Indian Journal of Biotechnology
Vol 10, April 2011, pp 202-206

Homology modeling deduced 3D structure of the *Cry*1Ab22 toxin

S Kashyap[1]*, B D Singh[2] and D V Amla[3]

[1]National Bureau of Agriculturally Important Microorganisms (ICAR), Kushmaur, Katholi, Mau Nath Bhanjan 275 101, India

[2]School of Biotechnology, Faculty of Sciences, Banaras Hindu University, Varanasi 221 005, India

[3]Molecular Biology & Genetic Engineering Division, National Botanical Research Institute Lucknow 226 001, India

Received 18 January 2010; revised 29 June 2010; accepted 9 September 2010

δ-Endotoxin *Cry*1Ab22 is produced by *Bacillus thuringiensis* BtS2491Ab. The toxic spectrum of this protein is reported to span Lepidopteron and Dipteran. Here, we predict the theoretical structural model of newly reported *Cry*1Ab22 toxin by homology modeling method on the structure of the *Cry*1Aa toxin. Proposed model resembles the target by sharing common three dimensional, three domain structure. The main differences being located in the length of loops, absence of helixes (α7b, α10a, α10b, α11a) and presence of additional components (β21, α9b). Few of the components like α9a, α9b and α12a are positioned spatially at different locations. A better understanding of the 3D structure will be helpful in designing the domain swapping and mutagenesis experiments aimed at improving toxicity, and will lead to a deeper understanding of the common mechanism of toxins.

Keywords: 3D structure, homology modelling, *Cry*1Ab22, *Bacillus thuringiensis*

Introduction

During the last few decades the use of synthetic insecticides has caused problems of insect resistance and environmental pollution worldwide[1]. Insecticidal crystal protein (ICP) produced by the soil bacterium *Bacillus thuringiensis* (Bt) belongs to a large toxin family with a target spectrum spanning insects, nematodes, flatworms and protozoa[2,3] and is considered harmless to mammals[3]. The *Cry*1A series toxins are produced as inactive protoxins during sporulation. These protoxins require solubilization and proteolysis in the insect midgut to form an active toxin, which further binds to a high affinity site on the gut membrane. The inserted toxin disturbs the electrolyte balance by creating pores in the cell membrane, leading to the cell lysis and finally to larval death[4]. Crystal structures of the active toxins have been analyzed for *Cry*1Aa[5], *Cry*2A protoxin[6], *Cry*3A[7], *Cry*3B[8], *Cry*1Ac[9], *Cry*4Ba[10] and *Cry*4Aa[11] by X-ray diffraction crystallography and *Cry*11Bb[12], *Cry*5Aa[13] and *Cry*5Ba[14] by homology modeling method. These reports have shown that the toxin has 3 structural domains. Domain I is α-helical bundle made up of seven α helices. Domain II is comprised of antiparallel β sheets, and Domain III is made up of a β sandwich. So far *Cry*1 toxins have been extensively used in studies of Lepidopteron insect control and has attracted less attention to their nematocidal activity either alone or in combination. But despite above, less-efforts have been focused on *Cry*1B structural studies. Therefore, a complete understanding of 3D structures of all the *Cry*1 family members is desirable for a comprehensive understanding of mechanisms underlying toxicity. Here, we report a model for structure of the *Cry*1Ab toxin based on the hypothesis of structural similarity[7] with *Cry*1A toxin. This model supports existing hypotheses and will further provide initiation point for the domain swapping experiments between *Cry*1 and other toxins for elucidation of possible nematocidal mode of toxin action.

Materials and Methods

The multiple alignment, among *Cry*1Ab (AWB87320), *Cry*1Aa1 (PDB:1CIYA) and *Cry*1Ac1 using MEGA software[15], was directly used to jumpstart HHpread Interactive server (http://protevo.eb.tuebingen.mpg.de/hhpred). The results obtained were manually narrowed down through selection of few high scoring entries and HHpread was rerun at local alignment and zero settings. The so obtained end alignment was directly fed to the MODELLER software[16]. The raw PDB was retrieved, suitably truncated and the model was validated with PROCHECK[17] (http://www.ebi.ac.uk).

*Author for correspondence:
Tel: 91-547-2530800; Fax: 91-547-2530358
E-mail- sudhanshukshyp@gmail.com

Results and Discussion

Alignment of Domain I was straightforward and the highly conserved nature of helix 5 in the *Cry*1Ab22 toxin made reliable placement of the other residues. The limits of Domain II could be determined because of the conserved nature of flanking Domains I and III. Domain III was well conserved on the N-terminal side, but on the C-terminal side, there was less correspondence in amino acids (only one identity present at Ile114). The final *Cry*1Ab22 model, corresponding to residues 33-609 of the primary structure (Figs 1a, b), was submitted to the PMDB database (http://www.caspur.it/PMDB/) to obtain PMDB identifier PM0076089. The superimposed backbone traces of *Cry*1Aa showed low RMS deviations while the Ramachandran plot showed that most (95.3%) of the residues have Φ and Ψ anglesin the core and allowed regions, except for some proline and glycine and few other residues located in the loop regions. Most bond lengths, bond angles, and torsion angles were between values expected for a naturally folded protein. The structural model in Fig. 1 shows that it contains all the general features of the *Cry* toxins (a α+β structure with three domains) with few differences located in the loops length of Domain II and III. Four of the five Cys residues present in the primary structure were included in the *Cry*1Ab22 model (Cys157, Cys186, Cys225 and Cys436), but it seems that they are not involved in structure stabilization through disulfide bonds, since the shortest S-S distance (Cys157-Cys225) among the six different possible bonds was 18.51 Å.

All the helices in the *Cry*1Ab22 model were slightly shorter than those in *Cry*1Aa. According to the amphiphilicity calculated with the Hoops and Woods values, the most exposed helices are α1, α2a,

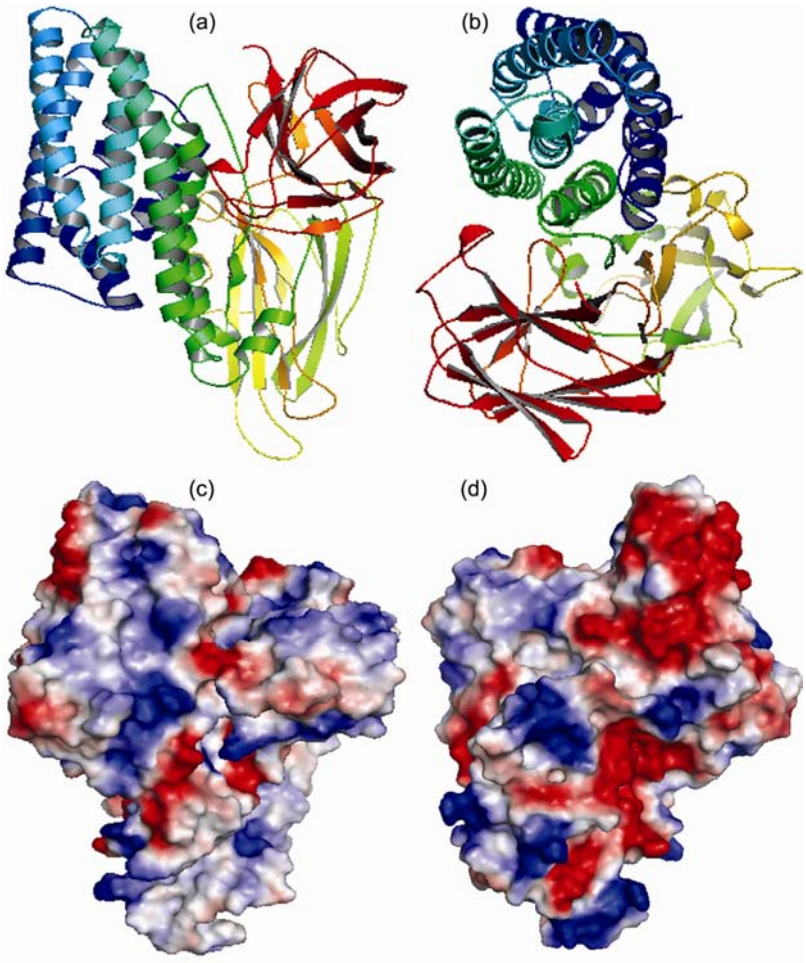

Fig. 1—3D, three domain structure of the *Cry*1Ab22 toxin molecule: (a) complete molecular view, (b) top view showing helical pore forming domain and lower antiparellel nature of sheets arrangement, (c) electrostatic potential distribution on the surface of *Cry*1Ab22 toxin molecule, & (d) same molecule after 90° rotation. Figures and electrostatic potential calculations were generated with PyMOL0.99rc6 (http://www.pymol.org/funding.html)

α2b, α3 and α6, which correspond well with the accessibility calculated with SwissPDB, except for α1, which is packed against Domain II. It is possible that this helix will show some mobility, if we consider that one of the cutting sites by gut proteases is located between Ser56 and Ile58, close to the middle of this helix[18]. The charge distribution pattern in the *Cry*1Ab22 theoretical model corresponds to a negatively charged patch along β4 and β13 (Fig. 1c, d) of domains II and III, respectively. The *Cry*1Ab22 Domain I model relates well with data from Gazzit *et al*[19], who suggested that α4 and α5 insert into the membrane in an antiparallel manner as a helical hairpin. It is possible that according to the surface electrostatic potential of helices 4 and 5 (Fig. 1c, d), there was a neutral region in the middle of the helices which probably shows, if the umbrella model is correct, that both helices cross the membrane with their polar sides exposed to the solvent, as it has been suggested by the results of mutagenesis experiments carried out by Kumar and Aronson[20] with the *Cry*1Ac toxin. This region is also the most conserved among the *Cry* toxins. They also demonstrated that mutation in the base of helix 3 and the loop between α3 and α4 that cause alterations on the balance of negative charged residues may cause loss of toxicity[20]. Mutations in helices α2, α6 and the surface residues of α3 have no important effect on toxicity; meanwhile, helices α4 and α5 seem to be very sensitive to mutation. Helix α1 probably does not play an important part in toxin activity after cleavage of the protoxin. It is possible that mutation aimed to an increase in amphilicity in these helices will improve the pore forming activity of *Cry*1Ab22 type of toxins. Domain II is the most variable domain among *Cry* toxins, and it has been shown that it is involved in receptor recognition and, therefore, considered as the specificity determining region. As for other *Cry* toxins, domain II of the *Cry*1Ab22 toxin consists of three Greek key β sheets arranged in a β prism topology. The main difference between the *Cry*1Ab22 theoretical model and *Cry*1Aa structures is the length of the two loops joining the apical β-strands (β2-β3 and β4-β5), absence of four α components (α7b, α10a, α10b, α11a) and presence of additional β21, α9b components. In addition to these, α9a, α9b and α1 2a are spatially located at downstream positions (Table 1). We propose that additional components have implications in receptor binding, and in specificity of the *Cry*1Ab22 toxin. Physical presence of such additional and displaced components has to be verified using small peptide NMR. Fragments swapping approach will certainly pin point their potential role in specificity and binding phenomenon. The first insertion between strands β2 and β3, the N-terminal part of this loop (Ile306-Thr311) is mostly hydrophobic, while the C-terminal half (Thr312-Thr318) is polar and has one positively charged residue (Glu316). This loop probably interacts with the receptor through both hydrophobic and electrostatic interactions; Gly315 probably helps in receptor binding by providing more mobility to

Table 1—The comparison between three domain structural components of *Cry*1Aa and *Cry*1Ab22 toxin molecules.

	*Cry*1Aa	*Cry*1Ab22		*Cry*1Aa	*Cry*1Ab22		*Cry*1Aa	*Cry*1Ab22
	Domain I			Domain II			Domain III	
α1	Pro35-Ser48	Pro35- Ser 48	α8a	Pro271-Glu274	Pro271-Asn275	α11a	Leu475-Lys477	--
α2a	Aln54-Ile63	Ala54-Trp65	α8b	Ala284-Gln289	Ser283-Ser290	β13a	Ser486-Val488	Thr480-Leu482
α2b	Pro70-Ile84	Pro 70-Ile84	β2	Asp298-His310	Ile299-His310	β13b	Ile498-Arg 501	Ser487-Val489
α3	Glu90-Ala119	Glu90-Ala119	β3	Phe313-Trp316	Glu313-Ser324	β14	Gly505-Asn513	Ile499-Arg502
α4	Pro124-Leu148	Pro124-Ala149	β4	Gly318-Pro325	Tyr359-Arg368	β15	Tyr522-Ser530	Gly506-Ile515
α5	Gln154-Trp182	Gln154-Trp182	α9a	Val326-Phe328	*Ser 409-Glu412*	β16	Leu534-Ile540	Tyr523-Ser531
α6	Ala186-Val218	Ala186-Val218	α9b	--	*Pro422-Gly425*	β17	Arg543-Phe550	Leu535-Ile541
α7a	Ser223-Thr239	Ser223-Tyr250	β5	Val348-Ser351	Leu380-Tyr390	α12a	Ser562-Ser564	*Ser563-Ser565*
α7b	Leu241-Tyr250	--	β6	Ile357-Arg367	Ala399-Tyr401	β18	Arg566-Gly569	Arg544-Phe551
β1a	Glu266-Thr269	Ile267-Thr269	β7	Leu380-Leu383	Thr406-Asp408	β19	Ser580-His588	Arg567-Gly570
			β8	Gly385-Phe390	His428-Val433	β20	Val569-Pro605	Ser581-His589
			β9	Thr400-Tyr402	Ile447-Arg449	β21	--	Val597-Pro606
			α10a	Ser410-Asp412	--			
			α10b	Pro423-Gly426	--			
			β10	His429-Val434	Phe453-His457			
			β11	Phe452-His456	Asn464-Ile466			
			β12	Thr471-Pro474	Thr472-Pro475			

(--: similar component not present, *Components in Italics are spatially present at downstream sites.)

Glu316 that may interact through salt bridges with the receptor. Loop β4-β5 is mostly hydrophilic, and the charged residues located at the tip of the loop (Lys353, Asp355 and His356) are probably important determinants of insect specificity. As in loop β2-β3, a glycine residue (354) is also present before a negatively charged residue (Asp355) supporting the view that orientation of charged residues in the specificity loops could be important in receptor recognition. The third insertion (Ser471- Lys476) probably does not play an important role in insect recognition as it is located in the region joining Domains II and III, but it probably gives more mobility to Domain III modulating its function in receptor binding. Mutations in defined regions of the *Cry*1Aa toxin have identified residues 365-371 (equivalent to residues in the *Cry*1Ab22 β6-β7 loop), as essential for binding to the membrane of midgut cells of *Bombyx mori*[21,22]. In the *Cry*1Ab22 model, this region is shorter than their counterparts in *Cry*1Aa. Loop β2-β3 seems to be able to modulate the toxicity and specificity of *Cry*1C[23]; in this region a five residue insertion (Gly308-Thr312) was found. The dual specificity of *Cry*2Aa for Lepidoptera and Diptera has been mapped to residues 307-382 that corresponds in the *Cry*1Ab22 theoretical model to sheet 1, strand ß6, and loop β6-β7, where most of the insertions/deletions and structural differences were located. The Domain III composed by residues 471-608 and showed high conservation of residues and the only important modification is a 3-residue deletion between β16 and β17. We propose the sequence RYRVRIRYAS as homolog to conserved block 4. For conserved block 5 (VYIDRIEFVP), there was only one identity and two conserved residues (Ile614, Phe611 and Asp615, respectively). Several studies showed that site mutations in this block reduce toxicity and alter channel properties in *Cry*1Ac[7] and *Cry*1Aa[24,25], divergence in block 5 possibly reveal an alternative mechanism of membrane permeabilization. In conclusion, evidences presented here, based on the identification of structural equivalent residues of *Cry*1Aa in *Cry*1Ab22 toxin through homology modeling show that, despite the differences in amino acid composition, absence-; presence-of structural components of which few are located at different positions. Still they share a common 3D structure. *Cry*1Ab22 contains the most variable regions in the loops of Domain II, which determines the specificity of these toxins. This is the first model of a *Cry*1Ab22 protein and its importance can be perceived since members of this group of toxins are potentially important candidates.

Acknowledgement

The authors are grateful to ICAR for RAship to SK. Infrastructure facility and encouragements by Director NBAIM are duly acknowledged.

References

1 Roush R T, Can we slow adaptation by pests to insect transgenic crops? in *Biotechanology and integrated pest management,* edited by G J Parsley (CAB International, Wallingford, UK) 1996, 242-263.

2 Roh J Y, Choi J Y, Li M S, Jin B R & Je Y H, *Bacillus thuringiensis* as a specific, safe and effective tool for insect pest control, *Microbiol Biotechnol*, 17 (2007) 547-559.

3 Schnepf E, Crickmore N, van Rie J, Lereclus D, Baum J *et al*, *Bacillus thuringiensis* and its pesticidal crystal proteins, *Microbiol Mol Biol Rev*, 62 (1998) 772-806.

4 Hofmann C, Vanderbruggen H, Hofte H & Van Mellaert H, Specificity of *Bacillus thuringiensis* δ-endotoxins is correlated with the presence of high-affinity binding sites in the brush border membrane of target insect midgets, *Proc Natl Acad Sci USA*, 85 (1988) 7844-7848.

5 Knowles B H & Ellar D J, Colloid osmotic lysis is a general feature of the mechanism of action of *Bacillus thuringiensis* δ-endotoxins with different insect specificities, *Biochim Biophys Acta*, 924 (1987) 509-518.

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

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

8 Galitsky N, Cody V, Wojtczak A, Ghosh D, Luft J R *et al*, Structure of the insecticidal bacterial endotoxin Cry3Bb1 of *Bacillus thuringiensis, Acta Cryst* D57, (2001) 1101-1109.

9 Derbyshire D J, Ellar D J & Li J, Crystallization of the *Bacillus thuringiensis* toxin Cry1Ac and its complex with the receptor ligand *N*-acetyl-D-galactosamine, *Acta Cryst* D57, (2001) 1938-1944.

10 Boonserm P, Davis P, Ellar D J & Li J, Crystal structure of the mosquito-larvicidal toxin Cry4Ba and its biological implications, *J Mol Biol*, 348 (2005) 363-382.

11 Boonserm P, Mo M, Angsuthanasombat C & Lescar J, Structure of the functional form of the mosquito larvicidal Cry4Aa toxin from *Bacillus thuringiensis* at 2.8-Angstrom resolution, *J Bacteriol*, 188 (2006) 3391-3401.

12 Gutierrez P, Alzate O & Orduz S A, Theoretical model of the 3D structure of *Bacillus thuringiensis* subsp. *medellin* Cry 11Bb toxin deduced by homology modeling, *Mem Inst Oswaldo Cruz*, 96 (2001) 357-364.

13 Min Z X-, Qui X L, Zhi D X & Xiang W F, The theoretical 3D structure of *Bacillus thuringiensis* Cry5Aa and its biological implications, *Protein J*, 28 (2009) 104-110.

14 Xia L Q, Zhao X M, Ding X Z, Wang F X & Sun Y J, The theoretical 3D structure of *Bacillus thuringiensis* Cry5Ba, *J Mol Model*, 14 (2008) 843-848.

15 Tamura K, Dudley J, Nei M & Kumar S, MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0, *Mol Biol Evol*, 24 (2007) 1596-1615.

16 Sali A, Potterton L, Yuan F, van Vlijmen H & Karplus M, Evaluation of comparative protein modeling by MODELLER, *Proteins*, 23 (1995) 318-326.

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

18 Segura C, Guzman F, Patarroyo M E & Orduz S, Activation pattern and toxicity of the Cry1Ab22Bb toxin of *Bacillus thuringiensis* subsp. *medellin*, *J Invertebr Pathol*, 76 (2000) 56-62.

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

20 Kumar A S M & Aronson A I, Analysis of mutations in the pore-forming region essential for insecticidal activity of a *Bacillus thuringiensis* ∂-endotoxin, *J Bacteriol*, 181 (1999) 6103-6107.

21 Ge A Z, Sivarova N D & Dean D H, Location of the *Bombyx mori* specificity domain of *Bacillus thuringiensis* ∂-endotoxin protein, *Proc Natl Acad Sci USA*, 86 (1989) 4037-4041.

22 Lu H, Rajamohan F & Dean D H, Identification of amino acid residues of *Bacillus thuringiensis* ∂-endotoxin CryIA(a) associated with membrane binding and toxicity to *Bombyx mori*, *J Bacteriol*, 176 (1994) 5554-5559.

23 Smith G P & Ellar D J, Mutagenesis of two surface exposed loops of the *Bacillus thuringiensis* CryIC ∂-endotoxin affects insecticidal specificity, *Biochem J*, 302 (1994) 611-616.

24 Chen X J, Lee M K & Dean D H, Site-directed mutations in a highly conserved region of *Bacillus thurngiensis* ∂-endotoxin affect inhibition of short circuit current across *Bombyx mori* midgets, *Proc Natl Acad Sci USA,* 90 (1993) 9041-9045.

25 Schwartz J L, Potvin L, Chen X J, Brousseau R, Laprade R *et al*, Single-site mutations in the conserved alternating-arginine region affect ion channels formed by CryIAa, a *Bacillus thuringiensis* toxin, *Appl Environ Microbiol,* 63 (1997) 3978-3984.