J Mol Model (2008) 14:843–848
DOI 10.1007/s00894-008-0318-8

ORIGINAL PAPER

The theoretical 3D structure of *Bacillus thuringiensis* Cry5Ba

Li-Qiu Xia · Xin-Min Zhao · Xue-Zhi Ding · Fa-Xiang Wang · Yun-Jun Sun

Received: 5 December 2007 / Accepted: 16 April 2008 / Published online: 27 May 2008

Abstract Cry5Ba is a δ-endotoxin produced by *Bacillus thuringiensis* PS86A1 NRRL B-18900. It is active against nematodes and has great potential for nematode control. Here, we predict the first theoretical model of the three-dimensional (3D) structure of a Cry5Ba toxin by homology modeling on the structure of the Cry1Aa toxin, which is specific to Lepidopteran insects. Cry5Ba resembles the previously reported Cry1Aa toxin structure in that they share a common 3D structure with three domains, but there are some distinctions, with the main differences being located in the loops of domain I. Cry5Ba exhibits a changeable extending conformation structure, and this special structure may also be involved in pore-forming and specificity determination. A fuller understanding of the 3D structure will be helpful in the design of mutagenesis experiments aimed at improving toxicity, and lead to a deep understanding of the mechanism of action of nematicidal toxins.

Keywords Three-dimensional structure · Homology modeling · Cry5Ba · *Bacillus thuringiensis*

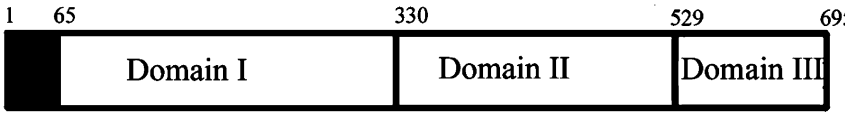
L.-Q. Xia (✉) · X.-M. Zhao · X.-Z. Ding · F.-X. Wang · Y.-J. Sun
Key Laboratory for Microbial Molecular Biology of Hunan Province, College of Life Science, Hunan Normal University, Changsha 410081, China
e-mail: xialq@hunnu.edu.cn

X.-M. Zhao
Department of Chemistry and Environmental Engineering, Hunan City University, Yiyang 413000, China

Introduction

In 1911, the German scientist Ernst Berliner isolated a bacteria that had killed a Mediterranean flour moth. He named it *Bacillus thuringiensis*, after the German town Thuringia where the moth was found. Cry toxins—sometimes referred to as insecticidal crystal proteins (ICP)—produced by the soil bacterium *B. thuringiensis* (*Bt*) are selectively toxic to different species from several invertebrate phyla: arthropods (mainly insects), nematodes, flatworms and protozoa [1]. The mode of action of Cry toxins is still a matter of investigation; generally, following ingestion by insects, they are activated by gut proteases and by binding to specific receptors on midgut epithelial cells [2]. Receptor binding induces the conformational change in the toxin necessary for membrane insertion, where it forms ion selective channels via oligomerization of toxin monomers; insects die from colloid osmotic lysis [3, 4].

Crystal structures of Cry toxins have been elucidated for the Coleoptera-specific Cry3Aa [5] and Cry3Bb1[6], Lepidoptera-specific Cry1Aa [7] and Cry1Ac [8], Lepidoptera/Diptera-specific Cry2Aa [9], and Diptera-specific Cry4Ba [10] and Cry4Aa [11] toxins. Pablo Gutierrez and co-workers predicted the structure of Cry11Bb by homology modeling on the structures of Cry1Aa and Cry3Aa [12]. The three dimensional (3D) structures of these toxins are remarkably similar in spite of their different insect specificities, in that they are all composed of three structurally conserved domains. The seven α-helices that form the N-terminal domain I have been implicated in pore formation [13]. Domain II consists of three antiparallel β-sheets with exposed loop regions, which vary significantly in length and amino acid sequence in different toxins. These loops are therefore thought to participate in receptor binding and hence in determining the specificity of the toxin for insect larvae




```
Cry5Ba    65                                                        GKLDYF
Cry1Aa    33                                                        YTPIDI
                                                                    .
Cry5Ba    71    ALTKASISLI GFIPGAEAAV PFINMFVDFV WPKLFGANTE GKDQQLFNAI
Cry1Aa    39    SLSLTQFLLS EFVPGAGFVL GLVDIIWGIF GPSQ------ ------WDAF
                .*. . . *   *.***   .  ..... .   *              .*.
Cry5Ba   121    MDAVNKMVDN KFLSYNLSTL NKTIEGLQGN LGLFQNAIQV AICQGSTPER
Cry1Aa    77    LVQIEQLINQ RIEEFARNQA ISRLEGLSNL YQIYAESFRE WE--------
                .   .. ... ..   .  .     .***     .. ....
Cry5Ba   171    VNFDQNCTPC NPNQPCKDDL DRVASRFDTA NSQFTQHLPE FKNPWSDENS
Cry1Aa   119    ---------A DPTNPALREE MRI--QFNDM NSALTTAIPL LAVQNYQVP-
                         .*  *   .   *  .*.  ** .*  .*. .     .
Cry5Ba   221    TQEFKRTSVE LTLPMYTTVA TLHLLLYEGY IEFMTKWNFH NEQYLNNLKV
Cry1Aa   157    ---------- -LLSVYVQAA NLHLSVLRDV SVFGQRWGF- DAATINSRYN
                           *..*   *  *** .       * .* *  .   .*.
Cry5Ba   271    ELQQLIHSYS ETVRTSFLQF LPTLNNRSKS SVNAYNRYVR NMTVNCLDIA
Cry1Aa   195    DLTRLIGNYT DYAVRWYNTG LERVWGPDSR DWVRYNQFRR ELTLTVLDIV
                .* .** .*. .      .   * .           **. * ..*.  ***
Cry5Ba   321    ATWPTFDTHN YHQGGKLDLT RIILSDTAGP IEEYTTGDKT SGPEHSNITP
Cry1Aa   245    ALFSNYDSRR YPIRTVSQLT REIYTNPVLE NFDGSFRGMA QRIE-QNIRQ
                *  . .*.  *     .** * * ..      . .   ..    *  **
Cry5Ba   371    NNILDTPSPT YQHS------ -FVSVDSIVY SRKELQQLDI ATYSTNNSNN
Cry1Aa   294    PHLMDILNSI TIYTDVHRGF NYWSGHQITA SPVGFSGPEF AFPLFGNAGN
                 ...** ..    .        . *  *  *  .  .. *    * *
Cry5Ba   414    CHPYGLRLSY TDGSRYDYGD NQPDFTTSNN NYCHNSYTAP ITLVNARHLY
Cry1Aa   344    AAP------- --PVLVSLTG LGIFRTLSSP LYRRIILGSG PNNQELFVLD
                  *                      * *.   * .    .      .   *
Cry5Ba   464    NAKGSLQNVE SLVVSTVNGG SGSCICDAWI NYLRPPQTSK NESRPDQKIN
Cry1Aa   385    GTEFSFASLT TNLPSTIY-- ---RQRGTVD SLDVIPPQDN SVPPRAGFSH
                .  *. ..  . . **.          .  .    *  . . .      .
Cry5Ba   514    VLYPITETVN --KGTGGNLG VISAYVPMEL VPENVIGDVN ADTKLPLTQL
Cry1Aa   430    RLSHVTMLSQ AAGAVYTLRA PTFSWQHRSA EFNNIIPS-S QITQIPLTK-
                * .*       .    .    .         .*.*   .   * .***
Cry5Ba   562    KGFPFEKYGS EYNNRGISLV REWINGNNAV KLSNSQ---S VGIQITNQTK
Cry1Aa   478    --------ST NLGSGTSVVK GPGFTGGDIL RRTSPGQIST LRVNITAPLS
                       .. .  .    .   . * . . ...   . . . **
Cry5Ba   609    QKYEIRCRYA SKGDNNVYFN VDLSENPFRN SISFGSTESS VVGVQGENGK
Cry1Aa   520    QRYRVRIRYA STTNLQFHTS ID---GRPIN QGNFSATMSS GSNLQSGSFR
                *.* .* *** *  .     . .*      *   .* .*.**    .*. . .
Cry5Ba   659    YILKSITTVE IPAGSFYVHI T---NQGSSD LFLDRIEFVP
Cry1Aa   567    TVGFT-TPFN FSNGSSVFTL SAHVFNSGNE VYIDRIEFVP AEVT
                  . . *  . .. **   . .    .... ...*******
```

Fig. 1 Schematic representation of the three domains present in mature Cry5Ba, and amino acid sequence alignment of Cry5Ba with the template Cry1Aa. *Asterisks* Identical residues, *dots* conserved residues

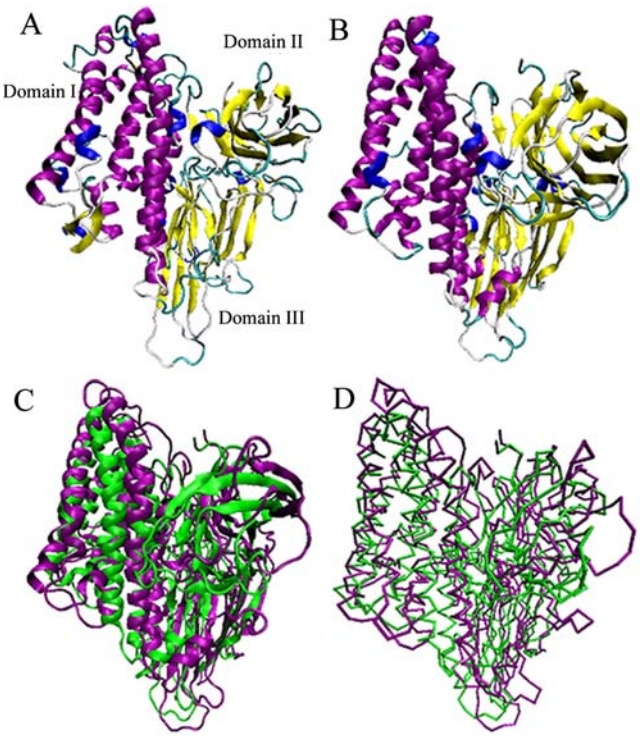


Fig. 2 Ribbon representations of Cry5Ba (**a**) and Cry1Aa (**b**), colored in structure. **c** Superimposition of the overall ribbon structures of Cry1Aa (*green*) and Cry5Ba (*purple*). **d** Superimposition of α-carbon traces of Cry1Aa (*green*) and Cry5Ba (*purple*)

[14], whereas domain III is a β-sandwich. Domains II and III are important in receptor recognition [15, 16].

Previously, the extent to which Cry toxins might also target the invertebrate phylum Nematoda has been largely ignored. Several *B. thuringiensis* strains with significant activity in inhibiting larval development of several nematode species have been identified [17, 18], and Cry toxins from *B. thuringiensis israelensis* were lethal to eggs of the nematode*Trichostrongylus colubriformis* in vitro [19]. The nematodes *Caenorhabditis elegans* and *Pristionchus pacificus* were found to be very susceptible to Cry5B [20]. Recently, purified Cry5B was found to be highly toxic in vitro and in vivo to early stage larvae of the hookworm parasite *Ancylostoma ceylanicum*, a blood-feeding gastrointestinal nematode for which humans are permissive hosts [21]. Cry5B thus warrants further clinical development for human and veterinary use.

Nematicidal activity has been found in families Cry1, Cry5, Cry6, Cry12, Cry13, Cry14 and Cry21. However, in comparison with insecticidal Cry toxins, the structure and mode of action of nematicidal Cry toxins is not fully understood. In addition, almost all the nematicidal Cry toxins registered in GenBank are protected by related patents [22, 23]. Only sparse data have been presented [24].

Here we report a model for the structure of the Cry5Ba δ-endotoxin based on a hypotheses of structural similarity with Cry1Aa toxin. A more complete understanding of the 3D structure of nematicidal Cry5Ba will be important in addressing the question of how Cry toxins target nematodes. Such insights will lead to a better understanding of the basis of specificity and the practical application of improved toxins in agriculture.

Methods

Homology modelling was performed as described previously [12]. An alignment of the amino acid sequences of Cry1Aa (PDB entries 1CIY) and Cry5Ba was produced with the ClustalW program (http://www.ebi.ac.uk/c-lustalw/#), and then corrected manually with the structural alignment tool of the program Swiss-PdbViewer until a satisfactory placement of conserved blocks and amino acid identities was obtained [25]. Cry5Ba contains four of the five protein motifs conserved among the main family of Cry toxins [1]. This alignment project file was submitted to Swiss-Model via the Expasy server (http://www.expasy.ch/spdbv/) and a preliminary model for Cry5Ba was retrieved. The model was validated with Procheck [26] by submitting the coordinates to the EMBL server (http://www.ebi.ac.uk). Sequence identities were calculated with ClustalW. The illustrations shown in the figures, and electrostatic potential calculations were generated with VMD [27] and the Pymol program [28]. The final model was submitted to the PMDB database (http://www.caspur.it/PMDB/); the PMDB identifier is PM0075036.

Results and discussion

Overall architecture

The sequence identity of Cry5Ba and Cry1Aa is 21.1%, However, Cry5Ba has four of the five blocks of amino

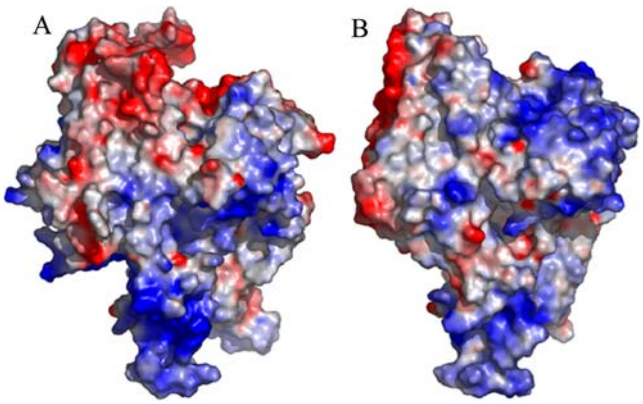


Fig. 3 Surface representations of the electrostatic potential of Cry5Ba (**a**) and Cry1Aa (**b**). *Blue* Positive electrostatic potential, *red* negative electrostatic potential

acids conserved among most Cry toxins. It is possible and reasonable to build a theoretical model by manual alignment. The final model comprises 631 amino acid residues spanning amino acids 65 to 695 (Fig. 1).

A Ramachandran plot (data not shown) indicated that most (95%) residues have ϕ and ψ angles in the core and allowed regions. Cry5Ba toxin is a rather compact molecule composed of three distinct domains, and has approximate overall dimensions of 85 × 65 × 45 Å (Fig. 2a). Domain I is composed of several α helices, domain II is of β-sheet. This structure resembles the previously reported Cry1Aa toxin structure but shows some distinctions (Fig. 2b–d); Cry5Ba has several insertions in the three domains compared to the Cry1Aa sequence. The surface electrostatic potential distribution of Cry5Ba and Cry1Aa is also different (Fig. 3).

Domain I

Domain I is composed of residues 65–330, and consists of seven α-helices and four small β-strands. The most hydrophobic helix, α5, is located centrally and is surrounded by the six remaining helices. Two loops, α3–α4 and α4–α5, are much longer than those of Cry1Aa, and are connected to the α helices by four small β-strands. Helix α2 is interrupted by two long insertions, leading to a different surface electrostatic potential distribution from that of Cry1Aa. The solvent-accessible surfaces of α-helices α1, α6, and α7 show relatively higher potential and clear charge separations from α-helices α3 and α4. In the present water-soluble structure of Cry5Ba toxin, electrostatic charges exposed at the surface of α-helices α1, α6, and α7 are largely neutralized by opposite charges located at the surface of the interacting domain II, thus the overall surface of the compact toxin is neutral. This has implications for the way that domain I approaches the target cell membrane.

An "umbrella model" has been proposed to account for the toxicity of known Cry toxins [3]. In this model, helices α4 and α5 are inserted into the membrane as a helical hairpin structure, with the remaining helices lying at the membrane surface. There are neutral regions in the middle of α4 and α5, which probably indicates, if the umbrella model is correct, that both helices cross the membrane, with their polar sides exposed to the solvent. Cry5Ba has the same most-conserved region as most Cry toxins. It is reasonable to assume that domain I of Cry5Ba plays the same pore-forming role as domain I of Cry1Aa (Fig. 4).

However, it was reported that Cry toxins were lethal to the eggs of nematodes [19]. It is possible that different mechanisms of action exist that we have overlooked because it is highly unlikely that Cry toxins can penetrate the impervious eggshell of nematodes. We can hypothesize that those specific long loops, i.e. α3–α4 and α4–α5, have unique roles in nematicidal activity. In addition to pore-forming, Domain I may also participate in receptor binding and hence in determining the specificity of the toxin nematodes. Further investigation of the functional importance of those regions of domain I in Cry5Ba are necessary and should prove very interesting.

Domain II

Receptor binding domain II, comprising residues 331–529, consists of three antiparallel β-sheets packed via formation of a central hydrophobic core as in Cry1Aa. Comparison of the known Cry toxin crystal structures pointed out their structural diversities and traced the most variable part of the Cry toxin family to domain II, especially the apical loops in domain II. Their surface accessibility, added to their

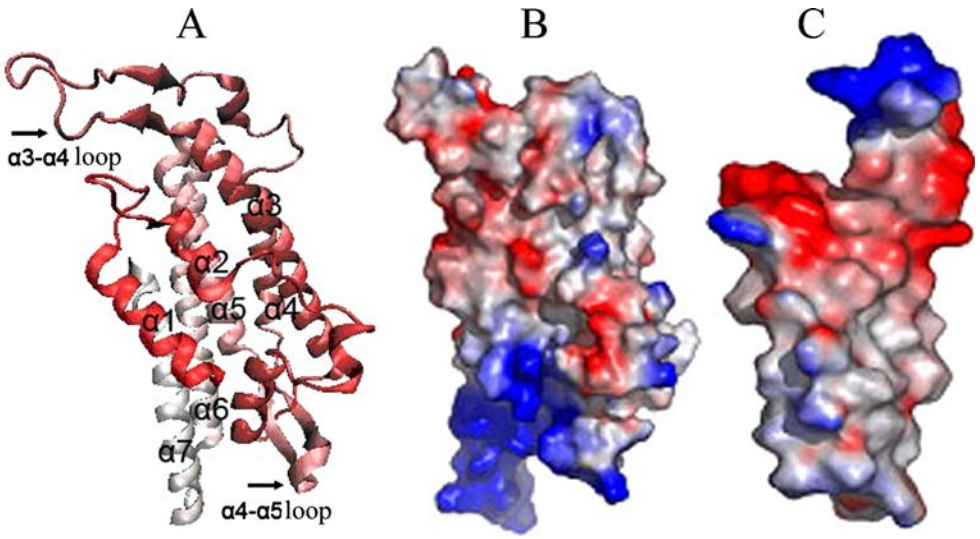


Fig. 4 **a** Cartoon representations of domain I of Cry5Ba. **b**, **c** Surface electrostatic potentials of domain I of Cry5Ba (**b**) and Cry1Aa (**c**). *Blue* Positive electrostatic potential, *red* negative electrostatic potential. High positive potential at the top of the insertion is suggested to facilitate membrane contact. This segment is thought to play an important role in membrane insertion and may be involved in determining specificity (see text)

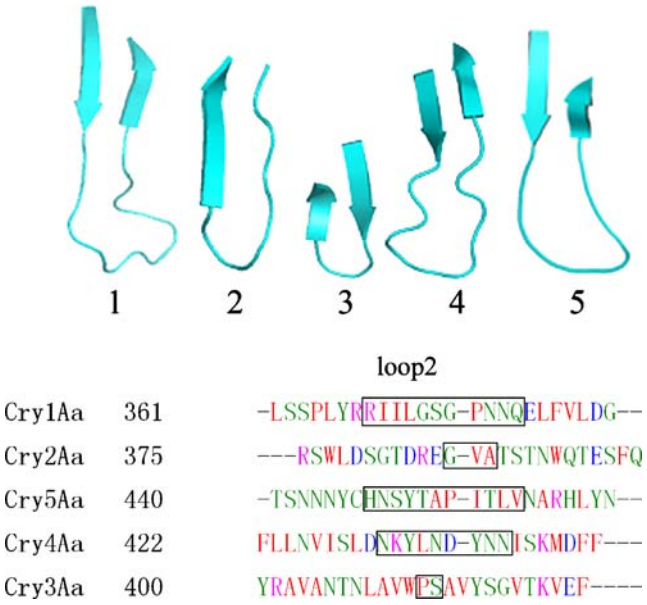


Fig. 5 Three-dimensional (3D) structure comparison and sequence alignment of the apical loop2 of Cry1Aa(1), Cry2Aa(2), Cry3Aa(3), Cry4Aa(4) and Cry5Ba(5). Loop2 of Cry5Ba is very hydrophobic, with three tyrosines in and nearby it, and those loops share very low sequence identity

variability, favours a receptor-recognition role for these loops. Site-directed mutagenesis of the loop residues in related toxins was reported to affect binding affinity and toxicity [29, 30]. Among Cry1Aa, Cry2Aa, Cry3Aa and Cry4Aa toxins, the apical loops of domain II are highly variable in length and amino acid sequence. Domain II is the most divergent domain and Cry2Aa is the most divergent member. Interestingly, the apical loops of domain II of Cry5Ba and Cry1Aa can be superimposed very well (Fig. 4a). The 11-residue loop2 of Cry5Ba is very hydrophobic, with one aromatic amino acid (Tyr450) and two aromatic amino acids (Tyr 445, Tyr463) nearby (Fig. 5). Hydrophobic patches on protein surfaces are generally determinants of protein–protein or protein–ligand interactions [31]. A large number of hydrophobic residues exposed to solvent are also found in other pore-forming toxins, including hemolysin E from *Escherichia coli* [32] and aerolysin [33]. These residues were proposed to interact with hydrophobic lipid tails. Aromatic Trp and Tyr residues have been reported to tend to interact specifically with the outer envelope of the lipid membrane, as was previously shown structurally for the fusion loops of class II viral envelope glycoproteins [34]. These three aromatic amino acids form a potential binding site with dimensions that could accommodate a short oligosaccharide [11]. The architecture suggests that domain II probably binds to the carbohydrate moiety of a glycoprotein receptor of the target insect membrane (Fig. 6). This notion is further reinforced by the finding that *C. elegans* resistance to Cry5B toxicity is linked to the loss of a gene encoding a galactosyltransferase [35]**.** We can hypothesize that mutation in this section may change the toxicity to nematodes and thus alter specificity.

Domain III

The C-terminal domain III, extending from residues 530 to 695, contains two antiparallel β-sheets that adopt a β-sandwich fold and show a jelly-roll-like topology. Domain III stacks on top of domain II and against the side of domain I. The outer sheet is composed of strands exposed to the solvent. The inner sheet, containing seven strands, faces the other two domains. Domains II and III are associated via the intersheet connection through hydrogen bonds and hydrophobic interactions. Superimposition of domain III of Cry5Ba and domain III of Cry1Aa revealed close structural similarity except for some loops of Cry5Ba exposed to the solvent. Mutations in domain III of Cry1Aa toxin had an effect on both ion channel activity and membrane permeability [36]. Domain III could play a role in protecting the toxin against further cleavage by gut proteases [37]. Domain swapping experiments suggested that domain III can function as a specificity determinant [38].

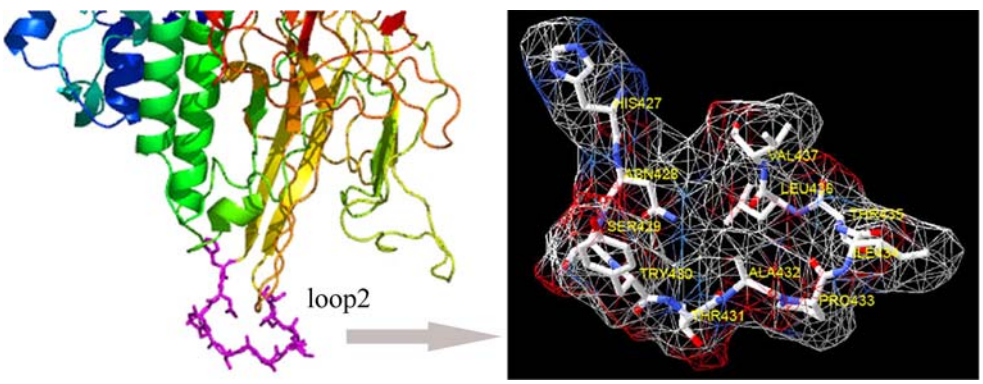


Fig. 6 View of the apical loop 2 (*purple*) in domain II of Cry5Ba. This loop is thought to play an important role in receptor recognition and binding. The right panel shows an electron density map of loop 2. The final electron density map is displayed at a contour of 1σ

Conclusions

Based on the template of Cry1Aa, we have built a 3D structural model for Cry5Ba and used the model to study the possible binding mechanism responsible for nematicidal activity. Despite the low amino acid homology between Cry5Ba and Cry1Aa, the two toxins share a common 3D structure. Compared with that of Cry1Aa, domain I of Cry5Ba has two long loops and its α2 is interrupted by two long insertions, which has implications for the way that domain I approaches the target cell membrane. The interrupted α2 may also be involved in pore-forming and specificity determination. Apical loop2 of domain II of Cry5ba is very hydrophobic, with three aromatic acids in and nearby that may be crucial to specificity determination. Domain III of both toxins superimposed very well. Some loops in domains I and II of Cry5Ba are exposed to the solvent and need to be investigated in depth. This is the first model of a Cry5Ba toxin. The model provides valuable structural information indicating the mechanism of nematicidal activity. The accumulating knowledge of Cry5Ba toxin structure has led, and will lead by experimentation, to a better understanding of the structural basis for receptor binding and pore formation, as well as to designing efficient bio-nematicides.

Acknowledgments We thank Dr. You-min Zhang (Gene Bridge GmbH, Dresden, Germany) for a critical reading of the manuscript. This research was supported by grants from the National Natural Science Foundation of China (No.30670052, 30570050) and 863 Program of China (2006AA02Z187, 2006AA10A212).

References

1. Schnepf E, Crickmore N, van Rie J, Lereclus D, Baum J, Feitelson J, Zeigler DR et al (1998) Microbiol Mol Biol Rev 62:772–806
2. Hofmann C, Vanderbruggen H, Hofte H, Van Mellaert H (1988) Proc Natl Acad Sci USA 85:7844–7848
3. Gazit E, La Rocca P, Sansom PM, Shai Y (1998) Proc Natl Acad Sci USA 95:12289–12294
4. Knowles BH, Ellar DJ (1987) Biochim Biophys Acta 924:509–518
5. Li JD, Carroll J, Ellar DJ (1991) Nature 353:815–821
6. Galitsky N, Cody V, Wojtczak A, Ghosh D, Luft JR, Pangborn W, English L (2001) Acta Crystallogr Sect D 57:1101–1109
7. Grochulski P, Masson L, Borisova S, Pusztai-Carey M, Schwartz JL, Brousseau R et al (1995) J Mol Biol 254:447–464
8. Derbyshire DJ, Ellar DJ, Li J (2001) Acta Crystallog Sect D 57:1938–1944
9. Morse RJ, Yamamoto T, Stroud RM (2001) Structure 9:409–417
10. Boonserm P, Davis P, Ellar DJ, Li J (2005) J Mol Biol 348:363–382
11. Boonserm P, Mo M, Angsuthanasombat C, Lescar J (2006) J Bacteriol 188:3391–3401
12. Gutierrez P, Alzate O, Orduz S (2001) Mem Inst Oswaldo Cruz 96:357–364
13. Jurat-Fuentes JL, Adang MJ (2001) Appl Environ Microbiol 67:323–329
14. Burton SL, Ellar DJ, Li J, Derbyshire DJ (1999) J Mol Biol 287:1011–1022
15. Masson L, Tabashnik BE, Mazza A, Prefontaine G, Potvin L, Brousseau, R, Schwartz JL (2002) Appl Environ Microbiol 68:194–200
16. Ciordia H, Bizzell WE (1961) J Parasitol 47:411–416
17. Kotze AC, O'Grady J, Gough JM, Pearson R, Bagnall NH, Kemp DH, Akhurst RJ (2005) Int J Parasitol 35:1013–1022
18. Bone LW, Bottjer KP, Gill SS (1987) J Parasitol 73:295–299
19. Wei JZ, Hale K, Carta L, Platzer E, Wong C, Fang SC, Aroian RV (2003) Proc Natl Acad Sci USA 100:2760–2765
20. Cappello M, Bungiro RD, Harrison LM, Bischof LJ, Griffitts JS, Barrows BD, Aroian RV (2006) Proc Natl Acad Sci UA 103:15154–15159
21. Sick AJ, Schwab GE, George E, Payne JM (1994) US Patent 5281530[P]
22. Payne J. Narva KE, Fu J (1998) US Patent 5831011[P]
23. Schnepf HE, Schwab GE, Payne J, Narva KE, Foncerrada L (1998) US Patent 6632792[P]
24. Zhao XM, Xia LQ, Wang FX, Ding XZ (2007) Agrochemicals 46:296–299
25. Guex N, Peitsch MC (1987) Electrophoresis 18:2714–2723
26. Laskowski RA, MacArthur MW, Moss DS, Thornton JM, PROCHECK (1993) J Appl Cryst 26:283–291
27. Humphrey W, Dalke A, Schulten K (1996) J Mol Graphics 14:33–38
28. DeLano WL (2002) The PyMOL user's manual. DeLano Scientific, San Carlos, CA
29. Fernandez LE, Perez C, Segovia L, Rodriguez MH, Gill SS, Bravo A, Soberon M (2005) FEBS Lett 579:3508–3514
30. Tuntitippawan T, Boonserm P, Katzenmeier G, Angsuthanasombat C (2005) FEMS Microbiol Lett 242:325–332
31. Lijnzaad P, Berendsen HJ, Argos P (1996) Proteins 25:389–397
32. Wallace AJ, Stillman, TJ, Atkins, A, Jamieson, SJ, Bullough, PA, Green, J, Artymiuk, PJ (2000) Cell 100:265–276
33. Parker MW, Buckley JT, Postma JP, Tucker AD, Leonard K, Pattus F, Tsernoglou D (1994) Nature 367:292–295
34. Bressanelli S, Stiasny K, Allison, SL, Stura EA, Duquerroy S, Lescar J, Heinz FX, Rey FA (2004) EMBO J 23:728–738
35. Griffitts JS, Haslam SM, Yang T, Garczynski SF, Mulloy B, Morris H, Cremer PS et al (2005) Science 307:922–925
36. Schwartz JL, Potvin L, Chen XJ, Brousseau R, Laprade R, Dean DH (1997) Appl Environ Microbiol 63:3978–3984
37. Masson L, Tabashnik BE, Mazza A, Prefontaine G, Potvin L, Brousseau R, Schwartz JL (2002) Appl Environ Microbiol 68:194–200
38. de Maagd RA, Weemen-Hendriks M, Stiekema W, Bosch D (2000) Appl Environ Microbiol 66:1559–1563