International Research Journal of Biological Sciences _____ ISSN 2278-3202
Vol. **1(8),** 7-18, December **(2012)** *Int. Res. J. Biological Sci.*

Molecular Modeling and Docking Studies of PirB Fusion Protein from *Photorhabdus Luminescens*

Maithri S.K., Ramesh K.V.*, Dieudonné Mutangana and Deshmukh Sudha
Department of Biotechnology, Center for Postgraduate studies, Jain University, Bangalore – 560011 INDIA

Available online at: www.isca.in
Received 4th September 2012, revised 10th September 2012, accepted 15th September 2012

Abstract

Genetic engineering of Cry proteins from Bacillus thuringiensis (BT) has resulted in the synthesis of various novel toxin proteins which exhibits increased insecticidal activity and highly specificity towards different insect pests. The present study focused on computational studies on PirB sequence from Photorhabdus luminescens. The consensus tree generated by PHYLIP for the PirB sequence revealed that this toxin sequence does not share any ancestral relationship with other Cry toxins from Bacillus thuringiensis considered in this study. Molecular modeling of PirB was followed by construction of two fusion proteins: Type I (PirB-Cry2AaII-Cry2AaIII) and Type II (PirB-Cry2AaII-Garlic lectin). Comparison of the 3D model of PirB with X-ray structure of N-terminal domain 1I5P_A revealed both the structures shared similar architecture. Validation of the tertiary structure of PirB by the structural assessment tools such as ProSA, ERRAT and PROCHECK suggested that the predicted structure was of reasonable quality. Docking studies carried out onto the cadherin receptor showed that Type II fusion protein had a greater affinity, suggesting the possibility of using this fusion protein as a potential bio-pesticide.

Keywords: PirB, fusion proteins, modeling, docking, cry toxins..

Introduction

Photorhabdus luminescens, a gram-negative bacterium belonging to the Enterobacteriaceae family, and has been used extensively to control large community of insects[1]. Insecticidal activity is due to the high molecular weight insecticidal toxins produced by this bacterium. Insect resistance towards this bacterial toxin has not been reported to date[2,3]. The genome of the insect pathogen *Photorhabdus luminescens subsp. laumondii* strain TT01 contains numerous genes predicting toxins, hemolysins, and proteases, which may be important for insect pathogenicity[4]. While the loci plu4093 to plu4092 encodes the protein PirA, plu4437 to plu4436 codes for the PirB protein in *Photorhabdus luminescens*[5].

The Photorhabdus insect related (Pir) toxins act as binary proteins. Both proteins are necessary for injectable but not for oral activity[6,7]. Compared to other toxins of *Photorhabdus luminescens,* PirA shows little similarity to known proteins, whereas PirB shows high sequence homology with N-terminal region of the pore-forming domain of the Cry2A insecticidal toxin, thus making them a putative substitute for *Bacillus thuringiensis.* PirB also has similarities with a developmentally regulated protein from the beetle *Leptinotarsa decemlineata* which appears to have juvenile hormone esterase (JHE) activity [8,9]. There are several studies reporting that glycosyl-phosphatidyl-inositol (GPI) anchored amino peptidase N (APN) and cadherin-like protein functions as receptors of Cry1A toxins. The APN protein belongs to the Zn-binding metalloprotease family of proteins [10,11]. The C-terminal stalk of the APN binding site is rich in N-acetylgalactosamine (GalNAc) and acts as the binding site of the Cry1Ac toxin[10,12].

The studies conducted by Lee et al clearly show that lectins also exhibit good insecticidal activity and therefore transgenic plants engineered to express lectin confers protection against different insect pests. Lectins are group of sugar-binding proteins which recognizes specific carbohydrate structures and are known to agglutinate various animal cells by binding to their cell- surface glycoproteins and glycolipids. The insecticidal activities of plant lectins against a wide range of insect pests belonging to homoptera, lepidoptera, coleoptera and diptera have been well documented[13,14].

Construction of fusion proteins offers the possibility of exploring toxins having higher insecticidal properties. For instance, different Bt fusion proteins, such as Cry1Ab–Cry1B, Cry1Ac–Cry1Ab and Cry1Ac–GFP have been synthesized and expressed in different plant systems. Transferring of carbohydrate binding domain-III of Cry1C to Cry1Ab resulted in a fusion protein (Cry1Ab–Cry1C), which was highly toxic to the army worm (*Spodoptera exigua*), which was earlier resistant to Cry1A toxin[15-18].

There are several studies conducted in the recent past to demonstrate the development of resistance of insects towards BT toxin[19,20]. Therefore, it is very much essential to design a novel biopesticide, which can circumvent the current problem faced by agricultural crops. In the present study, an attempt has been made to predict theoretical model of fusion protein using

PirB structure as one of the substituent to Cry toxin complex (1I5P_A), followed by docking studies with cadherin receptor.

Material and Methods

Phylogenetic analysis of PirB from *Photorhabdus luminescens:* Fasta sequence of PirB along with the insecticidal cry toxins from *Bacillus thuringiensis* such as 1DLC_A [21], 1JI6_A [22], 3EB7 [23] , 1CIY_A [24], 2C9K_A [25] (, 1W99_A [26] and 1I5P_A [27] were submitted to the CLUSTALW [28]. Output generated by CLUSTALW tool was saved as 'allign.phy'. Then SEQBOOT program in Phylip [29] was activated. Outfile of the SEQBOOT was renamed as infile and run in "promlk" by analyzing 100 data sets with 5 times jumble. Then outtree of this was renamed as intree and served as input for CONSENSE program.

Molecular modeling of PirB from *Photorhabdus luminescens:* Based on the sequence identity between the PirB sequence and the PDB template as suggested by the PSI Blast tool [30], tertiary structure of PirB was predicted using homology modeling (SWISS MODEL server [31]) or threading (I-TASSER server[32]) approach. The model quality factor suggested by the ERRAT server[33] (was taken into account for selecting the appropriate loop regions for refinement using MOODLOOP server[34]. Minimization of the model was carried out using the Deep View software[35]. Structurally homologous proteins were obtained by submitting the coordinated of the loop refined model to DALI server[36]. The refined models were later validated with PROCHECK [37] and ProSA server[38].

***In silico* fusion protein:** Two types of fusion proteins were modelled using computational approach. For all docking exercises, HEX (v6.3)[39] was used. In the first type, the N-terminal domain model of PirB (PirBI) was initially docked on to the edited X-ray crystal structure of Cry2Aa (PDB id 1I5P)[27] which contained the coordinates information only for 2nd and 3rd domain. The second type of fusion protein generated by docking the N terminal domain of PirBI to the 2nd domain of Cry2Aa was once again docked onto the PDB structure of garlic lectin (1KJ1)[40]. This was done because this protein is reported to have a folding pattern comparable to domain III of Cry toxin[25],[41] and moreover it has insecticidal property[42].

Docked conformations and interaction energies were recorded at the end of the docking exercise. During the dock operation, the total energies were calculated based on shape as well as electrostatics using a default grid spacing of 0.6 Å. The large number of conformations generated for both the fusion proteins (type I and II) by the HEX package were subjected to rigorous screening, using the X-ray crystal structure of Cry2Aa (1I5P) as the reference, for deciding on the correct orientation of the docked complexes.

Docking studies with cadherin receptors: The two fusion proteins generated through HEX were docked onto the X-ray crystal structure of cadherin receptor from *Drosophila melanogaster* (PDB ID 3UBH)[43] to evaluate the docking affinity. Among the 100 different orientations that were generated by HEX, the best one was selected based on total dock energy.

Results and Discussion

The templates that were selected to generate the consensus tree for PirB from *Photorhabdus luminesces* were Cry toxins from *Bacillus thuringiensis*. The output shows PirB toxin sequence as an out group in the tree, suggesting this sequence is evolutionary unrelated to the remaining Cry toxins *Bacillus thuringiensis* (figure-1). For the PirB sequence for *Photorhabdus luminescens*, the PSI BLAST tool was able to retrieve wide array of toxins present in diverse group of organisms from protein data bank (table-1 A). Among these toxins structures, the crystal structure of N-terminal domain of cry2Aa (1I5P) was selected as the PDB template for modeling purpose as this template sequence got aligned to the PirB sequence of *Photorhabdus luminescens* with maximum sequence identity (22%) compared to other toxin sequences (table-1 B). Homology modeling usually starts by searching the database of known protein structures such as protein data bank. Sometimes, the availability of many sequences related to the target makes it necessary to do more sensitive searching with profile methods and Hidden Markov Models [44-46]. Fiser and Sali (2000) have recommended the use of PSI BLAST tool 30 as one of the methods to construct a profile from a multiple alignment of the highest scoring hits from an initial BLAST search.

Submission of the PirB sequence threaded onto the PDB template 1I5P to the SWISS MODEL server failed in generating a 3D model based on homology modeling approach. However, I-TASSER server, using the same template structure, was able to generate a 3D model for PirB. All the top 10 PDB templates fetched by the server were cry toxin structures displaying sequence similarity to the PirB toxin sequence, which included 1I5P_A also as one of the template (figure-2b). Sequence alignment report generated by I-TASSER output showed that glu56 of PirB is well conserved with rest of the sequences. Alignment report also shows that, gln55 and arg100 of PirB were conserved with all the templates except 1I5P_A, which has tyrosine at the corresponding position. Comparison between PirB and 1I5P_A sequences revealed that serine (ser5,34,101,156,218) was maximally conserved followed by glu57,152,215, asn47, 104, 180, val11,15, leu36,129, ile52,219, phe117,132, tyr182,216, pro42, gln80, met132, asp151 and trp220. Among these residues, glu215 and tyr216 of PirB sequence are of significance, because same type of residues are present at the active site of the crystal structure of Cry2Aa-I (domain I) and are involved in binding to Cry2Aa-II (domain II; 1I5P_A).

Comparison of the 3D model of PirB with X-ray structure of N-terminal domain 1I5P_A revealed, both the structures shared

similar architecture (figure-3). Orientation of all the seven helices of PirB model closely matched with that of template (1I5P_A) structure. In addition to these seven helices, I – TASSER was able to predict additional 2 helices (197 - 211 and 122- 126) for PirB model. Further examination of the PirB model and 1I5P_A, revealed that, both of them had common loop segments (designated as L1 to L6) which connected all the seven helices (figure-3). While the number of residues in loop segments L3 and L6 of PirB model and 1I5P_A remained the same (table-2), the number of residues in L1, L2, L4 and L5 varied marginally. However, there was large variation in the number of residues of L5 connecting helix 5 and 6 of PirB and 1I5P_A (table-2). Except pro42 of L2 and ser155 of L6, none of the residues of the loop regions were conserved. The output generated by DALI server revealed that Cα back bone of the PirB model got super-imposed with top 7 structural homologs (1CIY_A, 1JI6_A, 2QKG_A, 3EB7_A, 2QKGB, 3EB7B, 1I5P_A), belongs to different cry toxins of *Bacillus thuringiensis* (figure-4). The z-score of these 7 structural homologs varied from 24.3 to 22.4 and RMSD varied from 2.0 to 1.0.

Validation of the tertiary structure of PirB by the structural assessment tools such as ProSA, ERRAT and PROCHECK showed that there was an improvement in the quality of the predicted structure upon loop refinement. Energy minimization of the final PirB model after loop refinement resulted in a decrease of total energy from -4130.422 to -6698.85 KJ / mol. The Z score value computed by the ProSA tool decreased from -7.21 to -7.23 of PirB model and was lying within the Z score values calculated for all experimentally determined protein structures of similar size, deposited in the protein data bank (Figure-5A). Energy profile plot, also generated by the ProSA server, revealed negative energy distribution pattern scored by the amino acid residues of PirB model (Figure-5B), suggesting an improvement in the model quality. The quality factor of the loop refined PirB model calculated by ERRAT reached 74.81 from 57.55. Ramachandran plot analysis performed by PROCHECK tool revealed that 89.4% of the residues of PirB structure were distributed in the core region, followed by 8.0% in the allowed region, 1.0% in the general region and remaining 1.5% in the disallowed region (Table-3). However Ramachandran plot analysis of X-ray crystal structure of 1I5P_A showed that 91.4% of the residues were present in the core, 8.1% of the residues in the allowed region, 0.5% of the residues in the generously allowed region and 0.0% in the disallowed region (table-3B). The overall G-score calculated by the tool for the model was -0.06 which was above the threshold value indicating the predicted model was satisfactory. Kashyap et al (2010) [47] have used similar kind of tools for validating the 3D structure of Cry1Ab17 from *Bacillus thuringiensis* predicted using X-ray structure of Cry1Aa1 from *Bacillus thuringiensis.*

Although the two fusion proteins, type I (PirBI-Cry2AaII-Cry2AaIII) and II (PirBI-Cry2AaII-Garlic lectin) (figure-6) that were generated using HEX package, displayed similar orientation, type II had better affinity (dock energy = -679.0 KJ / mol) towards the cadherin receptor (figure-7), compared to type I (dock energy = -616.0 KJ / mol). Francis et al. have reported cadherin-like protein functions as a receptor for Cry1A toxins. However, Tanje et al.[48] who have modeled "Cry-Garlic lectin" fusion protein using *in silico* approach, studied the interaction of this toxin protein with aminopeptidase N receptor from *Manduca sexta.* Efficacy of the fusion protein predicted in the present study cannot be compared with the studies of Tanje et al.[48], due to the difference in the type of receptor protein considered. Using PirB in the construction of fusion proteins as one of the component can be beneficial, because of the development of resistance by insects towards the Cry toxins of *Bacillus thuringiensis* [19,20].

Residues interacting at the interface of 3 domains of Type II fusion protein model and 1I5P_A: Comparisons were made between the residues interacting at the interface of 3 domains of type II fusion protein model with 1I5P_A (table-4). The results suggest that total number of residues present at the interface of the 3 domains were more in number in case of type II fusion protein (97 residues) compared to 1I5P_A (78 residues). While polar residues were dominant at the interface of type II fusion protein model, on the contrary non- polar types were more in number at the interface of 1I5P_A. Among the 97 residues present at the interface of the type II fusion protein model, 49 of them were associated with PirB model (i. e., domain I) and remaining 48 shared between II and III domain of Type II fusion protein. The residues of domain I that were involved in the interaction with remaining 2 domains of both type II fusion protein model and 1I5P were predominantly polar in nature, with acidic type of residues less in number (table-4).

Based on the data generated for residues interacting between Type II fusion protein and crystal structure of cadherin receptor from *Drosophila melnogaster*, it is seen that polar residues of Cry2Aa –II (domain II of Type II fusion protein) and non-polar residues of receptor (Cadherin) were maximally involved in the interaction (table-5) (figure-8).

Conclusion

Molecular modeling studies of PirB toxin resulted in the construction of 2 types of fusion protein, built using the Cry2Aa toxin complex from *Bacillus thuringiensis*. Among these two types of fusion protein, type II showed better affinity with the Cadherin receptor and therefore can be considered as probable alternative to the currently used BT toxins for controlling the agricultural crop pests.

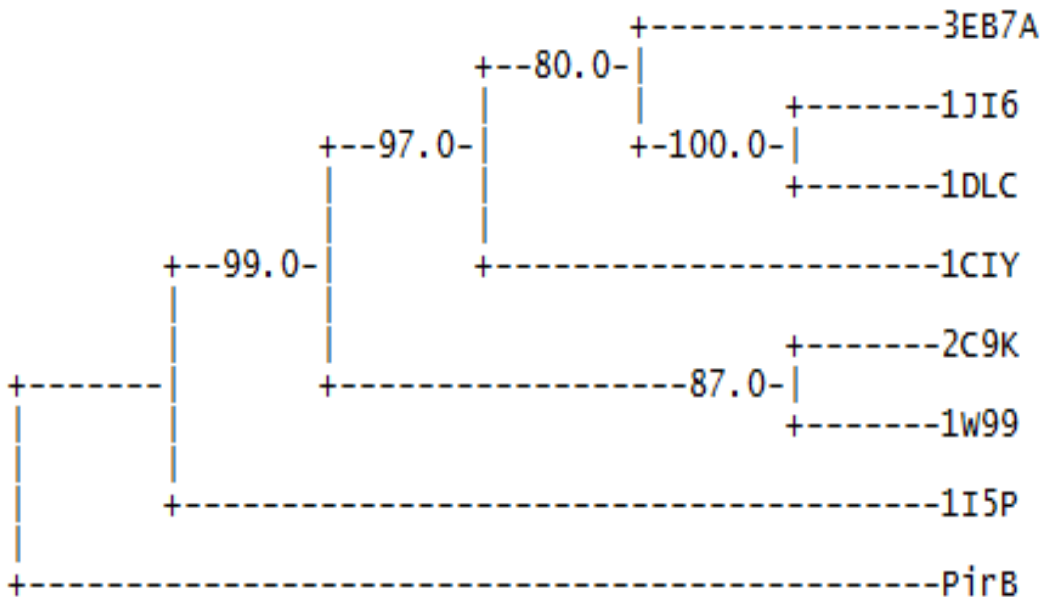
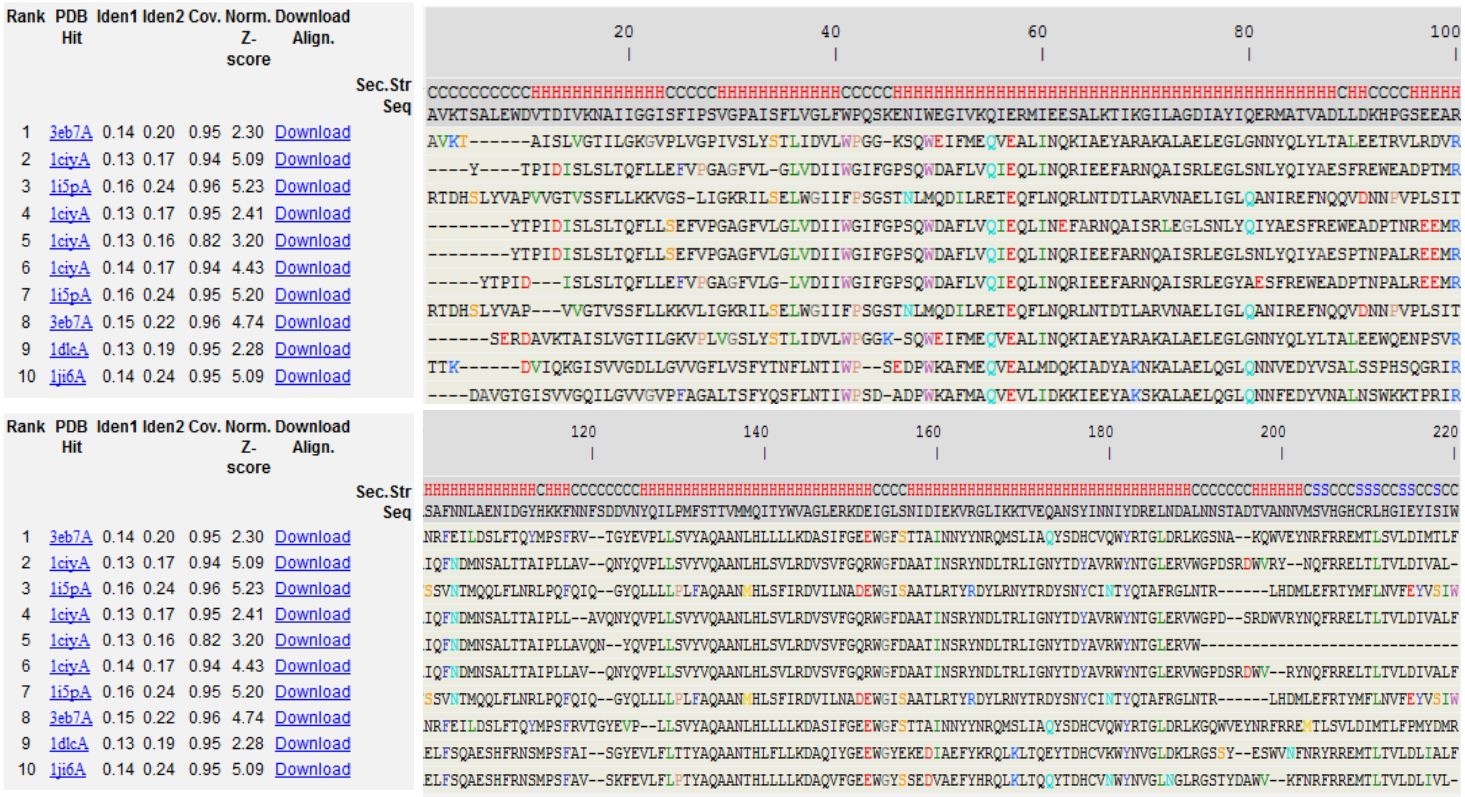


Figure-1

Consensus tree generated for PirB from *Photorhabdus luminescens* using Promlk program of PHYLIP. The number on the branches indicate the number of times the partition of the species into the two sets which are separated by that branch occurred among the trees, out of 100 trees. The tree is an unrooted one.



Rank	PDB Hit	Iden1	Iden2	Cov.	Norm. Z-score	Download Align.
1	3eb7A	0.14	0.20	0.95	2.30	Download
2	1ciyA	0.13	0.17	0.94	5.09	Download
3	1i5pA	0.16	0.24	0.96	5.23	Download
4	1ciyA	0.13	0.17	0.95	2.41	Download
5	1ciyA	0.13	0.16	0.82	3.20	Download
6	1ciyA	0.14	0.17	0.94	4.43	Download
7	1i5pA	0.16	0.24	0.95	5.20	Download
8	3eb7A	0.15	0.22	0.96	4.74	Download
9	1dlcA	0.13	0.19	0.95	2.28	Download
10	1ji6A	0.14	0.24	0.95	5.09	Download

Rank	PDB Hit	Iden1	Iden2	Cov.	Norm. Z-score	Download Align.
1	3eb7A	0.14	0.20	0.95	2.30	Download
2	1ciyA	0.13	0.17	0.94	5.09	Download
3	1i5pA	0.16	0.24	0.96	5.23	Download
4	1ciyA	0.13	0.17	0.95	2.41	Download
5	1ciyA	0.13	0.16	0.82	3.20	Download
6	1ciyA	0.14	0.17	0.94	4.43	Download
7	1i5pA	0.16	0.24	0.95	5.20	Download
8	3eb7A	0.15	0.22	0.96	4.74	Download
9	1dlcA	0.13	0.19	0.95	2.28	Download
10	1ji6A	0.14	0.24	0.95	5.09	Download

Figure-2

Alignment of N-terminal domain of PirB from *Photorhabdus luminescens* with the top 10 PDB templates generated by the I-TASSER server

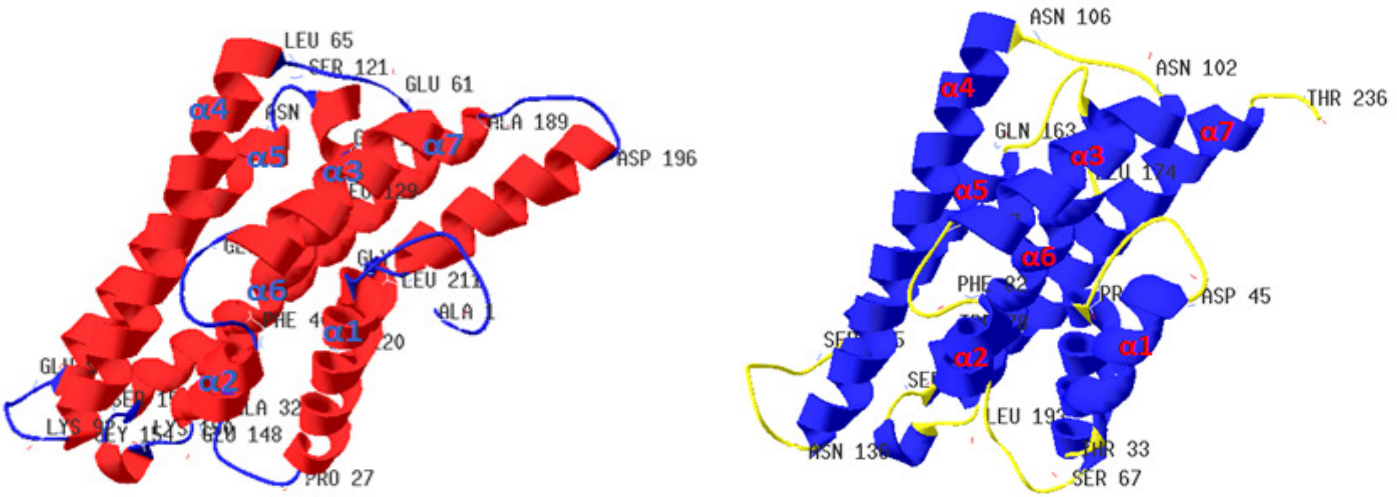
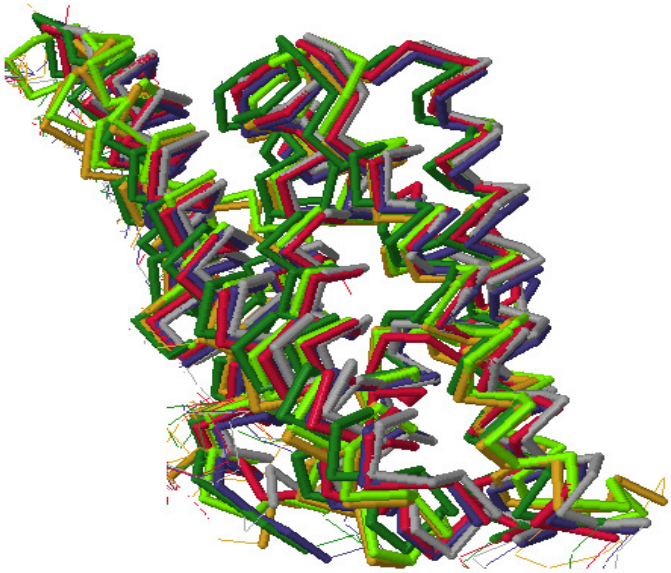


Figure-3

Tertiary structure of PirB from *Photorhabdus luminescens* generated by the I-TASSER server. The N –terminal region of PirB was predicted by furnishing the PDB coordinates of the X-ray structure of cry toxin "cry2Aa" from *Bacillus thuringiensis* in the template option of I-TASSER server

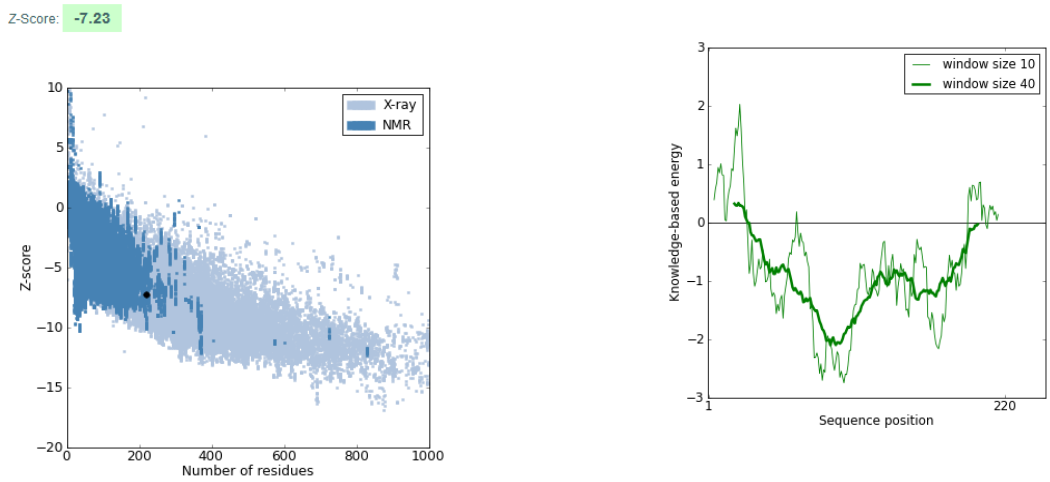


No:	Chain	Z	rmsd	lali	nres	%id	PDB	Description
1:	1civ-A	24.3	2.0	206	577	12	PDB	MOLECULE: CRYIA(A);
2:	1ji6-A	23.6	2.1	206	589	15	PDB	MOLECULE: PESTICIDIAL CRYSTAL PROTEIN CRY3BB;
3:	2qkg-A	23.4	2.1	208	589	13	PDB	MOLECULE: INSECTICIDAL DELTA-ENDOTOXIN CRY8EA1;
4:	3eb7-A	23.4	2.1	208	589	13	PDB	MOLECULE: INSECTICIDAL DELTA-ENDOTOXIN CRY8EA1;
5:	2qkg-B	23.0	2.1	208	589	13	PDB	MOLECULE: INSECTICIDAL DELTA-ENDOTOXIN CRY8EA1;
6:	3eb7-B	23.0	2.1	208	589	13	PDB	MOLECULE: INSECTICIDAL DELTA-ENDOTOXIN CRY8EA1;
7:	1i5p-A	22.4	1.0	209	633	16	PDB	MOLECULE: PESTICIDIAL CRYSTAL PROTEIN CRY2AA;

Figure-4

Dali server output showing super-imposed structural homologs for PirB from *Photorhabdus luminescens*

Green: PirB; *red*: 1CIY_A; *grey*: 1JI6_A; *orange*: 2QKG_A; *dark blue*: 3EB7_A; *light blue*: 2QKGB; *green*: 3EB7B; *yellow*: 1I5P_A



(A) **(B)**

Figure-5

ProSA server output generated for the N-terminal domain of PirB from *Photorhabdus luminescens*. Graphical display of (A) Z score plot (B) Energy profile plot

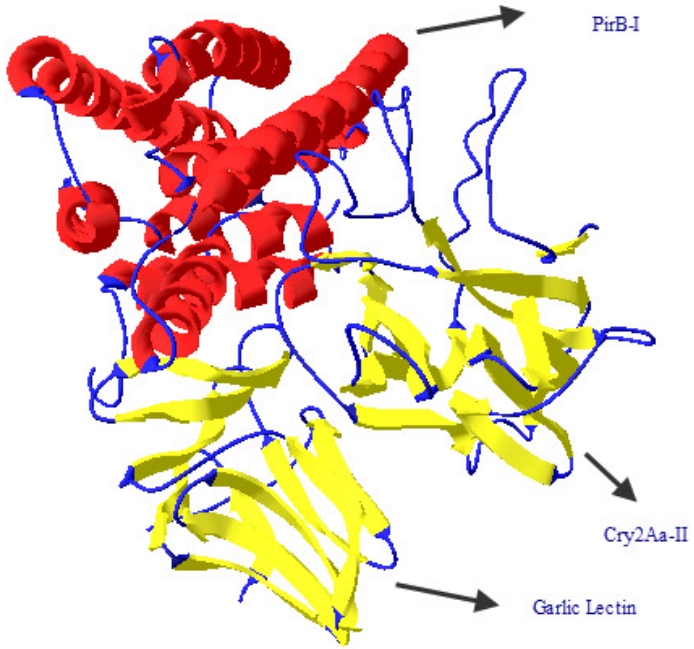


Figure-6

3D model of fusion protein (Type II) generated by docking PirB model from *Photorhabdusluminescens* onto the crystal structures of cry2Aa (PDB ID: 1I5P; *Bacillusthuringiensis*) and garlic lectin (PDB ID: 1KJ1; *Allium sativum*) using HEX software. Image was generated using SPDV package

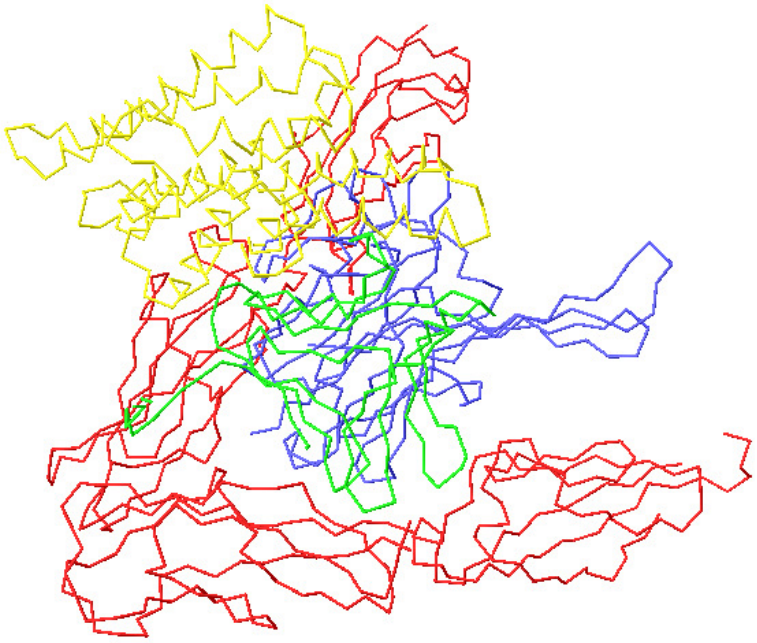


Figure-7
Docking of Type II fusion protein onto X-ray crystal structure of N-cadherin from *Drosophila melanogaster* (PDB ID 3UHB_A)using HEX software (Total energy = –679.0 KJ / mol)

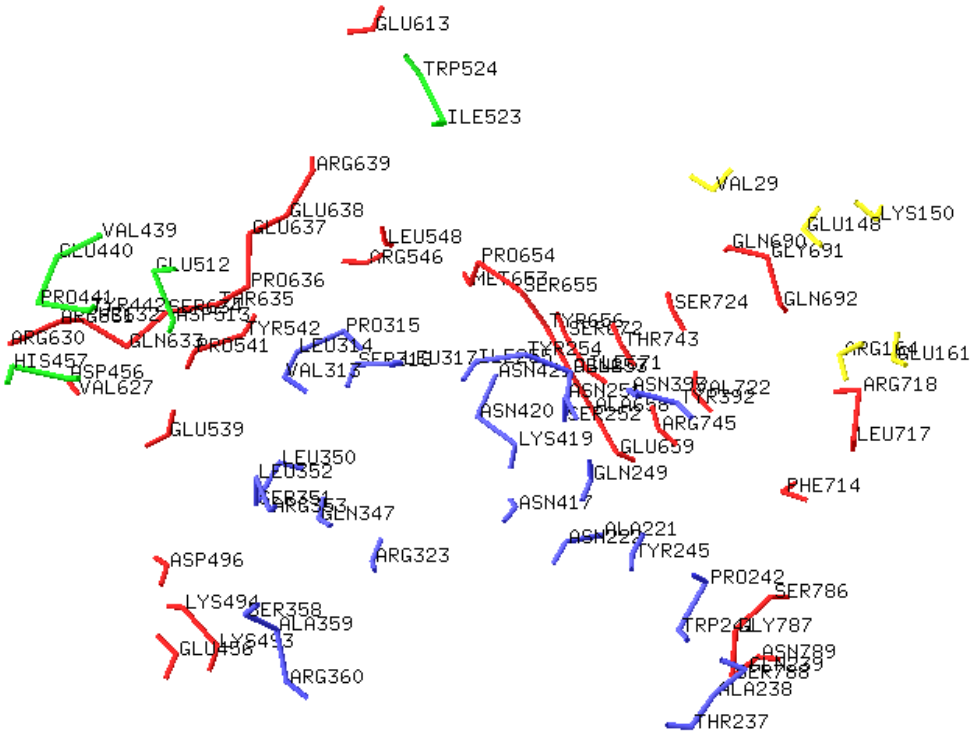


Figure-8
Amino acid residues of Type II fusion protein (Green: PirB; Blue: Cry2Aa-II; Yellow: Garlic lectin) interacting with crystal structure of cadherin receptor from *Drosophila melagaster* (Red) within 5Å distance

Table-1
Summary of (A) PDB templates generated by PSI blast results and (B) sequence alignment of PirB sequence from *Photorhabdus luminescens* at the end of 20th iteration

(A)E-value BETTER than threshold		
Sequences producing significant alignments	Score Bits	E-Value
pdb|1JI6|A Chain A, Crystal Structure Of The Insecticidal Bac...	226	6e-67
pdb|1DLC|A Chain A, Crystal Structure Of Insecticidal Delta-E...	221	2e-65
pdb|3EB7|A Chain A, Crystal Structure Of Insecticidal Delta-E...	213	5e-62
pdb|1I5P|A Chain A, Insecticidal Crystal Protein Cry2aa	**201**	**3e-57**
pdb|2IZW|A Chain A, Crystal Structure Of Ryegrass Mottle Viru...	175	2e-51
pdb|1CIY|A Chain A, Insecticidal Toxin: Structure And Channel...	179	7e-50
pdb|1W99|A Chain A, Mosquito-Larvicidal Toxin Cry4ba From Bac	176	7e-49
pdb|2C9K|A Chain A, Structure Of The Functional Form Of The M... 170 1e-46		
pdb|2QNT|A Chain A, Crystal Structure Of Protein Of Unknown F...	126	3e-34
pdb|3VDG|A Chain A, Crystal Structure Of Enolase Msmeg_6132 (.	81.4	1e-16
pdb|3VC5|A Chain A, Crystal Structure Of Enolase Tbis_1083(Ta...	77.9	2e-15
pdb|3VA8|A Chain A, Crystal Structure Of Enolase Fg03645.1 (T...	77.1	3e-15
pdb|3S8Y|A Chain A, Bromide Soaked Structure Of An Esterase F...	56.0	1e-08
pdb|3I6Y|A Chain A, Structure Of An Esterase From The Oil-Deg...	55.6	2e-08
pdb|3LS2|A Chain A, Crystal Structure Of An S-Formylglutathio...	50.6	8e-07

(B) ALIGNMENTS
>pdb|1I5P|A Chain A, Insecticidal Crystal Protein Cry2aa
Length=633
Score = 201 bits (510), Expect = 3e-57, Method: Composition-based stats.
Identities = 52/240 (22%), Positives = 100/240 (42%), Gaps = 33/240 (14%)

```
Query  18   AVKTSALEWDVTD--IVKNAIIGGIS--FIPSVGPAI-----SFLVGLFWPQSKENIWEG  68
            ++  +EW  TD +   ++G +S  + VG  I    S L G+ +P   N+ +
Sbjct  33   TIQKEWMEWKRTDHSLYVAPVVGTVSSFLLKKVGSLIGKRILSELWGIIFPSGSTNLMQD  92
Query  69   IVKQIERMIEE----SALKTIKGILAGDIAYIQERMATVADLLD--KHPGSEEARSAFNN  122
            I+++ E+ + +     L +  L G  A I+E    V + L+  ++P     S+ N
Sbjct  93   ILRETEQFLNQRLNTDTLARVNAELIGLQANIREFNQQVDNFLNPTQNPVPLSITSSVNT  152
Query  123  LAENIDGYHKKFNNFSDDVNYQIL--PMFSTTVMMQITYWVAGLERKDEIGLSNIDIEKV  180
            + +      +F       YQ+L  P+F+    M +++   +  DE G+S  +
Sbjct  153  MQQLFLNRLPQF----QIQGYQLLLLPLFAQAANMHLSFIRDVILNADEWGISAATLRTY  208
Query  181  RGLIKKTVEQANSYINNIYD---RELNDALNNSTADTVANNVMSVHGHCRLHGIEYISIW  237
            R ++     ++Y  N Y    R LN  L        ++++   + L+  EY+SIW
Sbjct  209  RDYLRNYTRDYSNYCINTYQTAFRGLNTRL---------HDMLEFRTYMFLNVFEYVSIW  259
```

Table-2
Summary of the loop segments involved in the connectivity of various helices of 3D model of PirB and PDB template 1I5P_A

PirB	**Helices**	**1I5P_A**
L1(pro^{27}-ala^{32})	α1- α 2	L1(asp^{45}-pro^{52})
L2(phe^{40}-glu^{46})	α2- α3	L2(ser^{67}-ile^{73})
L3(glu^{61}-leu^{65})	α3- α4	L3(phe^{82}-thr^{87})
L4(lys^{92}-glu^{98})	α4- α5	L4(asn^{102}-asn^{106})
L5(asn^{119}-ser^{121})	α5- α6	L5(asn^{136}-ser^{145})
L6(gly^{154}-ser^{156})	α6- α7	L6(leu^{193}-ala^{195})

Table-3
Summary of the PROCHECK analysis for the N-terminal domain of PirB from (A) *Photorhabdus luminescens* (B) 1I5P_A from *Bacillus thuringiensis*

```
 +----------<<<  P  R  O  C  H  E  C  K     S  U  M  M  A  R  Y  >>>----------+
 |                                                                             |
 | input_atom_only.pdb   2.5                                    220 residues   |
 |                                                                             |
*| Ramachandran plot:   89.4% core    8.0% allow    1.0% gener    1.5% disall  |
 |                                                                             |
*| All Ramachandrans:   11 labelled residues (out of 218)                      |
*| Chi1-chi2 plots:      8 labelled residues (out of 143)                      |
 | Main-chain params:    6 better     0 inside      0 worse                    |
 | Side-chain params:    5 better     0 inside      0 worse                    |
 |                                                                             |
*| Residue properties: Max.deviation:     5.5             Bad contacts:    3   |
*|                     Bond len/angle:    8.9    Morris et al class:  1  2  2   |
 |                                                                             |
 | G-factors           Dihedrals:  -0.06  Covalent:   0.00    Overall:  -0.01  |
 |                                                                             |
 | M/c bond lengths: 99.9% within limits   0.1% highlighted                    |
*| M/c bond angles:  90.9% within limits   9.1% highlighted       1 off graph  |
*| Planar groups:    81.9% within limits  18.1% highlighted       1 off graph  |
 |                                                                             |
 +-----------------------------------------------------------------------------+
```

(a)

```
 +----------<<<  P  R  O  C  H  E  C  K     S  U  M  M  A  R  Y  >>>----------+
 |                                                                             |
 | input_atom_only.pdb   2.5                                    227 residues   |
 |                                                                             |
+| Ramachandran plot:   91.4% core    8.1% allow    0.5% gener    0.0% disall  |
 |                                                                             |
+| All Ramachandrans:    5 labelled residues (out of 225)                      |
+| Chi1-chi2 plots:      1 labelled residues (out of 156)                      |
 | Main-chain params:    6 better     0 inside      0 worse                    |
 | Side-chain params:    5 better     0 inside      0 worse                    |
 |                                                                             |
+| Residue properties: Max.deviation:     4.1             Bad contacts:    2   |
+|                     Bond len/angle:    4.1    Morris et al class:  1  1  1   |
 |                                                                             |
 | G-factors           Dihedrals:   0.38  Covalent:   0.58    Overall:   0.46  |
 |                                                                             |
 | M/c bond lengths:100.0% within limits   0.0% highlighted                    |
 | M/c bond angles:  99.2% within limits   0.8% highlighted                    |
+| Planar groups:    96.0% within limits   4.0% highlighted       1 off graph  |
 |                                                                             |
 +-----------------------------------------------------------------------------+
```

(b)

Table-4
Comparison of the interacting residues at the interface of 3 domains of Type II fusion protein model and crystal structure of 1I5P

		Cry2Aa_II		Total
Acidic	**PirB** asp$^{88, 111, 122, 183, 188}$ glu$^{46, 50, 81, 108, 172, 185}$ (types=2; total = 11)	asp$^{16, 25}$ (types=1; total = 2)	**GL*** asp^{17,35}, glu$^{19, 27}$ (types = 2; total = 4)	**17**
	Cry2Aa_I asp^{22,45} (types=1, total =2)	asp^{420} (types=1; total = 1)	**Cry2Aa_III** asp^{578,582,589} (types = 1; total =3)	**6**
Basic	**PirB** arg^{184}, lys^{45,92,115,116,169} his^{114} (types=3; total = 7)	arg^{26,32,}, lys^{30,142}, his^{144} (types=3; total =5)	**GL** arg^{1}, his^{22,36} (types=2; total =3)	**15**
	Cry2Aa_I arg^{9,209,213,217,232,237,245} , his^{21,239} (types=2; total =9)	arg^{375} , his^{468} (types=2; total = 2)	**Cry2Aa_III** arg^{548} (types=1; total = 1)	**12**
	PirB		**GL**	
Polar	asn^{47,118,119,125,175,179,180,187}, gln^{43,173}, ser^{44,121,176} thr^{67,85} , tyr^{78,113,117} (types=5; **total =18**)	asn^{11,}gln^{27,34}ser^{19,21}, thr^{22,28,37},tyr^{20} (types=5; total =9)	gln^{14,26}, ser^{15,37}, tyr^{11,21,34} (types=3; total =7)	**34**
	Cry2Aa_I asn^{3,6,221,235} , ser^{6,7,25,260,266,270,271} thr^{11,236} , tyr^{264} (types=4; total=14)	asn^{274,341,416,469,472,473} , gln^{283,} ^{284,286,399}, ser$^{278,280.287,309,363,370,376,415}$, thr^{285,364} , tyr^{417,421} (types=4; total = 22)	**Cry2Aa_III** asn^{576,577,581} , ser^{588,621} , thr^{492,495,573} , tyr^{475,633} (types=4; total = 10)	**22**
Non polar	**PirB** ala^{73,77,84,} , gly^{70,74,112}, ile^{68,71,181}, phe^{120}, pro^{42}, trp^{41} , val^{171} (**types=7**; total =13)	ala^{29,31}, gly^{146}, ile^{13,33}, leu^{18,36}, phe^{23,143,}, val^{9,15,145} (types=6; total =12)	**GL** ala^{12}, gly^{13}, ile24, leu^{16}, pro^{20}, val^{18} (types=6; total =6)	**31**
	Cry2Aa_I gly^{8,272} , leu^{5}, met^{268}, trp^{41} (types=4; total = 5)	ala^{273,277},cys^{362}, gly^{279,281,310}, leu^{275,308,366,369}, phe^{288,418,422}, pro^{282,367,368,419}, val^{365,374} (types=7; total =19)	**Cry2Aa_III** gly^{493,579}, ile^{474,484,496,590}, leu^{547,632}, met^{620}, phe^{494,624}, pro^{631}, phe^{494,624} (types=7; total = 14)	**38**

Note: * GL = Garlic lectin

Table-5
Summary of residues interacting between Type II fusion protein and crystal structure of cadherin receptor from *Drosophila melnogaster* (PDB ID: 3UHB_A)

Acidic	Basic	Polar	Non Polar
Type II Fusion protein PirB (Domain I)			
alu^{148,161}	ly^{s150}, arg^{164}	NIL	ala^{221}
Cry2Aa-II (Domain II)			
NIL	arg^{323,353,360}, his^{457}, lys^{419}	asn$^{222,251,253,393,417, 420,421}$, gln^{239,249,347,} ser^{252,351,358}, thr^{237}, tyr^{245,254,392}	ala^{238,359}, ile^{255,}
Garlic Lectin (Domain III)			
asp^{456,513}, glu^{440,512}	NIL	tyr^{442,}	ile^{523},pro^{441}, trp^{524}, val^{439}
Cadherin receptor (PDB ID: 3UHB_A)			
glu^{456,539,613,637,638,657,659,791}	arg$^{546,630,631,639,718, 745}$, lys^{492,493,494,619}	gln^{690,692}, ser^{634,655,724,786,788}, thr^{635,743}, tyr^{542,656}	gly^{691,787}, ile^{671}, leu^{548,717}, met^{653}, phe^{632,714}, pro^{540,541,572,636,654,720}, val^{583,615,627,722}

References

1. Poinar G.O., Thomas G.M. and Hess R., Characteristics of the specific bacterium associated with Heterorhabditis bacteriophora (Heterorhabditidae: Rhabditida), *Nematologica.,* **23**, 97–102 **(1977)**

2. Chattopadhyay A. Bhatnagar N. and Bhatnagar R., Bacterial insecticidal toxins, *Crit. Rev. Microbiol.,* **30,** 33–54 **(2004)**

3. Daborn P.J., Waterfield N.R., Silva C.P., Au C.P., Sharma S. and Ffrench-Constant R.H., A single Photorhabdus gene, makes caterpillars floppy (mcf), allows Escherichia coli to persist within and kill insects, *Proc. Natl. Acad. Sci USA,* **99**, 10742–10747 **(2002)**

4. Duchaud E et al., The genome sequence of the entomopathogenic bacterium Photorhabdus luminescens, *Nat. Biotechnol,* **21**, 1307–1313 **(2003)**

5. Waterfield N.R., Hares M., Yang G., Dowling A. and ffrench-Constant R.H., Potentiation and cellular phenotypes of the insecticidal toxin complexes of Photorhabdus bacteria, *Cell Microbiol* **7,** 373–382 **(2005 a)**

6. ffrench-Constant R. H., Dowling A. and Waterfield N. R., Insecticidal toxins from Photorhabdus bacteria and their potential use in agriculture, *Toxicon* **49,** 436–451**(2007)**

7. Yang G., Dowling J., Gerike U., ffrench-Contant R. H. and Waterfield N. R., Photorhabdus virulence cassettes confer injectable insecticidal activity against the wax moth, *J. Bacteriol* **188,** 2254–2261 **(2006)**

8. Waterfield N. R., Kamita S. G., Hammock B. D. and ffrench-Constant R. H., The Photorhabdus Pir toxins are similar to a developmentally regulated insect protein but show no juvenile hormone esterase activity, *FEMS Microbiol. Lett* **245,** 47–52 **(2005 b)**

9. Wilkinson P., Waterfield N. R., Crossman L., Corton G., Sanchez-Contreras M. et al. Comparative genomics of the emerging human pathogen Photorhabdus asymbiotica with the insect pathogen Photorhabdus luminescens. *BMC Genomics* **10,** 302-323 **(2009)**

10. Knight P. J., Crickmore N. and Ellar D. J., The receptor for Bacillus thuringiensis CrylA(c) delta-endotoxin in the brush border membrane of the lepidopteran Manduca sexta is aminopeptidase N, *Mol. Microbiol* **11,** 429–436 **(1994)**

11. Francis B.R. and Bulla Jr L.A., Further characterization of BT-R1, the cadherin- like receptor for Cry1Ab toxin in tobacco hornworm (Manduca sexta) midgets, *Insect Biochem. Mol. Biol*, **27,** 541–550 **(1997)**

12. Knight P.J., Carroll J. and Ellar D.J., Analysis of glycan structures on the 120 kDa aminopeptidase N of Manduca sexta and their interactions with Bacillus thuringiensis Cry1Ac toxin, *Insect Biochem. Mol. Biol* **34,** 101–112 **(2004)**

13. Lee M.K., You T.H., Gould F.L. and Dean D.H., Identification of residues in domain III of Bacillus thuringiensis Cry1Ac toxin that affect binding and toxicity, *Appl. Environ. Microbiol* **65,** 4513–4520 **(1999)**

14. Bharathi Y., Reddy T.P., Reddy V.D. and Rao K.V., Plant lectins and their utilization for development of insect resistant transgenic crop plants, in: Pests and Pathogens: Management Strategies, *BS Publications*, 457–489 **(2010)**

15. Ho N.H., Oliva B.N., Datta K., Frutos R. and Datta S.K., Translational fusion hybrid Bt genes confer resistance against yellow stem borer in transgenic elite vietnamese rice (Oryza sativa L.) cultivars, *Crop Sci.* **46**, 781–789 **(2006)**

16. Honee G., Vriezen W. and Visser B., A translation fusion product of two different insecticidal crystal protein genes of Bacillus thuringiensis exhibits an enlarged insecticidal spectrum, *Appl. Environ. Microbiol.* **56**, 823–825 **(1990)**

17. Harper B.K., Mabon S.A., Leffel S.M., Halfhill M.D., Richards H.A. and Moyer K.A. et al., Green fluorescent protein as a marker for expression of a second gene in transgenic plants, *Nat. Biotechnol* **17,** 1125–1129 **(1999)**

18. de Maagd R.A., Kwa M.S., van der Klei H., Yamamoto T., Schipper N., Vlak J.M. et al., Domain III substitution in Bacillus thuringiensis delta-endotoxin CryIA(b) results in superior toxicity for Spodoptera exigua and altered membrane protein recognition, *Appl. Environ. Microbiol* **62**, 1537–1543 **(1996)**

19. Hoy M.A., Myths, models and mitigation of resistance to pesticides, *Phil. Trans. R. Soc. Lond. B* **353**, 1787–1795 **(1998)**

20. Michaud D., Avoiding protease mediated resistance in herbivorous pests. *Trends Biotech* **15**, 4–6 **(1997)**

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

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

23. Guo S., Ye S., Liu Y., Wei L., Xue J., Wu H., Song F., Zhang J., Wu X., Huang D. and Rao Z., Crystal structure of Bacillus thuringiensis Cry8Ea1: An insecticidal toxin toxic to underground pests, the larvae of *Holotrichia parallela*, *J struct Biol* , **168,** 259-266 **(2009)**

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

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

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

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

28. Thompson J. D., Higgins D. G. and Gibson T. J., CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, *Nucleic Acids Res* , **22,** 4673-4680 **(1994)**

29. Felsenstein J. *Inferring phylogeny*, Sinauer Associates, Sunderland, MA **(2003)**

30. Altschul S.F., Madden T.L., Schäffer1 A. A., Zhang J., Zhang Z., Miller W. and Lipman D. J., Gapped BLAST Nucleic Acids Res and PSI-BLAST: a new generation of protein database search programs, *Nucleic Acids Res*, **17,** 3389-3402 **(1997)**

31. Schwede T., Kopp J., Guex N. and Peitsch M. C., SWISS-MODEL: an automated protein homology-modeling serve,. *Nucl. Acids Res*, **31,** 3381-3385 **(2003)**.

32. Zhang Y., I-TASSER server for protein 3D structure prediction, *BMC bioinformatics*, **9,** 40-47 **(2008)**

33. Colovos C. and Yeates T. O., Verification of protein structures: Patterns of nonbonded atomic interactions, *Protein Science,* **2,**1511-1519 **(1993)**

34. Fiser A., Do R. K. and Sali A., Modeling of loops in protein structures. *Protein Science* **9,** 1753-1773**(2000)**

35. Guex N. and Peitsch M. C., SWISS-MODE: and the Swiss-PdbViewer: An environment for comparative protein modeling, *Electrophoresis*, **18,** 2714-2723 **(1997)**

36. Holm L. and Rosenstrom P. Dali server: conservation mapping in 3D, *Nucl Acids Res* **38,** 545-549 **(2010)**

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

38. Wiederstein M. and Sippl M. J., ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins, *Nucleic Acids Res* , **35**, W407-W410 **(2007)**

39. Ritchie D.W., Evaluation of protein docking predictions using Hex 3.1 in CAPRI rounds 1 and 2. Proteins, *Structure, Function, and Bioinformatics* **52,** 98–106 **(2003)**

40. Ramachandraiah G., Chandra N. R., Surolia A. and Vijayan M., Re-refinement using reprocessed data to improve the quality of the structure: a case study involving garlic lectin, *Acta Crystallogr D Biol crtstallogr*, **58,** 414-420 **(2002)**

41. Burton S. L., Ellar D. J., Li J. and Derbyshire D. J., N-acetylgalactosamine on the putative insect receptor aminopeptidase N is recognised by a site on the domain III lectin-like fold of a Bacillus thuringiensis insecticidal toxin, *J. Mol. Biol.* **287,** 1011–1022 **(1999)**

42. Hofmann C., Vanderbruggen H., Hofte H., Van Rie J., Jansens S. and Van Mel- laert H., Specificity of Bacillus thuringiensis delta-endotoxins is correlated with the presence of high-affinity binding sites in the brush border membrane of target insect midgets, *Proc. Natl. Acad. Sci* **85,**7844–7848 **(1988)**

43. Jin X., Walker M. A., Felsovalyi K., Vendome J., Bahna F., Mannepalli S., Cosmanescu F., Ahlsen G., Honig B. and Shapiro L., Crystal structures of Drosophila N-cadherin ectodomain regions reveal a widely used class of Ca^{2+}-free interdomain linkers, *Proc Natl Acad Sci*, U S A **3,** 127-134 **(2012)**

44. Gribskov M., McLachlan A. D. and Eisenberg D., Profile analysis: Detection of distantly related proteins, *Proc. Natl. Acad. Sci.* **84,** 4355-4358 **(1987)**

45. Krogh A., Brown M., Mian I. S., Sjolander K. and Haussler D., Hidden Markov models in computational biology: Applications to protein modeling, *J. Mol. Biol.* **235,** 1501-1531 **(1994)**

46. Eddy S. R., Hidden Markov models, *Curr. Opin. Struct. Biol.* **6,** 361-365 **(1996)**

47. Kashyap S., Singh B.D. and Amla D.V., Homology modeling deduced structure of the Cry1Ab22 toxin, *Indian journal of biotechnology,***10,** 202-206 **(2011)**

48. Tanje S., Sanam R., Gundla R., Gandhi N.S., Mancera R.L., Boddupally D., Vudem D. R. and Khareedu V. R., Molecular modeling of Bt Cry1Ac (DI-DII)-ASAL (Allium sativum lectin)-fusion protein and its interaction with aminopeptidase N (APN) receptor of Manduca sexta, *J. Mol. Graphics Modell*, **33**, 61-76 **(2012)**