Structure of Ptr ToxA: An RGD-Containing Host-Selective Toxin from Pyrenophora tritici-repentis

Ganapathy N. Sarma, Viola A. Manning, Lynda M. Ciuffetti, and P. Andrew Karplus

INTRODUCTION

Pyrenophora tritici-repentis is a fungal pathogen that causes tan spot of wheat (Triticum aestivum) in sensitive cultivars. This disease is found throughout the major wheat-growing regions of the world and is characterized by either necrotic lesions, chlorosis, or both. As for a number of fungi, disease is associated with the production of host-selective toxins (HSTs; reviewed in Walton, 1996; Wolpert et al., 2002). P. tritici-repentis produces various HSTs, each of which uniquely leads to progression of disease on wheat cultivars that are sensitive to that particular toxin (reviewed in De Wolf et al., 1998; Ciuffetti and Tuori, 1999; Strelkov and Lamari, 2003). Of the three HSTs thus far characterized from P. tritici-repentis, two—Ptr ToxA (Ballance et al., 1988; Tomas et al., 1990; Tuori et al., 1995; Zhang et al., 1997) and Ptr ToxB (Strelkov et al., 1998)—have been shown to be proteins, whereas Ptr ToxC is a low molecular weight compound (Effertz et al., 2002). For Ptr ToxA (hereafter called ToxA), the first protein HST to be isolated (Ballance et al., 1989; Tomas et al., 1990; Tuori et al., 1995; Zhang et al., 1997), sensitivity is conditioned by the presence of a single host locus, Tsn1, on the 5BL chromosome (Faris et al., 1996; Stock et al., 1996; Gamba et al., 1998; Anderson et al., 1999) and the ToxA gene in the fungus (Ballance et al., 1996; Ciuffetti et al., 1997); disease occurs only when both the genes are present.

As an HST, not only is ToxA required for toxicity but infiltrating leaves of sensitive wheat cultivars with ToxA alone results in symptom development (Ballance et al., 1989; Tomas et al., 1990; Tuori et al., 1995). As shown in the accompanying article (Manning and Ciuffetti, 2005), green fluorescent protein tagged ToxA exerts its toxic effect via internalization into sensitive wheat mesophyll cells. After internalization, it appears to localize to chloroplasts, but the mechanisms involved in cell death remain unknown. Because ToxA itself plays a pivotal role in toxicity, its structure is of great interest.

The ToxA gene encodes a pre-pro-protein (Ballance et al., 1996; Ciuffetti et al., 1997; Figure 1). The pre-region (i.e., the signal peptide; residues 1 to 22) targets the protein to the secretory pathway. The pro-region (also known as the N-domain; residues 23 to 60) is cleaved prior to secretion of the mature, 13.2-kD ToxA (i.e., the C-domain; residues 61 to 178) (Tuori et al., 1995). Evidence suggests that the N-domain, much like the pro-peptide of insulin, is required during folding for the proper formation of the disulfide in the C-domain that stabilizes the active, native conformation of ToxA (Tuori et al., 2000).

A database search for proteins similar to ToxA did not yield any potential homologs. In the absence of structural or sequence conservation information to guide mutagenesis experiments, sequence motif searches were performed to find potential functional sites. The searches revealed putative sites of phosphorylation (Thr residues 115, 123, 126, 132, 139, and 167) (Ciuffetti et al., 1997; Zhang et al., 1997), of myristoylation (Gly residues 62 and 93) (Ciuffetti et al., 1997), and of cell attachment (an Arg-Gly-Asp [RGD]-containing motif; residues Asn136 through Phe147) (Ciuffetti et al., 1997; Zhang et al., 1997; Meinhardt et al., 2002). Mutagenesis studies have identified a number of residues, including the RGD motif, that may be important for function (Tuori et al., 2000; Meinhardt et al., 2002; Manning et al., 2004), but without structural information it was not possible to conclusively deduce which residues were important for structural reasons and which might be directly involved in recognition events. Here, we...
RESULTS AND DISCUSSION

Structure Solution

The structure of mature ToxA was solved in two different crystal forms. The structure in form-I crystals (with one molecule in the asymmetric unit) was determined at 3.0-Å resolution using in-house sulfur single-wavelength anomalous dispersion (SAD) data collected to 1.9-Å resolution. Refinement against 1.65-Å resolution data collected at a synchrotron (see Methods) led to a model with a final R_free of 18.3% and good geometry (Table 1). The form-II crystal structure (with three molecules in the asymmetric unit) was determined by molecular replacement using the form-I structure as the search model. The structure was refined at 1.9-Å resolution to a final R_free of 22.6%. In both form-I and -II crystals, ToxA is present as a trimer (the trimer in form-I is generated by the crystallographic threefold axis). Having four independently refined molecules provides insight into the variability of structures. Based on analysis of Luzzati (Luzzati, 1952; Kleywegt and Brunger, 1996), the estimated coordinate accuracies of the well-ordered parts of the models are 0.2 and 0.3 Å for the form-I and form-II structures, respectively. The root mean square deviation (RMSD) between the Cα atoms of the three chains of form-II is 0.3 Å and is thus consistent with the estimated coordinate accuracy. The RMSDs between the form-I structure and the three monomers of form-II are 0.6 Å, indicating that small differences exist. Even though crystal form-I is at a higher resolution, we have prepared figures using the form-II structure because the structures are very similar and the form-II structure has a better ordered RGD-containing loop.

Table 1. Data Collection and Refinement Statistics

<table>
<thead>
<tr>
<th>Data Collection and Phasing</th>
<th>Sulfur SAD, Form-I (In House)</th>
<th>Native, Form-I (Synchrotron)</th>
<th>Native, Form-II (Synchrotron)</th>
<th>Native, Form-II (In House)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resolution</td>
<td>100–1.90 (1.97–1.90)</td>
<td>100–1.65 (1.71–1.65)</td>
<td>100–1.90 (1.97–1.90)</td>
<td>100–3.00 (3.11–3.00)</td>
</tr>
<tr>
<td>Reflections (total/unique)</td>
<td>808,412/12,883</td>
<td>147,079/19,469</td>
<td>191,032/31,109</td>
<td>60,556/8,241</td>
</tr>
<tr>
<td>Completeness</td>
<td>99.9 (100.0)</td>
<td>99.9 (99.8)</td>
<td>99.4 (100.0)</td>
<td>99.9 (99.9)</td>
</tr>
<tr>
<td>R_meas</td>
<td>8.2 (50.7)</td>
<td>8.0 (50.2)</td>
<td>6.8 (56.8)</td>
<td>10.3 (39.5)</td>
</tr>
<tr>
<td>R_mrgd-F</td>
<td>2.8 (12.5)</td>
<td>7.9 (40.9)</td>
<td>7.2 (43.2)</td>
<td>8.9 (27.5)</td>
</tr>
<tr>
<td>V/σ</td>
<td>62.9 (56.0)</td>
<td>13.9 (2.6)</td>
<td>11.3 (2.2)</td>
<td>19.0 (5.4)</td>
</tr>
<tr>
<td>FOM/ΔΨ^(−)–Mlphare^(c)</td>
<td>0.2/66</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FOM/ΔΨ^(−)–DM^(c)</td>
<td>0.7/51</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Refinement</th>
<th>Form-I (Synchrotron)</th>
<th>Form-II (Synchrotron)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resolution (Å)</td>
<td>100–1.65 (1.69–1.65)</td>
<td>100–1.90 (1.95–1.90)</td>
</tr>
<tr>
<td>Number of reflections</td>
<td>19,430</td>
<td>31,030</td>
</tr>
<tr>
<td>Number of amino acids</td>
<td>103</td>
<td>315</td>
</tr>
<tr>
<td>Number of solvent atoms</td>
<td>105</td>
<td>165</td>
</tr>
<tr>
<td>Number of Ni2⁺ ions</td>
<td>–</td>
<td>8</td>
</tr>
<tr>
<td>Total number of atoms</td>
<td>968</td>
<td>2,668</td>
</tr>
<tr>
<td>Average B (all atoms) (Å²)</td>
<td>28</td>
<td>45</td>
</tr>
<tr>
<td>R_crytal (%)</td>
<td>15.7 (21.6)</td>
<td>18.3 (24.5)</td>
</tr>
<tr>
<td>R_free (%)</td>
<td>18.3 (26.3)</td>
<td>22.6 (31.3)</td>
</tr>
<tr>
<td>RMSD bonds (Å)</td>
<td>0.012</td>
<td>0.008</td>
</tr>
<tr>
<td>RMSD angles (degrees)</td>
<td>1.435</td>
<td>1.097</td>
</tr>
<tr>
<td>ϕ,ψ Core region (%)</td>
<td>100</td>
<td>98</td>
</tr>
</tbody>
</table>

*a Numbers in parentheses correspond to values in the highest resolution shell.

*b R_mrgd-F is the multiplicity weighted merging R-factor, and R_mrgd-F is an indicator of the quality of reduced data (Diederichs and Karplus, 1997).

*c Figure of merit (FOM): average absolute phase difference between final refined phases and experimental phases both before (Mlphare) and after density modification (DM).

*d R_crytal = ∑ |Fobs – Fcal|/ ∑ |Fobs|; R_free, R_crytal calculated using 10% of the data that were excluded from the refinement.

*e Ramachandran plot quality as defined by Kleywegt and Jones (1996). In chain B of the form-II structure, Asn77 (ϕ,ψ = −111°, 73°) and Asn127 (ϕ,ψ = −135°, −159°) lie in the noncore region.
The majority of the main chain in both crystal forms has clear electron density (Figure 2). The exception is ~15 residues near the N terminus that, due to disorder, have little or no density and have not been modeled (see Methods). One last feature requiring comment is the disulfide bond between Cys64 and Cys160 (Tuori et al., 2000). The electron density seen for Cys64 and Cys160 in form-II shows that the sulfurs are unexpectedly 3.5 Å apart in a nonbonded interaction (and in form-I, no electron density is seen for Cys64). A control data set collected on our in-house x-ray source (Table 1) reveals an intact disulfide, so the cleavage is an artifact of synchrotron data collection as has been seen for other proteins (Burmeister, 2000; Dunlop et al., 2005). For this reason, we will treat it in this article as a disulfide.

Overall Structure
ToxA is a single domain protein having a β-sandwich fold with two antiparallel β-sheets composed of four strands each enclosing the hydrophobic core (Figure 3A). At the N-terminal region of the mature protein that includes pyroglutamate-61 (Pca61) (Tuori et al., 2000), only the first two residues in form-I and the first four in form-II show interpretable density (see Supplemental Figure 1 online) and then come 15 to 16 disordered residues before the first β-strand. The strands are numbered sequentially from β1 to β8 with sheet I composed of strands β1, β8, β5, and β6 and sheet II containing strands β2, β3, β4, and β7. A single one-turn α-helix (α1) is present between β1 and β2. The topologies of the sheets are shown in Figures 3A and 3B. As seen in Figure 4, the β-strands form a rather rigid protein core (low B-factors) and the loops are more mobile (high B-factors).

Among the seven loops connecting the secondary structural elements, three are tight turns (three to four residues) connecting adjacent antiparallel strands (loops β2-β3, β3-β4, and β6-β7), and the other four are longer (four to six residues), bridging strands and helices on opposite sides of the sandwich. The RGD residues (140 to 142) hypothesized to be functionally important are located in the β5-β6 loop (Figure 3A). In form-I, the loop conformation is influenced by the binding of a sulfate ion to the side chain of Arg140 and main chain amides, but it is nevertheless highly mobile (Figure 4). In form-II, Arg140 interacts with another chain in the trimer (see below), resulting in lower mobility when compared with the cubic form (Figure 4).

Trimer Structure
Although ToxA is present as a strongly interacting trimer in both the crystal forms (Figure 5A), during gel filtration it runs at an apparent Mr of 13 kD (see Methods and Supplemental Figure 2 online), suggesting that it is a monomer in solution (Tuori et al., 1995; Meinhardt et al., 1998). Sedimentation equilibrium results were consistent with the existence of a monomer-trimer equilibrium in solution with an association constant (K_a) \( \approx 10^9 \) M^{-2} (see Methods and Supplemental Figure 3 online). This K_a implies that at ToxA concentrations showing toxicity during infusion experiments (~1 μM) (Tuori et al., 2000), pure ToxA would be >99% in the monomeric form. Furthermore, the recombinant proToxA molecule that is fully active behaves as a stable monomer according to both gel filtration and sedimentation equilibrium (see Supplemental Figures 2 and 4 online). Because ToxA would appear to act as a monomer, it is important to describe the trimer interactions in order to understand how the monomer structure may differ from the trimer that is seen in the crystal. Also, since the interaction of ToxA with other proteins may enhance trimer formation, it is still possible that ToxA acts as a trimer sometime during its function.

---

Figure 2. Electron Density Map Quality and the RGD Loop Structure.
A stereodiagram of the 2F_o-F_e electron density map contoured at 1.0 \( \sigma_{\text{rms}} \) shows clear density for the main chain atoms and most of the side chain atoms of the RGD loop of chain B in the form-II crystals. The residues are labeled, and four water sites observed in all four molecules of ToxA are shown. Hydrogen bonds are indicated by dashed lines. This figure was prepared using BobScript (Esnouf, 1999) and Raster3D (Merritt and Bacon, 1997).
Figure 3. Tertiary Structure of ToxA.
The overall shape of the trimer is that of a three-blade pinwheel (Figure 5A). The trimer interface buries 1200 Å² of surface area per monomer and involves a concave broad surface of ToxA binding like a vise around a narrow spine of another monomer that is created largely by strand β6 (Figure 5B). In all, 37 residues participate to create a network of salt bridges near the threefold (involving Arg128, Asp149, and Glu177 from all three chains), a β-sheet interaction between strand β6 (residues Val143 through Leu146) of one chain and strand β7 (Arg156 through Phe159) of a neighboring chain, and the burial of many nonpolar side chains (see Supplemental Figure 5 online for details). In the form-II structure, hydrogen bonds involving the side chains of Arg140 and Arg156 further tie down the tip of the RGD-containing loop (Asn136 through Val143).

What are the implications of this for the monomer structure? First, because the β-ribbon involving the ends of the strands β5 and β6 (Trp134 to Asn136 and Asp142 to Tyr144) extend beyond the main β-sandwich core (see Figures 3A and 3B), we predict that in the absence of the interchain vise-like interactions stabilizing this conformation, these residues will form a much more mobile and larger (up to 11 residues) loop. A second expectation is that the N-terminal region (residues 61 to 64) will also be much less ordered in a monomer.

Rationalization of Existing Mutant Data
Combining the mature ToxA structure with existing mutagenesis data allows us to gain insight into ToxA structure–function interations.

Figure 4. Backbone Mobilities of ToxA.
The average main chain B-factors of form-I (thick line) and form-II chains A (thin line), B (dashed line), and C (dashed-dotted line) are plotted against residue numbers. Locations of the secondary structure elements and the RGD tripeptide are shown below the plot. As expected, loops and regions not stabilized by crystallographic packing interactions have B-factors higher than average, while β-strands forming the protein core have lower B-factors. For the loops with B-factors approaching 90 Å², the most mobile residues have weak main chain density (<1 ρrms), but the chain path is still easily interpretable.
relationships. To date, published mutation analyses of ToxA include 20 single mutations (Tuori et al., 2000; Meinhardt et al., 2002; Manning et al., 2004) that targeted the residues in the RGD-containing segment (Asn136 through Phe147) and residues that represented potential phosphorylation sites (Thr residues 115, 123, 126, 132, 139, and 167) or myristoylation sites (Gly residues 62 and 93) (Table 2). In addition to these mutants, we present here data for a further three mutated residues (see Table 2 and Methods). Of all the mutants, eight are active, two are partially active, and 13 are inactive. All of the mutation sites have unambiguous density in one or both crystal forms, and Figure 6 shows how these mutation sites map onto the structure.

Table 2. Correlation of ToxA Mutations with Structure

<table>
<thead>
<tr>
<th>Mutation</th>
<th>Structural Location</th>
<th>Accessible Side Chain Area (Å²)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S86G</td>
<td>α1</td>
<td>60</td>
<td>This study&lt;sup&gt;p&lt;/sup&gt;</td>
</tr>
<tr>
<td>G96A</td>
<td>β2</td>
<td>11</td>
<td>Meinhardt et al. (2002)</td>
</tr>
<tr>
<td>T115A</td>
<td>β5</td>
<td>32</td>
<td>Manning et al. (2004)</td>
</tr>
<tr>
<td>T123A</td>
<td>β4-β5 loop</td>
<td>65</td>
<td>Manning et al. (2004)</td>
</tr>
<tr>
<td>T126A</td>
<td>β4-β5 loop</td>
<td>61</td>
<td>Manning et al. (2004)</td>
</tr>
<tr>
<td>V138A</td>
<td>β5-β6 loop</td>
<td>124</td>
<td>Manning et al. (2004)</td>
</tr>
<tr>
<td>T167A</td>
<td>β7-β8 loop</td>
<td>15</td>
<td>Manning et al. (2004)</td>
</tr>
<tr>
<td>G168A</td>
<td>β7-β8 loop</td>
<td>34</td>
<td>Manning et al. (2004)</td>
</tr>
<tr>
<td>Impaired&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Structurally disruptive</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G93A</td>
<td>β2</td>
<td>25</td>
<td>This study&lt;sup&gt;p&lt;/sup&gt;</td>
</tr>
<tr>
<td>T132A</td>
<td>β5</td>
<td>3</td>
<td>Manning et al. (2004)</td>
</tr>
<tr>
<td>F147A</td>
<td>β6</td>
<td>33</td>
<td>Manning et al. (2004)</td>
</tr>
<tr>
<td>Structurally nondisruptive</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G62A</td>
<td>N terminus loop</td>
<td>26</td>
<td>This study&lt;sup&gt;p&lt;/sup&gt;</td>
</tr>
<tr>
<td>G64A</td>
<td>N terminus loop</td>
<td>15</td>
<td>Tuori et al. (2000)</td>
</tr>
<tr>
<td>N136A</td>
<td>β5</td>
<td>73</td>
<td>Manning et al. (2004)</td>
</tr>
<tr>
<td>T137A</td>
<td>β5</td>
<td>45</td>
<td>Tuori et al. (2000)</td>
</tr>
<tr>
<td>T139A</td>
<td>β5-β6 loop</td>
<td>92</td>
<td>Manning et al. (2004)</td>
</tr>
<tr>
<td>R140A</td>
<td>β5-β6 loop</td>
<td>137</td>
<td>Manning et al. (2004)</td>
</tr>
<tr>
<td>G141A</td>
<td>β5-β6 loop</td>
<td>21</td>
<td>Meinhardt et al. (2002); Manning et al. (2004)</td>
</tr>
<tr>
<td>D142E</td>
<td>β6</td>
<td>83</td>
<td>Meinhardt et al. (2002); Manning et al. (2004)</td>
</tr>
<tr>
<td>V143A</td>
<td>β6</td>
<td>25</td>
<td>Manning et al. (2004)</td>
</tr>
<tr>
<td>Y144A</td>
<td>β6</td>
<td>109</td>
<td>Manning et al. (2004)</td>
</tr>
<tr>
<td>E145A</td>
<td>β6</td>
<td>91</td>
<td>Manning et al. (2004)</td>
</tr>
<tr>
<td>L146A</td>
<td>β6</td>
<td>66</td>
<td>Manning et al. (2004)</td>
</tr>
</tbody>
</table>

<sup>a</sup> All reported mutants were created and tested in the context of proToxA and behaved well during expression and purification.

<sup>b</sup> Residues Gly62 and Gly93 are potential myristoylation sites; S86G was a PCR-induced mutation.

<sup>c</sup> T132A and L146A are partially active; all other mutants in this category were scored as inactive.

Figure 5. The Crystallographic Trimer of ToxA.

(A) A ribbon diagram of the ToxA trimer as seen in the asymmetric unit of the form-II structure looking down the threefold axis is shown. The three chains are labeled and shown in green, blue, and purple. The four N-terminal residues and the RGD residues are shown as stick models.

(B) The two trimerization surfaces of ToxA. For this image, the green molecule from (A) is separated from the others and rotated 90° around the vertical. The narrow spine-like interface interacting with the purple molecule is facing the reader. The broad vise-shaped interface interacting with the blue molecule faces down. For trimer assembly, the blue surface of one molecule meshes with the purple surface of another. Residues having ≥40-Å² surface area buried at the interface are labeled. The threefold axis is indicated and labeled. This figure was generated using Pymol (www.pymol.org).
Eight residues (Ser86, Gly96, Thr115, Thr123, Thr126, Val138, Thr167, and Gly168) when mutated to Ala or Gly resulted in variants that retained toxicity. The side chain of each of these residues is solvent exposed, where the mutation would not be expected to disrupt ToxA folding or stability. Their unperturbed activity indicates that these surface regions (green in Figure 6) are not crucially involved in receptor recognition or other interactions required for ToxA function.

In the case of partially and fully inactive variants, the location of each mutated residue allows us to distinguish whether that mutation would be expected to destabilize the protein (yellow in Figure 6) or not (purple in Figure 6). If destabilization is unlikely, then that site is implicated as a potential point of direct involvement in receptor recognition or another interaction required for ToxA function. Among the partially and completely inactive mutants, three (Gly93, Thr132, and Phe147) are in environments where the mutation would be expected to disrupt the structure. The structure reveals that Gly93, located at the start of strand $\beta_2$, adopts main chain torsion angles ($\phi, \psi = 105^\circ, -22^\circ$) that are unfavorable for non-Gly residues, so the mutation to Ala would destabilize the local conformation. Thr132, one of the potential phosphorylation sites, is mostly buried in the ToxA monomer by nearby side chains of Trp134 and Gln175, and we expect that mutation to Ala would disrupt these interactions and destabilize and/or alter ToxA structure. We conclude that Thr132 plays a structural role, but since it is partially exposed, we cannot rule out that it may play a direct role, such as to be phosphorylated in planta. For Phe147, the side chain is buried in the core of the monomer where mutation would likely disturb the structure.

The remaining 12 residues have partly or fully exposed side chains where, based on the folded protein structure, the mutations would not be expected to disrupt the structure. As seen in Figure 6, these sites are mostly located on one protrusion of the fold, which is the RGD-containing loop between $\beta_5$ and $\beta_6$. The only exceptions are Gly62 and Cys64. Interestingly, it is possible that these two residues represent a special case where the mutations would be compatible with the folded conformation, but they alter the folding process so that the native fold is never reached. Evidence that Cys64 is an important player in the folding process is that the disulfide involving Cys64 is properly formed during in vivo folding of proToxA but does not properly reform during in vitro refolding of reduced ToxA (Tuori et al., 2000). Thus, it is plausible that mutations of Cys64 and its neighbors (like Gly62) would interfere with folding.

The clustering of the inactive, nondisruptive mutations on the RGD-containing loop is consistent with the original hypothesis (Ciuffetti et al., 1997; Zhang et al., 1997; Meinhardt et al., 2002; Manning et al., 2004) that ToxA may interact with plant cell integrin-like receptors in a manner analogous to mammalian proteins such as vitronectin and fibronectin. In mammals, recognition of RGD residues is highly specific and exquisite in a variety of contexts (Xiong et al., 2002; Takagi et al., 2003; Springer and

**Figure 6.** Mapping of Site-Directed Mutagenesis Results onto the ToxA Structure.

The molecular surface of ToxA is colored to highlight the sites of fully active mutants (green), partially active or inactive mutants that are expected to disrupt the structure (yellow), and partially active or inactive mutants that are expected to be nondisruptive (purple). All residues except Thr132, Gly141, and Val143 are visible in these views. The surface on the left is in the same orientation as Figure 3A, while the surface on the right is rotated 180° around the vertical to show the other side of the protein. The figure was generated using Pymol (www.pymol.org).
Figure 7. Circular Permutation of ToxA.

(A) Topological comparison of ToxA and the tenth domain of fibronectin (FnIII10) shown as a representative for the FnIII domain using the strand nomenclature proposed by Leahy et al. (1992). Open arrows and rectangles represent elements that are common to both of the structures, and striped arrows represent elements unique to each structure. The positions of the conserved RGD residues are also indicated.

(B) Stereoview of the ToxA (pink)/FnIII10 (blue) comparison, with the unique elements of each chain indicated (striped). The Cα positions of RGD residues are shown as semitransparent spheres. The view is identical to that of Figure 3A. The secondary structural elements are labeled in both panels. The connection of the N and C termini (dashed line) and the cleaving of the β6-β7 loop (arrowhead) define the changes involved in the circular permutation.
Wang, Xiao et al., 2004). ToxA is similar in that even a subtle Asp142Glu mutation results in loss of recognition, but it is different in that not just the RGD but also many surrounding residues are also required for recognition. The interest of the RGD residues for ToxA function is further highlighted by the fact that interaction in animal cells of some RGD-containing proteins with integrins results in their internalization. In particular, this route is used by certain viruses as a cell entry mechanism (Wickham et al., 1993; Memmo and McKeown-Longo, 1998; Hynes, 2002; Marjomaki et al., 2002; Zubieta et al., 2005). While this phenomenon has not yet been seen in plants, that RGD-containing peptides are known to disrupt the plasma membrane–cell wall continuum (Canut et al., 1998; Mellersh and Heath, 2001; Senchou et al., 2004) provides circumstantial evidence for the existence of integrin-like proteins in plants.

Comparison with Mammalian RGD Proteins: A Possible Evolutionary Relationship

In order to gain further insight from the ToxA structure, a search for structurally similar proteins was performed using the program DALI (Holm and Sander, 1998). The top matches were all proteins with β-sandwich folds, but the Z-scores (highest 4.2) were well below the cutoff of 6.0 suggested by DALI as indicative of a reliable structural homolog. Interestingly, among the top scores were known RGD-containing mammalian proteins that are based on a fibronectin type III (FnIII) fold (also a β-sandwich structure). Therefore, the structure FnIII1 was compared with that of ToxA. A study of the topology diagrams revealed that the proteins could be made to have a similar topology through a simple circular permutation of ToxA (Figure 7A). The circular permutation event connects the N and C termini of ToxA, and new termini are formed by the cleavage of the β6–β7 loop. In this comparison, strands β2 and β3 of ToxA are insertions relative to FnIII, and strand A of FnIII is an insertion relative to ToxA (Figure 7A). Thus, after circular permutation, ToxA and FnIII can be aligned using a common six-stranded core of the same topology.

A new DALI search with a circularly permuted ToxA yielded the RGD-containing structures of the tenth domain of fibronectin (FnIII10; Dickinson et al., 1994) and the third domain of tenascin (FnIII3; Leahy et al., 1993; Memmo and McKeown-Longo, 1998; Hynes, 2002; Marjomaki et al., 2002; Zubieta et al., 2005). While this phenomenon has not yet been seen in plants, that RGD-containing peptides are known to disrupt the plasma membrane–cell wall continuum (Canut et al., 1998; Mellersh and Heath, 2001; Senchou et al., 2004) provides circumstantial evidence for the existence of integrin-like proteins in plants.

RGD functions in terms of binding to a cell surface receptor provide an attractive case for the existence of a common ancestor, especially given the evidence that the plant cell wall is the functional equivalent of the mammalian extracellular matrix (Kohorn, 2000; Baluska et al., 2003). Isolation of the plant protein that serves as a ToxA receptor and structural and functional characterization of fibronectin-like proteins from the cell wall, such as the putative one recently isolated (Pellenc et al., 2004), may provide clues to the evolutionary origins and mode of action of ToxA.

METHODS

Purification and Crystallization

ToxA was purified from culture filtrates as previously described (Tuori et al., 1995). A procedural modification involved dialyzing the protein against Epure water (Barnstead) instead of 30 mM TRIS, pH 9.0, and the concentration was determined using an extinction coefficient of ε280 = 18 mM−1 cm−1 (Zhang et al., 1997). Prior to crystallization, the protein was diluted to 1.5 mg/mL in 0.01 M HEPES, pH 7.0, and stored at 4°C.

Two crystal forms of ToxA were grown at room temperature using the hanging-drop vapor diffusion method. Form-I crystals exhibited cubic morphology and grew from a 1:1 drop using a reservoir solution containing 0.5 M (NH₄)₂SO₄, 15% dioxane, and 0.1 M MES, pH 6.5. Crystals grew over 2 weeks to a final size of ~0.15 mm³. For data collection, the cubic crystals were flash-frozen in liquid nitrogen after a 2-min incubation in a cryoprotectant consisting of the reservoir plus 15% glycerol. Form-II crystals were obtained by mixing equal volumes of the protein stock with a reservoir of 7.5% polyethylene glycol 400 (νν), 0.2 M NiCl₂, and 0.1 M HEPES-Na, pH 7.0. These tetragonal-shaped crystals appeared within 24 h and grew in 1 week to a size of ~0.15 × 0.15 × 0.10 mm³. Divalent cations were essential for crystallization with Co²⁺, Ca²⁺, and Cd²⁺ ions able to substitute for Ni²⁺. For data collection, the form-II crystals were progressively soaked for 5 min each in a stepwise manner through solutions of 15, 22.5, and 30% (νν) final concentration polyethylene glycol 400 plus reservoir and then flash-frozen in liquid nitrogen.

Data Collection

For structure determination, highly redundant SAD data were collected using a form-I crystal and in-house Cu-Kα radiation (Rigaku RU-H3R rotating anode operating at 50 kV and 100 mA and an R-Axis IV image plate detector: λ = 1.54 Å, Δλ = 1°, 540 15-min images) (Table 1). For structure refinements, 1.65-Å resolution native data were collected at beamline 8.2.1 of the Advanced Light Source (ALS; Berkeley National Laboratory). High-resolution pass, λ = 1.0 Å, Δλ = 1°, 40 20-s images; low-resolution pass, λ = 1.0 Å, Δλ = 1°, 405-s images) for the form-I crystals, and 1.90-Å data (λ = 1.0 Å, Δλ = 1°, 160 15-s images) were collected for the form-II crystals. A low-resolution, in-house data set extending to 3 Å (λ = 1.54 Å, Δλ = 1°, 99 15-min images) was also collected on a form-II crystal to confirm the presence of a disulfide bond. All data sets were processed using the HKL suite of programs (Otwinowski and Minor, 1997) (Table 1).

The form-I crystals belong to the cubic space group P2₁3 with unit cell dimensions a = b = c = 78.2 Å, with one molecule in the asymmetric unit and a solvent content of 60%. Form-II crystals belong to the tetragonal space group P4₁2₁2 with unit cell dimensions a = b = c = 74.5 Å and c = 137.3 Å and have three molecules in the asymmetric unit with a solvent content of 50%.

Structure Determination and Refinement

The 13.2-kD ToxA structure was determined from form-I crystals at 3.0-Å resolution using the weak anomalous signal arising from inherently
present sulfur atoms. The expected Bijvoet ratio (Hendrickson and Teeter, 1981) from the four sulfur atoms (two Met and two Cys) present in the 13.2-kD protein at 1.54-Å wavelength is ~0.8%. It was later observed that two of the four sulfurs (one Cys and one Met) were disordered, but the presence of a single well-ordered sulfate ion in the asymmetric unit led to three sulfurs contributing to the phasing of the 13.2-kD protein.

In order to determine the low-resolution cutoff for structure solution, the 540-image data set was split into two consecutive 270-image data sets (images 1 through 270 and 271 through 540), and the anomalous correlation coefficients between the two sets were calculated using the program XPREP (Bruker Analytical X-ray Systems). At 3-Å resolution, the correlation coefficients fell below 20%, and this was chosen as a cutoff. The program SHELX (Schneider and Sheldrick, 2002) was used to locate the positions of the sulfurs. The 100 search trials resulted in a clear solution (~50% success rate) with a correlation coefficient of 39% (all data) (Debreczeni et al., 2003). Heavy atom refinement (positions and occupancies) and phasing were performed using the program Miphare (Winn, 1991). Density modification by DM (Cowtan and Zhang, 1999) resulted in experimental maps that were of superior quality. A more detailed analysis assessing how much data were necessary to solve the structure will be presented elsewhere (G.N. Sarma and P.A. Karplus, unpublished data).

Prior to model building and refinement, 10% of the data were set aside for cross-validation. The automated model building program ARP/wARP (Perrakis et al., 1999) was used to extend the 3-Å phases to 1.65 Å and to build ~90% of the model. Furthermore, the program was used to pick waters using the following criteria: (1) a peak of ≥3 ρmax (RMSD electron density of map often reported as ρ) in the F_obs − F_calc map and a peak of ≥1 ρmax in the 2Fobs − Fcalc map (2) a distance of ≥2.5 and ≤3.5 Å to nearby hydrogen bond donor or acceptor. Manual rebuilding was performed using O (Jones et al., 1991). Anisotropic (TLS) refinement as incorporated in REFMAC (Murshudov et al., 1997) followed by positional and individual B-factor refinements resulted in a final R and Rfree of 15.7 and 18.3%, respectively. The final model has 103 residues (61 to 62 and 78 to 178), 105 waters, and 2 sulfates, with alternate conformations seen for Ser68, Ile95, Arg117, Ile131, Gln154, Arg173, and Glu177; and for waters 28, 35, 37, 40, 55, 62, 64, and 65. Residues 63 through 77 did not have any visible density and were not modeled.

For form-I structure solution, a crystallographic trimer generated using the TRUNC program (without waters) was used as a search model for molecular replacement at 4 Å using the direct Patterson search as implemented in the program CNS (Brüger et al., 1998). Cross-rotation and translation searches yielded a clear solution with a correlation coefficient of 0.550 versus 0.114 for the next best solution. The trimer was refined to 1.9 Å using REFMAC (Murshudov et al., 1997), with 10% of the data set aside for cross-validation. No noncrystallographic symmetry restraints were used during refinement. Waters were picked (using the criteria reported above), and alternate conformations were modeled for Ile95 in chain A, Arg113 in chain B, and waters 94 and 126. In one of the chains (chain B), a six-residue stretch (residues 74 to 79) forms an additional β-strand that interestingly lies on a crystallographic twofold axis and shows weak but clear main chain electron density. Thus, in the crystal, this loop shares the space with the exact same segment from another trimer in the crystal. Accordingly, it was modeled with half occupancy (see Supplemental Figure 6 online). The weak density for the loop is consistent with the high temperature factors for this region of the protein (Figure 4). Strong electron density peaks for Ni2+ ions were seen on the crystallographic twofold axes. Additional Ni2+ ions were located using anomalous difference maps calculated from another data set optimized for Ni anomalous signal (λ = 1.487 Å; data not shown).

Occupancies for the Ni ions not on special positions were estimated by occupancy test refinement runs with occupancies of 0.2, 0.4, 0.6, and 0.8, and for each Ni, the occupancy that yielded a B-factor most similar to the surrounding protein atoms was chosen. The eight Ni2+ ions ranged in occupancy from 0.2 to 0.5. Unsuccessful attempts to solve the structure of the form-II crystal using Ni multiple wavelength anomalous dispersion (data not shown) could be explained by the low occupancies and high temperature factors of the Ni2+ ions.

Iterative cycles of model building and refinement yielded an R and Rfree of 20.8 and 26.4%. TLS refinement, with each monomer as a rigid body, lowered the R and Rfree to 18.3 and 22.6%, respectively. The final model has 315 amino acids (103 chain A, 109 chain B, and 103 chain C), 165 waters, and 8 Ni2+ ions, with residues 65 to 79 (chains A and C) and residues 65 to 73 (chain B) not modeled due to weak electron density.

Since discretely ordered solvent molecules are an integral part of the structure and function of the protein, the preferred hydration sites (Kaprus and Faerman, 1994) in each crystal form were numbered according to their electron density strength (water 1 being the highest) from 1 to 105 in form-I and to 165 in form-II. In form-I and -II crystals, 84 and 141 of the sites, respectively, are present in the first solvent shell (making at least one hydrogen bond with the protein).

**Structural Comparison and Analyses**

Secondary structure assignments were made using DSSP (Kabsch and Sander, 1983) and also by visual inspection. Hydrogen bonding geometries were calculated using HBPLUS (McDonald and Thornton, 1994). Structure-based sequence alignments were done using HOCORE (P.A. Karplus, unpublished data) with a distance cutoff of 3.0 Å.

**Site-Directed Mutagenesis**

Of the three mutants newly reported in this study, G62A and G93A were generated by site-directed mutagenesis using the following primers (one of the complementary pair): G62A, 5’-CTACAGGAGCCGACGAGC-CTGCATGTCAATC-3’; G93A, 5’-CTCGGACAGCTTGTGCTATAGGC-TCGTGG-3’. The S86G mutation was generated inadvertently between PCR cloning of the proToxA into pTRC99-HPro (Tuori et al., 2000). All mutant proteins were expressed in *Escherichia coli* and assayed for bioactivity as described by Manning et al. (2004).

**Gel Filtration**

The existence of the ToxA monomer in solution (0.5 mg/mL in 100 mM NaCl and 10 mM sodium phosphate, pH 7.0) was confirmed by gel-filtration chromatography using a Superdex-75 10/300 GL size-exclusion column (Amersham Biosciences) at a flow rate of 0.5 mL/min. Using identical conditions, proToxA at 0.17 mg/mL was also chromatographed. Fractions were collected every minute, and activity was assayed by infiltrating peak fractions into the sensitive wheat (*Triticum aestivum*) cultivar Katepwa. For ToxA, the major active fraction and >80% of the A280 absorbing material ran at an elution volume corresponding to a molecular mass of ~13 kD, and for proToxA, >95% of the absorbing material ran at an elution volume corresponding to a molecular mass of ~19 kD (see Supplemental Figure 2 online).

**Analytical Ultracentrifugation**

Sedimentation equilibrium analytical ultracentrifugation experiments were performed with a Beckman Optima XL-A analytical ultracentrifuge at 20°C using a four-hole AN-60Ti rotor and double sector charcoal/Epon-filled centriprces. Buffer densities and viscosity corrections were made according to data published by Laue et al. (1992). The partial specific volume of ToxA was estimated from the protein sequence to be 0.727 cm³/g, using the method of Perkins (1986).

The centrifugation experiments were performed according to the protocol described by Ausio et al. (1992) in a solution containing 100 mM NaCl, 10 mM sodium phosphate, pH 7.0.
NaCl and 50 mM sodium phosphate, pH 7.0. Four 120-µL samples of ToxA were sedimented to equilibrium: 3.4 and 10.1 µM samples were run at 22,000 and 30,000 rpm, and 7.8 and 23.4 µM were run at 22,000, 26,000, and 30,000 rpm. Scans were collected with absorbance optics at wavelengths between 230 and 280 nm. The radial step size was 0.001 cm, and each c versus r data point was the average of 15 independent measurements. Wavelengths were chosen so that no points exceeded an absorbance of 1. Equilibrium data spanning this concentration range were examined by global fitting using UltraScan software for the Linux operating system (www.ultrascan.uthscsa.edu). Equilibrium data were fit to multiple models using global fitting of all data (Russell et al., 2004). The monomer-trimer equilibrium fit resulted in a monomer of molecular mass of ~13.4 kDa, which is in good agreement with the molecular mass of 13.2 kDa calculated from the ToxA sequence. Fitting the data to monomer-dimer and monomer-dimer-trimer models also resulted in reasonable fits giving monomer molecular masses of 12.9 and 13.6 kDa, respectively. The monomer-trimer model had the best statistics and the lowest systematic errors based on inspection of the residual patterns, and the monomer molecular weight is closest to the true value and is consistent with the formation of trimer in the crystal.

For proToxA, one sample at a concentration of 33.8 µM was analyzed. The speeds and scans were done as for ToxA. The data fit well to a single-component ideal model that yielded a molecular mass of 19.7 kDa, exactly matching the calculated molecular mass of a monomer.

**Accession Numbers**

The coordinates and the structure factors have been deposited in the Protein Data Bank (www.rcsb.org/pdb/) as entries 1ZLD (form-I) and 1ZLE (form-II).

**Supplemental Data**

The following materials are available in the online version of this article.

- Supplemental Figure 1. N-terminal Residues of ToxA.
- Supplemental Figure 2. Gel Filtration Analysis.
- Supplemental Figure 3. Sedimentation Equilibrium Analysis of ToxA.
- Supplemental Figure 4. Sedimentation Equilibrium Analysis of proToxA.
- Supplemental Figure 5. The Trimerization Interface of ToxA.
- Supplemental Figure 6. Residues on a Crystallographic Twofold Axis.

**ACKNOWLEDGMENTS**

We would like to thank Rick Faber for help in data collection and useful discussions and Sonia Anderson and Dean Malencik (Oregon State University) for performing the analytical ultracentrifugation experiments. The invaluable help rendered by the staff at the ALS (beamline 8.2.1), particularly Christine Trame, is also appreciated. The ALS is supported by the Director, Office of Science, Office of Basic Energy Sciences, Materials Sciences Division of the U.S. Department of Energy under Contract DE-AC03-76SF00098 at Lawrence Berkeley National Laboratory. This work was supported by a National Science Foundation grant (MCB-0488665) to L.M.C. and P.A.K. and in part by the National Research Initiative of the USDA Cooperative State Research, Education, and Extension Service (Grant 2001-35319-10017 to L.M.C.). This work was made possible in part by the Proteins and Nucleic Acids Core Facility of the Environmental Health Sciences Center at Oregon State University (National Institute of Environmental Health Science Grant P30 ES00210).

Received June 10, 2005; revised August 12, 2005; accepted September 7, 2005; published October 7, 2005.

**REFERENCES**


Structure of Ptr ToxA: An RGD-Containing Host-Selective Toxin from *Pyrenophora tritici-repentis*

Ganapathy N. Sarma, Viola A. Manning, Lynda M. Ciuffetti and P. Andrew Karplus

*Plant Cell* 2005;17;3190-3202; originally published online October 7, 2005;
DOI 10.1105/tpc.105.034918

This information is current as of June 22, 2017

<table>
<thead>
<tr>
<th>Supplemental Data</th>
<th>/content/suppl/2005/09/30/tpc.105.034918.DC1.html</th>
</tr>
</thead>
<tbody>
<tr>
<td>References</td>
<td>This article cites 65 articles, 13 of which can be accessed free at: /content/17/11/3190.full.html#ref-list-1</td>
</tr>
<tr>
<td>eTOCs</td>
<td>Sign up for eTOCs at: <a href="http://www.plantcell.org/cgi/alerts/ctmain">http://www.plantcell.org/cgi/alerts/ctmain</a></td>
</tr>
<tr>
<td>CiteTrack Alerts</td>
<td>Sign up for CiteTrack Alerts at: <a href="http://www.plantcell.org/cgi/alerts/ctmain">http://www.plantcell.org/cgi/alerts/ctmain</a></td>
</tr>
<tr>
<td>Subscription Information</td>
<td>Subscription Information for <em>The Plant Cell</em> and <em>Plant Physiology</em> is available at: <a href="http://www.aspb.org/publications/subscriptions.cfm">http://www.aspb.org/publications/subscriptions.cfm</a></td>
</tr>
</tbody>
</table>

© American Society of Plant Biologists
ADVANCING THE SCIENCE OF PLANT BIOLOGY