratory (1986) *CCP4. A Suite of Programs for Protein Crystallography*, SERC Daresbury Laboratory, Warrington, England.

Shen, W. H., Choe, S., Eisenberg, D., & Collier, R. J. (1994) *J. Biol. Chem. 269*, 29077−29084.

Tsuneoka, M., Nakayama, K., Hatsuzawa, K., Komada, M., Kitamura, N., & Mekada, E. (1993) *J. Biol. Chem. 268*, 26461−26465.

vanderSpek, J. C., Howland, K., Friedman, T., & Murphy, J. R. (1994) *Protein Eng. 7*, 985−989.

Van Ness, B. G., Howard, J. B., & Bodley, J. W. (1980) *J. Biol. Chem. 255*, 10710−10716.

BI942174X