Bacillus thuringiensis Cry4A and Cry4B Mosquito-larvicidal Proteins: Homology-based 3D Model and Implications for Toxin Activity

Chanan Angsuthanasombat*, Panapat Uawithya, Somphob Leetachewa, Walairat Pornwiroon, Puey Ounjai, Teerakiat Kerdcharoen[†], Gerd Katzenmeier and Sakol Panyim

Laboratory of Molecular Biophysics, Institute of Molecular Biology and Genetics,
Mahidol University, Salaya Campus, Nakornpathom 73170, Thailand
[†]Department of Physics, Faculty of Science, Mahidol University, Bangkok, 10400, Thailand

Three-dimensional (3D) models for the 65-kDa activated Cry4A and Cry4B δ-endotoxins from *Bacillus thuringiensis* subsp. *israelensis* that are specifically toxic to mosquito-larvae were constructed by homology modeling, based on atomic coordinates of the Cry1Aa and Cry3Aa crystal structures. They were structurally similar to the known structures, both derived 3D models displayed a three-domain organization: the N-terminal domain (I) is a seven-helix bundle, while the middle and C-terminal domains are primarily comprise of anti-parallel β-sheets. Circular dichroism spectroscopy confirmed the secondary structural contents of the two homology-based Cry4 structures. A structural analysis of both Cry4 models revealed the following: (a) Residues Arg-235 and Arg-203 are located in the interhelical 5/6 loop within the domain I of Cry4A and Cry4B, respectively. Both are solvent exposed. This suggests that they are susceptible to tryptic cleavage. (b) The unique disulphide bond, together with a proline-rich region within the long loop connecting α4 and α5 of Cry4A, were identified. This implies their functional significance for membrane insertion. (c) Significant structural differences between both models were found within domain II that may reflect their different activity spectra. Structural insights from this molecular modeling study would therefore increase our understanding of the mechanic aspects of these two closely related mosquito-larvicidal proteins.

Keywords: *Bacillus thuringiensis*, δ-Endotoxin, Disulfide bond, Homology modeling, 3D model, Mosquito-larvicidal activity, Tryptic cleavage site

*To whom correspondence should be addressed.
Tel: 662-441-9003 ext. 1237; Fax: 662-441-9906
E-mail: stcas@mahidol.ac.th

Introduction

The Gram-positive endospore-forming bacterium *Bacillus thuringiensis* (*Bt*) synthesizes highly specific larvicidal proteins in large quantities as different forms of parasporal crystalline inclusions during sporulation (Aronson *et al.,* 1986; Höfte and Whiteley, 1989). These cytoplasmic inclusions are composed of one or several insecticidal proteins (δ-endotoxins) that have been classified into two major families - Cry and Cyt toxins - based on the similarity of their deduced amino acid sequences (Höfte and Whiteley, 1989; Crickmore *et al.,* 1998). Currently, the Cry δ-endotoxins have been shown to be highly active against a wide variety of insect larvae in the Diptera (mosquitoes and flies), Lepidoptera (moths and butterflies), Coleoptera (beetles and weevils), and Hymenoptera (wasps and bees) (Schnepf *et al*., 1998; de Maagd *et al*., 2001) orders. For example, the 130-kDa Cry4A and Cry4B toxins that are produced from *Bt* subsp. *israelensis* are specifically toxic to mosquito-larvae of the genus *Aedes*, *Anopheles*, and *Culex* (Höfte and Whiteley, 1989; Schnepf *et al.,* 1998). These mosquitoes are serious human disease vectors that transmit dengue virus, malaria, and filarial parasites (Becker and Margalit, 1993).

Most of the native *Bt* δ-endotoxins are synthesized as inactive protoxins that are found within inclusion bodies (Aronson *et al.*, 1986). After ingestion by the susceptible insect larvae and dissolution in the midgut lumens (generally alkaline pH), the solubilized protoxins are activated by larval gut proteases to yield toxic fragments that are relatively resistant to further proteolysis (Aronson *et al.*, 1986; Höfte and Whiteley, 1989). The active toxins then bind to specific receptors on the midgut epithelial cell brush border membrane, followed by membrane insertion of the toxins to form cation permeable pores. These ion-leakage pores cause the target midgut epithelial cells to swell and lyse by colloid-osmotic lysis (Knowles and Ellar, 1987), resulting in extensive

damage to the midgut and eventually larval death (Knowles, 1994).

To date, the X-ray crystal structures of several Cry δ-endotoxins [Cry3Aa (Li *et al*., 1991), Cry1Aa (Grochulski *et al*., 1995), Cry1Ac (Derbyshire *et al*., 2001), Cry2Aa (Morse *et al*., 2001) and Cry3Bb (Galitsky *et al*., 2001)] have been elucidated. Although these Cry toxins exert their specific toxicity against different target insect larvae, they display a high degree of overall 3D structural similarity, consisting of three distinct domains. The N-terminal domain I, a seven-helix bundle (six amphipathic helices around a central core helix), is clearly equipped for membrane insertion and pore formation (Li *et al*., 1991; Grochulski *et al*., 1995; Morse *et al*., 2001). This suggestion is supported by various studies demonstrating that the isolated helical fragment from different Cry toxins is responsible for pore-forming activity (Walters *et al*., 1993; Von Tersch *et al*., 1994; Puntheeranurak *et al*., 2001; 2004). In addition, the helical hairpin α4-α5 within this pore-forming domain has been shown to be an important determinant for Cry toxin activity (Schwartz *et al*., 1997; Gazit *et al*., 1998; Uawithya *et al*., 1998; Masson *et al*., 1999; Gerber and Shai, 2000; Sramala *et al*., 2000). Domain II, a β-prism, is composed of three anti-parallel β-sheets that are packed around a hydrophobic core, each terminating in a loop at the apex of the molecule (Li *et al*., 1991; Grochulski *et al*., 1995). This domain participates in receptor recognition and hence insect specificity (Smedley and Ellar, 1996; Rajamohan *et al*., 1996; Ballester *et al*., 1999; Jurat-Fuentes and Adang, 2001). Domain III is a β-sandwich that is comprised of two anti-parallel β-sheets. The role of this C-terminal domain is still unclear although it has been implicated in the structural integrity of toxin molecules (Li *et al*., 1991), membrane permeabilisation (Schwartz *et al*., 1997; Masson *et al*., 2002), or receptor binding that can influence insect specificity (Lee *et al*., 1995; de Maagd *et al.*, 1996; Burton *et al*., 1999).

Currently, the precise mechanism of toxicity of the *Bt* Cry δ-endotoxins is still not completely understood, although the understanding of how these larvicidal proteins work at the molecular level has substantially increased over the last decade (de Maagd *et al*., 2001). To gain molecular structural knowledge of the two closely related *Bt* mosquito-larvicidal proteins, Cry4A and Cry4B, we generated plausible 3D models for the 65-kDa activated forms of these two toxins by homology modeling. Implications for the molecular basis of proteolytic activation, membrane-pore formation, and specificity determination of these mosquito-larvicidal proteins are discussed.

Materials and Methods

Alignment of protein sequences Amino acid sequences of the three known protein structures [Cry1Aa (D17518), Cry2Aa (U32369) and Cry3Aa (U10985)] and the two closely related mosquito-larvicidal proteins [Cry4A (Y00423) and Cry4B (X07423)] that were used in this study are based on the holotypes

that were used previously by Höfte and Whiteley (1989) and retrieved from GenBank (NCBI). The ClustalX program (Thompson *et al*., 1997) was used to generate a multiple sequence alignment of the selected protein sequences. Manual adjustment of the alignment, based on inter-sequence checks, was performed to generate the final sequence alignment.

Construction and evaluation of homology models The atomic coordinates of the 65-kDa activated Cry1Aa toxin (PDB entry 1DLC) and the Cry3Aa toxin (PDB entry 1CIY) were used as structural templates for constructing 3D modeled structures of the bio-active N-terminal half (65 kDa) of Cry4A (Glu-69 to Thr-678) and Cry4B (Ile-40 to Thr-634), respectively. The homology modeling program, WHATIF (Vriend, 1990), running under a Dec600 workstation, was used for constructing the models. Side-chain orientations were checked using the ROTCHECK option, and the unfit side-chains were replaced by the more favorable rotamers based on local backbone conformations that were provided in the database. Insertion regions in the model relative to the template protein was accomplished by extracting from a short fragment database using loop search algorithm (DGINS) in the WHATIF program.

The entire model structures were initially subject to energy minimization using a CHARMM force field of the TINKER program package (Ponder and Richards, 1987). Refinements were based on nonlinear conjugate gradient algorithm with root mean square (RMS) gradient criteria of 0.1 kcal mol^{-1} Å^{-2}. Energy minimization was done *in vacuo* to relieve steric interactions within the model structure. Stereochemical parameters of the model structures were analyzed with the PROCHECK program (Laskowski *et al*., 1993). Packing environment, bond lengths, and bond angles were analyzed using the CHECK option in the WHATIF program (Vriend, 1990).

Toxin expression and preparation *Escherichia coli* cells expressing the 130-kDa Cry4A toxin as a soluble form or the 130-kDa Cry4B toxin as a cytoplasmic inclusion (Angsuthanasombat *et al*., 1987) were harvested by centrifugation, resuspended in distilled water, and then disrupted in a French Pressure Cell at 16,000 psi. For preparation of the 130 kDa Cry4A toxin, the crude *E. coli* lysates were centrifuged at $8,000 \times g$ for 10 min and the supernatant containing the soluble protoxin was collected for further purification *via* the size-exclusion FPLC system (Superose 12 HR10, Amersham Pharmacia Biotech, Uppsala, Sweden) with a linear gradient of 10 mM sodium phosphate (NaH$_2$PO$_4$/Na$_2$HPO$_4$) buffer, pH 9.0 at 0.5 ml min^{-1}. The 1-ml fractions were collected and analyzed by SDS-15% w/v polyacrylamide gel electrophoresis. In the protoxin preparation for Cry4B, the sediment inclusions that were obtained from the disrupted cells were washed 3 times in distilled water. The inclusions (3 mg ml^{-1}) were subsequently solubilized in a 10 mM sodium phosphate buffer, pH 9.0, at 37°C for 1 h. Protein concentrations were determined by using a protein microassay (Bio-Rad, Hercules, USA), with bovine serum albumin fraction V (Sigma, St. Louis, USA) as a standard.

The FPLC purified Cry4A protoxin and the solubilized Cry4B protoxin (1-2 mg ml^{-1}) were further digested with trypsin (*L*-1-tosylamide-2-phenylethylchloromethyl ketone treated, Sigma) at a toxin:enzyme ratio of 20 : 1 (w/w). Trypsin-treated fractions were
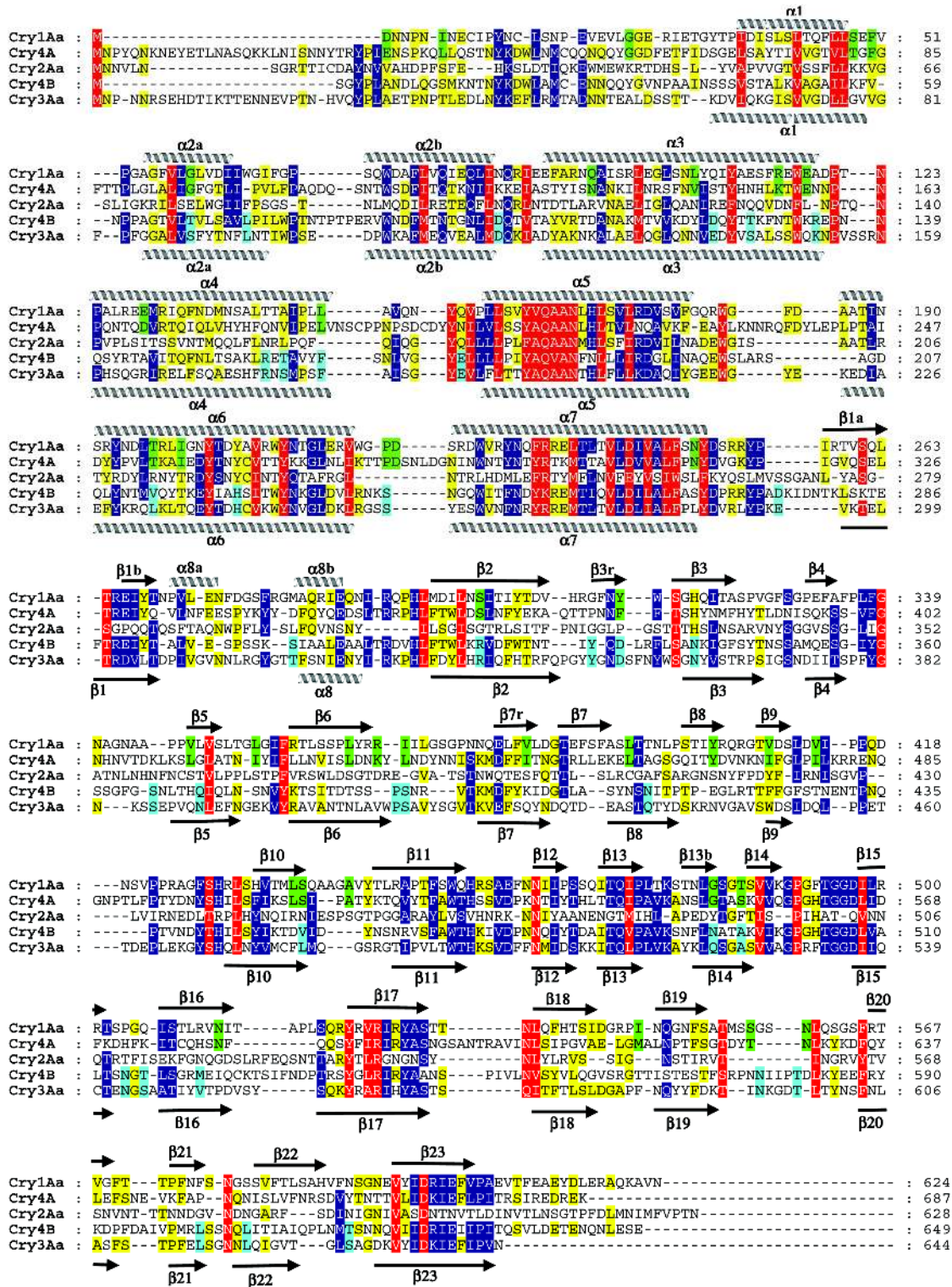
Fig. 1. Multiple sequence alignment of the 65-kDa active forms of Cry4A and Cry4B with the crystal structures of Cry1Aa, Cry2Aa, and Cry3Aa. The degree of conservation among the five sequences is represented by background shading of the residues with red, blue, and yellow shading for 100%, 80%, and 60% sequence homology, respectively. More regions of homology between Cry4A and Cry1Aa, and between Cry4B and Cry3Aa, are highlighted green and cyan, respectively. The positions of secondary structure elements of Cry1Aa and Cry3Aa are illustrated over and under the alignments, respectively.

then purified on a size-exclusion FPLC system, as previously described.

Identification of disulphide-containing peptides   Mobility-shift assayed on SDS-PAGE was performed to identify cysteine or disulphide containing protein fragments by the modified method of Gray (1997). Prior to analysis on 13% SDS-polyacrylamide gel, the trypsin-treated toxin fragments of Cry4A and Cry4B were prepared with 0.25 vol. of sample loading buffers (100 mM Tris, pH 7.5, 40% glycerol, 4% SDS, 4 mM EDTA, 0.1% Bromophenol blue, with/without 144 mM β-mercaptoethanol). The samples were incubated at 100ºC for 5 min.

Circular dichroism (CD) measurements   CD measurements of the FPLC purified 65-kDa Cry4A and Cry4B proteins were performed on a Jasco J-715 CD spectropolarimeter (Jasco Inc., Easton, USA) in the far UV region at 25ºC using a rectangular quartz cuvette (0.2-mm optical path-length). Samples were prepared in a 10 mM sodium phosphate buffer, pH 9.0, with protein concentrations of 0.11 mg ml^{-1} (Cry4A) and 0.30 mg ml^{-1} (Cry4B), as determined by far UV absorbance. The CD spectra were recorded at a scanning rate of 20 nm min^{-1} with a spectral bandwidth of 2 nm and response

times of 2 ms. Five accumulations were taken and the results were averaged. After background subtraction, the CD data were converted from a CD signal (mdegree) into mean residue ellipticity (deg cm^2 dmole^{-1}). The secondary structure content was estimated from the CD spectra using the variable selection method (Manavalan and Johnson, 1987).

Results and Discussion

Modeling of Cry4A and Cry4B and their validation   The C-terminal regions (ca. 465 deduced amino acid residues) of the 130-kDa Cry4A and Cry4B protoxins were almost identical, but their N-terminal proteolytic activated portions (ca. 65 kDa) shared only about 55% sequence similarity (Ward and Ellar, 1988; Chungjatupornchai *et al*., 1988). In this report, the sequence alignment of the deduced amino acid sequences of both the 65-kDa activated Cry4A and Cry4B toxins with those of the three known coordinates [Cry1Aa (Grochulski *et al*., 1995), Cry2Aa (Morse *et al*., 2001) and Cry3Aa (Li *et al*., 1991)] was initially performed. When gaps
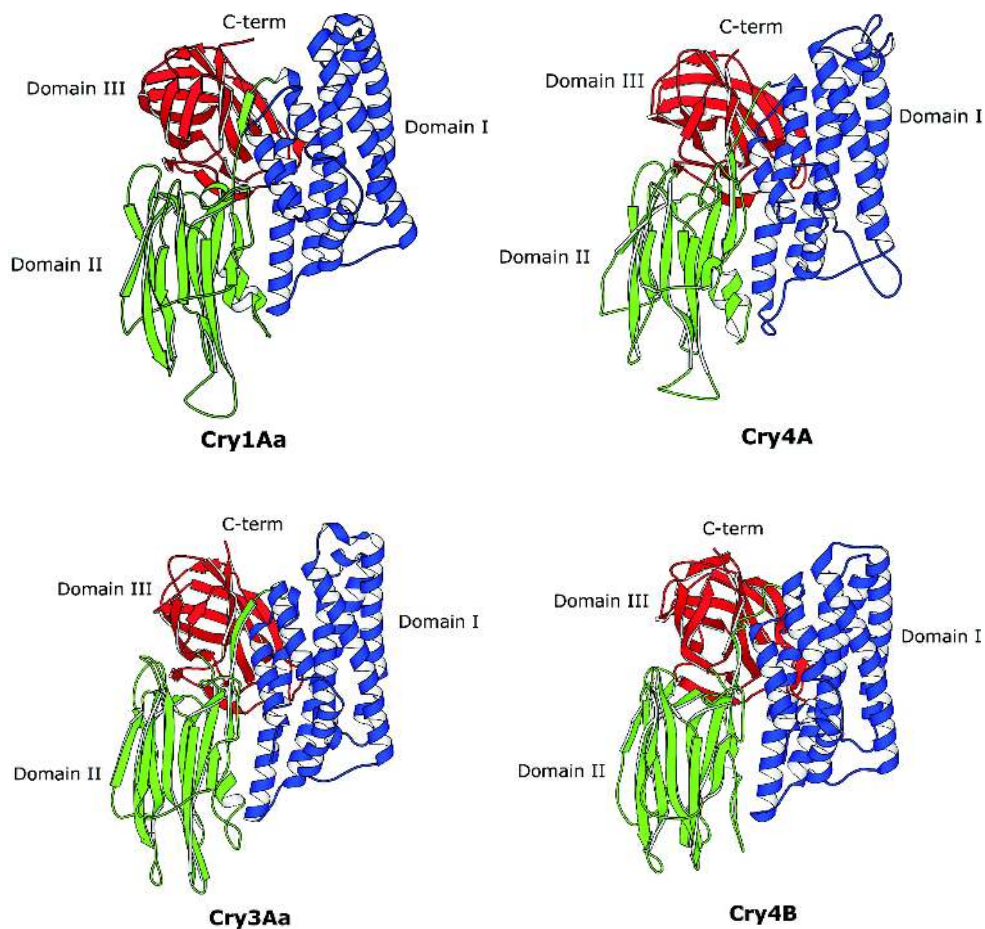
Fig. 2. Schematic ribbon representations of homology-based models for Cry4A and Cry4B in comparison with the known crystal structures of Cry1Aa and Cry3Aa, illustrating the three-domain organizations: domain I, α-helix bundle (blue); domain II, β-prism (green); domain III, β-sandwich (red).
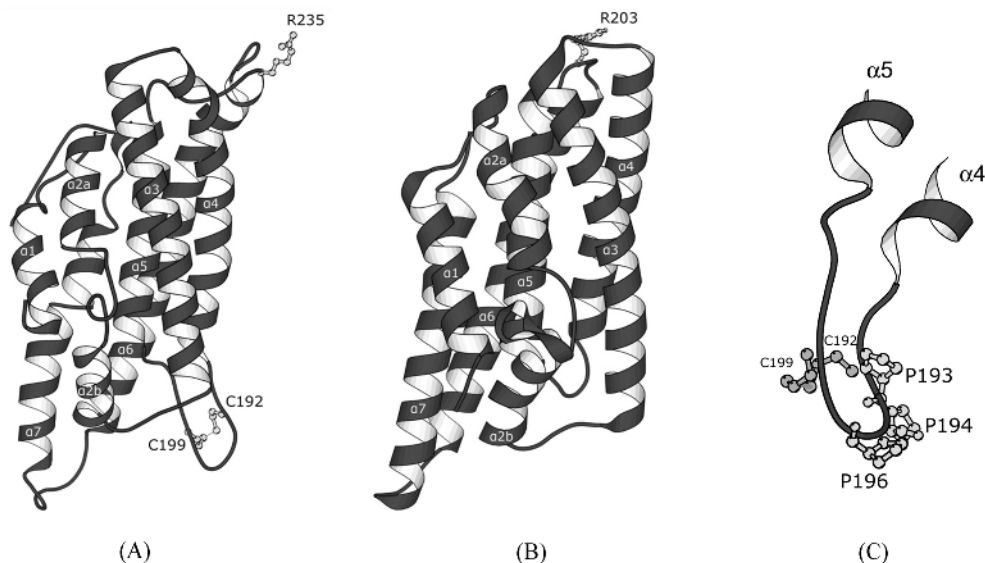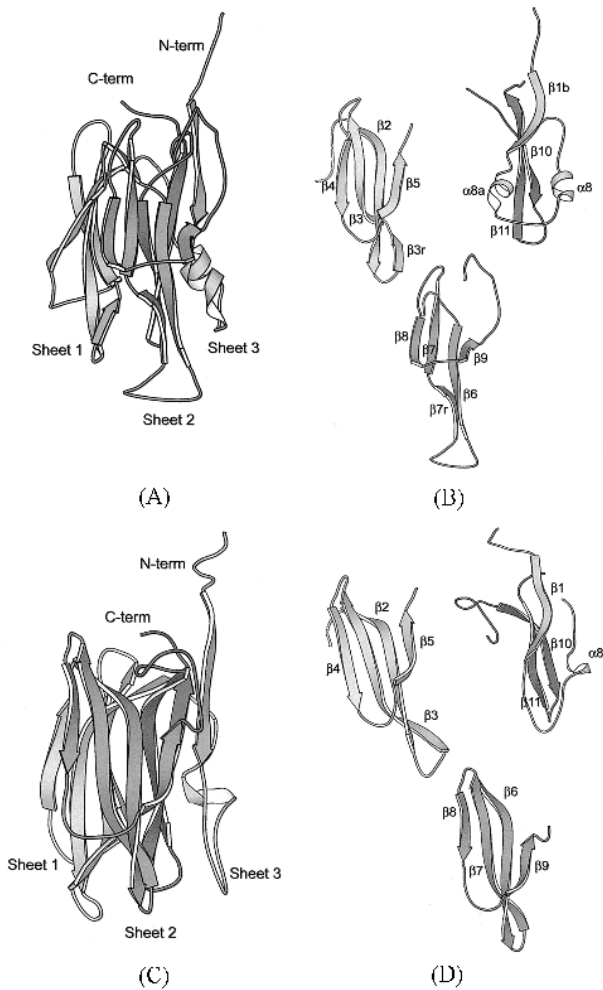
Fig. 3. Ribbon representations of the seven-helix bundle of domain I of the Cry4A and Cry4B models. Balls and sticks illustrate Arg-235 and Arg-203 in the interhelical 5/6 loop of Cry4A and Cry4B, respectively, and the disulphide bond between Cys-192 and Cys-199 in the α4/α5 loop of Cry4A. Helical structure assignments are: (A) Cry4A-domain I; α1 = L70-L81, α2a = G91-L100, α2b = N112-I125, α3 = S131-N160, α4 = P164-L188, α5 = V206-F227, α6 = L243-I275, α7 = N286-Y313; (B) Cry4B-domain I; α1 = S43-F58, α2a = A63-I76, α2b = R87-I100, α3 = A106-R135, α4 = Q140-F164, α5 = L174-N195, α6 = A205-L235, α7 = N240-S266. (C) Ribbon diagram of the loop linking helices 4 and 5 of Cry4A that illustrate balls and sticks for proline residues (P193, P194, and P196) and the disulphide bond between C192 and C199.

corresponding to the regions where polypeptide loops connecting the secondary structural elements were inserted or deleted, then identical or topologically equivalent (*i.e.* very similar shape, electronic, hydrogen bonding, and hydrophobic properties) amino acids were found at many positions in the aligned sequences (see Fig. 1). In addition, the alignment result revealed a clear trend of the previously described structurally-conserved regions that mainly correspond to the secondary structures (Höfte and Whiteley, 1989; Li *et al.*, 1991) for the two Cry4 sequences. This suggests that the 65-kDa activated forms of both Cry4A and Cry4B share a common 3D fold with the three known crystal structures. It also revealed that Cry4A has a comparatively higher sequence identity to Cry1Aa than to Cry2Aa and Cry3Aa, while Cry4B showed a higher sequence identity to Cry3Aa than to the two others. The 3D models for the two activated Cry4 toxins were thus generated using the corresponding known structures with the higher sequence similarity as templates, although their amino acid sequence homology (ca. 25%) was below the threshold for confidently identifying the overall structural homology (Sander and Schneider, 1991). Cry1Aa and Cry3Aa have little sequence identity to Cry2Aa (ca. 20% and ca. 17%, respectively), but the three-dimensional structures of these toxins are all similar in their overall topology (Morse *et al.*, 2001).

After the initial refinement *via* energy minimization of the homology-based 3D models for the activated Cry4A and Cry4B toxins, an examination to determine their structural flaws were performed automatically. The packing environment of the model structures, as analyzed by using the CHECK option, displayed Z-scores of −1.20 and −1.34 for Cry4A and Cry4B, respectively. These values were in an acceptable range of a correctly folded structure (Vriend, 1990). The stereochemical quality, evaluated by the PROCHECK program, showed overall G-factor values of −0.70 and −0.58, respectively, for the Cry4A and Cry4B model structures, which is equivalent to the expected value of a crystal structure resolved at 3.0 Å (Laskowski *et al.*, 1993). Additionally, the resulting Ramachandran plots of both Cry4 models revealed that over 95% of the non-glycine and non-proline residues have the φ/ψ backbone-dihedral angles in the energetically favorable and allowed regions (plots not shown). This indicates that the derived Cry4 structures that possess most of the sterically favorable main-chain conformations or stereochemistry of the model polypeptide backbones are correlated with their side-chain conformations.
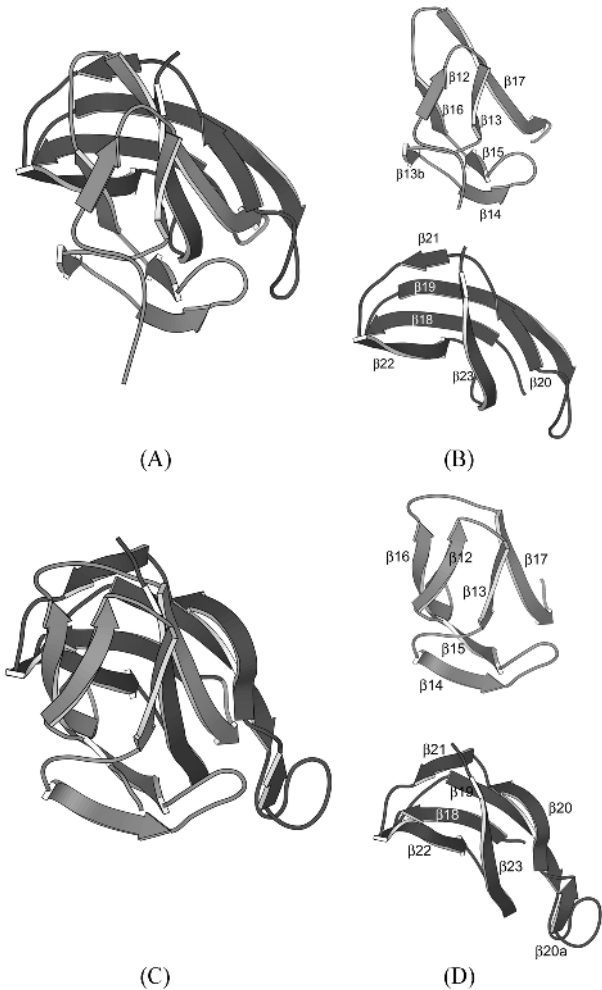
CD spectroscopy was employed in an attempt to elucidate the extent of secondary structures of the 65-kDa FPLC-purified Cry4A and Cry4B toxins when dissolved in a 10 mM sodium phosphate buffer, pH 9.0. The activated Cry4A protein exhibited secondary structural contents of 40% helix, 19% β-strand, and 42% turn and random coil to a total of 1.01 with RMSE of 0.370. The contents of the 65-kDa activated Cry4B protein were 33% helix, 22% β-strand, and 49% turn and random coil to a total of 1.04 with RMSE of 0.125. These values are comparable to those estimated from the derived models by the Swiss PDB software (35% helix and 23% β-strand for Cry4A; 31% helix and 20% β-strand for Cry4B).

Fig. 4. Ribbon representations of the three-sheet assembly and individual β-sheets of domain II, as viewed from the exterior of the domain of the Cry4A (A and B) and Cry4B (C and D) models. Residue numbers in the secondary structural elements are: Cry4A-domain II; β1 = I320-L326, α8a = L334-L337, α8b = F346-E350, β2 = F360-A372, β3r = P376-F379, β3 = T381-F387, β4 = S395-S399, β5 = K411-N418, β6 = L423-N431, β7r = D444-T448, β7 = T451-K456, β8 = Q464-D468, β9 = N472-G475, β10 = F503-S508, β11 = K514-T523; Cry4B-domain II; β1 = L282-L293, α8 = I304-A310, β2 = F318-T320, β3 = A340-T348, β4 = M353-G357, β5 = L368-N375, β6 = K380-S388, β7 = K395-D402, β8 = S407-T413, β9 = T421-G425, β10 = I444-V452, β11 = N458-T465.

Fig. 5. Ribbon representations of the intact β-sandwich and separate sheets of domain III of the Cry4A (A and B) and Cry4B (C and D) models. Residue numbers in the β-strands are: Cry4A-domain III; β12 = N531-Y534, β13 = T538-P542, β13b = N547-G550, β14 = K554-Q557, β15 = L566-K570, β16 = I575-F582, β17 = Y586-S594, β18 = L606-A612, β19 = L618-S623, β20 = F635-E639, β21 = K645-P648, β22 = I652-S659, β23 = V667-L675; Cry4B-domain II; β12 = N473-T477, β13 = I480-P484, β14 = F490-K496, β5 = D507-T512, β16 = L517-Q524, β17 = T534-A545, β18 = V553-G559, β19 = I566-F572, β20a = L583-Y585, β20 = F588-P593, β21 = P599-L602, β22 = Q606-I612, β23 = N621-I633.

This supports the validity of both homology-based Cry4 structures.

Description of the model structures and larvicidal implications The 3D structures of both derived Cry4 models revealed a common folding pattern that was similar to those of the known crystal structures that were also built from three structurally distinct domains and the corresponding domains that possessed the same topological folds (Fig. 2). (I) The N-terminal domain (Cry4A, residues Glu-69 to Lys-317; Cry4B, residues Ile-40 to Arg-271) formed a seven anti-

parallel helix bundle in which the central helix 5 was entirely encircled by six outer helices (Fig. 3). (II) The middle domain of Cry4A (residues Tyr-318 to Val-527) and Cry4B (Tyr-272 to Val-469) was a three-fold symmetric assembly that primarily consisted of β-sheets with sheets 1 and 2 being anti-parallel with a Greek key-like topology (Fig. 4). (III) The C-terminal domain (Cry4A, residues Asp-528 to Thr-678; Cry4B, residues Asp-470 to Thr-634) consisted of two twisted, anti-parallel β-sheets that formed a face-to-face sandwich or a jelly-roll topology (Fig. 5).

Like all the known Cry structures, the core and domain

(A)

(B)
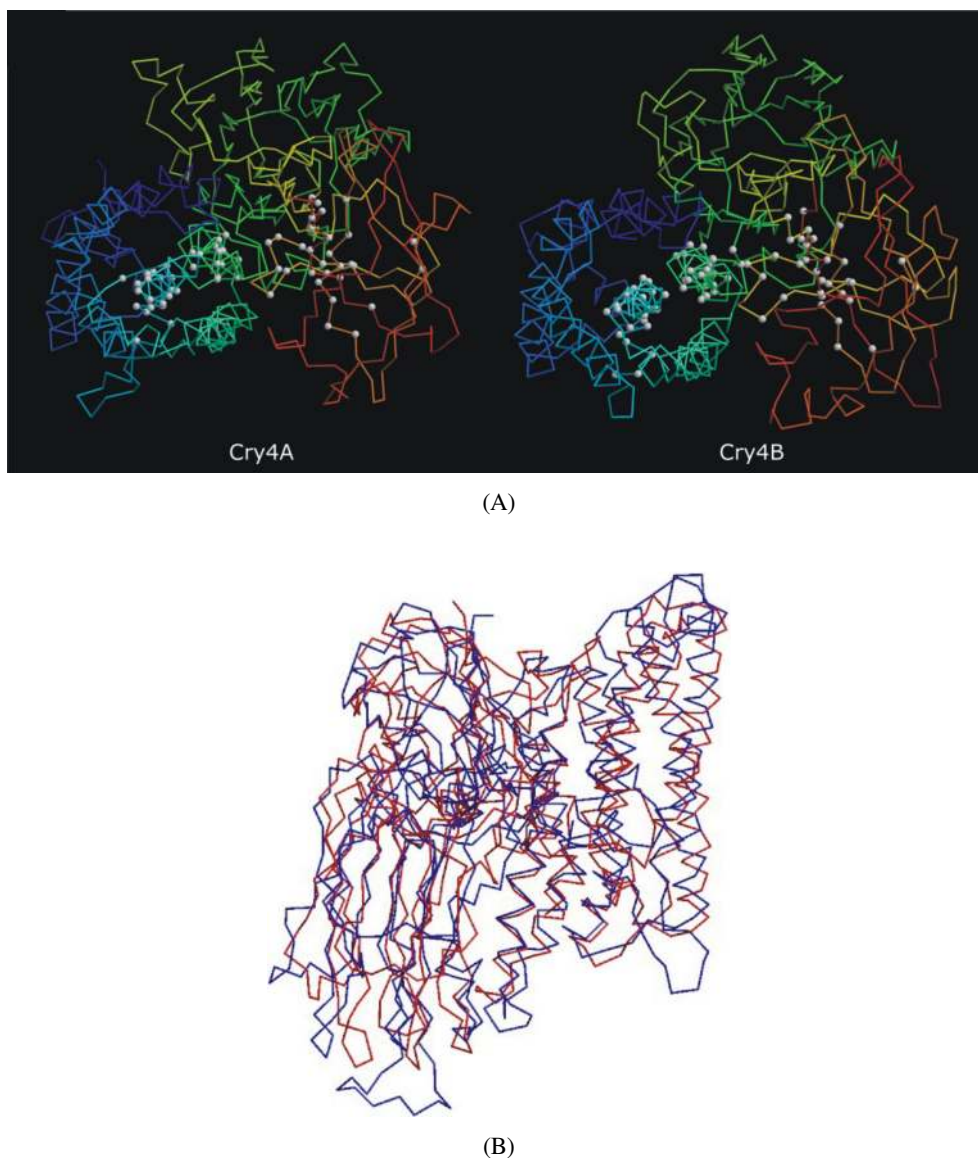
Fig. 6. (A) Cα traces of the Cry4A and Cry4B models with the conserved sequence blocks indicated by white beads at their Cα, illustrating the three-domain organizations as viewed down the central helix of the bundle from the bulky end of the molecules. The polypeptide chains are colored in the rainbow order, from red at the N-terminus to blue at the C-terminus. (B) Superposition of the Cα traces of all three domains of the Cry4A (blue) and Cry4B (red) models, illustrating the level of structural similarity.

interfaces of both model molecules are built from sequence blocks (as indicated by white beads in Fig. 6a) that are a highly conserved feature of all Cry toxins (Höfte and Whiteley, 1989; Crickmore *et al.*, 1998). In addition, the superimposed C$_\alpha$ traces of the two model molecules (all three domains) showed a 0.59 Å RMS deviation for about 550 equivalent C$_\alpha$ atoms out of 600 C$_\alpha$ atoms, indicating a high level of similarity in their structural folds (Fig. 6b). With the exception of the loop linking α4 and α5 (see below), the N-terminal domain of both models displays the highest similarity, followed by the C-terminal domain. The main divergences in both computer-derived molecules are located particularly in the surface-exposed loops in domain II (see below). They may be an important determinant of receptor recognition and hence target insect specificity (Rajamohan *et al.*, 1996; Smedley and Ellar, 1996; Ballester *et al.*, 1999; Jurat-Fuentes and Adang, 2001). We previously showed that both Cry4A and Cry4B toxins exhibit comparable activity against *Aedes aegypti* and *Anopheles gambiae* mosquito-larvae, but the Cry4A toxin is 25-fold more toxic to *Culex quinquefasciatus* larvae (Angsuthanasombat *et al.,* 1992). Structural differences between the domain II of both Cry4 models, which are essentially observed at the molecular apex of the structures where loops protruding from each sheet differ in length and conformation (Fig. 4), could therefore reflect their different activity spectra.
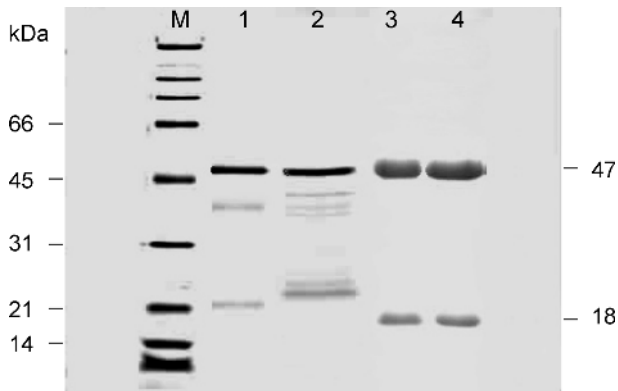
The domain I topology of both derived Cry4 structures is clearly equipped for membrane insertion and pore formation (Fig. 3), as was observed for the known structures of Cry1Aa, Cry2Aa, and Cry3Aa (Li et al., 1991; Grochulski et al., 1995; Morse et al., 2001). The role of domain I has been explored by a number of studies which suggest that the toxin-induced pore/channel is initiated by the insertion of the α4-α5 hairpin into the lipid membranes with subsequent association of this transmembrane hairpins from several other molecules to form an oligomeric helical bundle pore with a 10-20 Å diameter (Schwartz et al., 1997; Gazit et al., 1998; Masson et al., 1999; Gerber and Shai, 2000). However, the fact that the orientation of the hydrophobic faces of the amphipathic helices of the pore-forming domain is the reverse of that needed to form an aqueous pore/channel in the membranes. This indicates that perhaps substantial conformational changes may be occurring in the membrane-insertion and pore-formation stages. One interesting possibility for making this conformational change is the proteolysis of one or more of the solvent-exposed loops connecting the helices in the pore-forming bundle. This would possibly impart greater flexibility to the toxin molecule and might allow protrusion or release of the helical hairpin, i.e. α4-loop-α5 from the bundle to act as an initiator of membrane penetration. Proteolysis in the interhelical loops upon exposure to proteases in vitro has been detected in the α3-α4 loop of Cry2Aa (Nicholls et al., 1989) and Cry3Aa (Carroll et al., 1989), and in the α5-α6 loop of Cry4A and Cry4B (see below). This implies that this interhelical cleavage may be indeed essential to reorientate all or part of the helix bundle as a necessary step in the mechanism of action of the Bt Cry δ-endotoxins. We previously showed that both the 130-kDa Cry4A and Cry4B toxins were proteolytically processed to stable fragments of 46-48 kDa and 18-20 kDa upon trypsin activation, in addition to the removal of the C-terminal half of the protoxins (Angsuthanasombat et al., 1992; 1993). These two trypsin-resistant fragments were produced by the cleavage

at Arg-235, and Arg-203 located in the exposed loop linking helices 5 and 6 within the pore-forming domain of Cry4A and Cry4B, respectively (Angsuthanasombat et al., 1992; 1993). As expected, each arginine residue in the α5-α6 loop of both Cry4 models is essentially exposed to the solvent (Figs. 3a and b). This could allow the access of trypsin or larval gut trypsin-like proteases to attack.

Interestingly, the unique disulphide bond, which is formed between Cys-192 and Cys-199 within the long loop connecting helices 4 and 5, has also been observed in the Cry4A model (Fig. 3a). The presence of this putative disulphide bridge within the α1-α5 fragment of the trypsin-treated Cry4A product was verified by assaying the mobility of the 20-kDa fragments during reducing and non-reducing conditions on SDS-PAGE. This assay would take advantage of the fact that the electrophoretic mobility of the 20-kDa α1-α5 fragments is faster in the presence of a loop-disulphide bond. Figure 7 shows that the α1-α5 fragments of Cry4A, but not those of Cry4B, ran differently in the absence and presence of β-mercaptoethanol. This confirms that the disulphide bridge is present in the Cry4A α1-α5 fragment, conceivably being localized within the long loop of the helical hairpin α4-α5. Figure 3c shows that this long loop in the Cry4A toxin also contains an irregular proline-rich region (-P_{193}PNP-) which has been demonstrated to stabilize the conformation of the proline-rich loop in the icosahedral repeat unit within the human hepatitis B virus capsid (Wynne et al., 1999). We therefore believe that this proline-rich region, perhaps together with the disulphide bridge, is involved in forming a potential conformation of this particular long loop for membrane insertion of the proposed transmembrane helical hairpin. This supports the notion that the loop in the α4-α5 hairpin of the Cry toxins is required for efficient insertion into the lipid membranes (Gerber and Shai, 2000). Also, this critical loop is comprised of one highly conserved tyrosine residue (Cry4A: Tyr-202; Cry4B: Tyr-170; Fig. 1) whose aromatic structure plays a crucial role in mosquito-larvicidal activity, conceivably being involved in an interaction with the phospholipid head groups for stabilizing the oligomeric pore structure (Kanintronkul et al., 2003; Pornwiroon et al., 2004).

References

Angsuthanasombat, C., Chungjatupornchai, W., Kertbundit, S., Luxananil, P., Settasatian, C., Wilairat, P. and Panyim, S.

(1987) Cloning and expression of 130-kDa mosquito-larvicidal δ-endotoxin gene of *Bacillus thuringiensis* var *israelensis* in *Escherichia coli*. *Mol. Gen. Genet.* **208**, 384-389.

Angsuthanasombat, C., Crickmore, N. and Ellar, D. J. (1992) Comparison of *Bacillus thuringiensis* subsp. *israelensis* CryIVA and CryIVB cloned toxins reveals synergism *in vivo*. *FEMS Microbiol. Lett.* **94**, 63-68.

Angsuthanasombat, C., Crickmore, N. and Ellar, D. J. (1993) Effects on toxicity of eliminating a cleavage site in a predicted interhelical loop in *Bacillus thuringiensis* CryIVB delta-endotoxin. *FEMS Microbiol. Lett.* **111**, 255-261.

Aronson, A. I., Beckman, W. and Dunn, P. (1986) *Bacillus thuringiensis* and related insect pathogens. *Microbiol. Rev.* **50**, 1-24.

Ballester, V., Granero, F., Tabashnik, B. E., Malvar, T. and Ferre, J. (1999) Integrative model for binding of Bacillus thuringiensis toxins in susceptible and resistant larvae of the diamondback moth (*Plutella xylostella*). *Appl. Environ. Microbiol.* **65**, 2503-2507.

Becker, N. and Margalit, J. (1993) Use of *Bacillus thuringiensis* subsp. *israelensis* against mosquitoes and black flies; in *Bacillus thuringiensis*, an Environmental Biopesticide: Theory and Practice, Entwistle, P. F., Cory, J. S., Bailey, M. J. and Higgs, S. (eds.), pp. 147-170, John Wiley and Sons, Chichester, UK.

Burton, S. L., Ellar, D. J., Li, J. and Derbyshire, D. J. (1999) N-acetylgalactosamine on the putative insect receptor aminopeptidase N is recognized by a site on the domain III lectin-like fold of a *Bacillus thuringiensis* insecticidal toxin. *J. Mol. Biol.* **287**, 1101-1122.

Carroll, J., Li, J. and Ellar, D. J. (1989) Proteolytic processing of a coleopteran-specific delta-endotoxin produced by *Bacillus thuringiensis var. tenebrionis*. *Biochem. J.* **261**, 99-105.

Chungjatupornchai, W., Hofte. H., Seurinck, J., Angsuthanasombat, C. and Vaeck, M. (1988) Common features of *Bacillus thuringiensis* toxins specific for Diptera and Lepidoptera. *Eur. J. Biochem.* **173**, 9-16.

Crickmore, N., Zeigler, D. R., Feitelson, J., Schnepf, E., Van Rie, J., Lereclus, D., Baum, J. and Dean, D. H. (1998) Revision of the nomenclature for the *Bacillus thuringiensis* pesticidal crystal proteins. *Microbiol. Mol. Biol. Rev.* **62**, 807-813.

Derbyshire, D. J., Ellar, D. J. and Li, J. (2001) Crystallization of the *Bacillus thuringiensis* toxin Cry1Ac and its complex with the receptor ligand N-acetyl-D-galactosamine. *Acta Crystallogr. D. Biol. Crystallogr.* **57**, 1938-1944.

de Maagd, R. A., Bravo, A. and Crickmore, N. (2001) How *Bacillus thuringiensis* has evolved specific toxins to colonize the insect world. *Trends Genet.* **4**, 193-199.

Galitsky, N., Cody, V., Wojtczak, A., Ghosh, D., Luft, J. R., Pangborn, W. and English, L. (2001) Structure of the insecticidal bacterial delta-endotoxin Cry3Bb1 of *Bacillus thuringiensis*. *Acta Crystallogr. D. Biol. Crystallogr.* **57**, 1101-1109.

Gazit, E., La Rocca, P., Sansom, M. S. and Shai, Y. (1998) The structure and organization within the membrane of the helices composing the pore-forming domain of *Bacillus thuringiensis* delta-endotoxin are consistent with an "umbrella-like" structure of the pore. *Proc. Natl. Acad. Sci. USA* **95**, 12289-12294.

Gerber, D. and Shai, Y. (2000) Insertion and organization within membranes of the delta-endotoxin pore-forming domain, helix 4-loop-helix 5, and inhibition of its activity by a mutant helix 4 peptide. *J. Biol. Chem.* **275**, 23602-23607.

Gray, W. R. (1997) Disulphide bonds between cysteine residues; in Protein Structure (A Practical Approach). 2nd ed., Creighton, T. E. (ed.), pp. 165-186, Oxford University Press Inc, New York, USA.

Grochulski, P., Masson, L., Borisova, S., Pusztai-Carey, M., Schwartz, J.-L., Brousseau, R. and Cygler, M. (1995) *Bacillus thuringiensis* CryIA(a) insecticidal toxin: crystal structure and channel formation. *J. Mol. Biol.* **254**, 447-464.

Höfte, H. and Whiteley, H. R. (1989) Insecticidal crystal proteins of *Bacillus thuringiensis*. M*icrobiol. Rev.* **53**, 242-255.

Jurat-Fuentes, J.-L. and Adang, M. J. (2001) Importance of Cry1 delta-endotoxin domain II loops for binding specificity in *Heliothis virescens* (L.). *Appl. Environ. Microbiol.* **67**, 323-329.

Kanintronkul, Y., Sramala, I., Katzenmeier, G., Panyim, S. and Angsuthanasombat, C. (2003) Specific mutations within the α4-α5 loop of the *Bacillus thuringiensis* Cry4B toxin reveal a crucial role of Asn-166 and Tyr-170. *Mol. Biotechnol.* **24**, 11-19.

Knowles, B. H. (1994) Mechanism of action of *Bacillus thuringiensis* insecticidal δ-endotoxins. *Adv. Insect. Physiol.* **24**, 275-308.

Knowles, B. H. and Ellar, D. J. (1987) Colloid-osmotic lysis is a general feature of the mechanism of action of *Bacillus thuringiensis* delta endotoxin with difference insect specificity. *Biochim. Biophys. Acta* **924**, 509-518.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283-291.

Li, J., Carroll, J. and Ellar, D. J. (1991) Crystal structure of insecticidal delta-endotoxin from *Bacillus thuringiensis* at 2.5 Å resolution. *Nature* **353**, 815-821.

Lee, M. K., Young, B. A. and Dean, D. H. (1995) Domain III exchanges of *Bacillus thuringiensis* CryIA toxins affect binding to different gypsy moth midgut receptors. *Biochem. Biophys. Res. Commun.* **216**, 306-312.

Manavalan, P. and Johnson, W. C. (1987) Variable selection method improves the prediction of protein secondary structure from circular dichroism spectra. *Anal. Biochem.* **167**, 76-85.

Masson, L., Tabashnik, B. E., Liu, Y. B., Brousseau, R. and Schwartz, J.-L. (1999) Helix 4 of the *Bacillus thuringiensis* Cry1Aa toxin lines the lumen of the ion channel. *J. Biol. Chem.* **274**, 31996-32000.

Masson, L., Tabashnik, B. E., Mazza, A., Prefontaine, G., Potvin, L., Brousseau, R. and Schwartz, J.-L. (2002) Mutagenic analysis of a conserved region of domain III in the Cry1Ac toxin of *Bacillus thuringiensis*. *Appl. Environ. Microbiol.* **68**, 194-200.

Morse, R. J., Yamamoto, T. and Stroud, R. M. (2001) Structure of Cry2Aa suggests an unexpected receptor binding epitope. *Structure (Camb.)* **9**, 409-417.

Nicholls, C. N., Ahmad, W. and Ellar, D. J. (1989) Evidence for two different types of insecticidal P2 toxins with dual specificity in *Bacillus thuringiensis* subspecies. *J. Bacteriol.* **171**, 5141-5147.

Nunez-Valdez, M.-E., Sanchez, J., Lina, L., Guereca, L. and Bravo, A. (2001) Structural and functional studies of α-helix 5 region from *Bacillus thuringiensis* Cry1Ab δ-endotoxins.

Biochim. Biophys. Acta **1546**, 122-133.

Ponder, J. W. and Richards, F. M. (1987) An efficient Newton-like method for molecular mechanics energy minimization of large molecules. *J. Comput. Chem.* **8**, 1016-24.

Pornwiroon, W., Katzenmeier, G., Panyim, S. and Angsuthanasombat, C. (2004) Aromaticity of Tyr-202 in the α4-α5 loop is essential for toxicity of the *Bacillus thuringiensis* Cry4A toxin. *J. Biochem. Mol. Biol.* **37**, 292-297.

Puntheeranurak, T., Leetachewa, S., Katzenmeier, G., Krittanai, C., Panyim, S. and Angsuthanasombat, C. (2001) Expression and biochemical characterization of the *Bacillus thuringiensis* Cry4B α1-α5 pore-forming fragment. *J. Biochem. Mol. Biol.* **34**, 293-298.

Puntheranurak, T., Uawithya, P., Potvin, L., Angsuthanasombat, C., and Schwartz, J.-L. (2004) Ion channels formed in planar lipid bilayers by the dipteran-specific Cry4B *Bacillus thuringiensis* toxin and its α1-α5 fragment. *Mol. Membr. Biol.* **21**, 67-74.

Rajamohan, F., Hussain, S. R., Cotrill, J. A., Gould, F. and Dean, D. H. (1996) Mutations at domain II, loop 3, of *Bacillus thuringiensis* CryIAa and CryIAb delta-endotoxins suggest loop 3 is involved in initial binding to lepidopteran midguts. *J. Biol. Chem.* **271**, 25220-25226.

Sander, C. and Schneider, R. (1991) Database of homology-derived protein structures and the structural meaning of sequence alignment. *Proteins* **9**, 56-68.

Schnepf, E., Crickmore, N., Van Rie, J., Lereclus, D., Baum, J., Feitelson, J., Zeigler, D. R. and Dean, D. H. (1998) *Bacillus thuringiensis* and its pesticidal crystal proteins. *Microbiol. Mol. Biol. Rev.* **62**, 775-806.

Schwartz, J.-L., Juteau, M., Grochulski, P., Cygler, M., Prefontaine, G., Brousseau, R. and Masson, L. (1997) Restriction of intramolecular movements within the Cry1Aa toxin molecule of *Bacillus thuringiensis* through disulfide bond engineering. *FEBS Lett.* **410**, 397-402.

Smedley, D. P. and Ellar, D. J. (1996) Mutagenesis of three surface-exposed loops of a *Bacillus thuringiensis* insecticidal toxin reveals residues important for toxicity, receptor recognition and possibly membrane insertion. *Microbiology* **142**, 1617-1624.

Sramala, I., Uawithya, P., Chanama, U., Leetachewa, S., Krittanai, C., Katzenmeier, G., Panyim, S. and Angsuthanasombat, C. (2000) Single proline substitutions of selected helices of the *Bacillus thuringiensis* Cry4B toxin affect inclusion solubility and larvicidal activity. *J. Biochem. Mol. Biol. Biophys.* **4**, 187-193.

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. and Higgins, D. G. (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **24**, 4876-82.

Uawithya, P., Tuntitippawan, T., Katzenmeier, G., Panyim, S. and Angsuthanasombat, C. (1998) Effects on larvicidal activity of single proline substitutions in alpha 3 or alpha 4 of the *Bacillus thuringiensis* Cry4B toxin. *Biochem. Mol. Biol. Int.* **44**, 825-832.

Von Tersch, M. A., Slatin, S. L., Kulesza, C. A. and English, L. H. (1994) Membrane-permeabilizing activities of *Bacillus thuringiensis* coleopteran-active toxin CryIIIB2 and CryIIIB2 domain I peptide. *Appl. Environ. Microbiol.* **60**, 3711-3717.

Vriend, G. (1990) WHAT IF: a molecular modeling and drug design program. *J. Mol. Graph.* **8**, 52-56.

Walters, F. S., Slatin, S. L., Kulesza, C. A. and English, L. H. (1993) Ion channel activity of N-terminal fragments from CryIA(c) delta-endotoxin. *Biochem. Biophys. Res. Commun.* **196**, 921-926.

Ward, S. and Ellar, D. J. (1998) Cloning and expression of two homologous genes of *Bacillus thuringiensis* subsp. *israelensis* which encode 130-kilodalton mosquitocidal proteins. *J. Bacteriol.* **170**, 727-735.

Wynne, S. A., Crowther, R. A. and Leslie, A. G. W. (1999) The crystal structure of the human hepatitis B virus capsid. *Mol. Cell* **3**, 771-780.