A Natively Unfolded Toxin Domain Uses Its Receptor as a Folding Template*

Received for publication, December 12, 2003, and in revised form, March 5, 2004
Published, JBC Papers in Press, March 5, 2004, DOI 10.1074/jbc.M313603200

Gregor Anderluh†‡§, Isa Gökçe‡, and Jeremy H. Lakey‡***

From the †School of Cell and Molecular Biosciences, University of Newcastle upon Tyne, Framlington Place, Newcastle upon Tyne, NE2 4HH, United Kingdom, the ‡Department of Biology, Biotechnical Faculty, University of Ljubljana, Večna pot 111, 1000 Ljubljana, Slovenia, and the ***Department of Chemistry, Gaziosmanpaşa University, 60240 Tokat, Turkey

Natively unfolded proteins range from molten globules to disordered coils. They are abundant in eukaryotic genomes and commonly involved in molecular interactions. The essential N-terminal translocation domains of colicin toxins from *Escherichia coli* are disordered bacterial proteins that bind at least one protein of the Tol or Ton family. The colicin N translocation domain (ColN-(1–90)), which binds to the C-terminal domain of TolA (TolA-(296–421)), shows a disordered far-UV CD spectrum, no near-UV CD signal, and non-cooperative thermal unfolding. As expected, TolA-(296–421) displays both secondary structure in far-UV CD and tertiary structure in near-UV CD. Furthermore it shows a cooperative unfolding transition at 65 °C. CD spectra of the 1:1 complex show both increased secondary structure and colicin N-specific near-UV CD signals. A new cooperative thermal transition at 35 °C is followed by the unchanged unfolding behavior of TolA-(296–421). Fluorescence and surface plasmon resonance confirm that the new unfolding transition accompanies dissociation of ColN-(1–90). Hence upon binding the disordered structure of ColN-(1–90) converts to a cooperatively folded domain without altering the TolA-(296–421) structure.

The study of natively unfolded proteins has provoked a re-examination of the structure-function paradigm and emphasized the parallels between protein folding and molecular recognition (1–3). The extent of unfolding varies, but in the most extreme cases the proteins exist as “natural random coils.” One protein family showing this trait is the bacteriocins, which are produced by about one-third of enteric Gram-negative bacteria (4). Many of these are colicins, a group of protein toxins that are encoded on specialized plasmids in *Escherichia coli* and whose function is to kill competing bacteria (5, 6). Other Gram-negative bacterial species have similar bacteriocins (e.g. *Klebsiella* have klebicins (7)). They are important for the colonization of new habitats (8), and in the plague bacterium *Yersinia pestis*, pesticins use a receptor that is also important for pathogenesis (9). Successful colicin function involves a complex series of highly specific protein-protein and protein-membrane interactions that are slowly being defined in vivo and in vitro (10–18).

All colicins have a natively unfolded N-terminal region of 70–90 residues (the translocation or T-domain), 1 which contains sequences that bind to periplasmic receptor proteins such as TonB, TolA, and TolB. X-ray crystallographic structure analysis of colicin N (19) failed to detect electron density for residues 1–90 of the T-domain. Similarly the N-terminal 83 residues of colicin E3 were not resolved in the x-ray-derived structure (20), while in the structure of colicin Ia the first 23 residues are not resolved, and the next 44 residues have no defined secondary structure (21). Furthermore the NMR characterization of intact colicin E9 revealed that its T-domain is in a state of dynamic disorder (22), and CD spectra of the relevant colicin A domain shows it to lack significant secondary structure (23). This means that colicins are among a growing number of recognized natively unfolded proteins that interestingly also share a general propensity to be involved in macromolecular interactions (1). Thus concerted folding and binding is being increasingly recognized as a fundamental molecular process occurring throughout biology (1, 3). The absolute reliance of the colicin family upon this form of interaction thus gives us a powerful system to explore why this sort of molecular recognition is often used in nature.

The killing by colicins always proceeds in three steps, each promoted by a different structural domain: the binding to one or more outer membrane proteins, translocation across the periplasm with the aid of Tol or Ton protein systems, and finally killing achieved by enzymatic cleavage of nucleic acids (24) or pore formation (6). However, since the nucleases colicins must cross the inner membrane to reach their target and have also been shown to form ion channels (25) this is not an absolute dichotomy of function. Colicins can be divided in two groups according to the proteins involved in translocation. Ton proteins are used by B group colicins (colicins Ia, Ib, and B) (26), and the Tol QRAB family is used by group A colicins, which include colicins A, N, E1, and K (27). Of these, colicin N is the smallest pore-forming colicin (42 kDa) (28). In the first step it binds to outer membrane receptor OmpF via receptor-binding domain (R-domain) (29). The next step involves interaction of the unfolded T-domain with the protein TolA in the bacterial periplasm. In the final step, the pore-forming domain binds to the inner membrane and inserts hydrophobic helices into the membrane to form an open channel (6).

Pore formation by colicin N and its relatives has been extensively studied (6, 30, 31), but less is known about the receptor

---

* This work was supported by the Biotechnology and Biological Sciences Research Council (BBSRC) and Wellcome Trust Grants 56232, 40422, and 55979. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† Recipient of a long term Federation of European Biochemical Societies fellowship.

‡ A BBSRC Research Development Fellow. To whom correspondence should be addressed. Tel.: +44-191-222-8865; Fax: +44-191-222-7424; E-mail: j.h.lakey@ncl.ac.uk.

** A BBSRC Research Development Fellow. To whom correspondence should be addressed. Tel.: +44-191-222-8865; Fax: +44-191-222-7424; E-mail: j.h.lakey@ncl.ac.uk.

†‡§ This work was supported by the Biotechnology and Biological Sciences Research Council (BBSRC) and Wellcome Trust Grants 56232, 40422, and 55979. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† Recipient of a long term Federation of European Biochemical Societies fellowship.

‡ A BBSRC Research Development Fellow. To whom correspondence should be addressed. Tel.: +44-191-222-8865; Fax: +44-191-222-7424; E-mail: j.h.lakey@ncl.ac.uk.

The abbreviations used are: T-domain, translocation domain; FRET, fluorescence resonance energy transfer; SPR, surface plasmon resonance; H/S/T/R, helix/strand/turn/random; PII, polyproline II helix; FPLC, fast protein liquid chromatography.

---

22002 This paper is available on line at http://www.jbc.org
binding or translocation steps. Since requirements for receptor binding and translocation of colicin N are relatively simple, it is a good model for studying mechanisms of colicin translocation, and this has permitted a particularly detailed analysis of the T-domain (32-34). TolA is a three-domain protein anchored to the inner membrane by an N-terminal transmembrane helix (TolAI). A large central domain (TolAIII) links TolAI to a small C-terminal globular domain (TolAII) (35). The N-terminal domains of both group A colicins and g3p protein of filamentous phage (35-37) bind to TolAIII, proving a common entry point for two apparently unrelated invading proteins.

It is possible to prepare isolated colicin N T-domain and TolAIII in an E. coli expression system and measure their interaction in vitro (32–34, 38). The disorder expected of the isolated T-domain was confirmed by CD, NMR, and fluorescence spectroscopy (33, 34). However, like full-length colicin N, the isolated T-domain binds TolAIII with a Kd in the micromolar range and has an association rate of 105 s⁻¹ (32–34). The minimum colicin N TolA-binding box was defined as residues Asn42–Gly68 using surface plasmon resonance (SPR), isothermal titration calorimetry, fluorescence, and pull-down methods, and a high proportion of these residues are required for binding (32). Recently we showed by using steady state and time-resolved fluorescence resonance energy transfer (FRET) binding (32). Recently we showed by using steady state and time-resolved fluorescence resonance energy transfer (FRET) and NMR that the T-domain undergoes gross conformational changes upon binding to TolA (34). This indicated that the structure of complexed T-domain is likely to be significantly different to that in solution. In this report we show by the use of CD spectroscopy, fluorescence, and SPR that the complexed T-domain has acquired cooperatively folded secondary and tertiary structure.

**EXPERIMENTAL PROCEDURES**

**Preparation of the Proteins**—Colicin N, T-domain of colicin N, and all Tol variants were prepared using an E. coli BL21(DE3) expression system using modified pET5c vector, as described previously (39), that adds an N-terminal purification tag with the sequence MHSHg–~-~T-domain comprised colicin N (Swiss-Prot accession number P08083) amino acid residues 1-90, TolA-(329–421) comprised amino acids 329–421 of the TolA protein (Swiss-Prot accession number P19934), and TolA-(296–421) comprised amino acids 296–421. TolAT is a fusion partner that is used for preparation of fusion proteins by using pTol expression system (38). This variant comprises amino acids 329–421 of the TolA protein and, in addition, contains a linker of 12 amino acids at its C terminus that includes a thrombin cleavage site (GGGSVLPRSGST). Proteins were expressed and isolated from the bacterial cytoplasm exactly as described before (34). Colicin A was prepared as in Ref. 40. The purity and homogeneity of the proteins was checked by SDS-PAGE.

Protein concentrations were determined from UV absorption at 280 nm with extinction coefficients as follows: colicin N, 49.500 M⁻¹ cm⁻¹; T-domain, 12,660 M⁻¹ cm⁻¹; TolA-(329–421), 5120 M⁻¹ cm⁻¹; and TolAT, 5120 M⁻¹ cm⁻¹.

**CD Spectroscopy**—CD spectroscopy was carried out on a Jasco J-810 spectropolarimeter. For far-UV CD spectra (250–190 nm) bandwidth was set to 2 nm, scanning speed was set to 50 nm/min, and either 0.2-mm-path length demountable or, for thermal denaturation, 1-mm-path length standard cuvettes (all cuvettes were from Hellma) were used. The concentration of the proteins was 15–30 μM in 10 mM NaH₂PO₄, pH 7.0. Near-UV CD spectra (320–250 nm) were measured in the same buffer in 1-cm-path length stopped cuvette at 70 μM protein concentration. All spectra shown are averages of 10 scans with bandwidth set to 1 nm. Thermal unfolding was carried out in the range 20–90 °C at a rate of 1 °C/min. A full spectrum was taken every 5 °C. Samples were allowed to equilibrate for at least 20 min at 20 °C before starting each scan.

**Fluorescence Spectroscopy**—Fluorescence was measured in a 5-mm (Hellma 101.016-QS) cuvette at 20 °C using a Cary Eclipse (Varian) spectrofluorometer. Concentration of T-domain was 1 μM in 40 mM HEPES, 100 mM NaCl, pH 7.0. Excitation wavelength was 285 nm, and fluorescence was measured between 300 and 400 nm wavelength with a scan rate of 120 nm/min. Excitation and emission slits were set to 5 nm. TolA variants were added in aliquots and mixed, and spectra were measured again. Measured spectra were corrected for the contribution of the sample without T-domain and the dilution factor. Thermal unfolding of the T-domain-TolA-(296–421) 1:5 (molar) complex was done at a rate of 1 °C/min in the range 15–90 °C. A full spectrum was taken every 5 °C. After mixing of T-domain and TolA-(296–421), the resulting complex was allowed to equilibrate for at least 15 min at 15 °C before measurement.

**Gel Filtration Chromatography**—Gel filtration chromatography was performed on a Superose 12 column (Amersham Biosciences) equilibrated in 20 mM NaH₂PO₄, 300 mM NaCl, pH 7.0, at a flow rate of 0.5 ml/min using an Amersham Biosciences FPLC system. Proteins or mixtures applied contained 2 nmol of colicin N or A, 6 nmol of T-domain of colicin N, and 20 nmol of TolA-(296–421), TolA-(329–421), or TolAT.

**Surface Plasma Resonance Measurements**—SPR was performed on a BIACORE-X apparatus. TolA-(296–421) was immobilized on a CM5 chip as described before (32). The chip was equilibrated for more than 30 min after each temperature change. Solutions of desired T-domain concentration in HBS (10 mM HEPES (Sigma), 150 mM NaCl (British Drug House), 3.4 mM EDTA (Amersham Biosciences)), 0.05% P20 (poloxymethylene sorbitan monolaurate, Amersham Biosciences, Biosensor AB), pH 7.4, were passed across the chip at 5 μl/min. The resulting sensograms were fitted separately to a 1:1 binding model to obtain kₐ, k₋, and Kd using BIAevaluation software. The detailed data analysis that was applied can be found in Gökç et al. (32).

**RESULTS**

**TolA C Terminus Is Involved in T-domain Binding**—For this study three different C-terminal variants of the TolA protein were prepared in E. coli. TolA-(296–421) is the predicted size of T. BIACORE-X apparatus. TolA-(296–421) was immobilized in 20 mM NaH₂PO₄, pH 7.0, at a flow rate of 0.5 ml/min. The resulting sensorgrams were fitted separately to a 1:1 binding model to obtain kₐ, k₋, and Kd using BIAevaluation software. The detailed data analysis that was applied can be found in Gökç et al. (32).

The CD spectra confirm the folded nature of the TolA constructs (Fig. 1A). The estimated secondary structure content for TolA-(329–421) is H38/S23/T10/R29 (percentage of helix/strand/turn/random) using the known three-dimensional structure (37). It was most accurately analyzed using the CDSTR program of the Dichroweb site² (41). There was, however, a consistent underestimation of the helix content with a fitted content of H28/S24/T20/R27. TolA-(296–421) shows a larger amount of random and β-strand (H20/S30/T22/R30) that is presumably due to the flexible extension not present in the solved structure. As expected the T-domain in solution is predicted to have more random structure (H4/S31/T23/R41), but the strand content is significant.

The binding of full-length TolA-(1–90) T-domain to these TolA variants was measured using fluorescence spectroscopy and gel filtration chromatography. The T-domain contains two tryptophans that become buried upon TolA binding causing both an increase and a blue shift of their fluorescence emission (32–34). Both TolA variants with natural C termini bound to T-domain with similar affinity. However, TolAT showed no binding (Table I) even after treatment with thrombin, which removes 4 of the additional 12 amino acids. Since this assay

² See www.cryst.bbk.ac.uk/cdweb/.
and predicted (A421) (experimental complex and TolA-(296–421) (not shown). The difference between the complex spectrum and TolA-(296–421) calculated as in B is shown as a dashed line, B, experimental (dashed line) and predicted (dotted line) spectra of 1:1 T-domain:TolA-(296–421) complex, 30 μM each. Predicted spectra were calculated by summing the individual spectra of T-domain and TolA-(296–421) using Δε calculated using the molar protein concentrations (Δε_{protein}). This allows the correct 1:1 addition of the original data from A since Δε_{1:1} will overestimate the contribution of the shorter T-domain. The TolA-(296–421) data from A is shown as the solid line. The difference between the experimental complex and TolA-(296–421) was calculated and is shown in A as Δε_{protein} (dashed line). The T-domain spectrum from A is shown for comparison as a dot-dashed line.

only detects interactions with K_{D} < 10 μM (32) we cannot eliminate a weakened interaction, but this is the first data to indicate the importance of a natural TolAIII C terminus in ColN binding. Supporting data were obtained by using gel filtration chromatography (Fig. 2). Here the binding of three TolA constructs to ColN T-domain was analyzed. A decrease in elution volume (increase in size) was observed when T-domain was mixed with TolA-(296–421) and TolA-(329–421) but not in the presence of TolAT. Similar results were obtained with full-length colicins N and A (not shown) and showed that the C-terminal extension in TolAT inhibits binding of colicin A as well (see “Discussion”). The molecular masses determined from the calibration curve corresponded to 1:1 complexes. Interestingly the 98-residue His-tagged ColN-(1–90) eluted earlier and gave a larger predicted size of 15.7 kDa compared with the 100-residue His-tagged TolA-(329–421) (predicted size, 13.4 kDa). This might be expected if its unfolded state leads to a slightly larger hydrodynamic radius.

Secondary Structure Increases upon Binding—By expressing the far-UV CD spectra in units of Δε where concentration is in moles of protein/liter rather than the more conventional moles of residues/liter we can generate the expected spectrum of a 1:1 mixture of TolA-(296–421) and T-domain (Fig. 1B). This can be compared with the measured far-UV CD spectrum of the 1:1 complex. The spectra are different, indicating that changes in secondary structure occur upon binding. The difference between TolA-(296–421) and the complex is also shown (Fig. 1A). This may correspond to the new structure of the T-domain with a reduction in the unstructured component and its replacement by a secondary structure signal (210–225 nm). A small difference was observed when T-domain was mixed with TolAT (not shown), which is in agreement with a weaker binding that is only detectable at the higher protein concentrations used in CD. Another control was performed where mutant F66C of T-domain was added to TolA-(296–421). This mutant does not bind to TolA at all (34), and in this case the difference between predicted and experimental spectra was negligible.

Template-induced Folding of a Disordered Toxin Domain—The near-UV CD spectra of T-domain is virtually featureless and consistent with its mobile and exposed aromatic residues (Fig. 3A) (42). The spectra of TolA-(296–421) and TolAT exhibit a single minimum at ~280 nm. This is almost certainly due to some of its 4 tyrosine residues being immobilized since it contains no tryptophan residues and phenylalanine contributions are weak (43). The near-UV CD spectrum of the 1:1 complex between TolA-(296–421) and T-domain is, like the far-UV CD spectra, not simply a sum of the individual spectra (Fig. 3B). The complex has less intensity at 280 nm and a small shoulder visible at ~290 nm. It is important to note that the contributions to the near-UV CD spectra of individual residues can be positive or negative, and therefore, a reduction in intensity on binding a new peptide does not necessarily mean unfolding. The difference between the predicted and experimental spectra is shown in Fig. 3C (solid line). This reveals an additional positive signal in the spectrum of the complex in the region 270–290 nm with clearly resolved peaks at 272, 279, and 290 nm. The latter is unique to tryptophan and thus can be directly linked to T-domain folding upon binding. The difference for the complex made using TolAT is much less pronounced (Fig. 3C) in agreement with its reduced affinity for T-domain. As expected experimental and predicted spectra are virtually identical when the non-binding mutant F66C and TolA-(296–421) are mixed (Fig. 3C).

Thermal Dissociation of the Complex Releases a Folded TolAIII Domain—The T-domain, TolA-(296–421), and their 1:1 complex were thermally denatured to characterize the stability of the acquired structure. Thermal denaturation was followed by far-UV and near-UV CD (Figs. 4–7). Upon heating T-domain to 90 °C the far-UV CD signal at 198 nm decreases (Fig. 4A) with a concomitant increase of the broad shoulder at ~222 nm. In the temperature range of 15 to 90 °C, a non-cooperative transition was observed (Figs. 4 and 6); however, an isodichroic point was observed at 209 nm indicating a two-state transition also observed in simple model peptides (44). The change in T-domain conformation upon heating is reversible since cooling down to 20 °C resulted in a spectrum identical to that recorded previously (Fig. 4A). In TolA-(296–421) the peak at 192 nm was the most sensitive to heating. There is a clear transition between 55 and 70 °C (Figs. 4 and 6 and Table II), and above 75 °C a new minimum at ~205 nm emerged. The unfolding of TolA-(296–421) is only partially reversible (Fig. 4B). There was no precipitate observed in the cuvette after the experiment at the concentrations used (the highest concentration used for thermal denaturation measurements in the far-UV CD was 30 μM), indicating that TolA-(296–421) remained soluble even at the highest temperature used. A two-state transition is also indicated here by the isodichroic point at 207 nm. The T-domain:TolA-(296–421) 1:1 complex showed two transitions upon heating. One can be seen in the range 25–40 °C, and the second one can be seen in the range 55–70 °C (Figs. 4 and 6 and Table II). The first transition can be seen as a steady decrease in the signal between 195 and 205 nm. As this one ends the second one starts and can be seen clearly at 220 nm. The second transition coincides with the thermal un-
folding of free TolA-(296–421) (Fig. 6). The first transition could be attributed to the changes in the secondary structure content of T-domain and/or TolA-(296–421) after dissociation of the 1:1 complex. To see whether this indeed might be the case, the measured spectra of T-domain/TolA-(296–421) 1:1 complex were compared with predicted spectra obtained by summing individual T-domain and TolA-(296–421) spectra for the whole temperature range (Fig. 7A). The difference between the experimental and predicted spectra is observed between 25 and 40 °C. After 45 °C the difference was much less. Since any difference must be due to binding, this indicates a dissociation of the complex between 25 and 40 °C.

Thermal unfolding was also followed in the near-UV CD.

Here the T-domain alone did not show any changes if heated to 90 °C; spectra at 20 and 90 °C were almost identical (Fig. 3). TolA-(296–421) underwent a single transition between 60 and 70 °C indicated by the increase of the signal at 280 nm (Figs. 5 and 6). After the heating, a visible precipitate was observed in the cuvette, indicating that thermally unfolded TolA-(296–421) aggregated at the higher concentrations used in near-UV CD conditions. The T-domain/TolA-(296–421) 1:1 complex again showed two transitions. The first one could be monitored at both 280 (where the minimum in the TolA-(296–421) spec-

### Table I

<table>
<thead>
<tr>
<th>Protein</th>
<th>$K_d$</th>
<th>$k_a$</th>
<th>$k_d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>TolAII-III</td>
<td>0.28 ± 0.014</td>
<td>4.78 ± 0.41</td>
<td>0.016 ± 0.0002</td>
</tr>
<tr>
<td>TolA-(296–421)</td>
<td>1.25 ± 0.17</td>
<td>5.00 ± 0.71</td>
<td>0.042 ± 0.002</td>
</tr>
<tr>
<td>TolA-(329–421)</td>
<td>0.19 ± 0.007</td>
<td>4.25 ± 1.96</td>
<td>0.092 ± 0.006</td>
</tr>
<tr>
<td>TolAT</td>
<td>NB*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TolAT, thrombin-cleaved</td>
<td>NB</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TolA-(296–421), 20 °C</td>
<td>0.35 ± 0.02</td>
<td>4.78 ± 0.41</td>
<td>0.016 ± 0.0002</td>
</tr>
<tr>
<td>TolA-(296–421), 25 °C</td>
<td>0.83 ± 0.18</td>
<td>5.00 ± 0.71</td>
<td>0.042 ± 0.002</td>
</tr>
<tr>
<td>TolA-(296–421), 30 °C</td>
<td>2.72 ± 1.51</td>
<td>4.25 ± 1.96</td>
<td>0.092 ± 0.006</td>
</tr>
</tbody>
</table>

*NB, no binding was observed.

**Fig. 2.** Gel filtration chromatography. Gel filtration chromatography was performed on a Superose 12 column (Amersham Biosciences) equilibrated in 20 mM NaH2PO4, 300 mM NaCl, pH 7.0, at a flow rate of 0.5 ml/min using an Amersham Biosciences FPLC system. 6 nmol of T-domain were mixed with 20 nmol of either TolA-(296–421), TolA-(329–421), or TolAT and applied to the column. The molecular masses of proteins and complexes were compared with standards (Sigma molecular mass calibration kit) that are shown above the figure with triangles and thin dashed lines. Standards used were bovine serum albumin (66 kDa), carbonic anhydrase (29 kDa), cytochrome c (12.5 kDa), and aprotinin (6.5 kDa). Chromatograms shown are for T-domain (dotted line), Tol protein (dashed line), and complex of the two (solid line).

**Fig. 3.** Near-UV CD spectra of T-domain and TolA variants. CD spectra were measured in a 10-mm-path length cuvette at 20 °C. The concentration of proteins was typically 30 μM in 10 mM NaH2PO4, pH 7.0. A, near-UV CD spectra of T-domain (dotted line), T-domain at 90 °C (dashed line), TolA-(296–421) (solid line), and TolAT (dash-dotted line). B, experimental (solid line) and predicted (dashed line) spectra of 1:1 T-domain/TolA-(296–421) complex. C, the difference between experimental and predicted spectra is shown for T-domain in 1:1 complex with TolA-(296–421) (solid line), TolAT (dashed line), and T-domain mutant F66C in 1:1 complex with TolA-(296–421) (dotted line).
trum was observed) and at 292 nm (where a pronounced T-domain tryptophan peak was observed in the complex). At both wavelengths the intensity first decreased, and then the shoulder at 292 nm gradually disappeared from the spectra upon heating and was not visible above 45 °C. The signal at 292 nm then remained flat up to 90 °C, while that at 280 nm started to increase again due to TolA-(296–421) unfolding (Fig. 6). Transition temperatures \(T_m\) for both CD spectral regions and fluorescence spectroscopy (see below) are collected in Table II.

**Dissociation Measured by Fluorescence Spectroscopy**—T-domain tryptophans become buried upon complex formation, and their fluorescence exhibits an intensity increase and a 17-nm blue shift in the wavelength of maximum emission \(\lambda_{\text{max}}\) (33). The T-domain TolA-(296–421) 1:5 complex was heated as in the CD spectroscopy studies, and the \(\lambda_{\text{max}}\) of tryptophan fluorescence was followed as, unlike fluorescence intensity, it provides a temperature-insensitive measure of tryptophan environment upon heating (Fig. 7A). \(\lambda_{\text{max}}\) of the complex at 15 °C was 338 nm, consistent with tryptophans being buried. Upon heating \(\lambda_{\text{max}}\) reached its maximum of 355 nm indicating full water exposure by 45 °C (Fig. 7A).

**SPR Measurements**—Complex formation was also checked in the temperature range 20–40 °C by SPR (BIACORE) (Fig. 7B). TolA-(296–421) was immobilized, and increasing concentrations of T-domain were injected over it. The amount of stably bound T-domain at each temperature was determined by allowing the binding to reach equilibrium. The binding decreased with increasing temperature. Binding of TolA-(296–421) by 2 \(\mu\text{M}\) T-domain at 40 °C was 90% less than at 20 °C (Fig. 7B). The kinetic parameters were obtained from the sensograms and are reported in Table I. At 35 and 40 °C, the dissociation and association was too fast for meaningful estimation of rate and affinity constants (Fig. 7C). In the range 20–30 °C, the \(k_a\) was similar, but \(k_d\) increased with increasing temperature, leading to a 7-fold increase in \(K_d\) at 30 °C (Table I).

**DISCUSSION**

Our previous data from NMR and FRET hinted at great changes in the T-domain structure accompanying its binding to TolAIII (34). The FRET indicated that while the center of the binding box sequence moved away from the two tryptophan fluorescence donors, the distal end of the box moved closer. The heteronuclear single quantum correlation \(^{13}\text{N}-^{2}\text{H}\) NMR data on a 40-residue T-domain construct showed that all the backbone amide resonances shifted upon binding. Both methods indicated that the free T-domain, although unstructured, is not elongate but exists as a collapsed, hydrated globule. Finally FRET under stopped-flow conditions indicated that binding and folding were simultaneous. Nevertheless neither approach provided information on the structure of the complexed T-domain. The FRET provides too few distances, and the NMR lacks the residue assignment needed for high resolution structural work. Here we present the results of a comprehensive circular dichroism study to measure secondary and tertiary structural changes upon complex formation.

The N-terminal translocation domain of colicin N was first defined by El-Kouhen et al. (29) as the first 67 residues (ColN1–67) due to a chymotryptic cleavage site between residues 67 and 68. The remaining C-terminal 68–342 fragment exhibited receptor binding and pore formation but not translocation. Later the x-ray crystallographic study of colicin N (Protein Data Bank code 1A87) could not solve the structure of residues ColN1–90 due to their disorder (19), and this is our current working definition of the translocation domain (32, 33). We defined the minimal TolA-binding box of colicin N as 27 residues long (ColN40–66) (32), but this has a lower affinity for TolAIII than ColN1–90, and therefore, residues outside of this sequence enhance binding. A glutathione S-transferase fusion to residues 40–90 (32) and the binding peptide used for the NMR studies (Gly-Ser-Gly-[ColN40–76]) both showed wild type binding affinity (34). Residues 1–18 can be deleted without affecting toxicity (28), but only 1% activity is retained by a truncated colicin N starting at residue 36 (29). Hence residues 1–40 may not be important for TolAIII binding in vitro, but some do have a role in vivo. The T-domain used here (as in previous studies) is ColN1–90, which in fact shows the maximum affinity of any colicin N peptide yet produced since longer constructs, which also include receptor-binding (ColN1–190) and pore-forming domains (full-length colicin N, residues 1–387), show a decreased affinity (measured by isothermal titration calorimetry) (33) possibly due to steric effects. A very high proportion of the binding box residues (13 of a total of 15
alanine substitutions) inhibited binding (32).

Two features, however, suggest some ordering of the T-domain: first the collapsed state of the peptide and second the chemical shifts of the backbone NH resonances, which although narrowly distributed compared with regions with secondary structure are not closely grouped as in true random coils (34). The gel filtration data here further show that the hydrodynamic radius of free T-domain is only slightly larger than a folded protein of this size. The disordered nature of this domain is revealed by far-UV CD spectra (Fig. 1 and previous work (19, 33, 34)), and we show here that the T-domain also lacks two further features of folded proteins. The first is cooperative folding/unfolding as revealed by heat denaturation. The far-UV CD spectra recorded at increasing temperatures show a gradual almost linear change with temperature consistent with the isolated unfolding of individual structures rather than the sigmoidal unfolding curves found with proteins folded in cooperative domains (Fig. 4). The fact that the T-domain “unfolds” with temperature is supportive of the possibility that it contains some secondary structure (premolten globule) or polyproline II helix (PII), which can at higher temperatures adopt a different conformation (46). A comparable behavior has been well described for the C-terminal domain of the gp210 nuclear pore protein (47). The isodichroic point is supportive of a two-state non-cooperative transition, but since temperature can induce folding in some denatured polypeptides we cannot be sure of the nature of the high temperature form. The spectrum lacks a strong positive signal at 222 nm, which is considered characteristic of PII (47, 48), and using the intensity at 222 nm (–0.52 Δε or –1724 degrees cm²/dmol) in the equation of Park et al. (49) as utilized by Bienkiewicz (48) we obtain a possible PII content of 24%. Using the 200 nm intensity (–3.50 Δε or 11,570 degrees cm²/dmol) we obtain 7%. The 90-residue T-domain thus does not conform well to this model, and the predicted β-strand component from the CDSSTR analysis may partly explain this. Uversky (3) proposed two forms of natively unfolded protein, the native coil and the premolten globule. One way these can be distinguished is by their position on a plot of 222 versus 200 nm data. At 20 °C the weakness of the 200 nm minimum places it within the premolten globule group, and its negative 222 nm data places it in the coil region. At higher temperatures, where binding is reduced, the data resemble more the premolten globule state.

The second missing feature of folded protein domains is a near-UV CD signal caused by trapping of aromatic side chains within the tertiary structure. While some folded soluble globular proteins do not show strong signals in this region, the absence of such a signal is to be expected for any disordered structure with mobile aromatic residues not confined in a hydrophobic core. For the two tryptophans, which show water
exposure via their quenching and emission wavelength data, this is not surprising but it does indicate a lack of structure around the essential tyrosine (Tyr<sup>65</sup>).

It was initially surprising that the short C-terminal extension on the TolAT protein was capable of inhibiting the binding of both colicin N and colicin A translocation domains since Deprez<sup>et al.</sup> (23) have shown that an extended TolAIII construct (AEF-TolA-(335–421)-HHHHHHH) binds colicin A T-domain (ColA-(1–172)). However, the near-UV CD data at higher concentrations show that TolAT retains a reduced affinity, and thus the exact nature of the C-terminal extension may be important. Dubuisson<sup>et al.</sup> (50) showed that mutants expressing a TolA gene coding for a C-terminal P421S mutation are completely resistant to colicin A and phase f1. This mutation also destabilizes the outer membrane, and we need more data to know whether the resistance is solely due to a change in TolA-colicin A binding.

The CD data suggest large scale changes in the secondary structure of the T-domain during TolA binding. The measured spectrum of the complex differs significantly from the summed spectra of the two individual components, and subtraction of the TolA-(296–421) spectrum from the complex reveals a “T-domain like” signal, which, however, contains less random and more secondary structure. This “spectrum” shows a clear plateau region at 220 nm reminiscent of helix, but none of the CD analysis programs we tried modeled this accurately. In all cases the fitted curve was that of an unfolded structure and when compared with the data showed large differences in the region 215–230 nm. The retention of significant unfolded structure is to be expected if, as we suspect, only 27 of the 90 residues are firmly bound (32). Whether this difference spectrum is the spectrum of the bound T-domain or not depends upon whether the TolAIII spectrum changes upon binding to the colicin N T-domain. Deprez<sup>et al.</sup> (23) have studied the ColA-(1–172):TolAIII interaction by both CD and NMR. The CD spectra show a difference between the measured and predicted complex, indicating a reduction in structure rather than the increase shown here. Furthermore their NMR data show that TolAIII in the complex has the characteristics of a molten globule. The CD spectra presented here show that TolA-(296–421) complexed with colicin N T-domain is not a molten globule. In fact, the near-UV CD spectrum of the complex provides evidence of an additional tryptophan signal that can only come from ordering of the T-domain. At this point we must conclude that the interactions of these two colicin domains with TolA are possibly as unlike as their respective primary structures.

In the case of colicin N T-domain binding to TolA-(296–421) we have evidence that the TolA-(296–421) structure is largely unchanged upon binding to the disordered peptide. The isolated TolA-(296–421) domain has well defined near- and far-UV CD spectra indicating stable tertiary and secondary structure, respectively. These signals remain unchanged by increasing temperature until both collapse rapidly to provide a sigmoidal unfolding curve (T<sub>m</sub> = 65 °C) expected of a well folded single domain protein.

During the thermal denaturation of the complex the TolA-(296–421) behavior is maintained, but it is preceded by cooperative changes unseen in either T-domain or TolA-(296–421). At higher temperatures the data resemble the sum of the individual proteins and include the characteristic cooperative unfolding of TolA-(296–421). Fig. 7A shows that the major structural changes, when comparing summed far-UV CD spectra of both proteins with experimentally determined ones, are abolished by 45 °C. In the near-UV CD we observe a similar change in the adopted tertiary structure with the difference between the complex and TolA-(296–421) disappearing by 45 °C (Fig. 5). These changes are caused by the dissociation of T-domain from TolA upon heating that entirely precedes the unfolding of TolA. This dissociation can be observed by tryptophan fluorescence spectroscopy, which is particularly useful as it reports only the dissociation of T-domain and is insensitive to any changes in TolA-(296–421) (Fig. 7A). However, we also used the structure-independent method of surface plasmon resonance. The binding data were collected in the range 20–40 °C, and the reduction in the binding affinity of T-domain for TolA-(296–421) was clearly evident from these measurements (Fig. 7). The affinity of the T-domain for TolA-(296–421) is thus highly temperature-dependent and reversible, whereas the TolA-(296–421) protein is more thermostable. The dissociation of the T-domain may, however, be due to the formation of another structure in the free form at higher temperature for which the conversion to the folded form is less favorable. This might act to alter the equilibrium in favor of the unbound state by reducing the “on” rate, but the SPR data indicate that the “off” rate is the most altered, and thus the complex stability must be temperature-sensitive. The well defined complex can thus be dissociated by temperature with the end result being the same as one where the two proteins were simply mixed at that temperature. TolA-(296–421) shows no temperature sensitivity in this temperature range and only unfolds at its nor-

![Fig. 7. Dissociation of T-domain-TolA-(296–421) complex upon heating. A, the difference data (squares) are the average difference between experimental (measured far-UV CD spectra of T-domain/TolA-(296–421) 1:1 complex and predicted spectra (summed spectra of individual far-UV CD spectra of T-domain and TolA-(296–421)) for the whole temperature range measured. The wavelength data (circles) are the wavelengths of maximum emission intensity for the T-domain tryptophan emission spectra; 1 μM T-domain and 5 μM TolA-(296–421) were placed in a 1-cm-path length cuvette in 10 mM NaH<sub>2</sub>PO<sub>4</sub>, pH 7.0, and allowed to stand at 15 °C for 15 min. λ<sub>max</sub> was determined from emission spectra in the range 300–400 nm obtained by excitation at 295 nm. The complex was then heated at a rate of 1 °C/min, and after each 2.5 °C a spectrum was taken. The data are collected from two independent experiments. B, binding of T-domain to TolA-(296–421) was followed by SPR. TolA-(296–421) was immobilized as described in the text. Various concentrations of T-domain were passed at 5 μl/min over the chip, and the response was measured. C, sensitograms of 3 μM T-domain binding to TolA-(296–421) at different temperatures. Flow rate was 5 μl/min. R.U., response units.](image-url)
of a Disordered Toxin Domain

Acknowledgments—We thank Dr. W. Webster for TolA plasminoids and Geoff Moore, Ruth Boetzel, and Colin Macdonald for helpful discussions.

REFERENCES