Egyptian Journal of Biological Pest Control

REVIEW ARTICLE Open Access

Structural, functional, and evolutionary analysis of Cry toxins of *Bacillus thuringiensis*: an *in silico* study

Sujit Kumar Das[1,2], Sukanta Kumar Pradhan[1], Kailash Chandra Samal[3] and Nihar Ranjan Singh[4*]

Abstract

Background: *Bacillus thuringiensis* (*Bt*) is a gram-positive spore-forming soil bacterium that synthesizes crystalline (Cry) protein, which is toxic and causing pathogenicity against mainly three insect orders: Coleoptera, Diptera, and Lepidoptera. These crystalline protein inclusions, i.e., δ-endotoxins are successfully used as a bio-control agent against insect pests.

Main body: A total of 58 various Cry proteins belonging to these 3 insect orders were retrieved from SwissProt database and are categorized into different groups. Structural and functional analysis were performed to understand the functional domain arrangements at sequence level as well as at structural level involving both experimental and predicted 3-dimensional models. Besides, the analysis of evolutionary relationship involving all 58 observed Cry proteins at the sequence, domain, and structural levels were done using different bioinformatics tools. Evolutionary analysis revealed that some Cry proteins having toxicity for a specific insect order are found to be clustered for another different insect order, which concludes that they might have toxicity for more than one insect order. Three-dimensional (3D) structure analysis of both experimental and predicted models revealed that proteins might have toxicity for a specific insect order differ in their structural arrangements and was observed in Cry proteins belonging to 3 different insect orders.

Conclusions: It could be hypothesized that an inner-molecular domain shift or domain insertion/deletion might have taken place during the evolutionary process, which consequently causes structural and functional divergence of *Bt*. The study output may be helpful for understanding the diversity as well as specificity of the analyzed insecticidal proteins and their application as a biopesticide in the field of agriculture.

Keywords: *Bacillus thuringiensis*, Biopesticides, Cry toxins, Bioinformatics, Phylogenetic analysis

Background

Bacillus thuringiensis (*Bt*) is a gram-positive spore-forming soil bacterium causing pathogenicity in insects. *Bt* strains synthesize Crystal (Cry) and cytolytic (Cyt) toxins (also known as δ-endotoxins) that have a natural insecticidal effect on selective insect orders. δ-endotoxins of *Bt* are being used successfully as a biological control agent against some insect pests (Schnepf 1995). These endotoxins are exclusively active against larval stages of different insect orders such as Lepidoptera (Butterflies, Moths), Coleoptera (Flies and Mosquitoes), and Diptera (Beetles and Weevils) (Raymond et al. 2010). During sporulation phase, *Bt* produces these Cry or Cyt toxins that have hazardous effect on insects (Bravo et al. 2011). Cry and Cyt toxins are considered as parasporal inclusion proteins from *Bt* that exhibit toxic effects and hemolytic activity respectively (Crickmore et al. 1998).These two types of toxins belong to a class of pore-forming toxins (PFTs) that are secreted as water-soluble proteins and undergo conformational changes in order to insert into

* Correspondence: nihar.singh@gmail.com
[4]Department of Botany and Center of Excellence in Environment and Public Health, Ravenshaw University, Cuttack, Odisha 753003, India
Full list of author information is available at the end of the article

the host membrane. When Crystal (Cry) toxins ingested by insects are get solubilized in their midgut, then it gets proteolytically activated by midgut proteases and bind to specific receptors located in the insect cell membrane leading to cell disruption and cell death (Bravo et al. 2007). Most Cry proteins exist as inactive protoxins that can be converted into active toxins by certain kinds of insect midgut proteases (Höfte and Whiteley 1989). This activation process appears to involve a sequential series of proteolytic cleavages, starting at the C-terminus and proceeding toward the N-terminus until the protease-stable toxin is generated (Choma and Kaplan 1990). Besides these membrane proteins, other components have been identified due to their capacity to interact with 3d-Cry toxins (3 domain-Cry toxins) such as glycolipids, intracellular proteins, V-ATPase subunit A or actin (McNall and Adang 2003; Griffitts et al. 2005; Bayyareddy et al. 2009). The interaction of Cry1 toxins with different proteins present in lepidopteran midgut cells is a complex process involving multiple membrane proteins such as cadherin-like proteins (CADs), aminopeptidase N (APN), and alkaline phosphatase (ALP) (Pigott and Ellar 2007; Soberon et al. 2009).

Extensive screening of *Bt* strains and Cry gene sequencing has led to the identification of more than 700 Cry gene sequences (Höfte and Whiteley 1989; Van Frankenhuyzen 2009; Crickmore et al. 2011). These sequences have been classified according to their amino acid sequence identities in 70 different Cry gene groups (Cry1, Cry2...Cry 70) where toxins belonging to each group share less than 40% amino acid identity with proteins from other groups (Crickmore et al. 1998). Within each group, a capital letter (Cry1A, Cry1B, etc.) is given when they share less than 70% identity. A small letter (Cry1Aa, Cry1Ab, etc.) is given when toxins share more than 70% but less than 95% identity (Bravo et al. 2013). Many Cry proteins show pesticidal properties, while some protein produced as parasporal crystals that have unknown invertebrate targets called as parasporins such as Cry31A, Cry41A, Cry45A, Cry46A, Cry63A, and Cr64A exhibit strong and specific cytocidal activity against human cancer cell (Mizuki et al. 2000). Phylogenetic analysis of Cry toxins shows that the great variability in the biocidal activity of the 3d-Cry group has resulted from the independent evolution of the 3 structural domains and the domain III swapping among different toxins. These two processes have generated proteins with similar modes of action but with different specificities (de Maagd et al. 2001). Various 3-dimensional crystal structures of activated Cry toxins have been determined by experimental methods. These are Coleoptera-specific Cry3Aa (PDB entry: 1DLC) (Li et al. 1991) and Cry3Bb1 (PDB entry: 1JI6) (Galitsky et al. 2001), Cry 34 Ab1(PDB entry: 4JOX) and Cry 35 Ab1 (PDB entry: 4JPo) (Kelker et al. 2014), Lepidoptera-specific Cry1Aa (PDB entry: 1CIY) (Grochulski et al. 1995), Cry1Ac (PDB entry: 4ARX/4ARY/4W8J) (Evdokimov et al. 2014), Cry1Da (PDB entry: 6OVB) (Wang et al. 2019), Cry1Fa (PDB entry: 6DJ4) (Wang et al. 2018), Cry1Be (PDB entry: 6OWK) (Wang et al. 2019), Lepidoptera/Diptera-specific Cry2Aa (PDB entry: 1I5P) (Morse et al. 2001), Diptera-specific Cry4-Ba (PDB entry: 1W99) (Boonserm et al. 2005), and Cry4Aa (PDB entry: 2C9K) (Boonserm et al. 2006). Although these toxins exhibit markedly different insecticidal specificities, the overall folding patterns of their structures are quite similar, comprising 3 domains. The domain (I) is a bundle of 7–8 α helices involved in pore formation, domain (II) is a β-prism with exposed loops regions involved in receptor binding, and domain (III) is a β-sandwich and has influence on receptor binding, ion channel formation, and insect specificity (Li et al. 1991; Grochulski et al. 1995; Derbyshire et al. 2001; Galitsky et al. 2001; Morse et al. 2001; Guo et al. 2009, and Hui et al. 2012).

It has been observed that the Crystal toxins showed specificity to different insect orders which signify that they might have shared certain level of relationships at different levels such as sequence, domain, and structure. The present article aimed to retrieve different Cry proteins from various protein databases, followed by phylogenetic analysis, domain characterization, structure predictions, experimental structure analysis, and comparison.

Main text

Sequence retrieval and analysis of Cry proteins

Crystal (Cry) proteins of *Bt* having specificity to 3 broad insect orders: Lepidoptera, Coleoptera, and Diptera were retrieved from the UniprotKB protein sequence database (https://www.uniprot.org/). Selections were made only from manually annotated and reviewed sequences belonging to Swiss-Prot and stored in FASTA format for further bioinformatics analysis.

Domain identification, characterization, and functional analysis

Identification and annotation of genetically mobile domains and its architectures in the manually annotated and reviewed protein sequences were carried out using a web-based bioinformatics domain prediction tool, i.e., NCBI Batch CD-Search service (https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi). These domains are extensively annotated with respect to phyletic distributions, functional class, tertiary structures, and functionally important residues.

Sequence alignment and phylogenetic tree construction

Divergence analysis among the retrieved Crystal proteins and the predicted domains were conducted in Molecular

Table 1 List of cry proteins having toxicity for Coleoptera insect order

Entry	Protein	Gene	Organism	Length	PDB ID	Insect_Order
P17969	cry3Ba	cry3Ba	*Bacillus thuringiensis subsp. tolworthi*	659	*N/A*	Coleoptera
P0A379	cry3Aa	cry3Aa	*Bacillus thuringiensis subsp. tenebrionis*	644	1DLC,4QX0,4QX1,4QX2,4QX3	Coleoptera
P0A381	cry3Aa	cry3Aa	*Bacillus thuringiensis subsp. san diego*	644	*N/A*	Coleoptera
P0A380	cry3Aa	cry3Aa	*Bacillus thuringiensis subsp. morrisoni*	644	*N/A*	Coleoptera
Q45744	cry3Ca	cry3Ca	*Bacillus thuringiensis subsp. kurstaki*	649	*N/A*	Coleoptera
Q45704	cry8Aa	cry8Aa	*Bacillus thuringiensis subsp. kumamotoensis*	1157	*N/A*	Coleoptera
Q45705	cry8Ba	cry8Ba	*Bacillus thuringiensis subsp. kumamotoensis*	1169	*N/A*	Coleoptera
Q45708	cry7Ab	cry7Ab	*Bacillus thuringiensis subsp. kumamotoensis*	1138	*N/A*	Coleoptera
Q45706	cry8Ca	cry8Ca	*Bacillus thuringiensis subsp. japonensis*	1160	*N/A*	Coleoptera
O06014	cry9Da	cry9Da	*Bacillus thuringiensis subsp. japonensis*	1169	*N/A*	Coleoptera
Q45707	cry7Ab	cry7Ab	*Bacillus thuringiensis subsp. dakota*	1138	*N/A*	Coleoptera
Q939T0	cry34Ab1	*N/A*	*Bacillus thuringiensis*	123	4JOX	Coleoptera
Q03749	cry7Aa	cry7Aa	*Bacillus thuringiensis*	1138	*N/A*	Coleoptera
Q939S9	cry35Ab1	*N/A*	*Bacillus thuringiensis*	383	4JP0	Coleoptera
Q06117	cry3Bb1	cry3Bb1	*Bacillus thuringiensis*	652	1JI6	Coleoptera

Evolutionary Genetics Analysis X (MEGA X) (Kumar et al. 2018) using Neighbor-Joining (NJ) method (Saitou and Nei 1987). The predicted phylogenetic tree was evaluated using bootstrap reliability test for 1000 replicates. The tree was drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method (Zuckerkandl and Pauling 1965) and are in the units of the number of amino acid substitutions per site.

Structural diversity analysis and prediction

The selected Cry proteins belonging to the 3 major insect orders Lepidoptera, Coleoptera, and Diptera were searched in RCSB-PDB (Research Collaborator for Structural Bioinformatics-Protein Data Bank) (https://www.rcsb.org/) for availability of experimentally solved 3-dimensional structures and PDB files were downloaded. The protein sequences having no 3-dimensional structures in PDB were submitted to web based Protein Homology/analogY Recognition Engine V2.0 (Phyre^2) (http://www.sbg.bio.ic.ac.uk/~phyre2/html/page.cgi?id=index) for protein modeling and prediction of 3-dimensional structures.

Crystal (Cry) proteins of *Bacillus thuringiensis*

Total 58 Cry protein sequences of bacteria *Bt* were retrieved from UniProtKB database (only annotated and reviewed). Among all of these 58 sequences, total 15, 10, and 33 numbers of Cry proteins were observed in the insect orders: Coleoptera, Diptera, and Lepidoptera, respectively. Primary sequence variation in length of Cry proteins were noticed as 123–1169 amino acids in case of order Coleoptera (Table 1), 643–1180 amino acids in Diptera (Table 2), and mostly more than 1100 amino acids in Lepidoptera (Table 3).

Table 2 List of cry proteins having toxicity for Diptera insect order

Entry	Protein	Gene	Organism	Length	PDB ID	Insect_Order
Q9ZIU5	cry11Bb	cry11Bb	*Bacillus thuringiensis subsp. medellin*	750	*N/A*	Diptera
O32307	cry19Aa	cry19Aa	*Bacillus thuringiensis subsp. Jegathesan*	648	*N/A*	Diptera
Q45730	cry11Ba	cry11Ba	*Bacillus thuringiensis subsp. Jegathesan*	724	*N/A*	Diptera
P05519	cry4Ba	cry4Ba	*Bacillus thuringiensis subsp. Israelensis*	1136	1W99,4MOA	Diptera
P09662	cry10Aa	cry10Aa	*Bacillus thuringiensis subsp. Israelensis*	675	*N/A*	Diptera
P16480	cry4Aa	cry4Aa	*Bacillus thuringiensis subsp. Israelensis*	1180	2C9K	Diptera
P21256	cry11Aa	cry11Aa	*Bacillus thuringiensis subsp. Israelensis*	643	*N/A*	Diptera
Q9S597	cry27Aa	cry27Aa	*Bacillus thuringiensis subsp. higo*	826	*N/A*	Diptera
O86170	cry19Ba	cry19Ba	*Bacillus thuringiensis subsp. higo*	682	*N/A*	Diptera
O32321	cry20Aa	cry20Aa	*Bacillus thuringiensis subsp. fukuokaensis*	753	*N/A*	Diptera

Table 3 List of cry proteins having toxicity for Lepidoptera insect order

Entry	Protein	Gene	Organism	Length	PDB ID	Insect_Order
Q9ZAZ5	cry1Bd	cry1Bd	*Bacillus thuringiensis subsp. wuhanensis*	1231	*N/A*	lepidoptera
Q9ZAZ6	cry1Gb	cry1Gb	*Bacillus thuringiensis subsp. wuhanensis*	1169	*N/A*	lepidoptera
Q45733	cry9Ca	cry9Ca	*Bacillus thuringiensis subsp. tolworthi*	1157	*N/A*	lepidoptera
Q45729	cry15Aa	cry15Aa	*Bacillus thuringiensis subsp. thompsoni*	340	*N/A*	lepidoptera
P0A369	cry1Aa	cry1Aa	*Bacillus thuringiensis subsp. sotto*	934	*N/A*	lepidoptera
Q45715	cry1Ka	cry1Ka	*Bacillus thuringiensis subsp. morrisoni*	1215	*N/A*	lepidoptera
Q45718	cry1Hb	cry1Hb	*Bacillus thuringiensis subsp. morrisoni*	1155	*N/A*	lepidoptera
O66377	cry1Fb	cry1Fb	*Bacillus thuringiensis subsp. morrisoni*	1169	*N/A*	lepidoptera
P0A370	cry1Ab	cry1Ab	*Bacillus thuringiensis subsp. kurstaki*	1155	*N/A*	lepidoptera
P0A366	cry1Aa	cry1Aa	*Bacillus thuringiensis subsp. kurstaki*	1176	1CIY	lepidoptera
P21254	cry2Ab	cry2Ab	*Bacillus thuringiensis subsp. kurstaki*	633	*N/A*	lepidoptera
P05068	cry1Ac	cry1Ac	*Bacillus thuringiensis subsp. kurstaki*	1178	4ARX, 4ARY,4W8J	lepidoptera
Q57458	cry1Ea	cry1Ea	*Bacillus thuringiensis subsp. kenyae*	1171	*N/A*	lepidoptera
Q99031	cry9Aa	cry9Aa	*Bacillus thuringiensis subsp. galleriae*	1156	*N/A*	lepidoptera
P0A368	cry1Aa	cry1Aa	*Bacillus thuringiensis subsp. entomocidus*	1176	*N/A*	lepidoptera
P0A375	cry1Ca	cry1Ca	*Bacillus thuringiensis subsp. entomocidus*	1189	*N/A*	lepidoptera
Q03748	cry1Ae	cry1Ae	*Bacillus thuringiensis subsp. alesti*	1181	*N/A*	lepidoptera
P0A376	cry1Ca	cry1Ca	*Bacillus thuringiensis subsp. aizawai*	1189	*N/A*	lepidoptera
P0A367	cry1Aa	cry1Aa	*Bacillus thuringiensis subsp. aizawai*	1176	*N/A*	lepidoptera
Q03744	cry1Ad	cry1Ad	*Bacillus thuringiensis subsp. aizawai*	1179	*N/A*	lepidoptera
P19415	cry1Da	cry1Da	*Bacillus thuringiensis subsp. aizawai*	1165	6OVB	lepidoptera
Q03745	cry1Eb	cry1Eb	*Bacillus thuringiensis subsp. aizawai*	1174	*N/A*	lepidoptera
Q03746	cry1Fa	cry1Fa	*Bacillus thuringiensis subsp. aizawai*	1174	6DJ4	lepidoptera
Q9ZNL9	cry9Ea	cry9Ea	*Bacillus thuringiensis subsp. aizawai*	1150	*N/A*	lepidoptera
Q45746	cry1Ga	cry1Ga	*Bacillus thuringiensis*	1166	*N/A*	Lepidoptera
Q9XDL1	cry1Id	cry1Id	*Bacillus thuringiensis*	719	*N/A*	Lepidoptera
O85805	cry1Be	cry1Be	*Bacillus thuringiensis*	1227	6OWK	Lepidoptera
Q45748	cry1Ha	cry1Ha	*Bacillus thuringiensis*	1172	*N/A*	Lepidoptera
Q45716	cry1Jb	cry1Jb	*Bacillus thuringiensis*	1170	*N/A*	Lepidoptera
Q9S515	cry1Ag	cry1Ag	*Bacillus thuringiensis*	1176	*N/A*	Lepidoptera
Q45747	cry1Db	cry1Db	*Bacillus thuringiensis*	1160	*N/A*	Lepidoptera
Q45738	cry1Ja	cry1Ja	*Bacillus thuringiensis*	1167	*N/A*	Lepidoptera
Q45739	cry1Bb	cry1Bb	*Bacillus thuringiensis*	1229	*N/A*	Lepidoptera

Analysis of Cry proteins structural domains

Primary sequences of 58 Cry proteins with different specificity in the 3 insect orders were further analyzed to annotate different structural domains involved in various biological processes (Tables 4, 5, and 6). Domain analysis was revealed that most of the Cry proteins have 3 structural domains such as Endotoxin_N (PF03944), Endotoxin_M (PF00555), and Endotoxin_C (PF03945) (Tables 4, 5, and 6). It was also observed that Endotoxin_N, Endotoxin_M, and Endotoxins_C are presented in N terminus, middle, and C-terminal region of the protein (Figs. 1 and 2), respectively. Generally, N-terminal helical domain involves in membrane insertion and pore formation whereas middle and C-terminal domains have a vital role in receptor bindings.

Evolutionary analysis of Cry proteins and domains

Evolutionary tree was constructed for 58 Cry proteins using MEGA X tool after elimination of all residual positions containing gaps and was further analyzed (Fig. 3). The tree is clustered into 3 major groups such as clusters I, II, and III. Cluster I (red color) consists of Cry proteins

Table 4 List of cry protein domains having toxicity for Coleoptera insect order

Protein name	Uniprot accession No	Sequence length	Endotoxin_N		*E* value	Endotoxin_M		*E* value	Endotoxin_C		*E* value
			Start	End		Start	End		Start	End	
cry3Ba_coleoptera	P17969	659	73	296	4.4e−30	304	510	9.6e−62	520	659	2.5e−41
cry3Aa_coleoptera_1	P0A379	644	65	287	3.3e−30	295	499	2.1e−62	509	644	7.6e−41
cry3Aa_coleoptera_2	P0A381	644	65	287	3.3e−30	295	499	2.1e−62	509	644	7.6e−41
cry3Aa_coleoptera_3	P0A380	644	65	287	3.3e−30	295	499	2.1e−62	509	644	7.6e−41
cry3Ca_coleoptera	Q45744	649	64	285	2.9e−32	293	502	3.6e−58	512	649	1e−39
cry8Aa_coleoptera	Q45704	1157	69	289	1.1e−27	297	516	2.6e−44	526	662	6.8e−46
cry8Ba_coleoptera	Q45705	1169	80	290	8.8e−39	298	512	7.3e−54	522	658	9.3e−36
cry7Ab_coleoptera_1	Q45708	1138	61	278	1.8e−29	286	487	9.7e−59	497	637	1.5e−42
cry8Ca_coleoptera	Q45706	1160	77	290	3.5e−26	298	503	5.7e−50	513	659	2.6e−33
cry9Da_coleoptera	O06014	1169	73	293	5.6e−26	301	521	1.7e−57	531	668	2.7e−48
cry7Ab_coleoptera_2	Q45707	1138	61	278	2.6e−30	286	487	2.5e−57	497	637	3.7e−41
cry34Ab1_coleoptera[a]	Q939T0	123	–	–	–	–	–	–	–	–	–
cry7Aa_coleoptera	Q03749	1138	63	278	1.5e-27	286	487	2e-59	497	637	1.9e-42
cry35Ab1_coleoptera[b]	Q939S9	383	–	–	–	–	–	–	–	–	–
cry3Bb1_coleoptera	Q06117	652	65	288	1.2e−29	296	502	1.6e−63	512	651	1e−42

Toxin_10: domain position (174–348)
[a]Aegerolysin: domain position (3–118)
[b]RICIN: domain position (4–138)

having toxicity for insect order Lepidoptera. Interestingly, all Cry1A, Cry1D, Cry1E, Cry1H, Cry1F, Cry1G, Cry1C, and Cry1J are fallen in this cluster. However, in cluster II (green color), all Cry proteins show toxicity towards Coleoptera are joined together along with Cry11D, Cry9Ca, Cry1Be, Cry1Ka, Cry1Bd, Cry1Bb, and Cry9Ea. But, Cry9Aa and Cry2Ab targets for Lepidoptera and Cry8Ca targets Coleoptera were obtained as a member of cluster III (blue color) which shows toxicity towards Diptera. Interestingly, lepidopteran target Cry15Aa and coleopteran targets Cry34Ab1 and Cry35Ab1 were seemed as an out-group in the whole tree and were found to be distantly related from all the Cry protein sequences. The Cry proteins for the specific order were found to be diverged from the own group and placed in different insect orders indicate that they might have insecticidal property for more than one insect order. Besides, in order to understand the insect specificity of these Cry proteins, the functional domain regions were also taken for studying the divergence among them. The phylogenetic tree of the domain Endotoxin_N, Endotoxin_M, and Endotoxin_C was constructed using MEGA X software and presented (Fig. 4).

Table 5 List of cry protein domains having toxicity for Diptera insect order

Protein name	Uniprot accession no.	Sequence length	Endotoxin_N		*E* value	Endotoxin_M		*E* value	Endotoxin_C		*E* value
			Start	End		Start	End		Start	End	
cry11Bb_diptera	Q9ZIU5	750	56	240	2.7e−21	–	–	–	–	–	–
cry19Aa_diptera	O32307	648	69	292	4.3e−25	300	502	3.6e−49	512	648	1.2e−31
cry11Ba_diptera	Q45730	724	55	240	1.2e−18	–	–	–	–	–	–
cry4Ba_diptera	P05519	1136	52	268	1.1e−26	283	470	2.1e−40	480	634	2.2e−40
cry10Aa_diptera	P09662	675	95	300	2.9e−23	308	500	5.9e-37	510	650	2.8e−35
cry4Aa_diptera	P16480	1180	80	314	1.3e-21	322	528	2.6e-54	538	678	2.6e-40
cry11Aa_diptera	P21256	643	41	240	5.1e−23	–	–	–	–	–	–
cry27Aa_diptera	Q9S597	826	92	302	6.4e−23	311	499	5.1e−19	524	683	8.5e−34
cry19Ba_diptera	O86170	682	71	280	3.2e−32	293	496	2.2e−48	506	638	9.1e−34
cry20Aa_diptera	O32321	753	67	280	1.7e−22	288	485	3.2e−44	495	631	1.5e−26

Table 6 List of protein domains having toxicity for Lepidoptera insect order

Protein name	Uniprot accession no.	Sequence length	Endotoxin_N		*E* value	Endotoxin_M		*E* value	Endotoxin_C		*E* value
			Start	End		Start	End		Start	End	
cry1Bd_lepidoptera	Q9ZAZ5	1231	59	275	2.4e−26	283	495	2.1e−65	505	643	1.1e−38
cry1Gb_lepidoptera	Q9ZAZ6	1169	45	245	4.7e−22	253	449	4.8e−54	459	596	3.7e−48
cry9Ca_lepidoptera	Q45733	1157	70	290	1.6e−29	298	505	3.8e-56	515	658	6e−41
cry15Aa_lepidoptera[a]	Q45729	340	–	–	–	–	–	–	–	–	–
cry1Aa_lepidoptera_1	P0A369	934	46	251	1e−31	259	460	4.9e−53	470	607	8.9e−47
cry1Ka_lepidoptera	Q45715	1215	56	276	2.7e−25	284	490	6.1e−60	500	637	4.8e−40
cry1Hb_lepidoptera	Q45718	1155	48	248	3.6e−28	256	454	2.3e−58	464	596	3.2e−38
cry1Fb_lepidoptera	O66377	1169	44	249	5.2e−33	257	454	6.2e−54	464	600	1.6e−49
cry1Ab_lepidoptera	P0A370	1155	46	251	8.2e−31	259	461	1.6e−57	471	608	2.6e−46
cry1Aa_lepidoptera_2	P0A366	1176	46	251	2.3e−31	259	460	1e−52	470	607	2.6e−46
cry2Ab_lepidoptera	P21254	633	53	263	2.5e−31	–	–	–	–	–	–
cry1Ac_lepidoptera	P05068	1178	46	251	3.5e−31	259	461	3.6e−58	471	609	1.9e−36
cry1Ea_lepidoptera	Q57458	1171	44	250	3e−28	258	454	1.6e−60	464	601	2.2e−39
cry9Aa_lepidoptera	Q99031	1156	61	287	5.7e−17	295	510	5.3e−46	520	656	1.4e−36
cry1Aa_lepidoptera_3	P0A368	1176	46	251	2.3e−31	259	460	1e−52	470	607	2.6e-46
cry1Ca_lepidoptera_1	P0A375	1189	42	250	9.1e−28	260	457	1.4e−51	467	616	1.5e−37
cry1Ae_lepidoptera	Q03748	1181	46	251	5.5e−29	259	461	3.7e−57	471	608	2.8e−46
cry1Ca_lepidoptera_2	P0A376	1189	42	250	6.2e−28	260	457	5.1e−54	467	616	1.5e−37
cry1Aa_lepidoptera_4	P0A367	1176	46	251	3.4e−31	259	460	1e−52	470	607	2.6e−46
cry1Ad_lepidoptera	Q03744	1179	46	251	9.6e−31	259	460	5.8e−53	470	607	1.2e−44
cry1Da_lepidoptera	P19415	1165	45	250	1.9e−32	258	450	1.4e−54	460	592	7e−39
cry1Eb_lepidoptera	Q03745	1174	42	249	8.5e−31	257	453	2e−59	463	599	1.4e−39
cry1Fa_lepidoptera	Q03746	1174	44	249	8.7e−33	257	454	1.4e−53	464	601	2.2e−43
cry9Ea_lepidoptera	Q9ZNL9	1150	71	293	3.9e−22	301	505	2.6e−47	515	651	2.1e−42
cry1Ga_lepidoptera	Q45746	1166	46	245	1.7e−26	253	446	9.4e−51	456	593	2.9e−49
cry1Id_lepidoptera	Q9XDL1	719	64	279	1.8e−33	287	497	6e−65	507	644	2.3e−48
cry1Be_lepidoptera	O85805	1227	60	275	5.8e−25	283	493	6.9e−65	503	639	5.8e−39
cry1Ha_lepidoptera	Q45748	1172	48	249	8.2e−28	257	455	8.3e−58	465	599	5.6e−35
cry1Jb_lepidoptera	Q45716	1170	46	250	6.9e−27	258	449	8.6e−57	459	596	1.1e−43
cry1Ag_lepidoptera	Q9S515	1176	46	251	2.9e−31	259	454	2e−44	473	607	1.6e−37
cry1Db_lepidoptera	Q45747	1160	45	250	4e−31	258	450	4e−54	460	592	1.7e−38
cry1Ja_lepidoptera	Q45738	1167	46	250	6.8e−27	258	449	1.7e−54	459	595	8.1e−49
cry1Bb_lepidoptera	Q45739	1229	60	275	2.4e−26	283	495	2.1e−65	505	641	6.3e−46

[a]ETX_MTX2: domain position (40–265)

The Cry protein sequence Endotoxin_N domains (55 numbers) was clustered into 3 major groups. The evolutionary analysis (Fig. 4a) depicted that the divergence pattern of all the domain sequences as similar with that of all Cry protein sequences and presented in (Fig. 3). Evolutionary analysis of Endotoxin_M domains (51 numbers) present in Crystal protein target for all the 3 insect orders depicted in (Fig. 4b) showed a similar type of divergence in the phylogenetic tree as that of two predicted tree for Endotoxin_N and total Cry protein sequence. Phylogenetic tree involving Endotoxin_C domains (51 numbers) was revealed that all Cry1 and Cry9 targets for lepidopterans are grouped into a single cluster, except for Cry1C, Cry1B, Cry1E, and Cry1Ac, which are found to be clustered into the second group of Cry protein, which targets the coleopterans. However, all Cry proteins for dipteral were clustered into the same group, except for Cry8Ba, which targets the coleopteran (Fig. 4c).

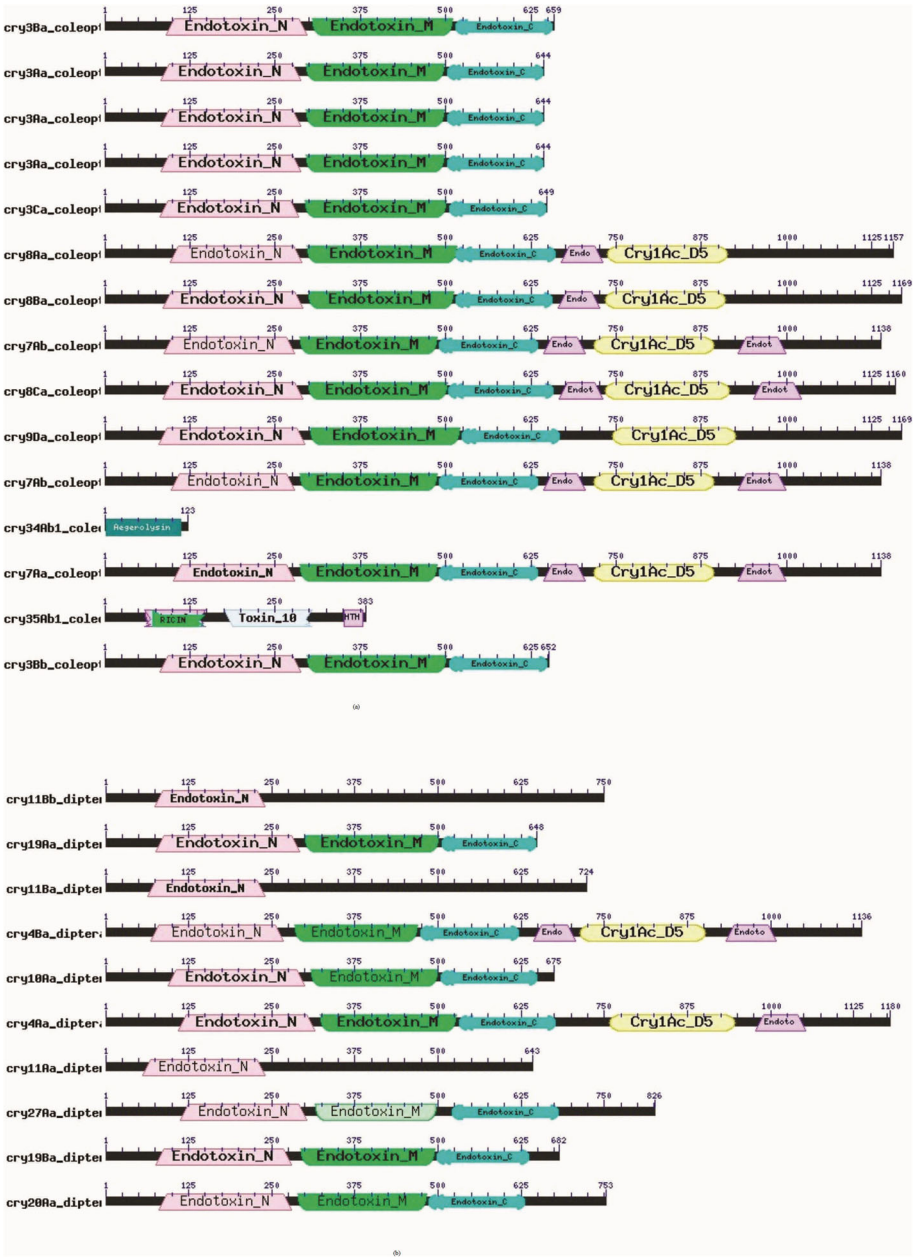

Fig. 1 **a** NCBI batch conserved domain-search result for Coleoptera insect order. **b** NCBI batch conserved domain-search for Diptera insect order

Analysis of Cry protein structure

Structural analysis of Cry protein sequences was performed to understand its function in a better way. It has been observed that most of the Cry proteins do not have experimentally solved 3-dimensional structures in the RCSB PDB, except for Cry3Aa, Cry34Ab1, Cry35Ab1, and Cry3Bb1 for Coleoptera (Fig. 5a), Cry4Ba and Cry4Aa for Diptera (Fig. 5b), and Cry1Aa, Cry1Ac, Cry1Da, Cry1Fa, and Cry1Be for Lepidoptera (Fig. 6a,b). The rest 47 unsolved 3-dimensional structures of Cry proteins belongs to 3 different insect orders predicted through Phyre2 homology modeling server, i.e., 11 numbers of model structure for coleopteran, 08 numbers of model structures for dipterans, and 28 numbers of model structure for lepidopterans. Structural analysis of Cry proteins for coleopteran's group revealed that the proteins having Endotoxin_N, Endotoxin_M, and Endotoxin_C at sequence level corresponding to domain I, domain II, and domain III, respectively, at their structural levels. Domain I regions found to be observed starting from 60 to 300 amino acids at the N-terminal region of the Cry proteins and consists of α-helices (~8

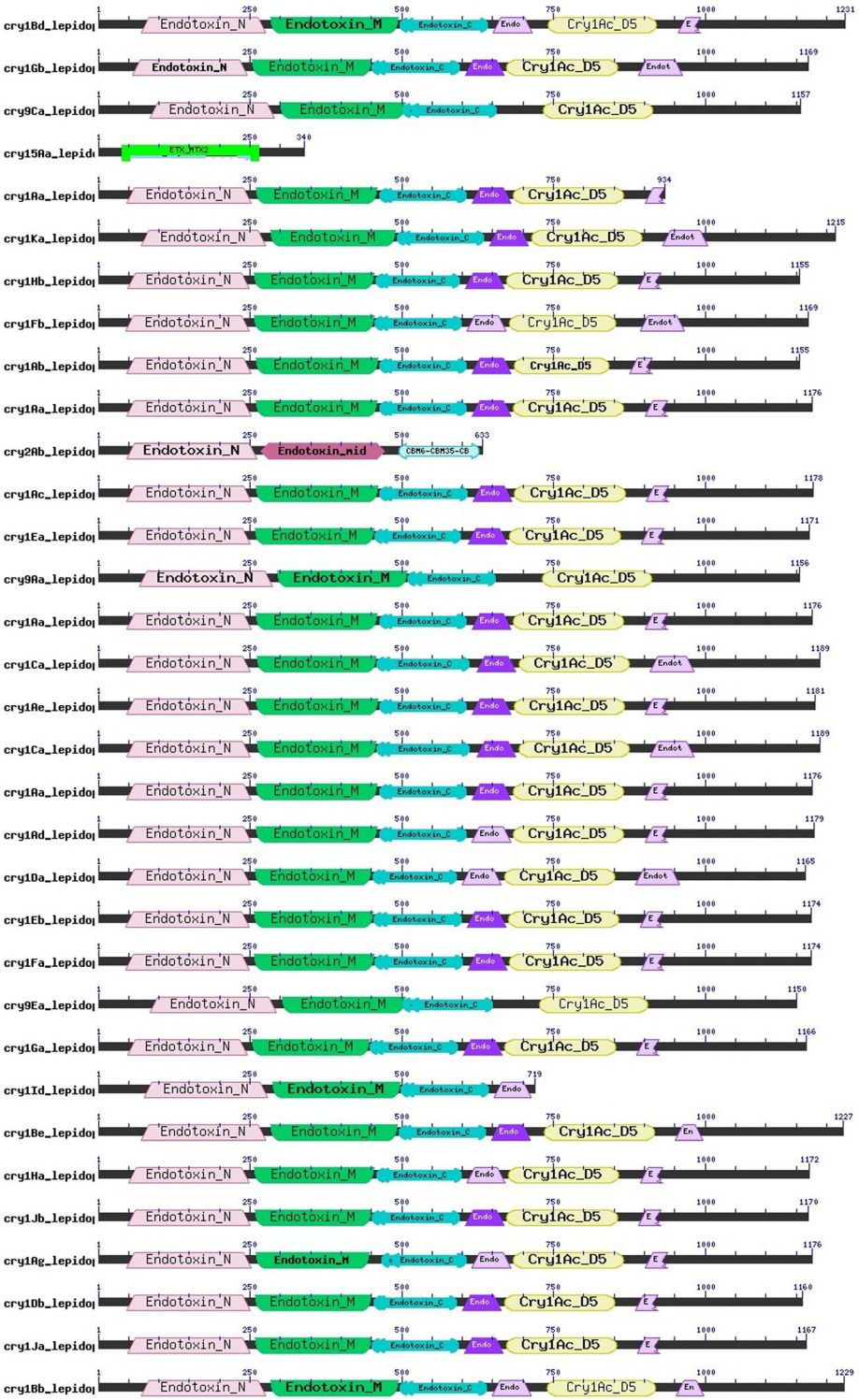

Fig. 2 NCBI batch conserved domain-search for Lepidoptera insect order

numbers) whereas amino acid position 280–510 form domain II which consists of ant parallel β-sheets and short helices. Domain III is a β-sandwich of two ant parallel highly twisted β-sheets and comprises 490–670 amino acid residues, corresponding to Endotoxin_C domain of the Cry protein at the C-terminal region. By doing structure-structure alignment involving both experimental and structural Cry proteins of Coleopteran

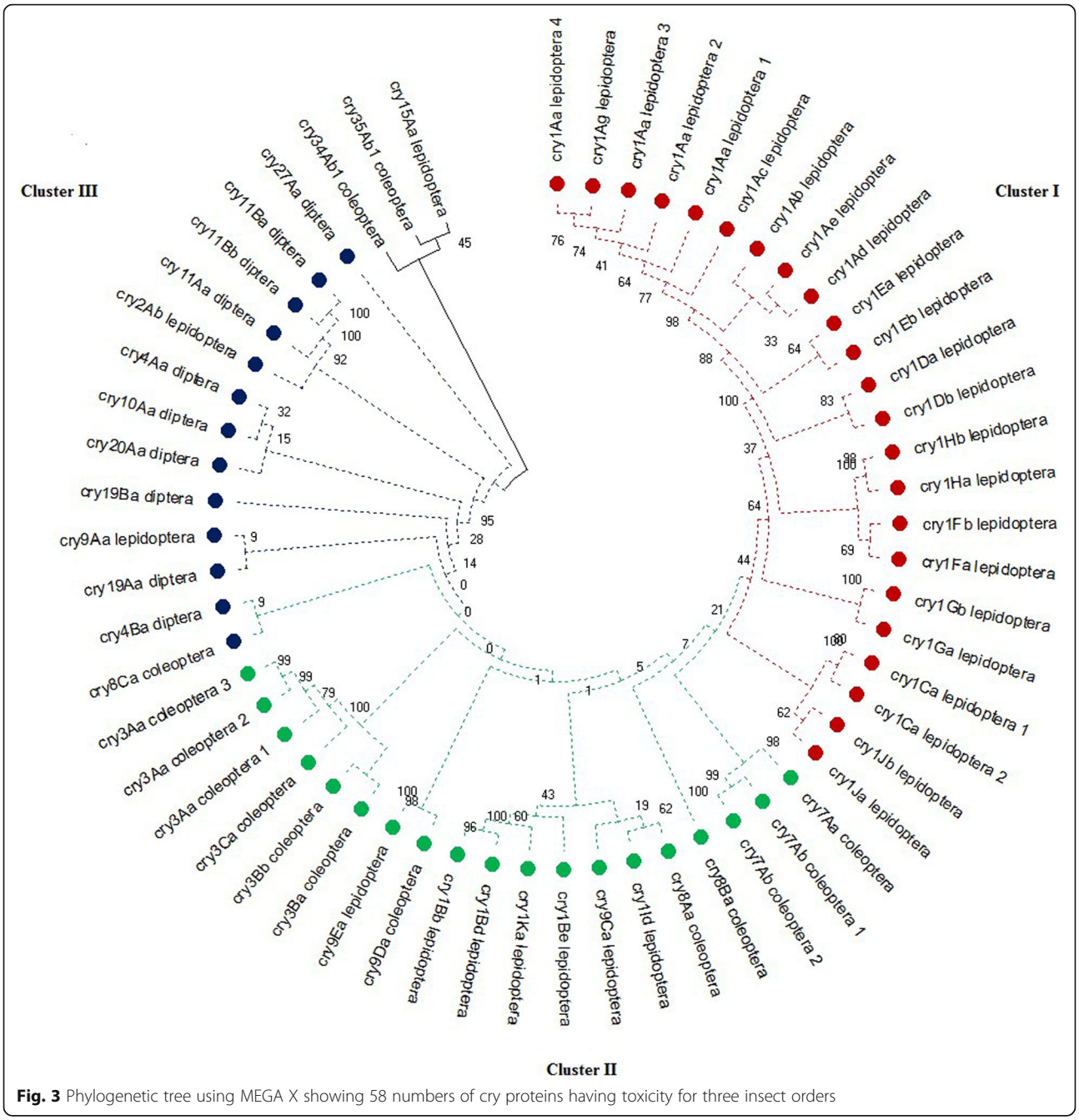

Fig. 3 Phylogenetic tree using MEGA X showing 58 numbers of cry proteins having toxicity for three insect orders

group depicted that all 15 three-dimensional structures for Coleopteran are clustered into two major groups comprising Cry7Aa, Cry9Da, Cry8Ba, Cry8Ca, Cry7Ab (2 numbers), and Cry8Aa in one group whereas Cry3Aa (3 numbers), Cry3Ca, Cry3Ba, and Cry3 in another group whereas Cry34Ab1 and Cry35Ab1 are observed as out-groups (Fig. 7a). The divergence level may be due to the absence of Endotoxin_N, Endotoxin_M, and Endotoxin_C domains. In case of Cry proteins having specificity for Diptera, 8 model structures were predicted such as Cry10Aa, Cry11Aa, Cry11Ba, Cry11Bb, Cry19Aa, Cry19Ba, Cry20Aa, and Cry27Aa. Three-dimensional structural analysis of both experimental (2 numbers) and model (8 numbers) structure revealed that domain I is present in all the protein which extends from residue 40–302 is totally α-helical and contains around 8 numbers of α-helices whereas domain II and domain III span over residue 280–530 and 480–685, respectively, except Cry11Ba and Cry11Aa whereas Endotoxin_M and Endotoxin_C domains are absent. In structural neighbor analysis, all are grouped in single group except Cry4Ba which is found to be an out-group for this group (Fig.

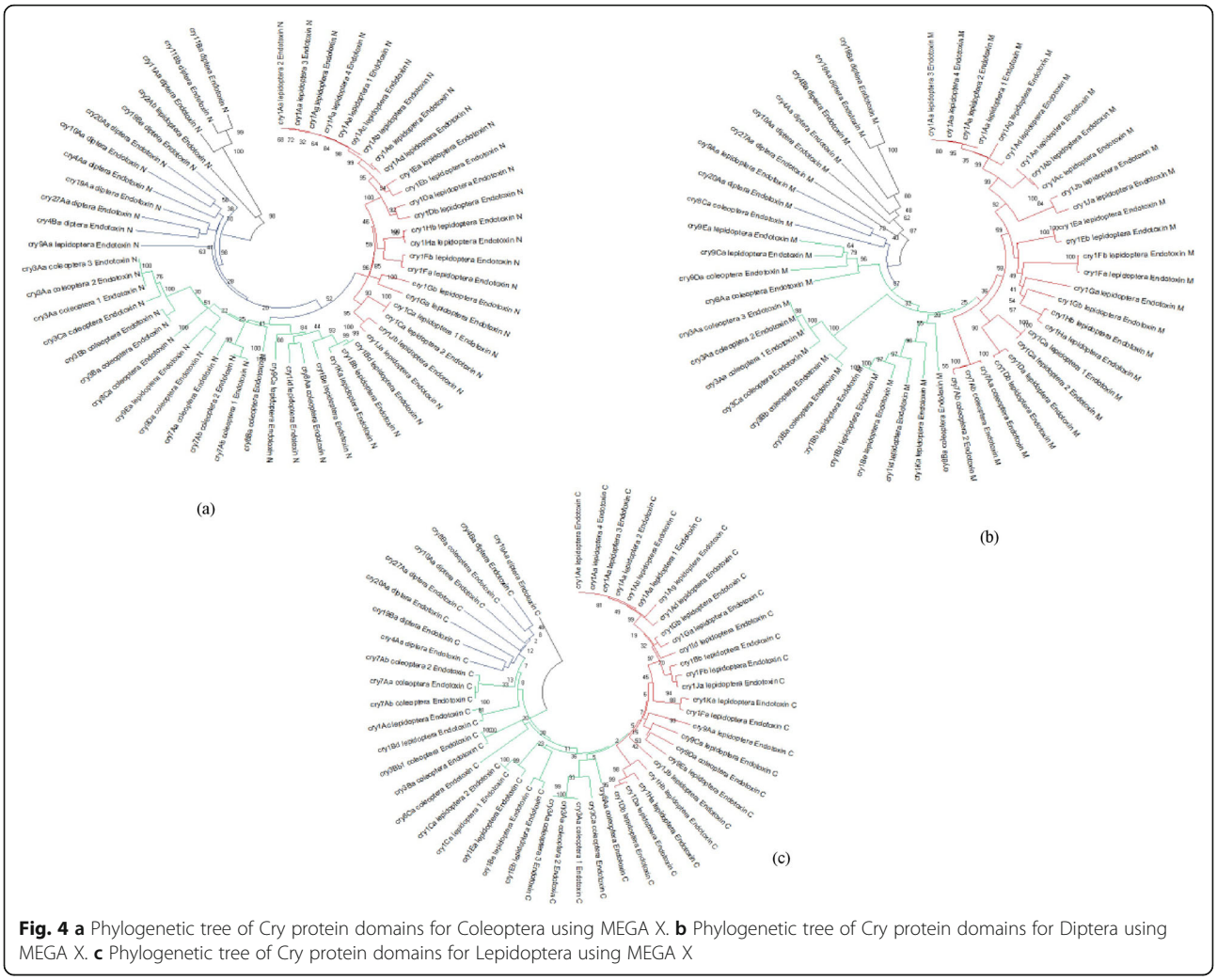

Fig. 4 **a** Phylogenetic tree of Cry protein domains for Coleoptera using MEGA X. **b** Phylogenetic tree of Cry protein domains for Diptera using MEGA X. **c** Phylogenetic tree of Cry protein domains for Lepidoptera using MEGA X

7b). A total of 28 numbers of predicted protein structures having specificity for lepidopteran's group were predicted. It was revealed that Cry1Aa, Cry1Be, Cry1Ba, and Cry1fa have the same type of folding pattern whereas totally different orientation was observed in case of Cry15Aa due to absence of Endotoxin_N, Endotoxin_M, and Endotoxin_C domains. Different domain regions such as domain I, domain II, and domain III extend from 40 to 295, 253–510, and 458–660, respectively, for lepidopteran's group (Table 6). Alignment of all 28 proteins structures was clustered into one major group leaving behind Cry2Ab as distantly related to them whereas Cry15Aa was observed as out-group (Fig. 7c).

B. thuringiensis (*Bt*) strains produce a wide variety of proteins having toxicity against diverse insect orders. These toxins classified into 2 major groups: crystal (Cry) and cytolytic (Cyt). More than 700 Cry gene sequences that code for crystal protein (Cry) have been identified in plasmids by several researchers (Höfte and Whiteley 1989; Schnepf et al. 1998; Van Frankenhuyzen 2009). Many Cry proteins are reported to have useful insecticidal properties for controlling insect pests in agriculture (Sanchis and Bourguet 2008). However, strong cytocidal activities have also been noticed against vertebrates (Palma et al. 2014). Primary protein sequence database search revealed that there are 58 numbers of Cry proteins showing specificity towards the 3 major insect orders: Coleoptera, Diptera, and Lepidoptera (Donovan et al. 2016; Sanchis and Bourguet 2008; Naimov et al. 2008). Sequences of different insect orders showed different sequence's length. All Cry1, Cry2, and Cry15 groups showed toxicity towards lepidopteran insects, whereas Cry9 group showed toxicity towards Lepidoptera and Coleoptera. Similarly, Cry proteins such as Cry3, Cry7, Cry8, Cry34, and Cry35 groups were found to have specificity for the coleopterans. Cry4, Cry10, Cry11, Cry19, Cry20, and Cry27 groups exhibit insecticidal activity against the insect belonging to dipteran's order. The classification of Cry proteins and their insecticidal activity against specific insect orders have been

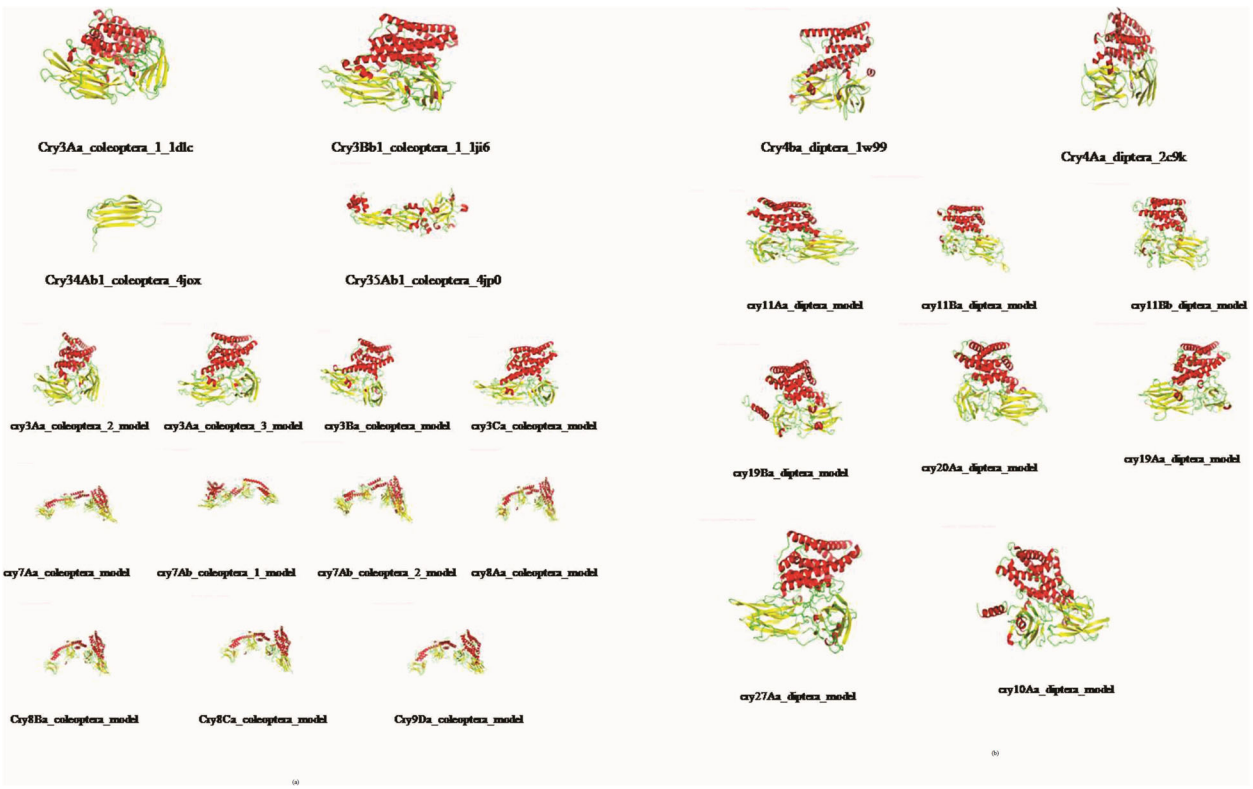

Fig. 5 a Experimental and predicted 3D structures (model) of Cry proteins having toxicity for Coleoptera insects. **b** Experimental and predicted 3D structures (model) of Cry proteins having toxicity for Diptera insects

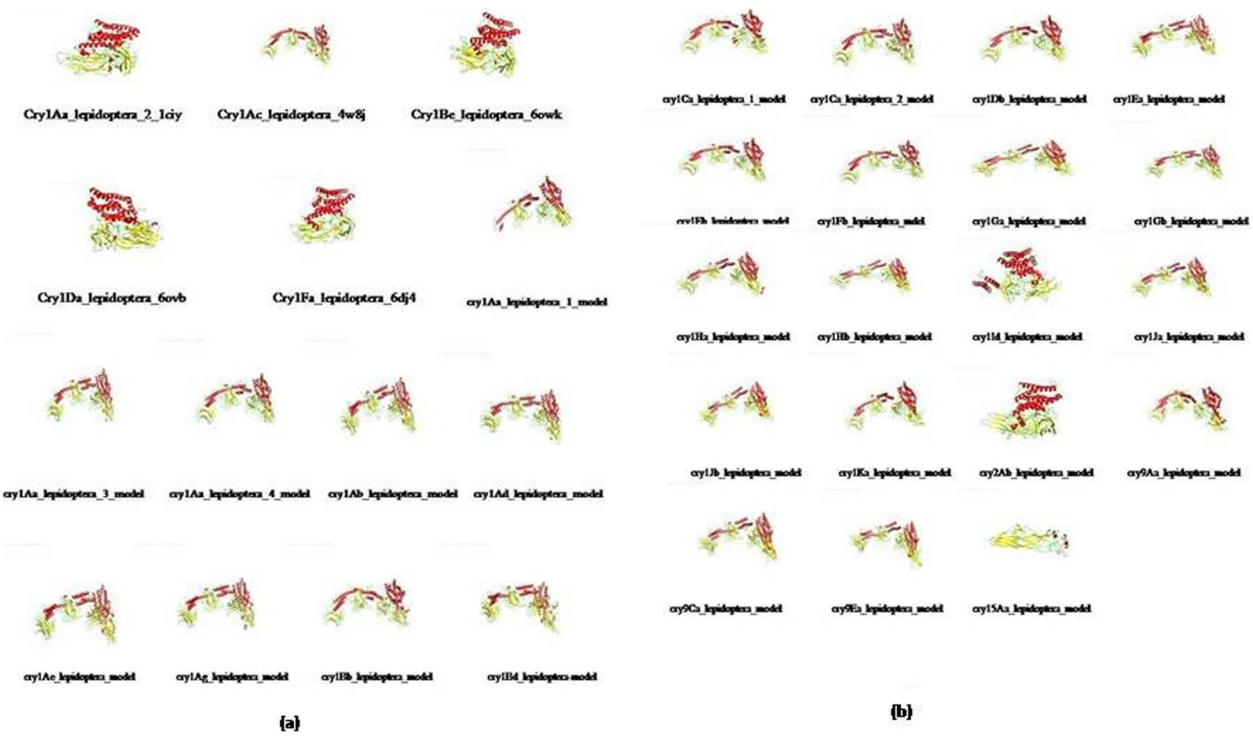

Fig. 6 a, **b** Experimental and predicted 3D structures (model) of Cry proteins having toxicity for Lepidoptera insects

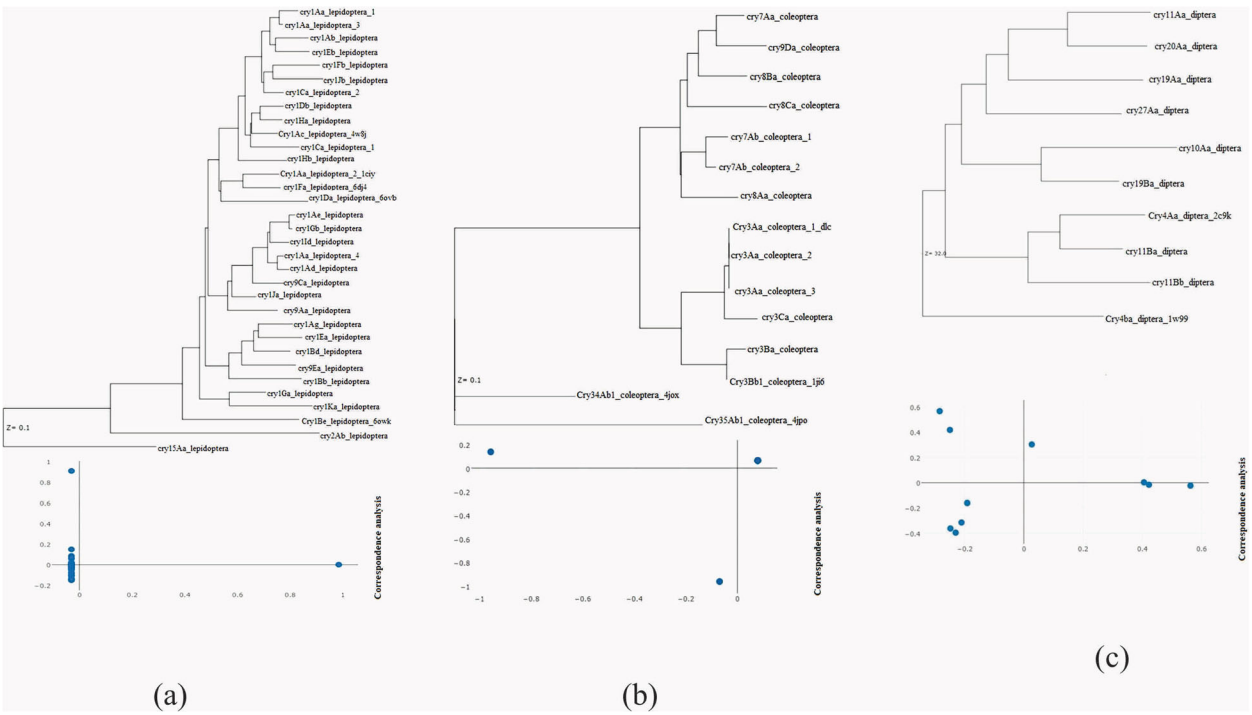

Fig. 7 **a** 3D Structure alignment result using Phyre^2 server of Cry proteins (Coleoptera). **b** 3D Structure alignment result using Phyre^2 server of Cry Proteins (Diptera). **c** 3D Structure alignment result using Phyre^2 server of Cry Proteins (Lepidoptera)

studied by several researchers (Crickmore 2000; de Maagd et al. 2001). Analysis of evolutionary relationship revealed that Cry proteins such as Cry1Id, Cry9Ca, Cry1Be, Cry1Ka, Cry1Bd, Cry1Bb, and Cry9Ea showed toxicity towards lepidopteran's insects also clustered together with the Cry proteins having toxicity for the coleopterans. This finding is supported by (Crickmore 2000). It has been also suggested that these proteins may have toxicity against both insect orders and was later showed for Cry1B toxin (López-Pazos et al. 2009). Similarly, Cry8Ca proteins of coleopteran's group and Cry2Ab and Cry9Aa for lepidopterans are found to be clustered in Diptera insect group. This phylogenetic relationship of whole Cry proteins could not able to reveal how Cry toxin involves in insect specificity. To validate further, phylogenetic analysis of the 3 structural domains such as domain I (Endotoxin_N), domain II (Endotoxin_M), and domain III (Endotoxin_C) were carried out independently. Divergence pattern were observed similar in case of both sequence level and structural level with minor fluctuation.

As per literature, the domain swapping of different Cry toxin is likely to be an active evolutionary process for determining insect specificity (de Maagd et al. 2001). Three-dimensional X-ray crystallography structures of several Cry toxins of *Bt* have been reported in this connection. A total of 11 numbers of three-dimensional structures of the Cry proteins were retrieved from RCSB PDB out of which 4 proteins for Coleoptera, 2 for Diptera, and 5 for Lepidoptera. Homology model structure for 47 Cry proteins were predicted using web based Phyre2 tool to study the structural arrangements of 3 different domains. All experimental and predicted models of Cry proteins revealed that the domain I is present in N-terminal region having α-helices, domain II consists of three antiparallel β-sheets, and domain III consists of two twisted anti-parallel β-sheets forming a sandwich. This observations and findings are also supported by different published reports (Grochulski et al. 1995; de Maagd et al. 2001, and Gouet et al. 2003). Structure-structure alignment was also carried out involving all the 58 numbers of Cry proteins (experimental and model 3-dimensional structures) in order to understand the divergence among each other. In case of Coleoptera insect order, 3-dimensional structures of Cry34Ab1 and Cry35Ab1 were observed as out-group. In case of Diptera, 3-dimensional structures of Cry4Ba were found diverged from other Cry proteins whereas Cry2Ab protein of Lepidoptera group noticed to be diverged from other group of proteins. However, Cry15aa was predicted as an out-group in compare to other Cry proteins.

Conclusions

Bacillus thuringiensis (Bt) synthesizes various insecticidal proteins and thus recommended as potential bio control agent against various insect pests in agriculture. The evolution and diversification of these Cry proteins have

been studied extensively by various researchers to discover the existence of important determinants, which confer insect specificity for improvement of its insecticidal activity. There are total 58 numbers of different Cry protein groups belong to major three insect orders: Coleoptera, Diptera, and Lepidoptera were retrieved and analyzed both at structural and sequence level. Structural and functional analysis was performed to understand the domain arrangements at sequence and structural level involving both experimental and predicted 3D models. Cry proteins having toxicity for a specific insect order are grouped accordingly. Three-dimensional structure analysis of both experimental and predicted models revealed that the Cry proteins might have toxicity for a specific insect order differ in their structural arrangements and is observed in 3 different groups. It could be hypothesized that an inner-molecular domain shift or domain insertion/deletion might have taken place during the evolutionary process, which consequently causes structural and functional divergence of *Bt*. These findings lead to understand the wide diversity of insecticidal proteins and their application as biopesticides in agriculture.

Abbreviations

Bt: *Bacillus thuringiensis*; PFTs: Pore-forming toxins; 3D: Three-dimensional; CADs: Cadherin-like proteins; APN: Aminopeptidase N; UniProt: Universal Protein Resource; NCBI : National Center for Biotechnology Information; CD-Search: Conserved Domain-Search; RCSB-PDB: Research Collaboratory for Structural Bioinformatics-Protein Data Bank; MEGA X: Molecular Evolutionary Genetics Analysis X; Phyre^2: Protein Homology/analogY Recognition Engine V 2.0

Acknowledgements

Financial assistance to the Center of Excellence in Environment and Public Health by the Higher Education Department, Government of Odisha under OHEPEE is grateful acknowledged (HE-PTC-WB-02017).

Authors' contributions

SKD: Research concept and design, collection and/or assembly of data, data analysis and interpretation, writing the article, and critical revision of the article. SKP: Research concept and design, and data analysis and interpretation. KCS: Data analysis and interpretation, and critical revision of the article. NRS: Research concept and design. The authors read and approved the final manuscript.

Funding

Not applicable.

Availability of data and materials

The datasets used and/or analyzed during the current study are available in the UniProtKB (The Universal Protein Resource Knowledgebase).

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]Department of Bioinformatics, Centre for Post Graduate Studies, Odisha University of Agriculture & Technology, Bhubaneswar, Odisha 751003, India. [2]School of Life Sciences, Ravenshaw University, Cuttack, Odisha 753003, India. [3]Department of Agricultural Biotechnology, College of Agriculture, Odisha University of Agriculture & Technology, Bhubaneswar, Odisha 751003, India. [4]Department of Botany and Center of Excellence in Environment and Public Health, Ravenshaw University, Cuttack, Odisha 753003, India.

Received: 27 August 2020 Accepted: 17 February 2021
Published online: 01 March 2021

References

Bayyareddy K, Andacht TM, Abdullah MA, Adang MJ (2009) Proteomic identification of *Bacillus thuringiensis* subsp. israelensis toxin Cry4Ba binding proteins in midgut membranes from *Aedes aegypti*Linnaeus (Diptera, Culicidae) larvae. Insect Biochem Mol Biol 39(4):279–286. https://doi.org/10.1016/j.ibmb.2009.01.002

Boonserm P, Davis P, Ellar DJ, Li J (2005) Crystal structure of the mosquito-larvicidal toxin Cry4Ba and its biological implications. J Mol Biol 348(2):363–382. https://doi.org/10.1016/j.jmb.2005.02.013

Boonserm P, Mo M, Angsuthanasombat C, Lescar J (2006) Structure of the functional form of the mosquito larvicidal Cry4Aa toxin from *Bacillus thuringiensis* at a 2.8-angstrom resolution. J Bacteriol 188(9):3391–3401. https://doi.org/10.1128/JB.188.9.3391-3401.2006

Bravo A, Gill SS, Soberón M (2007) Mode of action of *Bacillus thuringiensis* Cry and Cyt toxins and their potential for insect control. Toxicon 49(4):423–435. https://doi.org/10.1016/j.toxicon.2006.11.022

Bravo A, Gómez I, Porta H, García-Gómez BI, Rodrigue-Almazan C, Pardo L, Soberón M (2013) Evolution of *Bacillus thuringiensis* Cry toxins insecticidal activity. J Microbial Biotechnol 6(1):17–26. https://doi.org/10.1111/j.1751-7915.2012.00342.x

Bravo A, Likitvivatanavong S, Gill SS, Soberón M (2011) *Bacillus thuringiensis*: a story of a successful bioinsecticide. Insect Biochem Mol Biol 41(7):423–431. https://doi.org/10.1016/j.ibmb.2011.02.006

Choma CT, Kaplan H (1990) Folding and unfolding of the protoxin from *Bacillus thuringiensis*: evidence that the toxic moiety is present in an active conformation. Biochemistry 29(49):10971–10977. https://doi.org/10.1021/bi00501a015

Crickmore N (2000) The diversity of *Bacillus thuringiensis* δ-endotoxins. In: Entomopathogenic bacteria: from laboratory to field application. Springer, Dordrecht, pp 65–79

Crickmore N, Zeigler DR, Feitelson J, Schnepf ES, Van Rie J, Lereclus D, Baum J, Dean DH (1998) Revision of the nomenclature for the *Bacillus thuringiensis* pesticidal crystal proteins. Microbiol Mol Biol Rev 62(3):807–813. https://doi.org/10.1128/MMBR.62.3.807-813.1998

Crickmore N, Zeigler DR, Schnepf E, Van Rie J, Lereclus D, Baum J (2011) *Bacillusthuringiensis* toxin. [WWWdocument].URL http://www.lifesci.sussex.ac.uk/Home/Neil_Crickmore/Bt/.

de Maagd RA, Bravo A, Crickmore N (2001) How *Bacillus thuringiensis* has evolved specific toxins to colonize the insect world. Trends Genet 17(4):193–199. https://doi.org/10.1016/S0168-9525(01)02237-5

Derbyshire DJ, Ellar DJ, 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(12):1938–1944. https://doi.org/10.1107/S090744490101040X

Donovan WP, Engleman JT, Donovan JC, Baum JA, Bunkers GJ, Chi DJ, Clinton WP, English L, Heck GR, Ilagan OM, Krasomil-Osterfeld KC (2016) Discovery and characterization of Sip1A: a novel secreted protein from *Bacillus thuringiensis* with activity against coleopteran larvae. Appl Microbiol Biotechnol 72(4):713–719. https://doi.org/10.1007/s00253-006-0332-7

Evdokimov AG, Moshiri F, Sturman EJ, Rydel TJ, Zheng M, Seale JW, Franklin S (2014) Structure of the full length insecticidal protein Cry1 A c reveals intriguing details of toxin packaging into in vivo formed crystals. Protein Sci 23(11):1491–1497. https://doi.org/10.1002/pro.2536

Galitsky N, Cody V, Wojtczak A, Ghosh D, Luft JR, Pangborn W, English L (2001) Structure of the insecticidal bacterial δ-endotoxin Cry3Bb1 of *Bacillus thuringiensis*. Acta Crystallogr D Biol Crystallogr 57(8):1101–1109. https://doi.org/10.1107/S0907444901008186

Gouet P, Robert X, Courcelle E (2003) ESPript/ENDscript: extracting and rendering sequence and 3D information from atomic structures of proteins. Nucleic Acids Res 31(13):3320–3323. https://doi.org/10.1093/nar/gkg556

Griffitts JS, Haslam SM, Yang T, Garczynski SF, Mulloy B, Morris H, Cremer PS, Dell A, Adang MJ, Aroian RV (2005) Glycolipids as receptors for *Bacillus thuringiensis* crystal toxin. Science 307(5711):922–925. https://doi.org/10.1126/science.1104444

Grochulski P, Masson L, Borisova S, Pusztai-Carey M, Schwartz JL, Brousseau R, Cygler M (1995) *Bacillus thuringiensis* CryIA (a) insecticidal toxin: crystal structure and channel formation. J Mol Biol 254(3):447–464. https://doi.org/10.1006/jmbi.1995.0630

Guo S, Ye S, Liu Y, Wei L, Xue J, Wu H, Song F, Zhang J, Wu X, Huang D, Rao Z (2009) Crystal structure of *Bacillus thuringiensis* Cry8Ea1: an insecticidal toxin toxic to underground pests, the larvae of *Holotrichia parallela*. J Struct Biol 168(2):259–266. https://doi.org/10.1016/j.jsb.2009.07.004

Höfte H, Whiteley HR (1989) Insecticidal crystal proteins of *Bacillus thuringiensis*. Microbiol Mol Biol Rev 53(2):242–255

Hui F, Scheib U, Hu Y, Sommer RJ, Aroian RV, Ghosh P (2012) Structure and glycolipid binding properties of the nematicidal protein Cry5B. Biochemistry 51(49):9911–9921. https://doi.org/10.1021/bi301386q

Kelker MS, Berry C, Evans SL, Pai R, McCaskill DG, Wang NX, Russell JC, Baker MD, Yang C, Pflugrath JW, Wade M (2014) Structural and biophysical characterization of *Bacillus thuringiensis* insecticidal proteins Cry34Ab1 and Cry35Ab1. PLoS One 9(11):e112555. https://doi.org/10.1371/journal.pone.0112555

Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol 35(6):1547–1549. https://doi.org/10.1093/molbev/msy096

Li J, Carroll J, Ellar DJ (1991) Crystal structure of insecticidal δ-endotoxin from *Bacillus thuringiensis*at 2.5 Å resolution. Nature 353(6347):815–821. https://doi.org/10.1038/353815a0

López-Pazos SA, Gómez JE, Salamanca JA (2009) Cry1B and Cry3A are active against *Hypothenemus hampei* Ferrari (coleoptera: scolytidae). J Invertebr Pathol 101(3):242–245

McNall RJ, Adang MJ (2003) Identification of novel *Bacillus thuringiensis* Cry1Ac binding proteins in *Manduca sexta* midgut through proteomic analysis. Insect Biochem Mol Biol 33(10):999–1010. https://doi.org/10.1016/S0965-1748(03)00114-0

Mizuki E, Park YS, Saitoh H, Yamashita S, Akao T, Higuchi K, Ohba M (2000) Parasporin, a human leukemic cell-recognizing parasporal protein of *Bacillus thuringiensis*. Clin Diagn Lab Immunol 7(4):625–634. https://doi.org/10.1128/CDLI.7.4.625-634.2000

Morse RJ, Yamamoto T, Stroud RM (2001) Structure of Cry2Aa suggests an unexpected receptor binding epitope. Structure 9(5):409–417. https://doi.org/10.1016/S0969-2126(01)00601-3

Naimov S, Boncheva R, Karlova R, Dukiandjiev S, Minkov I, de Maagd RA (2008) Solubilization, activation, and insecticidal activity of *Bacillus thuringiensis* serovarthompsoni HD542 crystal proteins. Appl Environ Microbiol 74(23):7145–7151. https://doi.org/10.1128/AEM.00752-08

Palma L, Muñoz D, Berry C, Murillo J, Caballero P (2014) *Bacillus thuringiensis* toxins: an overview of their biocidal activity. Toxins 6(12):3296–3325. https://doi.org/10.3390/toxins6123296

Pigott CR, Ellar DJ (2007) Role of receptors in *Bacillus thuringiensis* crystal toxin activity. Microbiol Mol Biol Rev 71(2):255–281. https://doi.org/10.1128/MMBR.00034-06

Raymond B, Johnston PR, Nielsen-LeRoux C, Lereclus D, Crickmore N (2010) *Bacillus thuringiensis*: an impotent pathogen? Trends Microbiol 18(5):189–194. https://doi.org/10.1016/j.tim.2010.02.006

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Molecular biology and evolution 4(4):406–425. https://doi.org/10.1093/oxfordjournals.molbev.a040454

Sanchis V, Bourguet D (2008) *Bacillus thuringiensis*: applications in agriculture and insect resistance management. A review. Agronom Sustain Devs 28(1):11–20. https://doi.org/10.1051/agro:2007054

Schnepf E, Crickmore NV, Van Rie J, Lereclus D, Baum J, Feitelson J, Zeigler DR, Dean D (1998) *Bacillus thuringiensis* and its pesticidal crystal proteins. Microbiol Mol Biol Rev 62(3):775–806. https://doi.org/10.1128/MMBR.62.3.775-806.1998

Schnepf HE (1995) *Bacillus thuringiensis* toxins: regulation, activities and structural diversity. Curr Opin Biotechnol 6(3):305–312. https://doi.org/10.1016/0958-1669(95)80052-2

Soberon M, Gill SS, Bravo A (2009) Signaling versus punching hole: How do *Bacillus thuringiensis* toxins kill insect midgut cells? Cell Mol Life Sci 66(8):1337–1349. https://doi.org/10.1007/s00018-00-8330-9

Van Frankenhuyzen K (2009) Insecticidal activity of *Bacillus thuringiensis* crystal proteins. J Invertebr Pathol 101(1):1–6. https://doi.org/10.1016/j.jip.2009.02.009

Wang C, Li W, Kessenich CR, Petrick JS, Rydel TJ, Sturman EJ, Lee TC, Glenn KC, Edrington TC (2018) Safety of the *Bacillus thuringiensis*-derived Cry1A. 105 protein: evidence that domain exchange preserves mode of action and safety. Regul Toxicol Pharmacol 99:50–60. https://doi.org/10.1016/j.yrtph.2018.09.003

Wang Y, Wang J, Fu X, Nageotte JR, Silverman J, Bretsnyder EC, Chen D, Rydel TJ, Bean GJ, Li KS, Kraft E (2019) Novel receptor interaction of *Bacillus thuringiensis* Cry1Da_7 and Cry1B. 868 proteins allow control of resistant fall armyworm, *Spodoptera frugiperda* (JE Smith). Appl Environ Microbiol AEM-00579. https://doi.org/10.1128/AEM.00579-19

Zuckerkandl E, Pauling L (1965) Evolutionary divergence and convergence in proteins. Evolving Genes Proteins 1965:97–166. https://doi.org/10.1016/B978-1-4832-2734-4.50017-6

Publisher's Note