BIOINFORMATION
Discovery at the interface of physical and biological sciences

open access www.bioinformation.net Hypothesis

Volume 9(15)

In silico identification of cross affinity towards Cry1Ac pesticidal protein with receptor enzyme in *Bos taurus* and sequence, structure analysis of crystal proteins for stability

King Solomon Ebenezer[1], Ramesh Nachimuthu[2], Prabha Thiagarajan[3] & Rajesh Kannan Velu[1]*

[1]Rhizosphere Biology Laboratory, Department of Microbiology, Bharathidasan University, Tiruchirappalli - 620024, TamilNadu, India; [2]School of Biosciences and Technology, Vellore Institute of Technology, Vellore-632014, TamilNadu, India; [3]Department of Microbiology, JJ College of Arts and Science, Pudukkottai - 622422, TamilNadu, India; Rajesh Kannan Velu – Email: uvrajesh@gmail.com; *Corresponding author

Received July 09, 2013; Revised July 15, 2013; Accepted July 16, 2013; Published August 28, 2013

Abstract:
Any novel protein introduced into the GM crops need to be evaluated for cross affinity on living organisms. Many researchers are currently focusing on the impact of *Bacillus thuringiensis* cotton on soil and microbial diversity by field experiments. In spite of this, *in silico* approach might be helpful to elucidate the impact of *cry* genes. The crystal a protein which was produced by *Bt* at the time of sporulation has been used as a biological pesticide to target the insectivorous pests like Cry1Ac for *Helicoverpa armigera* and Cry2Ab for *Spodoptera* sp. and *Heliothis* sp. Here, we present the comprehensive *in silico* analysis of Cry1Ac and Cry2Ab proteins with available *in silico* tools, databases and docking servers. Molecular docking of Cry1Ac with procarboxypeptidase from *Helicoverpa armigera* and Cry1Ac with Leucine aminopeptidase from *Bos taurus* has showed the 125th amino acid position to be the preference site of Cry1Ac protein. The structures were compared with each other and it showed 5% of similarity. The cross affinity of this toxin that have confirmed the earlier reports of ill effects of *Bt* cotton consumed by cattle.

Keywords: *Bacillus thuringiensis*, Crystal protein, Docking, Phylogenetic analysis, Persistence, Cross affinity.

Background:
Cry proteins of *Bacillus thuringiensis*, a Gram positive soil bacteria has been used to control insects by transgenic *Bt* Cotton approach. There are numerous strains of *B. thuringiensis,* each toxin was highly selective to *Lepidoptera, Coleoptera and Diptera*. Genetically Modified crops have been accepted as animal feed in several countries like European countries and Japan **[1].** Prior to the release of Boll Bollgard cottons I & II in market, the food, feed and environmental safety were evaluated by standard regulatory agencies **[2].** But,*Bt* cotton has some risks on environment and to living organisms. A very few *in vitro*, *in vivo* and *in silico* research on the Cry proteins were reported so far based on safety assessment and impact of *Bt* crops over living organisms.

There has been a previous *in silico* study of the motif regions in aminopeptidase N which is a receptor for Cry1Ac **[3].** There were many subtypes of crystal proteins; out of these the three Cry proteins namely Cry1A, Cry1F and Cry1J formed a cluster for domain II. Moreover, it was found that these proteins have similarities in the domain II region **[4].** Here, we report inclusive and exhaustive sequence and structure based analysis on the Cry1Ac and Cry2Ab proteins of the *B. thuringiensis* for *in silico* based identification of cross affinity of Cry1Ac toxin with receptor enzyme in cow *(Bos taurus*) was performed.

Methodology:
Sequence
Cry1Ac (ABY83188.1) and Cry2Ab (ABC95996.1) protein sequences were retrieved from NCBI protein database. BlastP

was used to search most of the similar protein sequences. Protparam tool was used to predict physico & chemical and bio-chemical properties. PSORTb v3.0.2 was used to predict the cellular localization of the proteins. PRED-CLASSv 1.0 was used to classify the proteins. Conserved Domain Database (CDD), a protein annotation database was used to identify conserved regions, motifs and family of the proteins **[5].**

Structure

The biological macromolecular protein structures were obtained by sequences from RCSB protein databank. Protein structure server SWISS-MODEL was used for homology-based modeling. Mistral web server was used for determining the p-Value, Z_score and sequence identity of the compared structures **[6].**

Docking

To understand the cross affinity in the active site region for Cry1Ac with procarboxypeptidase from *H. armigera* and Leucine aminopeptidase from *Bos taurus* were performed using the Hex, Autodock and ClusPro, online docking protein-ligand servers **[7].**

Results & Discussion:

Sequence analysis

Meta-analyses was performed to compare non-target invertebrates which were exposed to *Bacillus thuringiensis* (Bt) Cry proteins under laboratory studies and the results were showed of non-target invertebrates from independent field studies **[8].** BlastP analysis of the Cry1Ac protein sequence (ABY83188.1) showed maximum identity with insecticidal delta endotoxin [*B.thuringiensis serovar kurstaki*] (AAA73077.1), pesticidal crystal protein cry1Ac [*B.thuringiensis serovar kurstaki str* T03a001], Cry1Ac [*B. thuringiensis*] and toxin crystal protein [*B.thuringiensis*] respectively. Cry2Ab protein (ABC95996.1) showed maximum identity with pesticidal crystal protein Cry2Ab [*B.thuringiensis serovar kurstaki str*]. There has been an earlier report of Blastp analysis of *Cry2* genes of *B.thuringiensis* isolated from soil **[9].**

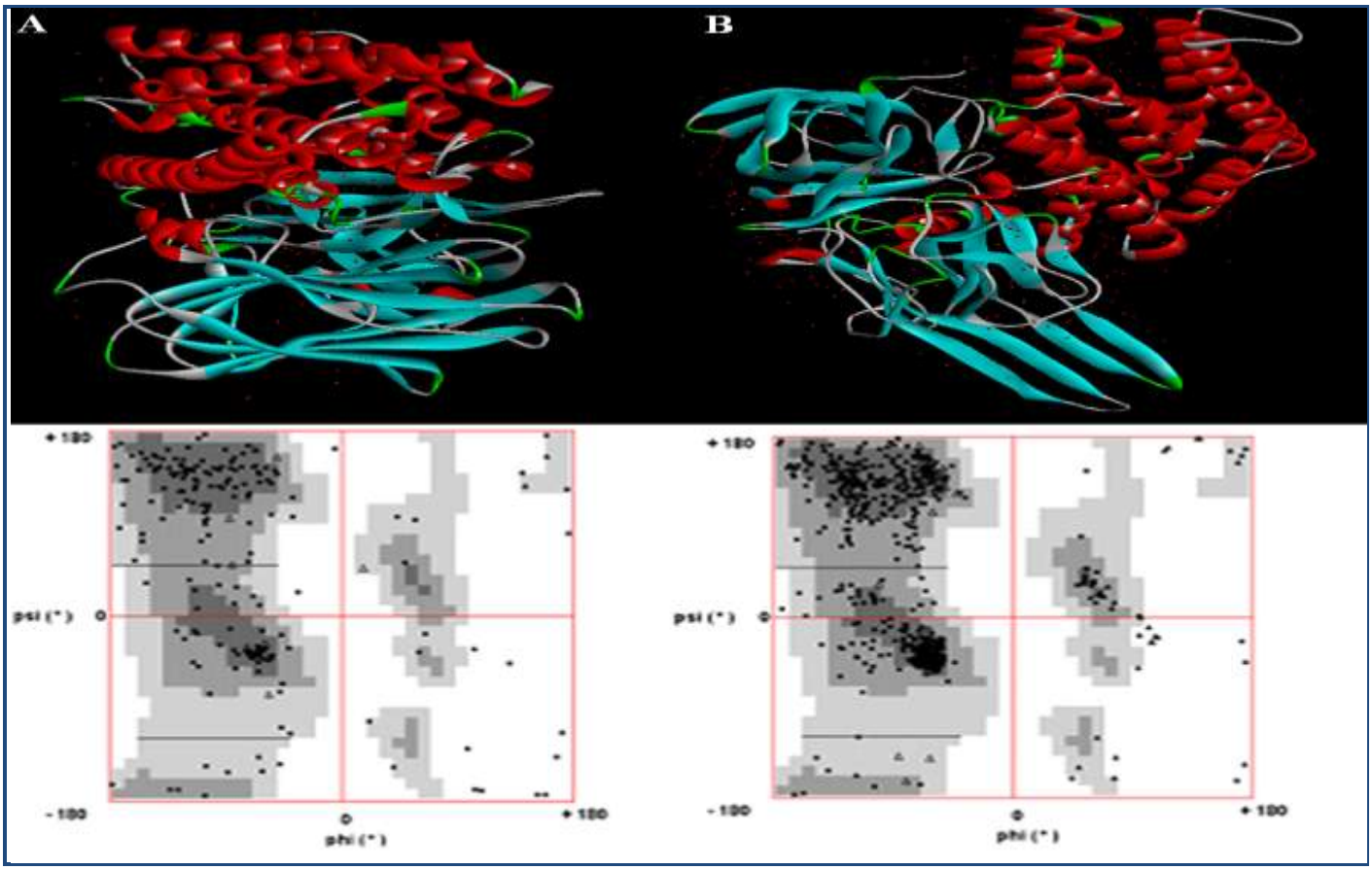

Figure 1: Three dimensional structures of Cry1Ac and Cry 2Ab. **A)** The template used for the structure prediction of Cry1Ac was 1CIYA- Cry1A insecticidal toxin**; B)** The template used for the structure predicition of Cry2Ab proteins was 1I5P- Cry2Aa insecticidal crystal protein. The structure visualization and modeling was done using the Discovery Studio from Accelerys software. A Ramachandran plot, against a background of phi-psi probabilities was constructed. Black shaded portion indicate fully Allowed region. Dark grey shaded portion indicate additionally allowed region, light grey shaded portion indicate generously Allowed region. Black triangles indicate glycine residues, white triangles indicate proline residues, and small black squares indicate other residues.

The stability index of Cry1Ac (35.80) was high compared to Cry2ab (31.42), but aliphatic index was high in Cry2ab (83.92) than Cry1ac (81.32) respectively. More or less both Cry proteins have very similar stability and aliphatic index value. Grand average of hydropathicity index values (-0.408 and -0.287) indicate their solubility and hydrophilic nature as disulphide bridges play an important role in determining the thermostability of the proteins, so it indicate that cry proteins

has less number of disulphide bonds in the structure. This may be the reason of insecticidal crystal proteins persist in the rhizosphere soil. The insecticidal activity of *B.thuringiensis serovar kurstaki* in soil even after persists for 234 days and their aliphatic index values of Cry proteins were high thermo stable **[10].** There has been reported earlier that the *Bt* crystal proteins were mostly located in cytoplasm and distribution density was larger in the cytoplasm than in the intercellular space of *Bt* transgenic cotton **[11].** Here in our study the sub-cellular localization prediction indicated that Cry1Ac is an extracellular protein. Cry1Ac and Cry2Ab were predicted to be membrane and globular proteins respectively. CDD analysis of Cry1Ac has showed the presence of putative metal binding site, a delta_endotoxin_C domain of Cry1Ac. Cry2Ab has at its C terminal, a delta_ endotoxin domain. This domain has been reported earlier with carbohydrate binding function. This region of delta endotoxin has 3 structural domains. Domain I is responsible for pore formation in the target insect, the remaining II and III domains are involved in receptor binding and Carbohydrate binding module region (CBM6). Functional partner analysis of Cry proteins indicated that β-Lactamase domain containing protein as their functional partner based upon the Neighborhood parameters in the STRING. There has been an earlier report of expression of β-lactamase in the culture supernatant of Cry negative strain of *B.thuringiensis* **[12].** Functional homologs search of Cry1Ac and Cry2Ab proteins were found to be present in *Brevibacillus brevis and Bacillus weihenstephanensis*.

Structure analysis

Structure of Cry1Ac and Cry2Ab proteins were predicted by homology based approach **(Figure1A & 1B).** The templates were used for the structure prediction of Cry1Ac and Cry2Ab proteins. The templates were used for the structure prediction of Cry1Ac and Cry2Ab proteins were 1CIYA- CryIA(a) insecticidal toxin **[13]** and 1I5P- Cry2Aa insecticidal crystal protein **[14].** Aligned region, percentage of identity and QMean Z_score were found to be significant. Ramachandran plot of Cry1Ac protein indicated fully aligned region of 110 residues (55.56%), additionally allowed region of 49 residues (24.75%), generously allowed region of 26 residues (13.13%) and outside region of 13 residues (6.57%). α-helix region comprise of 35 residues (31.02%) and β-sheet region comprise of 75 residues (68.18%). Ramachandran plot of Cry2Ab protein indicated fully aligned region of 467 residues (74.01%), additionally allowed region of 120 residues (19.02%), generously allowed region of 25 residues (3.96%) and outside region of 19 residues (3.01%). α-helix region comprise of 245 residues (52.46%) and β-sheet region comprise of 214 residues (45.82%). The predicted structures of Cry1Ac and Cry2Ab proteins were compared and the p-value was 5.04^{e-01} with Z_score as -0.2. The homology modeled phylogenetic tree of Cry1Ac and Cry2Ab proteins were constructed based on the PDB structure of Cry1Ac (1CIY_A). Cry1Ac forms a clade with 3EB7_A (Chain A, crystal structure of insecticidal delta-endotoxin Cry8eA1 from *B.thuringiensis*). Cry2ab forms a Clade with 3EB7_A (Chain A, crystal structure of insecticidal delta-endotoxin Cry8eA1 from *B.thuringiensis*).

Docking

A protein–protein docking study is based on the number of hydrogen bonds, hydrophobic interactions, aromatic–aromatic, aromatic–sulphur, cation–pi interactions and binding energy. A previous report had showed that Cry1Ac fusion proteins with the aminopeptidase N (APN) of *Manduca sexta* rationalized the higher binding affinity of the fusion protein with the APN receptor was compared **[15].** Docking of tertiary structure of Cry1Ac protein with 1JQG - crystal structure of procarboxypeptidase from *H. armigera* was performed **[16].** Here the docking of tertiary structure of Cry1Ac protein with 1LAM Leucine aminopeptidase (*Bos taurus*) **(Figure 2A & 2B)** was performed and the active site for former at Try 125 and latter at Leu 125. This indicated that procarboxypeptidase A and leucine amniopeptidase showed the 125th amino acid position to be the preference site of Cry1Ac protein.

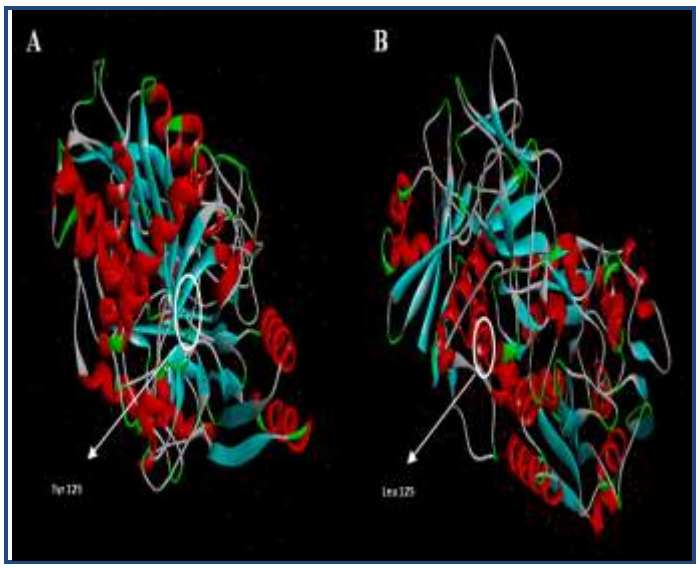

Figure 2: Docking of Cry1Ac protein tertiary structure. **A)** Cry1Ac protein with crystal structure of procarboxypeptidase A (Helicoverpa armigera) indicated that 125th aminoacid position (TYR-Tyrosine) to be preference site of the Cry1Ac protein; **B)** Cry1Ac protein with crystal structure of Leucine aminopeptidase (Bos taurus) indicated that 125th aminoacid position (Leu-Leucine) to be preference site of the Cry1Ac protein. Red color dots indicates water molecules.

Structure superposition

The structure of carboxypeptidase A (*H.armigera*) (1JQG) and Leucine aminopeptidase (*Bos taurus*) (1LAM) were compared. The 5 % of identity between these two structures in particular to the preference site could be the reason for the cross affinity nature of Cry1Ac protein. The structure of Cry1Ac and Cry2Ab, a protein was superimposed and shows the aligned regions in particular to the active site. Earlier report showed that fragments of recombinant *cry1Ab* gene were detected in the gastrointestinal contents of the Bt11 cotton fed pigs but not in the control pigs **[17].**

Conclusion:

This study was performed to find out the possibility of cross affinity Cry1Ac with leucine aminopeptidase of *Bos taurus* to prove the recent news of cattle grazing *Bt* cotton crops into the fields of Adilabad district of Andhra Pradesh, India (http://farmwars.info/wp-content/uploads/2011/09/BT-Cotton-Animal-Deaths). The results confirmed that even a 5% of identity would affect the cattle and genetically modified *Bt* cotton have an impact even to non-target animals. This work

may be helpful to animal and plant ecologist to do further research over on *Bt* cotton grazing cattles.

Acknowledgment:
Authors are greatly acknowledged to Prof. Dr. P. Gunasekaran, Vice chancellor, Thiruvalluvar University, Vellore and Former Head, Department of Genetics & Coordinator for UGC-NRCBS, Madurai Kamaraj University for giving visiting fellowship to complete the research. We also thank Dr. S. Shiek Asraf and Dr. K. Sujatha for assistance in the acquisition of research article.

References:

[1] Hino A, *Toxicol pathol.* 2002 **30**: 126 [PMID: 11890464]
[2] John P Purcell *et al*. *AgBioForum.* 2004 **7**: 27
[3] Siva Prasad CVS & Singh A, *SABionfo proceedings.* 2007 **28**: 30
[4] Tabashnik BE *et al*. *Appl Environ Microbiol.* 1996 **62**: 2839 [PMID: 8702276]
[5] Chothia C & Lesk AM, *EMBO J.* 1986 **5**: 823 [PMID: 3709526]
[6] Schwede T *et al. Nucleic Acids Res.* 2003 **13**: 3381 [PMID: 12824332]
[7] Al Karaawi MA, *Bioinformation.* 2013 **9**: 477 [PMID: 23847403]
[8] Duan JJ *et al. Biol Lett.* 2010 **6**: 74 [PMID: 19740894]
[9] Hongxia Liang *et al*. *Brazilian J Microbiol.* 2011 **42**: 140
[10] Tapp H & Stotzky G, *Appl Environ Microbiol.* 1995 **61**: 1786 [PMID: 16535020]
[11] Dong HZ *et al. J Agron Crop Sci.* 2006 **192**: 132
[12] Gohar M *et al*. *Proteomics.* 2005 **5:** 3696 [PMID: 16167365]
[13] Grochulski P *et al*. *J Mol Biol.* 1995 **254**: 447 [PMID: 7490762]
[14] Morse RJ *et al*. *Structure.* 2001 **9**: 409 [PMID: 11377201]
[15] Tajne S *et al. J Mol Graph Model.* 2012 **33**: 61 [PMID: 22182469]
[16] Estebanez Perpina E *et al.* J *Mol Biol.* 2001 **313:** 629 [PMID: 11676544]
[17] Chowdhury EH *et al*. *J Anim Sci.* 2003 **81:** 2546 [PMID: 14552382]

Edited by P Kangueane
Citation: **Ebenezer *et al.*** Bioinformation 9(15): 792-795 (2013)