Journal of Molecular Graphics and Modelling 33 (2012) 61–76

Contents lists available at SciVerse ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/JMGM

Molecular modeling of *Bt* Cry1Ac (DI–DII)–ASAL (*Allium sativum* lectin)–fusion protein and its interaction with aminopeptidase N (APN) receptor of *Manduca sexta*

Sunita Tajne[a,b], Ramadevi Sanam[b], Rambabu Gundla[b], Neha S. Gandhi[c], Ricardo L. Mancera[c,d], Dayakar Boddupally[a], Dashavantha Reddy Vudem[a], Venkateswara Rao Khareedu[a,*]

[a] *Centre for Plant Molecular Biology, Osmania University, Hyderabad 500 007, India*
[b] *Division of Informatics, GVK Biosciences Pvt. Ltd., S-1, Phase-1, T.I.E. Balanagar, Hyderabad 500037, India*
[c] *Western Australian Biomedical Research Institute, Curtin Health Innovation Research Institute, School of Biomedical Sciences, Curtin University, GPO Box U1987, Perth, WA 6845, Australia*
[d] *School of Pharmacy, Curtin University, GPO Box U1987, Perth, WA 6845, Australia*

ARTICLE INFO

Article history:
Accepted 4 November 2011
Available online 23 November 2011

Keywords:
Allium sativum lectin (ASAL)
Bacillus thuringiensis
Cry proteins
Fusion protein
Homology modeling
Insect resistance
Protein–protein docking
Molecular mechanics
Poisson–Boltzmann surface area (PBSA)
Generalized Born surface area (GBSA)

ABSTRACT

Genetic engineering of *Bacillus thuringiensis* (*Bt*) Cry proteins has resulted in the synthesis of various novel toxin proteins with enhanced insecticidal activity and specificity towards different insect pests. In this study, a fusion protein consisting of the DI–DII domains of Cry1Ac and garlic lectin (ASAL) has been designed *in silico* by replacing the DIII domain of Cry1Ac with ASAL. The binding interface between the DI–DII domains of Cry1Ac and lectin has been identified using protein–protein docking studies. Free energy of binding calculations and interaction profiles between the Cry1Ac and lectin domains confirmed the stability of fusion protein. A total of 18 hydrogen bonds was observed in the DI–DII–lectin fusion protein compared to 11 hydrogen bonds in the Cry1Ac (DI–DII–DIII) protein. Molecular mechanics/Poisson–Boltzmann (generalized-Born) surface area [MM/PB (GB) SA] methods were used for predicting free energy of interactions of the fusion proteins. Protein–protein docking studies based on the number of hydrogen bonds, hydrophobic interactions, aromatic–aromatic, aromatic–sulphur, cation–pi interactions and binding energy of Cry1Ac/fusion proteins with the aminopeptidase N (APN) of *Manduca sexta* rationalised the higher binding affinity of the fusion protein with the APN receptor compared to that of the Cry1Ac–APN complex, as predicted by ZDOCK, Rosetta and ClusPro analysis. The molecular binding interface between the fusion protein and the APN receptor is well packed, analogously to that of the Cry1Ac–APN complex. These findings offer scope for the design and development of customized fusion molecules for improved pest management in crop plants.

1. Introduction

The gram-positive bacterium *Bacillus thuringiensis* (*Bt*) has been used extensively to control major insect pests in plants. The insecticidal activity of *Bt* is mainly attributed to the crystal proteins encoded by the *cry* genes [1,2]. *Bt* synthesizes cytoplasmic inclusions during sporulation that contain one or more Cry proteins as inactive protoxins [3]. Ingestion of crystal proteins by insect larvae results in the conversion of protoxins into active toxins at alkaline pH. The activated toxins bind to insect-specific receptors present on the plasma membrane of the midgut epithelial cells and create transmembrane pores, leading to disruption of ionic balance, cell lysis and death of insects [4].

Diverse Cry fusion toxins have been built from the DI, DII and DIII conserved domains of *Bt* proteins since members of Cry protein family adopt a similar protein folding pattern [5]. The X-ray structures of various Cry toxins, such as Cry1Aa [6], Cry2Aa [7], Cry3Aa [8], Cry3Bb1 [9], Cry4Aa [10] and Cry4Ba [11] have been determined. These proteins exhibit similar organization into three domains. Domain-I consists of a seven-alpha-helical bundle and is involved in membrane insertion and pore formation [6,12]. Domain-II, containing three antiparallel beta-sheets with exposed loop regions, plays a dual role in binding to insect brush border membranes and receptor recognition [13,14]. Domain-III exists as a beta-sandwich and is involved in different functions, including the determination of insect specificity [15,16] and ion channel activity [17]. Among Cry proteins, Cry1Ac is the Cry toxin with highest

* Corresponding author. Tel.: +91 40 27098087; fax: +91 40 27096170.
E-mail address: rao_kv1@rediffmail.com (V.R. Khareedu).

1093-3263/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2011.11.001

insecticidal activity against different lepidopteron insects [18,19]. Glycosylphosphatidyl-inositol (GPI) anchored aminopeptidase N (APN) [20] and cadherin-like protein [21] have been studied extensively as receptors of Cry1A toxins. APN belongs to the zinc-binding metalloprotease family of proteins. The C-terminal stalk of the APN binding site has been found to be rich in N-acetylgalactosamine (GalNAc), and this region acts as the binding site of the Cry1Ac toxin [20,22].

Lectins constitute a unique group of sugar-binding proteins which can recognize 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 [23]. Transgenic crop plants expressing various lectins have been developed to confer protection against different pests. Plant lectins exhibit their insecticidal activity mostly through their carbohydrate-binding properties [24]. The 3D crystal structures of different plant lectins have revealed the presence of a beta-prism-II fold comprising three antiparallel four-stranded beta-sheets arranged as a 12-stranded beta-barrel [25]. It has been reported that the DIII domain of Cry1A proteins exhibits a lectin-like fold [10,26].

Different *Bt* fusion proteins, such as Cry1Ab–Cry1B [27], Cry1Ac–Cry1Ab [28] and Cry1Ac–GFP [29], have been synthesized and expressed in different plant systems. Transfer of the carbohydrate binding domain-III of Cry1C to Cry1Ab generated a hybrid (Cry1Ab–Cry1C) protein that proved highly toxic to the armyworm (*Spodoptera exigua*) and which was resistant to Cry1A toxin alone [30,31]. Transgenic potato, expressing a hybrid Cry protein containing domains I and III from Cry1Ba and domain II from Cry2a exhibited high-level resistance to the potato tuber moth (*Phthorimaea operculella*) and Colorado potato beetle (*Leptinotarsa decemlineata*) [32]. Based on these observations, it would be ideal to design new fusion molecules with higher insecticidal activities than that of native proteins through innovative approaches. Protein–protein docking is a powerful technique for predicting the structures of protein complexes, especially in cases where the crystal structures of these complexes are difficult to determine. The ZDOCK, RDOCK [33], Rosetta [34] and ClusPro [35] programs have been used successfully for making high accuracy predictions of the structures of multiple known protein–protein complexes in the CAPRI (Critical Assessment of PRedicted Interaction) challenge [33,36,37].

Earlier, we reported that *Allium sativum* lectin (ASAL)-expressing transgenic rice could exhibit high insecticidal activity against major sap-sucking pests of rice, viz., brown planthopper (BPH), green leafhopper (GLH) and whitebacked planthopper (WBPH) [23,38]. Garlic lectin not only reduced the growth and fecundity of the insects but also caused adverse affects in second generation insects when reared on an artificial diet supplemented with the lectin [39]. Furthermore, garlic is known to have beneficial effects on human health, and as such might find greater acceptance as a candidate gene for insect resistance in crop plants [40]. In the present *in silico* study we designed and constructed a fusion protein by replacing the DIII of Cry1Ac with the garlic lectin to evaluate the functional role of the fusion protein in terms of its toxicity and binding ability towards the APN receptor protein. Rigid body protein–protein docking simulations were performed employing Cry1Ac–lectin fusion protein structures extracted from molecular dynamics (MD) simulations to study its interactions with APN of *Manduca sexta*. The ZDOCK, RDOCK, Rosetta and ClusPro programs were used to analyze the binding interactions between the insect APN and the Cry1Ac-lectin fusion protein as well as the Cry1Ac protein. Docking of the Cry1Ac–lectin fusion complex with the APN receptor revealed higher binding affinity and an enhanced contact pattern with the receptor when compared to the Cry1Ac protein.

2. Materials and methods

2.1. Homology modeling of Cry1Ac (DI–DII–DIII) and M. sexta APN receptor

Sequences of the DI–DII–DIII domains of Cry1Ac (Swiss-Prot entry P05068) and APN (Swiss-Prot entry Q11001) were used to build up the 3D structures using the comparative protein modeling method of MODELER [41]. The BLAST [42] algorithm was used to identify homologous structures by searching the Protein Data Bank (PDB). Based on the highest sequence similarity crystal structures of Cry1Aa (PDB ID: 1CIY) and endoplasmic reticulum aminopeptidase 1 (PDB ID: 2XDT) were selected as templates for homology modeling of Cry1Ac and APN, respectively. The ClustalW program [43] was employed to identify the structurally conserved and variable regions between the query and the template proteins. Restraints, distances and dihedral angles were extracted from the template structures. Stereochemical restraints, viz., bond length and bond angle preferences, were obtained from the molecular mechanics force field CHARMM-22 [44]. Models of Cry1Ac and APN were further refined using 600 ps MD simulations in explicit water. Minimization was carried out using the consistent valence force field (CVFF) [45] with a van der Waals cutoff of 9.5 Å and a distance-dependent dielectric constant of $1 \times r$. One thousand steps of steepest descents were performed followed by 1000 steps of conjugate gradients until the root mean square (RMS) gradient reached a value of less than 0.001 kcal/mol/Å. The homology models were each solvated with a 10 Å water layer and optimized using MD simulations for 2 ns at a 300 K temperature. All molecular modeling studies were carried out using Discovery Studio (version 2.1, Accelrys Inc., USA). The stereochemical quality of the Cry1Ac and APN models was estimated using the program PROCHECK [46].

2.2. Construction of fusion protein (Cry DI–DII–lectin) by protein–protein docking and MD simulations

The amino acid sequence of the Cry1Ac DIII domain was aligned with the garlic lectin sequence to analyze the similarity between them. Their secondary structures were compared using ClustalW and JPred [47]. As a homology model of Cry1Ac was determined for all the three domains, the amino acid coordiantes of the DIII domain were deleted. The remaining structure containing the DI–DII domains of Cry1Ac was used to generate the fusion-protein comprising these domains and the ASAL protein. The rigid-body docking program (ZDOCK 2.1), which employs a fast Fourier transform (FFT) algorithm, was used to generate the fusion protein. Initially, to evaluate the accuracy of the ZDOCK program, the DIII domain was docked to the DI–DII domains. The DI–DII domains were treated as the fixed target protein (receptor protein) and the DIII domain as a probe protein (ligand protein) in 120 docking runs. The correct assembly of the DI–DII–DIII complex was observed in the top 10 poses, validating this approach. The garlic lectin (ASAL) was then docked to the DI–DII domains (fixed target protein) to identify the binding interface with ASAL in the fusion protein. The crystal structure of garlic lectin (PDB ID: 1KJ1-chain A) [48] was taken as the probe protein and the DI–DII domains were considered to be the target protein. The top 30 predicted complexes were grouped into different clusters based on their interactions. Complexes showing residues of the lectin protein interacting with the DI–DII domain of the Cry1Ac protein were selected and their non-bonded interactions examined. The stereochemical quality of

the DI–DII–lectin fusion protein model was analyzed using protein structure validation programs, viz., PROCHECK and Qualitative Model Energy Analysis (QMEAN) [49].

2.3. MD simulations in explicit water

Initial coordinates of the protein systems for simulations were obtained from the homology models and subsequent docking studies. For MM/PB(GB)SA calculations, initially DIDII was considered as a receptor while DIII and lectin were considered as ligands. For further MM/PB(GB)SA calculations, APN was considered as a receptor and DIDIIDIII and DIDII–lectin were considered as ligands. Proteins described in Supplementary information Table S1 were modeled with the Amber ff99SB [50] force field implemented in the AMBER 11 [51] MD simulation program. The complexes were solvated with TIP3P [52] water using a cubic periodic box extending at least 12 Å from the complex. Net charges in the receptor–ligand system were neutralized by adding an appropriate number of counter ions (Supplementary information Table S1). Complexes were first optimized by 500 steps of steepest descents followed by another 500 steps of conjugate gradients minimization, keeping all atoms of the complex restrained to their initial positions with a weak harmonic potential. After minimization, the system was heated from 50 K to 300 K with a restraining force constant of 2 kcal/mol/Å applied to the complex in order to equilibrate the solvent at 300 K without undesirable drifts of the structure. Later, a 50 ps MD simulation with a restraining force constant of 2 kcal/mol/Å was applied to the complex and was carried out at a pressure of 1 atm using Berendsen's barostat [53] to equilibrate the density. This was followed by equilibration for 7.5 ns without restraints. Finally, a 14 ns production simulation at constant pressure was conducted. Full flexibility of the receptor and ligands was allowed during the production stage of simulations. Coordinates were saved every 10 ps. All bond lengths involving hydrogen atoms were constrained using the SHAKE [54] algorithm with a 2 fs time step. The temperature was kept fixed at 300 K using the Langevin thermostat [55] with a collision frequency of 2 ps^{-1}. Electrostatic interactions were treated with particle-mesh Ewald (PME) scheme [56] with a fourth-order B-spline interpolation and a tolerance of 10^{-5} Å. The non-bonded cutoff applied was 8 Å and the nonbonded pair list was updated every 50 fs. Various properties such as density, temperature, pressure, kinetic, and potential energies were monitored during the simulations to ensure proper equilibration.

2.4. MM–PBSA calculations

In the MM–PBSA method, free energies were calculated for representative 'snapshot' structures taken from the MD trajectories of the system of interest. This method combines explicit solvation simulations with calculations of electrostatic contribution to the solvation free energy using the Poisson–Boltzmann approach and of the non-polar solvation free energy to estimate free energy of binding [57,58]. Free energies of binding are defined as:

$$\Delta G_{\text{binding}} = \Delta G_{\text{gas}} + \Delta G_{\text{sol-cmplx}} - [\Delta G_{\text{sol-prot}} + \Delta G_{\text{sol-lig}}] \quad (1)$$

where ΔG_{gas} is the interaction energy between protein and ligand in the gas phase, which was calculated using a molecular mechanics approach. $\Delta G_{\text{sol-prot}}$, $\Delta G_{\text{sol-lig}}$ and $\Delta G_{\text{sol-cmplx}}$ are solvation free energies of the protein, ligand and ligand–protein complex, respectively, which were estimated using a continuum Poisson–Boltzmann/surface area approach (or the generalized Born/surface approach in the case of GBSA calculations). In MM–PBSA, the free energy of binding is calculated as

$$\Delta G_{\text{binding}} = \Delta E_{\text{MM}} + \Delta G_{\text{PBSA}} - T\Delta S \quad (2)$$

where ΔE_{MM} is the difference in the average molecular mechanics energy, which was calculated as

$$\Delta E_{\text{MM}} = \Delta E_{\text{int}} + \Delta E_{\text{vdw}} + \Delta E_{\text{elec}}, \quad (3)$$

where $\Delta E_{\text{int}} = \Delta E_{\text{bond}} + \Delta E_{\text{angle}} + \Delta E_{\text{tors}}$.

ΔE_{int} corresponds to the sum of differences in average internal bond stretching, bond bending and torsional angle energies. ΔE_{vdw} is the difference in average van der Waals energy, while ΔE_{elec} is the difference in average electrostatic energy.

ΔG_{PBSA} is the free energy of solvation, given by:

$$\Delta G_{\text{PBSA}} = \Delta G_{\text{PB}} + \Delta G_{\text{SA}}. \quad (4)$$

where ΔG_{PB} is the electrostatic component of the free energy of solvation calculated by solving the Poisson–Boltzmann equation or the generalized Born equation in case of GB calculations [59]. ΔG_{SA} is the non-polar contribution to the free energy of solvation calculated from the solvent-accessible surface area (SASA) [60]. This term is computed with the equation $\Delta G_{\text{SA}} = \gamma \text{SA} + \beta$, where SA is the solvent-accessible surface area calculated by MSMS program [61], while γ and β are parameterized constants.

The final term, $T\Delta S$, is the change in entropy upon the formation of ligand–protein complex, and is approximated by analytical statistical mechanics equations for rotational and translational entropy and by performing normal modes calculations for the vibrational entropy. Alternatively, a quasi-harmonic approach can be applied to determine conformational entropies [62,63]. Here, entropy was calculated using normal modes analyses for DIDII–DIII and DIDII–lectin systems. Since the APN systems are too large for normal modes analyses, entropy calculations were performed using quasi-harmonic approximation. The MM–PBSA.*py*.MPI module of AMBER 11.0 [51] and AmberTools 1.5 [64] were used to compute the components of free energy. For 14.0 ns simulations, 1400 snapshots of the coordinates of the system were taken at 10 ps intervals. All solvent molecules and counter ions were removed prior to analysis. The snapshots were analyzed using generalized Born (GB) solvation model [65], modified for use with PARM99SB parameters to obtain free energies of solvation. Poisson–Boltzmann calculations were solved numerically by the PBSA program included with AmberTools to obtain free energies of solvation with a dielectric constant (ε) of 1 for the solute and 80 for the solvent. The "OBC" model [66] (igb = 5) was used, with values for α, β and γ constants of 1.0, 0.8, and 4.85, respectively. The surface parameters used for PB calculations were $\gamma = 0.00542$ kcal/Å^2 and $\beta = 0.92$ kcal/mol [60,61]. PBSA/GBSA calculations were performed assuming an ionic strength of 0.1 M.

2.5. Protein–protein docking of Cry1Ac (DI–DII–DIII) and fusion proteins with APN using ZDOCK, Rosetta and ClusPro

The DI–DII–lectin fusion complexes showing interactions similar to the DI–DII–DIII protein were selected for further docking to APN. Domain-II residues, Arg249-Met265 and His278-Ala312 [67], were selected as the APN interacting residues in the ZDOCK and ClusPro programs. For APN, C-terminal residues (Phe700-Asp894) rich in GalNAc were considered as the Cry1Ac binding site for both the programs. The Rosetta program was used to search the rigid-body and side-chain conformational space of Cry1Ac and the fusion protein in their interactions with APN. Main chain-main chain, main chain-side chain and side chain-side chain hydrogen bonds as well as hydrophobic, ionic, aromatic–aromatic, aromatic–sulphur and cation–pi interactions between the DI–DII–lectin fusion/Cry1Ac (DI–DII–DIII) proteins and APN were analyzed for the top 20 ZDOCK, 10 Rosetta

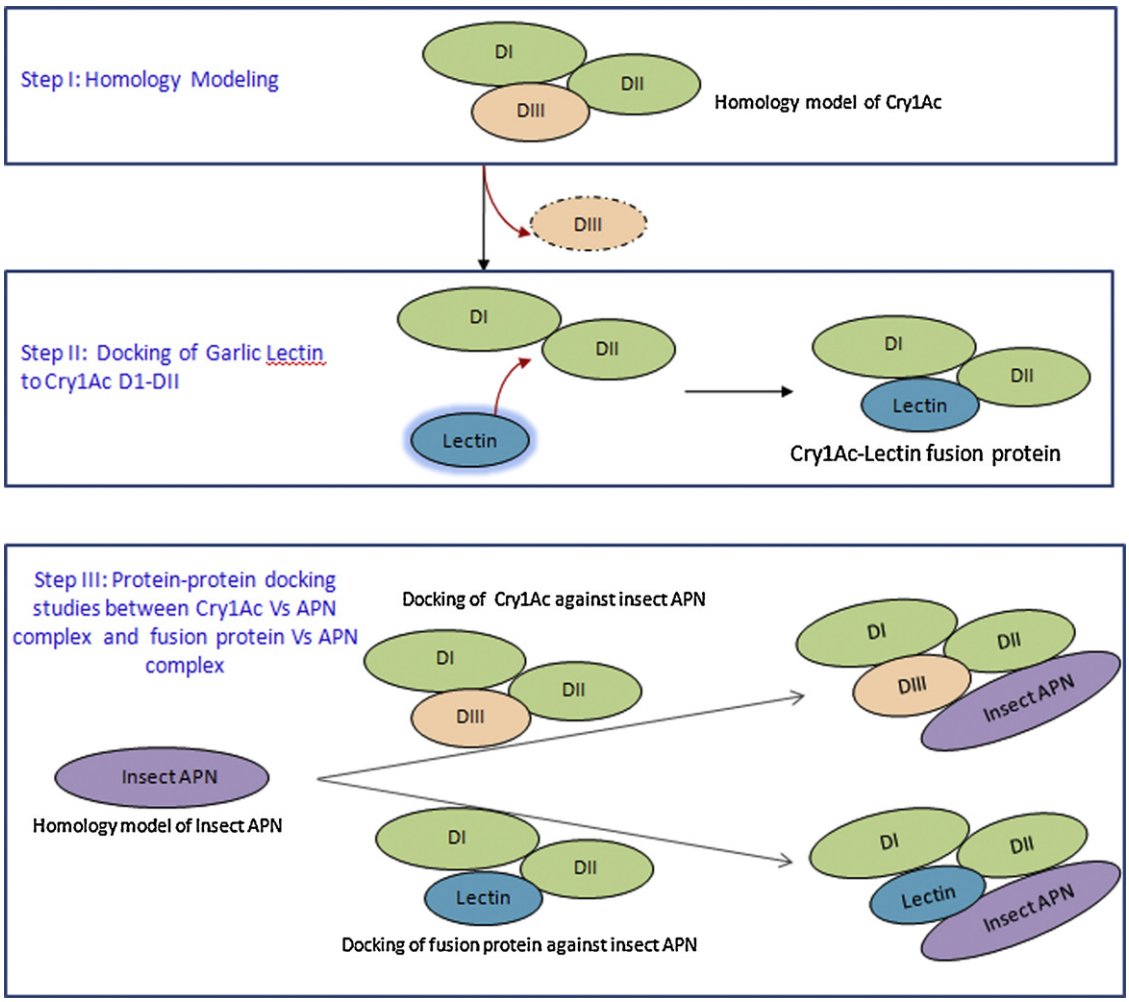

Fig. 1. Steps involved in the analysis of interactions between the fusion protein (DI–DII of Cry1Ac and garlic lectin) and APN of *M. sexta* predicted by protein–protein docking.

and 15 ClusPro predicted poses, employing the protein interaction calculator (PIC) [68]. For hydrogen bonds, the cutoff for the angle (acceptor–donor-hydrogen) and distance (acceptor-hydrogen) were set to 30° and 3.5 Å, respectively. The distance between a carbon atom and the closest nonpolar part of another residue was calculated with a cutoff of 5 Å was calculated for hydrophobic interactions. Aromatic–aromatic interactions were considered for pairs of phenyl ring centroids that are separated by a preferential distance between 4.5 and 7 Å. Similarly, based on the report of Nakanishi et al. [69], the N-terminal domain I of APN, involved in binding with Cry, was also subjected to docking with the fusion protein in another round of simulations using ZDOCK.

Free energies of binding were calculated for the protein–protein complexes using the Dcomplex program [70] to verify the stability of the protein complexes obtained using the docking programs and to understand the effects of protein conformational changes on binding affinity. Heat maps were generated for protein–protein interactions based on interacting residues in order to identify the critical residues involved in binding. Green color was used to indicate the presence of interactions with the given amino acid, while red color was used to indicate the absence of interactions. The work flowchart of molecular modeling and protein–protein docking of the insecticidal fusion protein (DI–DII–lectin) is shown in Fig. 1.

3. Results and discussion

3.1. Homology modeling of Bt Cry1Ac and M. sexta APN

The 3D structures of Cry1Ac (DI–DII–DIII) and *M. sexta* APN proteins were constructed by homology modeling. Using the crystal structure of Cry1Aa as a template, the 3D model of Cry1Ac was generated for the amino acid stretch Tyr33 to Thr609. The sequence identity and similarity between Cry proteins were ~73% and ~79%, respectively. As shown in Fig. 2a, the sequence identity between the DI domain of these proteins is very high (98%), while for the DII and DIII domains the identity was ~69% and ~38%, respectively. Superimposition of the homology model of Cry1Ac with the crystal structure of Cry1Aa revealed certain variations in the geometry of the DIII domain, while the DI and DII domains were found to be conserved (Fig. 2b). Several hydrogen bonds and van der Waals and electrostatic interactions were observed between the DI–DII–DIII domains of Cry1Ac (data not shown). In the DI and DII domains of Cry1Ac, four highly conserved salt bridges were identified between the amino acid pairs Arg201-Glu256, Arg202-Glu242, Asp190-Arg249 and Asp210-Arg233, similarly to Cry1Aa (Fig. 2b). These interdomain salt bridges might contribute significantly to the stability of Cry proteins. It is well known that the receptor affinity of Cry1Ac mainly lies in the DII and DIII domains. Different lepidopteran insects are known to show differential sensitivity

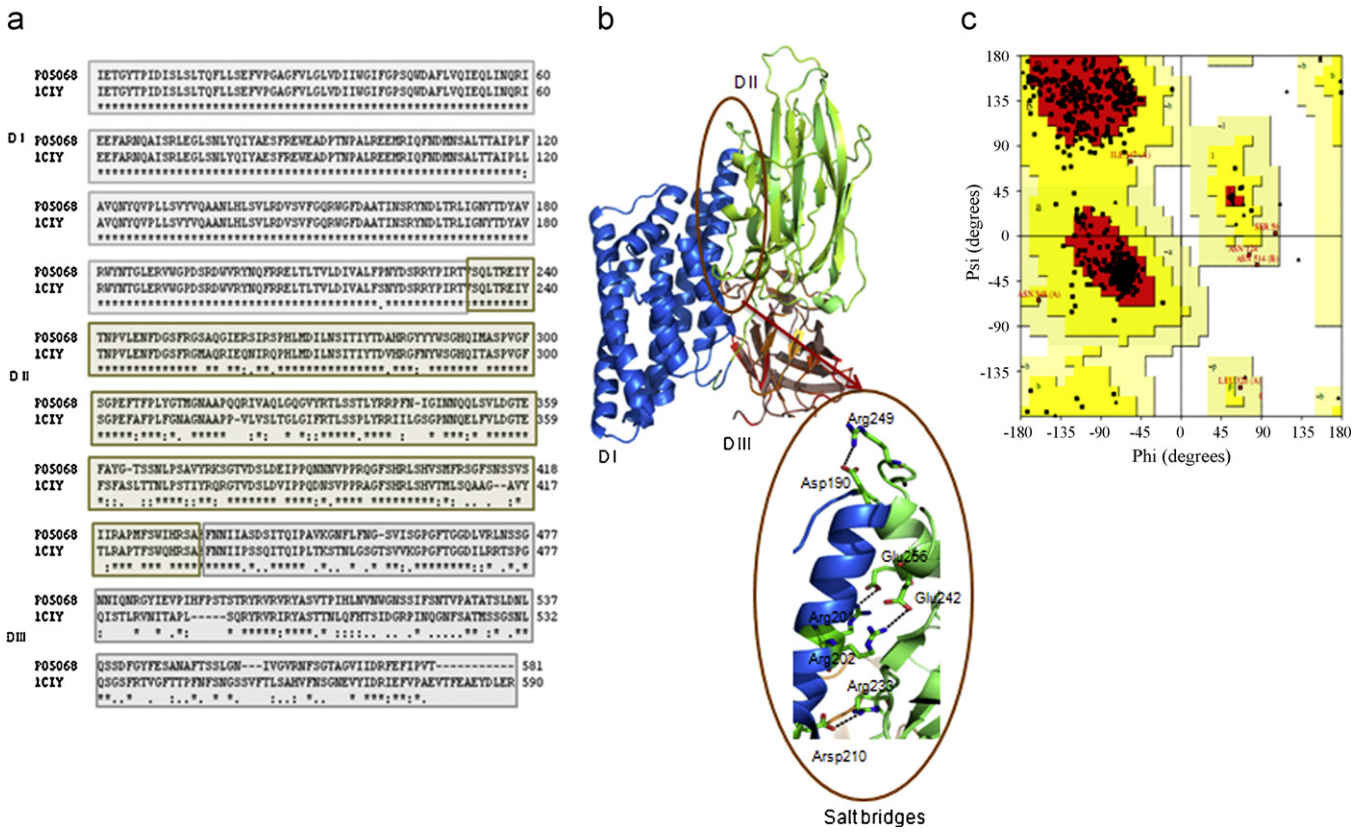

Fig. 2. Sequence alignment, 3D structure, salt bridges and Ramachandran plot analysis of Cry1Ac. (a) Sequence alignment between Cry1Ac (P05068) and Cry1Aa (1CIY). The three domains are shown in different boxes. (b) Structural superimposition of the Cry1Ac model with Cry1Aa (1CIY). Domain I is completely overlapped, domain II is slightly variable and domain III is structurally different. (c) Ramachandran plot for the Cry1Ac model determined by PROCHECK.

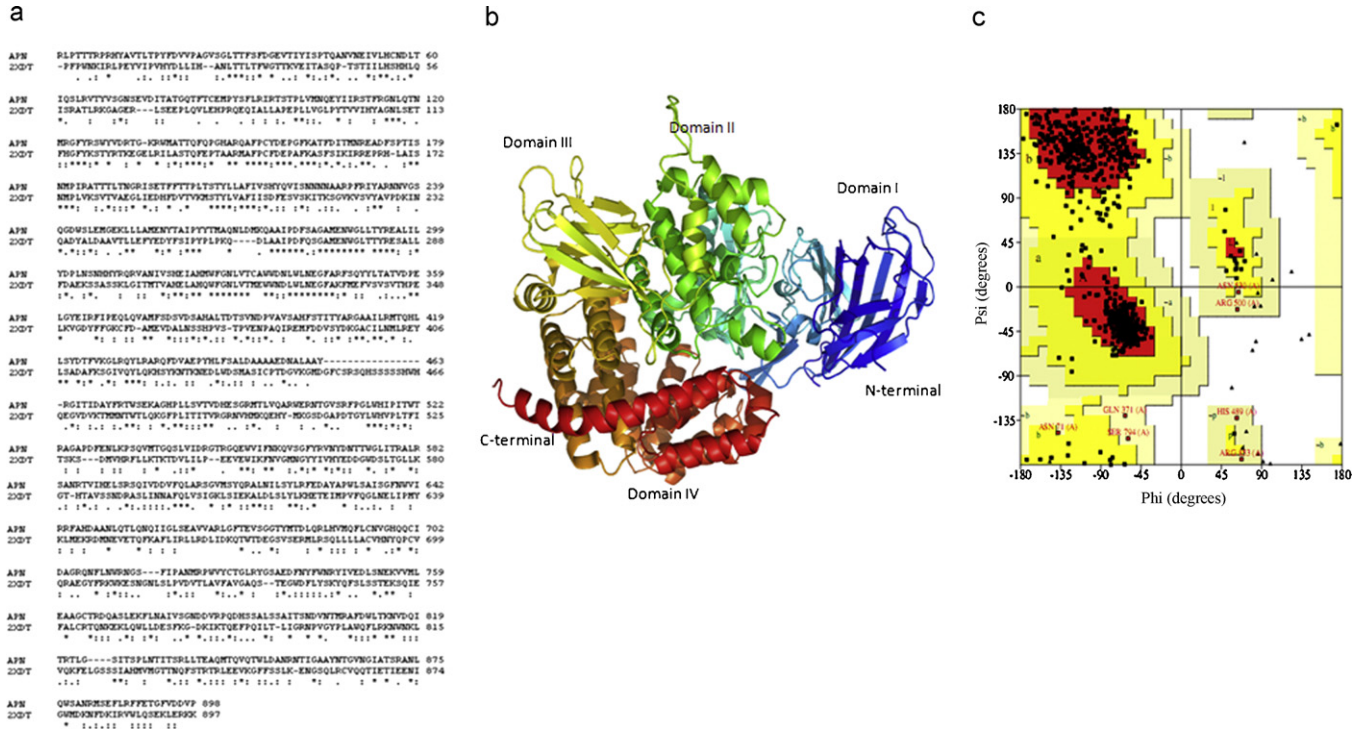

Fig. 3. Sequence alignment, 3D structure and Ramachandran plot analysis of APN of *M. sexta*. (a) Sequence alignment between APN and endoplasmic reticulum aminopeptidase 1. (b) Homology model of APN of *M. sexta*. (c) Ramachandran plot of the APN model determined by PROCHECK.

Table 1
Binding energy and ZDOCK score of Cry1Ac DI–DII with DIII and DI–DII with garlic lectin (ASAL) complex.

Ligand protein	Receptor protein	Predicted binding energy (kcal/mol)[a]	ZDOCK score[b]	E_RDOCK
DIII	DI–DII	−19.153	15.02	−8.291
Lectin	DI–DII	−21.532	15.92	−11.724

[a] Binding energy calculated by Dcomplex.
[b] ZDOCK and E_RDOCK score were calculated by ZDOCK program of Discovery Studio.

to various Cry toxins since various domain regions influence the receptor binding affinity and their activity [71].

The sequence similarity between the APN receptor protein (Leu2 to Asp894) and the template endoplasmic reticulum aminopeptidase 1 (Pro46 to Arg940) was found to be ~48% (Fig. 3a). The 3D model of the APN protein showed four distinct domains with the alpha/beta fold having a predominantly alpha helical central core and several peripheral domains of beta strands (Fig. 3b). Domain IV is exclusively composed of 19 alpha helices forming a right-handed super-helical solenoid structure which forms a top layer covering domains II and III containing a large cavity. Domain IV contains GalNAc and possesses all the ten O-glycosylation sites and serves as the Cry1Ac binding site [20,22]. The stereochemical quality of the homology models of Cry1Ac and APN was analyzed using the PROCHECK program. Ramachandran plot analysis revealed that ~98% of the amino acid residues are present in the most favored and additional allowed regions (Figs. 2c and 3c), indicating that the two models have good stereochemical properties.

3.2. Docking of garlic lectin/DIII protein of Cry1Ac (ligand proteins) to the DI–DII domains of Cry1Ac (receptor protein)

Low sequence similarity was observed between the DIII domain of Cry1Ac and lectin (Supplementary information Fig. 1A). The secondary structures of the DIII domain and lectin revealed that both are beta proteins with 9 and 10 beta strands, respectively (Supplementary information Fig. 1B). Superimposition of these proteins revealed a high degree of structural similarity between them (Supplementary information Fig. 1C). It has been previously reported that the DIII domain of Cry1Ac, which has a lectin-like fold and a carbohydrate binding domain similar to that of garlic lectin, contains essential binding sites for interaction with the APN protein [10,26]. To examine the binding mode, stability and interaction profiles of ligand and receptor proteins, the crystal structure of garlic lectin was docked to the DI–DII domains of Cry1Ac. The homology model of the DIII domain of Cry1Ac was also docked against the DI–DII domains. Rigid-body docking of these proteins employing the ZDOCK program generated 2000 potential structures of these complexes. These complexes were further reranked and reduced to 30 structures by employing the RDOCK program. The binding energies and ZDOCK scores for the best conformations of the DI–DII–DIII and DI–DII–lectin complexes are presented in Table 1. Interactions between the garlic lectin and the DI–DII domains showed a similar orientation to that of the DIII domain with the DI–DII domains (Fig. 4). Out of 18 hydrogen bonds observed in the fusion protein (Supplementary information Table S2), 5 bonds were found to be retained in the DI–DII–DIII domains of Cry1Ac (Table 2). In addition, ionic interactions (Table 3), hydrophobic, aromatic–aromatic and cation–pi interactions were also analyzed in the Cry1Ac and fusion proteins (data not shown).

The structural stability of the fusion protein was assessed by calculating the free energy of binding of the DI–DII–lectin complex, which provided a thermodynamic measure of the strength of interaction between the two proteins. MM/PB(GB)SA calculations for the predicted free energy of binding and its various components for all the force field/water potential combinations are presented in Tables 4 and 5. When using the MM/PBSA method, the free energy of binding between DI–DII domains and the DIII

Table 3
Protein–protein ionic interactions between DIII and lectin against DI–DII.

DI–DII residue	DIII residue	Lectin residue
ARG230	ASP441	–
ASP177	–	LYS54
ASP177	–	GLU69
ARG226	–	ASP60
HIS267	–	ASP17

Table 2
Hydrogen bond interactions between DIII of Cry1Ac and lectin (ASAL) against DI–DII of Cry1Ac.

DI–DII residues	DIII				Lectin			
	DI–DII atom	Residues	DIII atom	Length (Å)	DI–DII atom	Residues	Atom	Length (Å)
GLY173	–	–	–	–	O	LYS54	NZ	3.29
ASN174	–	–	–	–	OD1	ALA68	O	2.82
ASP177	OD1	LEU534	N	3.23	OD2	GLY52	O	3.08
ASP177	–	–	–	–	OD2	LYS54	NZ	2.63
VAL216	–	–	–	–	O	TYR11	OH	2.65
ALA217	–	–	–	–	N	TYR11	OH	3.47
ASN221	–	–	–	–	ND2	GLY7	O	3.07
ARG226	–	–	–	–	NH1	GLN58	OE1	2.93
ARG226	–	–	–	–	NH1	ASP60	OD2	2.98
ARG226	–	–	–	–	NH2	TYR66	OH	3.32
TYR227	OH	TYR499	OH	2.55	OH	LEU57	O	3.1
ARG230	O	ASP441	OD1	2.48	–	–	–	–
ARG230	O	SER442	N	2.89	–	–	–	–
THR231	N	ASP441	OD1	3.21	OG1	ASN6	O	2.96
VAL232	O	THR444	N	3.13	O	THR5	OG1	3.21
VAL232	N	SER442	O	3.13	–	–	–	–
SER233	–	–	–	–	OG	ASN6	O	3.4
GLN234	NE2	THR444	O	3.27	N	GLU8	OE2	3.11
GLN234	–	–	–	–	O	GLU8	OE2	2.83
THR236	N	GLY467	O	2.89	–	–	–	–
ARG237	N	GLY467	O	3.12	–	–	–	–
HIS267	–	–	–	–	O	ARG1	N	3.06
LEU268	–	–	–	–	N	ASP17	O	3.03
MET269	N	ALA448	O	2.71	–	–	–	–

Table 4
MM/PB(GB)SA energy component analysis of DIDII with DIII.

Energy components	Complex			Receptor			Ligand			Delta (complex–receptor–ligand)		
	Average	Std. dev.	Std. error	Average	Std. dev.	Std. error	Average	Std. dev	Std. error	Average	Std. dev.	Std. error
MM/GBSA energy component analysis of interactions of the DIDII with the DIII domain averaged over 1400 snapshots												
VDW	−4457.83	46.74	2.81	−3378.26	34.35	2.06	−938.53	16.89	1.01	−141.05	8.71	0.52
EEL	−37405.3	116.08	6.97	−27857.8	98.32	5.91	−9229.9	55	3.3	−317.58	38.14	2.29
EGB	−5131.46	82.27	4.94	−3948.44	72.78	4.37	−1614.19	40.22	2.42	431.17	33.31	2
ESURF	229.18	5.65	0.34	179.35	3.8	0.23	68.79	2.07	0.12	−18.96	0.94	0.06
Ggas	−41863.1	106.1	6.37	−31236.1	95.05	5.71	−10168.4	54.13	3.25	−458.63	37.49	2.25
Gsolv	−4902.28	81.14	4.88	−3769.09	71.91	4.32	−1545.4	39.27	2.36	412.21	32.93	1.98
GBTOTAL	−46765.4	65.56	3.94	−35005.2	53.08	3.19	−11713.8	38.38	2.31	−46.42	9.28	0.56
$T\Delta S$										−55.108	10.82	
$\Delta G_{\text{binding-GBSA}}$										8.69		
MM/PBSA energy component analysis of the interactions of the DIDII with the DIII domain averaged over 1400 snapshots												
VDW	−4457.83	46.74	2.81	−3378.26	34.35	2.06	−938.53	16.89	1.01	−141.05	8.71	0.52
EEL	−37405.3	116.08	6.97	−27857.8	98.32	5.91	−9229.9	55	3.3	−317.58	38.14	2.29
EPB	−5563.49	84.53	5.08	−4230.77	72.63	4.36	−1721.48	40.09	2.41	388.77	36.21	2.18
Ecavity	153.05	3.83	0.23	122.12	2.73	0.16	49.32	1.24	0.07	−18.38	0.69	0.04
Ggas	−41863.1	106.1	6.37	−31236.1	95.05	5.71	−10168.4	54.13	3.25	−458.63	37.49	2.25
Gsolv	−5410.43	83.3	5.01	−4108.65	71.87	4.32	−1672.17	39.51	2.37	370.39	35.91	2.16
PBTOTAL	−47273.6	71.88	4.32	−35344.7	55.17	3.32	−11840.6	41.9	2.52	−88.24	11.3	0.68
$T\Delta S$										−55.11	10.82	
$\Delta G_{\text{binding-PBSA}}$										−33.13		

VDW = van der Waals contribution from MM; EEL = electrostatic energy as calculated by the MM force field; EPB/EGB = the electrostatic contribution to the solvation free energy calculated by PB or GB respectively; ECAVITY = nonpolar contribution to the solvation free energy calculated by an empirical model; GBTOTAL/PBTOTAL = final estimated binding free energy calculated from the terms above; $T\Delta S$, T (temperature); ΔS (sum of rotational, translational and vibrational entropies); ΔG binding, total binding energy of the system. Entropy calculations were based on normal modes analysis using 80 snapshots. All energies are in kcal/mol.

domain was found higher (−33.1347 kJ/mol) than the total binding energy of DI–DII domains to the lectin (−60.6286 kJ/mol), indicating that the formation of fusion complex is more favorable. Calculation with MM/GBSA method of the interaction of the DI–DII domains with the lectin also predicted a lower free energy of binding (−17.1447 kJ/mol) (Table 5) compared to that of interaction of the DI–DII domains with the DIII domain (8.6924 kJ/mol) (Table 4), confirming that the formation of DI–DII–lectin complex is thermodynamically favored. The root mean square deviation (RMSD) of the coordinates in each snapshot with respect to the coordinates in the initial snapshot was monitored for simulation of DIDII with DIII and lectin ligand complex (Fig. 5). The simulations of DIDII–DIII and DIDII–lectin appear to converge more rapidly than that of the APN systems. A significant amount of backbone motion of ∼3.5 Å was observed in DIDII–DIII and DIDII–lectin complexes (Fig. 5). This conformation is mainly at N- and C-termini and do not disturb the protein–protein binding interface (data not shown).

The best conformation of the DI–DII–lectin fusion protein was selected based on the binding energy and ZDOCK. The low free energy of binding predicted by Dcomplex (−21.532 kcal/mol) of the fusion-protein suggests that the fusion-complex is more stable than the DI–DII–DIII complex (−19.153 kcal/mol) (Table 1). It is well known that hydrogen bonds play a prominent role in the binding affinity of protein–protein complexes [72,73]. The best predicted poses of the fusion protein and the DI–DII–DIII domains of Cry1Ac were evaluated in terms of the number of interchain

Table 5
MM/PB(GB)SA energy component analysis of DIDII with lectin.

Energy components	Complex			Receptor			Ligand			Delta (complex–receptor–ligand)		
	Average	Std. dev.	Std. error	Average	Std. dev.	Std. error	Average	Std. dev	Std. error	Average	Std. dev.	Std. error
MM/GBSA energy component analysis of the interactions of the DIDII with the lectin domain averaged over 1400 snapshots												
VDW	−4310.87	31.76	1.9	−3417.68	27.01	1.61	−745.53	12.67	0.76	−147.67	11.29	0.67
EEL	−35775.9	112.92	6.75	−27642.7	97.42	5.82	−7713.15	61.89	3.7	−420.04	48.22	2.88
EGB	−5151.12	88.39	5.28	−4152.97	68.92	4.12	−1514.06	44.37	2.65	515.92	40.3	2.41
ESURF	201.43	3.18	0.19	175.9	2.56	0.15	46.72	0.95	0.06	−21.19	1.17	0.07
Ggas	−40086.8	114.4	6.84	−31060.4	100.21	5.99	−8458.67	61.33	3.67	−567.7	46.16	2.76
Gsolv	−4949.69	87.07	5.2	−3977.07	67.81	4.05	−1467.35	43.91	2.62	494.73	40.21	2.4
GBTOTAL	−45036.5	66.25	3.96	−35037.5	64.04	3.83	−9926.02	29.61	1.77	−72.98	11.7	0.7
$T\Delta S$										−55.832	6.74	
$\Delta G_{\text{binding-GBSA}}$										−17.14		
MM/PBSA energy component analysis of the interactions of the DIDII with the lectin domain averaged over 1400 snapshots												
VDW	−4310.87	31.76	1.9	−3417.68	27.01	1.61	−745.53	12.67	0.76	−147.67	11.29	0.67
EEL	−35775.9	112.9	6.75	−27642.7	97.42	5.82	−7713.15	61.89	3.7	−420.04	48.22	2.88
EPB	−5529.47	90.27	5.39	−4410.03	69.57	4.16	−1589.7	43.22	2.58	470.25	38.34	2.29
Ecavity	139.1	1.84	0.11	119.09	1.42	0.08	39.02	0.5	0.03	−19.01	0.79	0.05
Ggas	−40086.77	114.4	6.84	−31060.4	100.21	5.99	−8458.67	61.33	3.67	−567.7	46.16	2.76
Gsolv	−5390.37	89.53	5.35	−4290.93	68.82	4.11	−1550.68	42.95	2.57	451.24	38.29	2.29
PBTOTAL	−45477.14	66.81	3.99	−35351.3	64.49	3.85	−10009.4	31.38	1.88	−116.46	15.07	0.9
$T\Delta S$										−55.83	6.74	
$\Delta G_{\text{binding-PBSA}}$										−60.63		

VDW = van der Waals contribution from MM; EEL = electrostatic energy as calculated by the MM force field; EPB/EGB = the electrostatic contribution to the solvation free energy calculated by PB or GB respectively; ECAVITY = nonpolar contribution to the solvation free energy calculated by an empirical model; GBTOTAL/PBTOTAL = final estimated binding free energy calculated from the terms above; $T\Delta S$, T (temperature); ΔS (sum of rotational, translational and vibrational entropies); ΔG binding, total binding energy of the system. Entropy calculations were based on normal modes analysis using 80 snapshots. All energies are in kcal/mol.

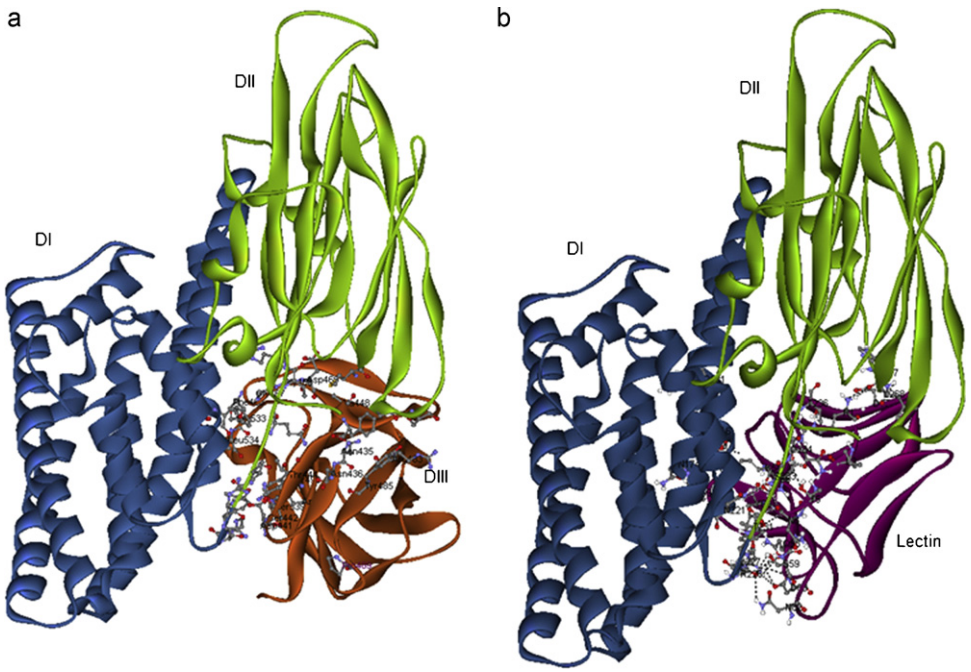

Fig. 4. Protein–protein docking of the DIII (Cry1Ac)/lectin (ASAL) fusion protein to the DI–DII domains of Cry1Ac. (a) Hydrogen bonds between the DI–DII and DIII domains. (b) Hydrogen bonds between DI–DII and the lectin.

hydrogen bonds, salt bridges and other interactions. The ZDOCK analysis revealed that the interactions of residues of the DI–DII domains at the interface with the DIII domain are similar to those of the DI–DII–DIII homology model (Fig. 4), thus suggesting the accuracy of the prediction. Two hydrogen bonds observed between Thr231 of the DI–DII domains and Asp441 of the DIII domain, and Gln234 of the DI–DII domains and Thr444 of the DIII domain, were also present in the DI–DII–lectin complex, indicating that the binding interface of garlic lectin is similar to that of the DIII domain with the DI–DII domains. In the DI–DII–lectin fusion protein, 18 hydrogen bonds were detected compared to 11 in the DI–DII–DIII complex (Table 2). Details of the hydrogen bonds observed between the DI–DII domains and the lectin protein are provided as Supplementary data. Hydrophobic, ionic and cationic–pi interactions were also observed between the DI–DII domains and the lectin protein, which probably contributes to the stability of the fusion complex (data not shown). Residue-based binding contributions were plotted in heat maps to highlight the most important residues for protein–protein interactions, revealing that Val180, Leu213, Val216, Phe219, Pro220, Tyr227, Thr231, Val232, Glu234, Leu268, Met269 and Val298 in the DI–DII domains are critical for their interaction with lectin/DIII (Fig. 6).

The interaction profile, binding energy and ZDOCK score suggest that the fusion protein (DI–DII–lectin) is as stable as the Cry1Ac protein (DI–DII–DIII). Furthermore, the Ramachandran plot obtained using the PROCHECK program revealed that ~99% of the amino acid residues are present in the most favored and additional allowed regions (Supplementary information Fig. S2). A high QMEAN score of 0.845 and a QMEAN Z-score of 1.24 were observed for the DI–DII–lectin complex (Supplementary information Fig. S3). Several other QMEAN scores, predicted local error plot and energy profile (Supplementary information Fig. S3) indicated that the DI–DII–lectin complex model is of good quality.

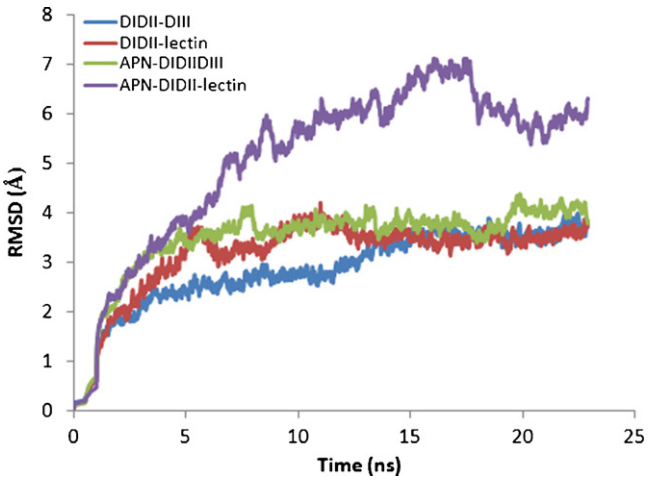

Fig. 5. Root-mean-square deviation (RMSD) calculated for all backbone atoms of DIDIIDIII, DIDII–lectin, APN–DIDIIDIII and APN–DIDII–lectin during MD simulations.

3.3. *Docking of DI–DII–DIII (Cry1Ac protein)/DI–DII–lectin (fusion protein) with APN*

The predicted high stability of the fusion protein prompted us to examine its binding affinity with the insect APN receptor and compare it with the Cry1Ac protein. Protein–protein docking studies were performed to analyze the binding affinity of the fusion protein with the APN receptor. The homology model of the APN of *M. sexta* was considered as the receptor protein, while the DI–DII–DIII domains of Cry1Ac and the fusion protein were treated as the ligand proteins. The top 200 poses obtained by ZDOCK were further refined using RDOCK to obtain the best 10 poses of the Cry1Ac–APN and fusion protein–APN complexes. The binding energy of the best pose of the complex of the fusion protein with APN was −23.146 kcal/mol, compared to −20.12 kcal/mol for the best pose of the Cry1Ac–APN complex (Table 6). This suggests that the fusion protein–APN interaction is energetically more favorable than that of the Cry1Ac–APN. Hydrogen bonds as well as hydrophobic, electrostatic and pi–pi interactions were calculated in both the complexes. A total of 19 hydrogen bonds were found between the fusion protein and the APN protein receptor, while 17 hydrogen

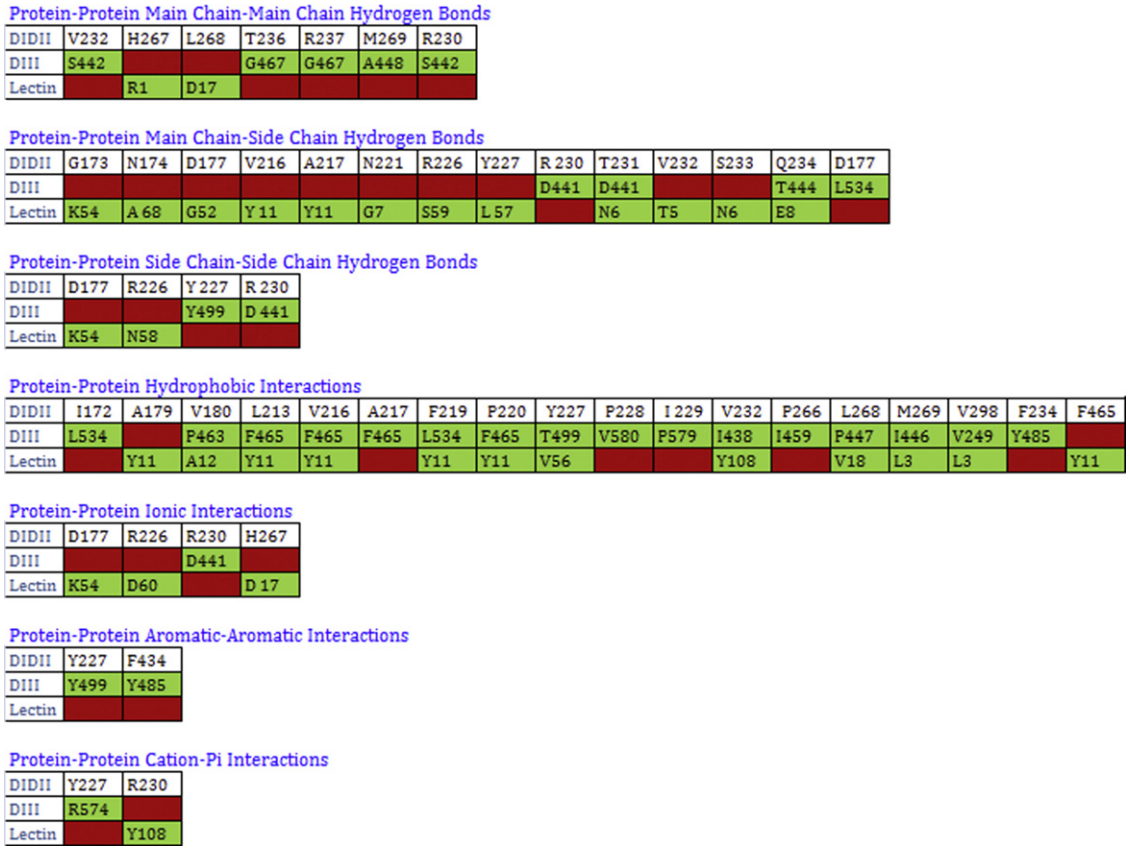

Fig. 6. Heat map showing the protein–protein interactions between Cry1Ac–DI–DII and DIII/lectin (ASAL) predicted using ZDOCK.

Table 6
Binding energy and ZDOCK score of the DI–DII–DIII and DI–DII–lectin with APN complex.

Ligand protein	Receptor protein	Predicted binding energy (kcal/mol)[a]	ZDOCK score[b]	E_RDOCK
DI–DII–DIII	APN	−20.12	15.34	−10.261
DI–DII–Lectin	APN	−23.146	17.58	−14.959

[a] Binding energy calculated by Dcomplex.
[b] ZDOCK and E_RDOCK score were calculated by ZDOCK program of Discovery Studio.

bonds were observed between Cry1Ac and APN (Table 7). Amino acids Leu334-Glu347 (loop 3) of the DII domain of the fusion protein were seen to play a crucial role in the formation of hydrogen bonds with the APN protein receptor (Fig. 7). It has been previously reported that loop 3 of the DII domain interacts with the APN receptor of *M. sexta* [74]. Superimposition of the DI–DII–DIII domains with the Cry–lectin fusion protein revealed that the amino acid residues Ser, Thr, Leu, Arg, Val and Asn (508–513) overlap with the lectin loop residues His466, Ser467, Thr468 and Thr472, which are involved in the binding with the APN protein receptor. This observation indicates that the predicted binding poses of Cry1Ac–APN and fusion protein–APN complexes are similar. It has been reported that residues Ser, Thr, Leu, Arg, Val and Asn (508–513) of the Cry1Ac DIII domain are involved in the binding with the APN protein receptor [67]. Details of the hydrogen bonds found between the DI–DII–lectin fusion protein and the APN protein receptor are given in Supplementary information Table S3.

Amino acid residues Val448, Tyr451, Pro450, Val470 and Trp471 of garlic lectin were found to be involved in hydrophobic interactions with the APN protein receptor. Compared to the complex of the Cry1Ac protein and the APN protein receptor, a larger number of hydrophobic interactions were observed between the fusion protein and the APN protein receptor (Table 8). Ionic interactions were observed between Asp465 and His466 residues of lectin with Arg731 and Asp801 of the APN protein receptor (Table 9). An aromatic–aromatic interaction was observed between Trp471 of the lectin fusion protein and Tyr732 of the APN protein receptor, while no such interaction was detected between the Cry1Ac protein and the APN protein receptor (Table 10a). However, aromatic–sulphur and cation–pi interactions were observed in both the Cry1Ac–APN and fusion protein–APN complexes (Table 10b and c).

The interaction profiles showed a larger number of non-bonded and hydrogen-bonded contacts in the fusion protein–APN complex compared to the Cry1Ac–APN complex (Fig. 8). APN protein receptor amino acids Phe644, Ala645, Ala649, Glu698, Arg731, Leu651, Leu654, Arg765, Asp766, Ile701, Tyr732, Val802 and Met805 showed favorable interactions with the Cry1Ac/fusion proteins, and hence these residues might serve as core contacts between the receptor and ligands.

The Rosetta server returned the 10 best-scoring structures for the fusion–APN and Cry1Ac–APN complexes. The binding energy and docking environment score for the best pose of the fusion protein–APN complex was −27.8 kcal/mol and −11.2, compared to −18.0 kcal/mol and −6.7 for the best pose of the Cry1Ac–APN complex. Residue-residue pair energies showed that amino acids Ser251-Thr437 of the fusion protein interact with amino acids Ala645-Glu889 of the APN protein receptor (Supplementary information Tables S4 and S5). Employing PIC, a total of 22 hydrogen bonds were observed between the fusion protein and the APN protein receptor, while only 15 hydrogen bonds were observed between Cry1Ac and the APN protein receptor. A larger number of non-bonded interactions were also observed in the fusion protein–APN complex compared to the Cry1Ac–APN complex (data

Table 7
Hydrogen bonds profile between Cry1Ac and fusion protein (DI–DII–lectin) against APN receptor of *M. sexta*.

APN residues	Cry1Ac				DI–DII–lectin (fusion protein)			
	APN atom	Residue[a]	Cry atom	Length	APN atom	Residue[a]	Fusion Cry atom	Length
ALA645	O	LEU309 (DII)	N	3.32	O	GLU449 (lectin)	OE1	2.91
ALA648	N	LEU309 (DII)	O	3.29	–	–	–	–
ARG731	NH1	GLU304 (DII)	O	2.93	NH1	ASP465 (lectin)	O	2.75
ARG731	NH2	PRO303 (DII)	O	3.04	NH2	THR468 (lectin)	O	3.07
ARG765	NH1	GLU304 (DII)	O	2.6	NE	GLY523 (lectin)	O	2.9
ARG886	–	–	–	–	NE	GLY407 (DII)	O	2.84
ARG886	–	–	–	–	NH2	PHE408 (DII)	O	2.84
ASN650	OD1	TYR310 (DII)	OH	3.46	ND2	GLU449 (lectin)	O	3.03
ASN800	–	–	–	–	OD1	TYR451 (lectin)	OH	3.36
ASP702	OD1	ARG483 (DIII)	NH2	2.8	–	–	–	–
ASP766	OD1	ARG321 (DII)	NH2	3.47	N	ASP522 (lectin)	OD2	3
ASP801	OD2	THR306 (DII)	OG1	3.21	–	–	–	–
ASP894	OD2	ASN456 (DIII)	ND2	2.79	–	–	–	–
CYS693	–	–	–	–	O	THR468 (lectin)	OG1	2.52
GLN657	–	–	–	–	NE2	HIS466 (lectin)	ND1	2.86
GLN698	NE2	GLY302 (DII)	N	2.83	OE1	SER467 (lectin)	O	3.11
GLN698	NE2	SER301 (DII)	OG	2.69	–	–	–	–
GLN767	NE2	GLN319 (DII)	OE1	3.23	N	ASP522 (lectin)	OD2	3.3
GLU889	–	–	–	–	OE1	ARG405 (DII)	NH2	2.91
GLY696	–	–	–	–	O	SER467 (lectin)	OG	2.96
HIS697	ND1	PHE453 (DIII)	O	3.04	–	–	–	–
HIS697	NE2	ASN452 (DIII)	OD1	3.02	–	–	–	–
MET805	SD	TYR278 (DII)	OH	3.54	–	–	–	–
PHE892	O	ARG472 (DIII)	NH2	2.78	–	–	–	–
SER882	–	–	–	–	OG	GLY307 (DII)	N	2.65
SER882	–	–	–	–	OG	LEU305 (DII)	O	2.67
TYR732	–	–	–	–	OH	THR472 (lectin)	O	2.99
VAL695	–	–	–	–	O	SER467 (lectin)	OG	2.75

[a] Parenthesis indicates the residue's domain.

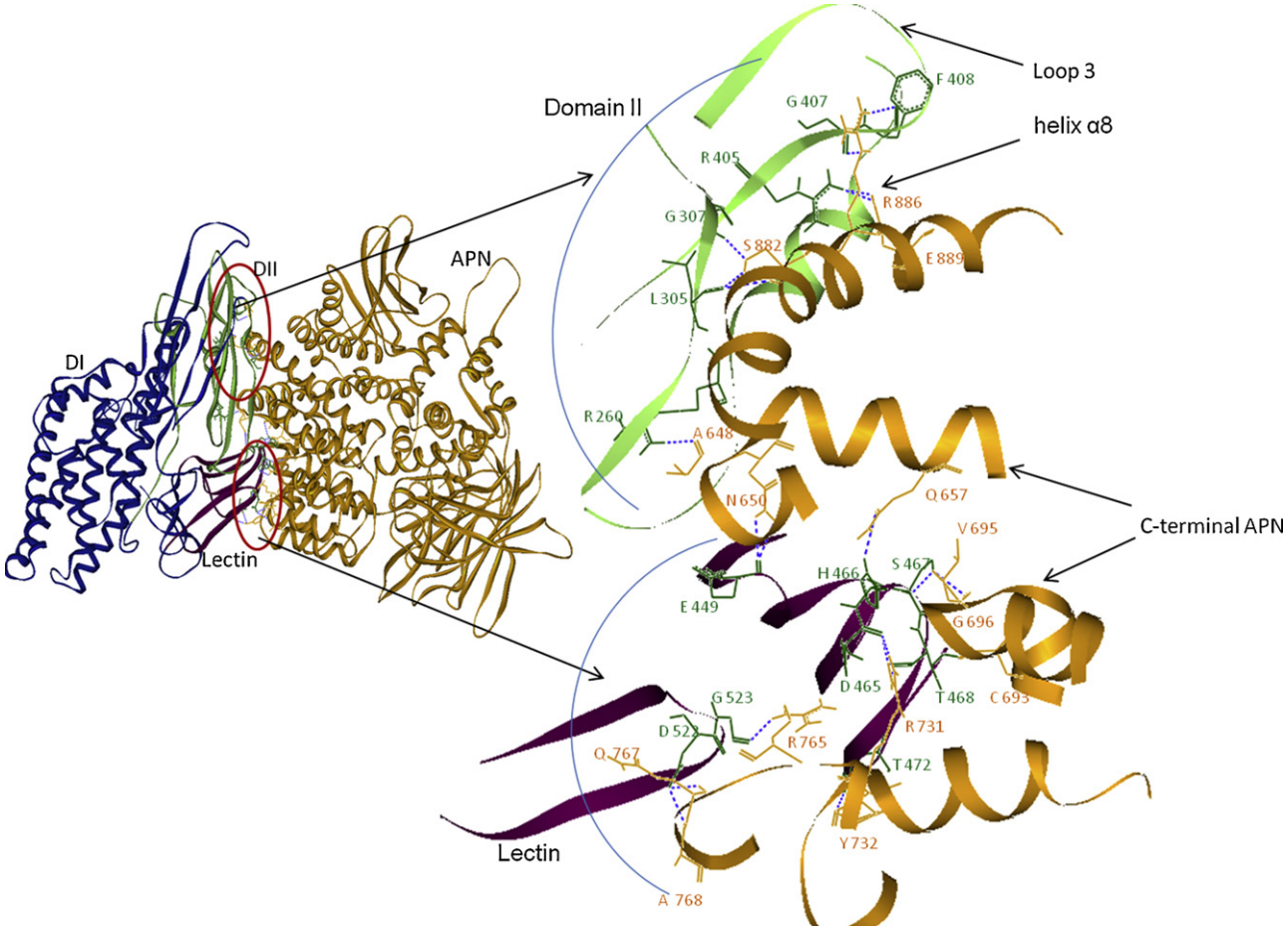

Fig. 7. Binding interface between the fusion protein (DI–DII–lectin) and APN of *M. sexta*.

Table 8
Hydrophobic interaction between Cry1Ac and fusion protein (DI–DII–lectin) against APN receptor of *M. sexta*.

APN residues	Cry1Ac residue[a]	DI–DII–lectin residue[a]
ALA645	PRO308 (DII)	PRO450 (Lec)
ALA648	PRO308 (DII), TYR310 (DII), MET407 (DII), PHE408 (DIII)	LEU264 (DII), ALA291 (DII), PHE301 (DII), PHE303 (DII)
ALA649	PRO308 (DII), TYR310 (DII)	PHE303 (DII), PRO304 (DII)
ALA664	PHE455 (DIII)	–
ALA840	–	PRO299 (DII), PHE301 (DII)
ALA878	–	PRO304 (DII), LEU305 (DII), MET309 (DII)
ILE641	PRO308 (DII)	–
ILE658	LEU268 (DII), PHE453 (DIII), VAL459 (DIII)	–
ILE659	PHE455 (DIII)	–
ILE701	PRO297 (DII), PRO303 (DII)	TYR464 (Lec), ALA469 (Lec)
LEU621	PHE455 (DIII)	–
LEU651	ILE263 (DII), PHE307 (DII), TYR310 (DII), VAL405 (DII), MET407 (DII), PHE425 (DII)	ILE259 (DII), PRO262 (DII), TYR306 (DII), MET403 (DII), PHE421 (DII)
LEU654	PHE305 (DII)	PRO262 (DII)
LEU661	PHE453 (DIII)	–
LEU669	LEU454 (DIII)	–
LEU837	–	PHE301 (DII), PRO304 (DII)
LEU885	–	TYR306 (DII), MET403 (DII), PHE404 (DII)
MET805	TYR278 (DII), LEU309 (DII), MET313 (DII)	PRO299 (DII)
MET881	–	PHE303 (DII), PRO304 (DII), LEU305 (DII), TYR306 (DII), PHE404 (DII)
PHE623	PHE564 (DIII)	
PHE644	PRO308 (DII)	PRO304 (DII)
PHE892	VAL459 (DIII), ILE460 (DIII)	–
TYR732	PHE300 (DII), PRO303 (DII)	VAL470 (Lec), TRP471 (Lec)
VAL665	PHE453 (DIII), LEU454 (DIII), PHE455 (DIII)	–
VAL695	PHE453 (DIII)	–
VAL802	MET294 (DII)	VAL448 (Lec), PRO450 (Lec), TYR451 (Lec)
VAL893	–	PHE408 (DII)

[a] Parenthesis indicates the residue's domain.

Table 9
Protein–protein ionic interactions between Cry1Ac and fusion protein (DI–DII–lectin) against APN receptor of *M. sexta*.

APN residues	Cry1Ac residues[a]	Fusion protein residues[a]
ARG765	GLU304 (DII)	ASP465 (lectin)
ASP766	ARG321 (DII)	ASP522 (lectin)
ASP702	ARG483 (DIII)	–
ARG731	–	ASP465 (lectin)
ASP801	–	HIS466 (lectin)
GLU839	–	ARG317 (DII)
GLU889	–	ARG405 (DII)
ASP894	ARG472 (DIII)	–

[a] Parenthesis indicates the residue's domain.

not shown). Amino acids Leu305, Gly307, Arg405 and Gly407 of loop 3 in the DII domain of the fusion protein appear to play a critical role in the formation of hydrogen bonds with the APN protein receptor. On the other hand, amino acids Glu449, Tyr451, Asp465, Ser467 and Gly523 of the lectin protein were involved in interactions with amino acids Ala645, Asn800, Arg731, Val695 and Arg767 of the APN protein receptor (Supplementary information Tables S4 and S5).

A ClusPro analysis revealed a larger number of interactions between DI–DII–lectin and the APN protein receptor compared to complex of DI–DII–DIII with the APN protein receptor. For the protein–protein interaction studies with the APN protein receptor, models with cluster sizes of 128 for the DI–DII–lectin and 123 for the DI–DII–DIII domains were selected. Ten main chain-main chain hydrogen bonds were seen between DI–DII–lectin and the APN protein receptor, whereas only one hydrogen bond was evident between the DI–DII–DIII domains and the APN protein receptor (Supplementary information Tables S6 and S7). Twenty-main chain-side chain hydrogen bonds were observed between the DI–DII–lectin and the APN protein receptor, while only six were detected between the DI–DII–DIII domains and the APN protein receptor. A total of ten side chain-side chain hydrogen bonds and seven ionic interactions were observed between DI–DII–lectin and the APN protein receptor, while six side chain-side chain hydrogen bonds and one ionic interaction were observed between the DI–DII–DIII domains and the APN protein receptor (Supplementary information Tables S3 and S4). Heat maps revealed that residues Gln698, Phe700, Val893 and Glu800 in the APN protein receptor are critical for its interaction with the DIII/lectin domain (Fig. 9).

Although the three models employed in this study showed variation in the number of hydrogen bonds, non-ionic interactions, binding site residues and interaction patterns between the fusion protein–APN and Cry1AC–APN complexes, the ZDOCK and Rosetta models are similar, and as such might serve as docking models for the fusion protein–APN and Cry1Ac–APN complexes.

The N-terminal domain I of the APN protein receptor was docked to the fusion protein using ZDOCK to identify any interaction sites in the N-terminal region of the APN protein receptor. A low level of interaction between the fusion protein and the N-terminal domain I of the APN protein receptor was observed. As shown in Supplementary information Fig. S4, the APN protein receptor interacted with either the DII domain of the fusion protein or the lectin of the fusion protein. No binding pose was observed in which the APN protein receptor could interact with the DII domain of Cry and lectin. Five hydrogen bonds were observed in the binding pose where the APN protein receptor interacted with the DII domain of the fusion protein, while three hydrogen bonds were observed between the APN protein receptor and the lectin of the fusion protein (Supplementary information Table S8). The energy of interaction between the APN-N terminus and the fusion protein complex was predicted to be higher (−12.321 kcal/mol) than that of APN-C-terminus and the fusion protein complex

Table 10
Different protein–protein interactions between fusion protein (DI–DII–lectin) and APN receptor.

Residue[a]	APN	*D* (centroid–centroid)	Dihedral angle
(a) Protein–protein aromatic–aromatic interactions			
TRP471 (lectin)–fusion protein	TYR732	6.61	54.23
(b) Protein–protein aromatic–sulphur interactions			
TYR278 (DII)– Cry1Ac	MET805	4.63	73.62
TYR306 (DII)– fusion protein	MET881	4.58	169.01
(c) Protein–protein cation–pi interactions			
ARG257 (DII)–fusion protein	PHE892	4.06	44.27
PHE455 (DIII)–Cry1Ac	ARG622	4.39	15.26
TYR451 (lectin)–fusion protein	ARG765	4.90	18.27

[a] Parenthesis indicates the residue's domain.

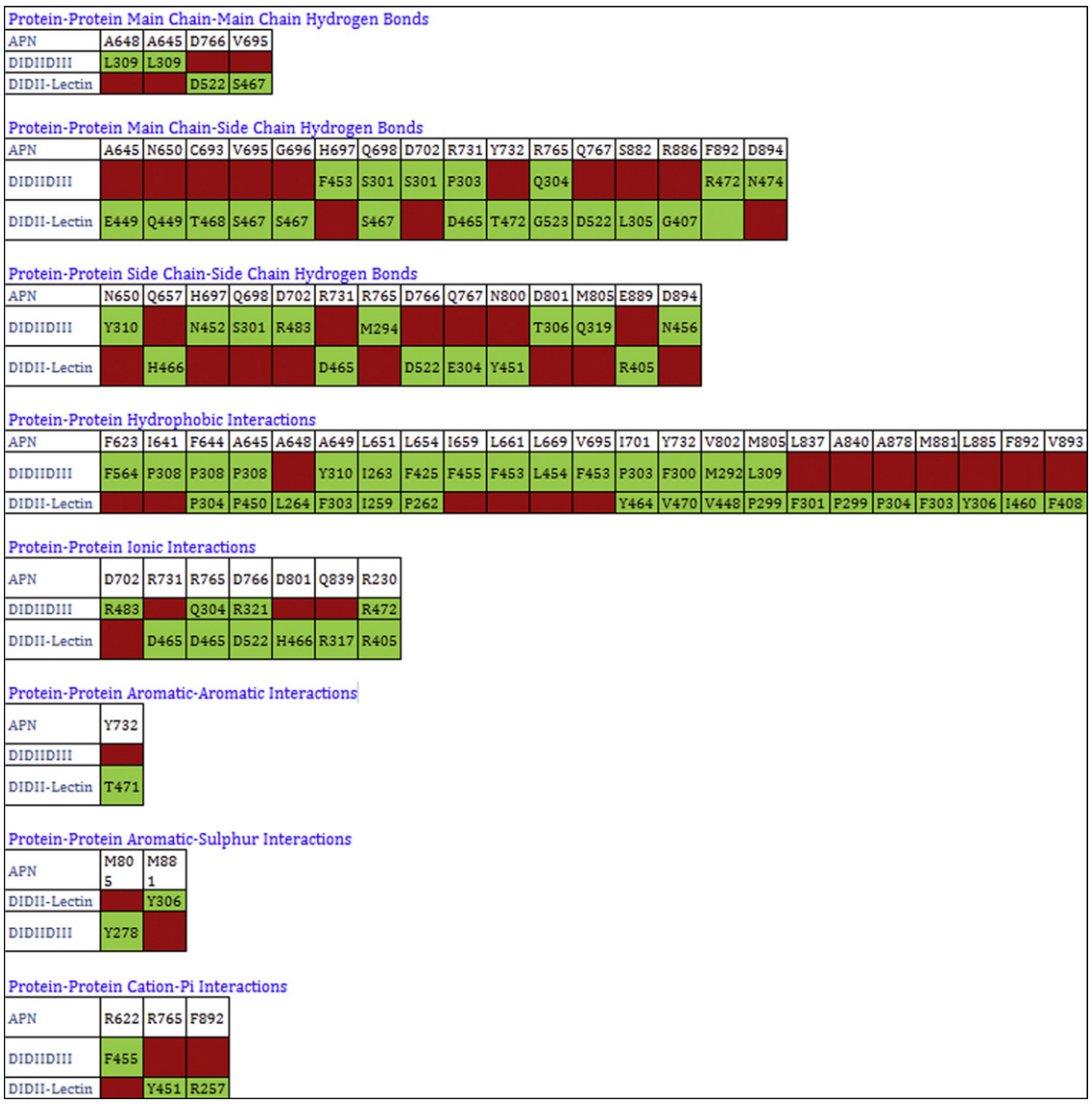

Fig. 8. Heat map showing the protein–protein interactions between the native Cry1Ac (DI–DII–DIII)/fusion protein (DI–DII–lectin) and APN of *M. sexta* predicted using ZDOCK.

(−23.146 kcal/mol). This indicates that the APN protein receptor interacts with the Cry protein through its C-terminal end rather than the N-terminal region.

The RMSD between the initial coordinates of the backbone atoms of the proteins complex and the coordinates along the MD simulations was computed for every snapshot. A significant amount of backbone motion (up to ∼8 Å) in APN–DIDII–lectin can be observed (Fig. 5), as measured from the separate simulations of the protein and ligand. The high RMSD values indicate the occurrence of a conformational change in the backbone of the DIDII–lectin upon binding to APN. MM/GB(PB)SA calculations showed that the entropy for APN–DIDIIDIII is −19.541 kcal/mol and for APN–lectin–DIDII is −19.540 kcal/mol (Tables 11 and 12). The entropy for APN complexes with DIDIIDIII and DIDII–lectin is calculated by quasi-harmonic methods. These entropy calculations are either overestimated or greatly underestimated [75,76]. There are some approximate methods which can give reasonable estimates of conformational entropies for inclusion in free energy calculations.

Interaction energies, hydrogen bonds, hydrophobic and pi–pi interactions revealed that the fusion protein–APN complex is more stable and energetically favorable than the Cry1Ac protein–APN complex. The molecular surface of the complexes of Cry1Ac and Cry1Ac–lectin proteins with the APN protein receptor were computed to visualize the shape complementarities at the interface. As illustrated in Fig. 10, the binding interface between the fusion protein and the APN protein receptor is well packed, similarly to that of the Cry1Ac–APN complex.

The primary objective of creating fusion proteins is to design more potent insecticides with enhanced properties than those of the parent proteins. It has been found that hybrid Cry proteins, developed by domain swapping, exhibit increased toxicities in comparison with the parent proteins [30–32]. When a hybrid protein comprising the DI and DII domains of Cry1Ia and the DIII domain of Cry1Ba was tested against the Colorado potato beetle larvae, an LC_{50} value of 22.4 μg/ml was observed, compared to 33.7 and 142 μg/ml for the parent proteins of Cry1Ia and Cry1Ba, respectively [32]. For the hybrid protein, consisting of the DI and DII domains of Cry1Ab and the DIII domain of Cry1C, an LC_{50} value of 1.66 μg/g was observed compared to >100 and 11.0 ± 1.13 μg/g for the Cry1Ab and Cry1C parent proteins against *S. exigua* larvae [30,31]. Ge et al. [13] reported that, by replacing the DIII domain of Cry1Aa (amino acids 450–612) with that of Cry1Ac, the toxicity of the fusion protein increased 300-fold and 10-fold against *Heliothis virescens* and *Trichoplusia ni*, respectively, as compared

Fig. 9. Heat map showing the protein–protein interactions between the native Cry1Ac (DI–DII–DIII)/fusion protein (DI–DII–lectin) and APN of *M. sexta* predicted by ClusPro.

Table 11
MM/PB(GB)SA energy component analysis of the interactions of the APN with the DIDIIDIII.

Energy components	Complex			Receptor			Ligand			Delta (complex–receptor–ligand)		
	Average	Std. dev.	Std. error	Average	Std. dev.	Std. error	Average	Std. dev.	Std. error	Average	Std. dev.	Std. error
MM/GBSA energy component analysis of the interactions of the APN with the DIDIIDIII averaged over 1400 snapshots												
VDW	−11492.6	62	3.643	−7208.66	42.9	2.519	−4166.24	33.31	1.956	−117.678	10.725	0.6298
EEL	−95376.4	275	16.17	−59845	192.3	11.289	−35435.9	123.5	7.254	−95.4216	76.917	4.5167
EGB	−12900.3	172	10.11	−7384.3	105.8	6.2111	−5727.3	95.21	5.591	211.3361	72.336	4.2477
ESURF	534.4193	7.02	0.412	326.7348	4.463	0.2621	223.9534	3.211	0.189	−16.269	1.6788	0.0986
Ggas	−106869	253	14.85	−67053.7	179.8	10.558	−39602.2	118.8	6.974	−213.099	79.118	4.646
Gsolv	−12365.8	173	10.18	−7057.57	105.7	6.2066	−5503.35	94.95	5.576	195.0671	71.68	4.2092
GBTOTAL	−119235	130	7.657	−74111.2	115.3	6.7718	−45105.5	61.97	3.639	−18.032	12.775	0.7502
$T\Delta S$										−19.54		
$\Delta G_{\text{binding-GBSA}}$										1.51		
MM/PBSA energy component analysis of the interactions of the APN with the DIDIIDIII averaged over 1400 snapshots												
VDW	−11492.6	62	3.643	−7208.66	42.9	2.519	−4166.24	33.31	1.956	−117.678	10.725	0.6298
EEL	−95376.4	275	16.17	−59845	192.3	11.289	−35435.9	123.5	7.254	−95.4216	76.917	4.5167
EPB	−13972.3	169	9.95	−8041.9	109.9	6.4539	−6100.17	93.77	5.506	169.7583	73.068	4.2907
Ecavity	341.6734	4.44	0.261	206.1185	2.484	0.1458	152.6744	2.26	0.133	−17.1195	1.1576	0.068
Ggas	−106869	253	14.85	−67053.7	179.8	10.558	−39602.2	118.8	6.974	−213.099	79.118	4.646
Gsolv	−13630.6	169	9.916	−7835.78	109	6.3985	−5947.5	93.65	5.499	152.6388	72.618	4.2643
PBTOTAL	−120500	170	10.01	−74889.5	146.2	8.5849	−45549.7	66.59	3.91	−60.4603	12.167	0.7145
$T\Delta S$										−19.54		
$\Delta G_{\text{binding-GBSA}}$										−40.92		

VDW = van der Waals contribution from MM; EEL = electrostatic energy as calculated by the MM force field; EPB/EGB = the electrostatic contribution to the solvation free energy calculated by PB or GB respectively; ECAVITY = nonpolar contribution to the solvation free energy calculated by an empirical model; GBTOTAL/PBTOTAL = final estimated binding free energy calculated from the terms above; $T\Delta S$: T (temperature); ΔS (sum of rotational, translational and vibrational entropies); $\Delta G_{\text{binding}}$: total binding energy of the system. Calculations were based on quasi-harmonic analysis; all energies are in kcal/mol.

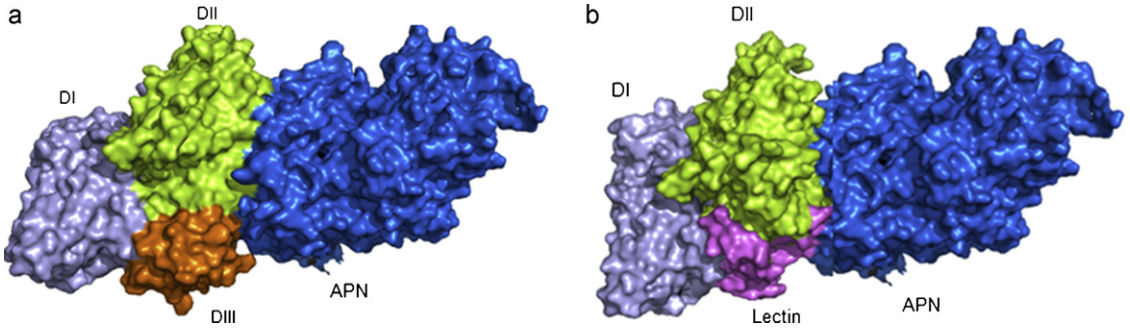

Fig. 10. Molecular surface representation of the binding interface of Cry1Ac/fusion proteins and the APN receptor of *M. sexta*: (a) Cry1Ac–APN, (b) fusion protein–APN.

Table 12
MM/PB(GB)SA energy component analysis of the interactions of the APN with the DIDII–lectin.

Energy components	Complex			Receptor			Ligand			Delta (complex–receptor–ligand)		
	Average	Std. dev.	Std. error	Average	Std. dev.	Std. error	Average	Std. dev.	Std. error	Average	Std. dev.	Std. error
MM/GBSA energy component analysis of the interactions of the APN with the DIDII–lectin averaged over 1400 snapshots												
VDW	−11162.3	231.33	13.82	−6429.13	226.94	13.562	−4517.26	30.091	1.7983	−215.868	12.218	0.7302
EEL	−97523.2	176.27	10.53	−59603.7	144.66	8.645	−37419.3	103.66	6.1946	−500.269	57.069	3.4105
EGB	−11587.8	147.85	8.84	−7315.67	116.05	6.935	−4966.44	70.719	4.2263	694.2951	55.665	3.3266
ESURF	513	5.88	0.35	318.9219	4.0366	0.2412	222.7478	2.6771	0.16	−28.6729	1.9889	0.1189
Ggas	−108685	301.41	18.01	−66032.8	271.92	16.25	−41936.5	101.8	6.0839	−716.137	63.8	3.8128
Gsolv	−11074.8	144.92	8.66	−6996.75	114.81	6.86	−4743.7	69.83	4.1731	665.6222	54.158	3.2366
GBTOTAL	−119760	255.78	15.28	−73029.5	242.54	14.495	−46680.2	66.975	4.0025	−50.5147	15.38	0.9192
$T\Delta S$										−19.541		
$\Delta G_{\text{binding-GBSA}}$										−30.97		
MM/PBSA energy component analysis of the interactions of the APN with the DIDII-lectin averaged over 1400 snapshots												
VDW	−11162.3	231.33	13.825	−6429.13	226.94	13.562	−4517.26	30.09	1.7983	−215.868	12.218	0.7302
EEL	−97523.2	176.27	10.534	−59603.7	144.66	8.645	−37419.3	103.7	6.1946	−500.269	57.069	3.4105
EPB	−12718.9	143.6	8.5815	−7961.98	118.53	7.0832	−5353.7	69.38	4.146	596.7987	52.36	3.1291
Ecavity	328.8574	3.6489	0.2181	206.0094	2.7127	0.1621	148.6719	1.548	0.0925	−25.8239	1.0199	0.061
Ggas	−108685	301.41	18.013	−66032.8	271.92	16.25	−41936.5	101.8	6.0839	−716.137	63.8	3.8128
Gsolv	−12390	142.09	8.4914	−7755.97	118.02	7.0529	−5205.03	68.85	4.1145	570.9748	51.63	3.0855
PBTOTAL	−121075	260.23	15.552	−73788.8	245.53	14.673	−47141.5	69.51	4.1537	−145.162	18.567	1.1096
$T\Delta S$										−19.541		
$\Delta G_{\text{binding-PBSA}}$										−125.62		

VDW = van der Waals contribution from MM; EEL = electrostatic energy as calculated by the MM force field; EPB/EGB = the electrostatic contribution to the solvation free energy calculated by PB or GB respectively; ECAVITY = nonpolar contribution to the solvation free energy calculated by an empirical model; GBTOTAL/PBTOTAL = final estimated binding free energy calculated from the terms above; $T\Delta S$: T (temperature); ΔS (sum of rotational, translational and vibrational entropies); $\Delta G_{\text{binding}}$: total binding energy of the system. Calculations were based on quasi-harmonic analysis; all energies are in kcal/mol.

to Cry1Aa. Furthermore, the Cry1Aa–Cry1Ac fusion protein proved to be ~100 times more toxic to *Bombyx mori* than that Cry1Aa alone [77]. The fusion of *cry*1C and *cry*1E by *in vivo* recombination resulted in a chimeric protein with a broader activity and binding to different receptors than the native proteins [78]. Transgenic rice and maize plants, expressing the fusion protein containing Cry1Ac and galactose-binding domain of the nontoxic ricin B-chain, exhibited enhanced toxicity to different insects compared to Cry1Ac alone, owing to the additional binding domains providing a number of potential interactions at the molecular level in target insects [79].

4. Conclusions

A novel fusion protein, comprising DI–DII domains of Cry1Ac and garlic lectin, has been constructed *in-silico*. Protein–protein docking studies revealed that various interactions observed between garlic lectin and DI–DII domains are similar to those of DIII domain with DI–DII. In the DI–DII–lectin fusion protein, a larger number of hydrogen bonds were found compared to those observed in the DI–DII–DIII domains. The calculated binding free energy of fusion proteins using MM/PB(GB)SA methods indicated that the formation of DIDII–lectin fusion protein is more favorable than DIDII–DIII. The interaction profile, binding energy and ZDOCK score suggest that the DI–DII–lectin fusion protein is as stable as the DI–DII–DIII Cry1Ac protein. Docking predictions of the fusion protein–APN and Cry1Ac–APN complexes indicate that the fusion protein–APN complex is more stable and energetically favorable than the Cry1Ac–APN complex. Binding energy and other parameters demonstrate that the binding affinity in fusion protein–APN complex is higher than that of the Cry1Ac–APN complex. The molecular surface of the binding interface between Cry1Ac–lectin and the APN protein receptor is well packed akin to that of Cry1Ac–APN complex. An overview of the results intimates the feasibility of designing and developing new potent toxins with novel specificities against different insect pests.

Acknowledgements

The authors extend their thanks to Prof. T. Papi Reddy of the Department of Genetics, Osmania University, for his helpful suggestions and critical evaluation of the manuscript.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmgm.2011.11.001.

References

[1] E. Schnepf, N. Crickmore, J. Van Rie, D. Lereclus, J. Baum, J. Feitelson, et al., *Bacillus thuringiensis* and its pesticidal crystal proteins, Microbiol. Mol. Biol. Rev. 62 (1998) 775–806.
[2] J. Van Rie, *Bacillus thuringiensis* and its use in transgenic insect control technologies, Int. J. Med. Microbiol. 290 (2000) 463–469.
[3] A.I. Aronson, W. Beckman, Transfer of chromosomal genes and plasmids in *Bacillus thuringiensis*, Appl. Environ. Microbiol. 53 (1987) 1525–1530.
[4] C. Hofmann, H. Vanderbruggen, H. Hofte, J. Van Rie, S. Jansens, H. Van Mellaert, Specificity of *Bacillus thuringiensis* delta-endotoxins is correlated with the presence of high-affinity binding sites in the brush border membrane of target insect midguts, Proc. Natl. Acad. Sci. U. S. A. 85 (1988) 7844–7848.
[5] A. Bravo, Phylogenetic relationships of *Bacillus thuringiensis* delta-endotoxin family proteins and their functional domains, J. Bacteriol. 179 (1997) 2793–2801.
[6] P. Grochulski, L. Masson, S. Borisova, M. Pusztai-Carey, J.L. Schwartz, R. Brousseau, et al., *Bacillus thuringiensis* CryIA(a) insecticidal toxin: crystal structure and channel formation, J. Mol. Biol. 254 (1995) 447–464.
[7] R.J. Morse, T. Yamamoto, R.M. Stroud, Structure of Cry2Aa suggests an unexpected receptor binding epitope, Structure 9 (2001) 409–417.
[8] O.I. Loseva, E.I. Tiktopulo, V.D. Vasiliev, A.D. Nikulin, A.P. Dobritsa, S.A. Potekhin, Structure of Cry3A delta-endotoxin within phospholipid membranes, Biochemistry 40 (2001) 14143–14151.
[9] N. Galitsky, V. Cody, A. Wojtczak, D. Ghosh, J.R. Luft, W. Pangborn, L. English, Structure of the insecticidal bacterial delta-endotoxin Cry3Bb1 of *Bacillus thuringiensis*, Acta Crystallogr. D: Biol. Crystallogr. 57 (2001) 1101–1109.
[10] P. Boonserm, M. Mo, C. Angsuthanasombat, J. Lescar, Structure of the functional form of the mosquito larvicidal Cry4Aa toxin from *Bacillus thuringiensis* at a 2.8-angstrom resolution, J. Bacteriol. 188 (2006) 3391–3401.
[11] P. Boonserm, P. Davis, D.J. Ellar, J. Li, Crystal structure of the mosquito-larvicidal toxin Cry4Ba and its biological implications, J. Mol. Biol. 348 (2005) 363–382.
[12] X.J. Chen, A. Curtiss, E. Alcantara, D.H. Dean, Mutations in domain I of *Bacillus thuringiensis* delta-endotoxin CryIAb reduce the irreversible binding of toxin

to *Manduca sexta* brush border membrane vesicles, J. Biol. Chem. 270 (1995) 6412–6419.
[13] A.Z. Ge, D. Rivers, R. Milne, D.H. Dean, Functional domains of *Bacillus thuringiensis* insecticidal crystal proteins. Refinement of *Heliothis virescens* and *Trichoplusia ni* specificity domains on CryIA(c), J. Biol. Chem. 266 (1991) 17954–17958.
[14] F. Rajamohan, E. Alcantara, M.K. Lee, X.J. Chen, A. Curtiss, D.H. Dean, Single amino acid changes in domain II of *Bacillus thuringiensis* CryIAb delta-endotoxin affect irreversible binding to *Manduca sexta* midgut membrane vesicles, J. Bacteriol. 177 (1995) 2276–2282.
[15] F. Coux, V. Vachon, C. Rang, K. Moozar, L. Masson, M. Royer, et al., Role of interdomain salt bridges in the pore-forming ability of the *Bacillus thuringiensis* toxins Cry1Aa and Cry1Ac, J. Biol. Chem. 276 (2001) 35546–35551.
[16] D.P. Smedley, D.J. Ellar, Mutagenesis of three surface-exposed loops of a *Bacillus thuringiensis* insecticidal toxin reveals residues important for toxicity, receptor recognition and possibly membrane insertion, Microbiology 142 (Pt. 7) (1996) 1617–1624.
[17] J.L. Schwartz, L. Potvin, X.J. Chen, R. Brousseau, R. Laprade, D.H. Dean, Single-site mutations in the conserved alternating-arginine region affect ionic channels formed by CryIAa, a *Bacillus thuringiensis* toxin, Appl. Environ. Microbiol. 63 (1997) 3978–3984.
[18] S.B. Hu, P. Liu, X.Z. Ding, L. Yan, Y.J. Sun, Y.M. Zhang, et al., Efficient constitutive expression of chitinase in the mother cell of *Bacillus thuringiensis* and its potential to enhance the toxicity of Cry1Ac protoxin, Appl. Microbiol. Biotechnol. 82 (2009) 1157–1167.
[19] R.J. McNall, M.J. Adang, Identification of novel *Bacillus thuringiensis* Cry1Ac binding proteins in *Manduca sexta* midgut through proteomic analysis, Insect Biochem. Mol. Biol. 33 (2003) 999–1010.
[20] P.J. Knight, N. Crickmore, D.J. Ellar, 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 (1994) 429–436.
[21] B.R. Francis, L.A. Bulla Jr., Further characterization of BT-R1, the cadherin-like receptor for Cry1Ab toxin in tobacco hornworm (*Manduca sexta*) midguts, Insect Biochem. Mol. Biol. 27 (1997) 541–550.
[22] P.J. Knight, J. Carroll, D.J. Ellar, 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 (2004) 101–112.
[23] Y. Bharathi, T.P. Reddy, V.D. Reddy, K.V. Rao, Plant lectins and their utilization for development of insect resistant transgenic crop plants, in: Pests and Pathogens: Management Strategies, BS Publications, 2010, pp. 457–489.
[24] M.K. Lee, T.H. You, F.L. Gould, D.H. Dean, Identification of residues in domain III of *Bacillus thuringiensis* Cry1Ac toxin that affect binding and toxicity, Appl. Environ. Microbiol. 65 (1999) 4513–4520.
[25] A. Barre, Y. Bourne, E.J. Van Damme, W.J. Peumans, P. Rouge, Mannose-binding plant lectins: different structural scaffolds for a common sugar-recognition process, Biochimie 83 (2001) 645–651.
[26] S.L. Burton, D.J. Ellar, J. Li, D.J. Derbyshire, 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 (1999) 1011–1022.
[27] N.H. Ho, B.N. Oliva, K. Datta, R. Frutos, S.K. Datta, Translational fusion hybrid Bt genes confer resistance against yellow stem borer in transgenic elite vietnamese rice (*Oryza sativa* L.) cultivars, Crop Sci. 46 (2006) 781–789.
[28] G. Honee, W. Vriezen, B. Visser, A translation fusion product of two different insecticidal crystal protein genes of *Bacillus thuringiensis* exhibits an enlarged insecticidal spectrum, Appl. Environ. Microbiol. 56 (1990) 823–825.
[29] B.K. Harper, S.A. Mabon, S.M. Leffel, M.D. Halfhill, H.A. Richards, K.A. Moyer, et al., Green fluorescent protein as a marker for expression of a second gene in transgenic plants, Nat. Biotechnol. 17 (1999) 1125–1129.
[30] R.A. de Maagd, M. Weemen-Hendriks, W. Stiekema, D. Bosch, *Bacillus thuringiensis* delta-endotoxin Cry1C domain III can function as a specificity determinant for *Spodoptera exigua* in different, but not all, Cry1–Cry1C hybrids, Appl. Environ. Microbiol. 66 (2000) 1559–1563.
[31] R.A. de Maagd, M.S. Kwa, H. van der Klei, T. Yamamoto, B. Schipper, J.M. Vlak, 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 (1996) 1537–1543.
[32] S. Naimov, M. Weemen-Hendriks, S. Dukiandjiev, R.A. de Maagd, *Bacillus thuringiensis* delta-endotoxin Cry1 hybrid proteins with increased activity against the Colorado potato beetle, Appl. Environ. Microbiol. 67 (2001) 5328–5330.
[33] K. Wiehe, B. Pierce, W.W. Tong, H. Hwang, J. Mintseris, Z. Weng, The performance of ZDOCK and ZRANK in rounds 6–11 of CAPRI, Proteins 69 (2007) 719–725.
[34] S. Lyskov, J.J. Gray, The RosettaDock server for local protein–protein docking, Nucleic Acids Res. 36 (2008) W233–W238.
[35] S.R. Comeau, D.W. Gatchell, S. Vajda, C.J. Camacho, ClusPro: a fully automated algorithm for protein–protein docking, Nucleic Acids Res. 32 (2004) W96–W99.
[36] S.J. Fleishman, J.E. Corn, E.M. Strauch, T.A. Whitehead, I. Andre, J. Thompson, Rosetta in CAPRI rounds 13–19, Proteins 78 (2010) 3212–3218.
[37] D. Kozakov, D.R. Hall, D. Beglov, R. Brenke, S.R. Comeau, Y. Shen, et al., Achieving reliability and high accuracy in automated protein docking: ClusPro, PIPER, SDU, and stability analysis in CAPRI rounds 13–19, Proteins 78 (2010) 3124–3130.
[38] B. Yarasi, V. Sadumpati, C.P. Immanni, D.R. Vudem, V.R. Khareedu, Transgenic rice expressing *Allium sativum* leaf agglutinin (ASAL) exhibits high-level resistance against major sap-sucking pests, BMC Plant Biol. 8 (2008) 102.
[39] A. Roy, S. Banerjee, P. Majumder, S. Das, Efficiency of mannose-binding plant lectins in controlling a homopteran insect, the red cotton bug, J. Agric. Food Chem. 50 (2002) 6775–6779.
[40] F. Clement, Y.P. Venkatesh, Dietary garlic (*Allium sativum*) lectins, ASA I and ASA II, are highly stable and immunogenic, Int. Immunopharmacol. 10 (2010) 1161–1169.
[41] A. Sali, T.L. Blundell, Comparative protein modelling by satisfaction of spatial restraints, J. Mol. Biol. 234 (1993) 779–815.
[42] S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman, Basic local alignment search tool, J. Mol. Biol. 215 (1990) 403–410.
[43] J.D. Thompson, T.J. Gibson, D.G. Higgins, Multiple sequence alignment using ClustalW and ClustalX, Curr. Protoc. Bioinform. (2002), Chapter 2, Unit 23.
[44] B.R. Brooks, C.L. Brooks 3rd, A.D. Mackerell Jr., L. Nilsson, R.J. Petrella, B. Roux, et al., CHARMM: the biomolecular simulation program, J. Comput. Chem. 30 (2009) 1545–1614.
[45] P. Dauber-Osguthorpe, V.A. Roberts, D.J. Osguthorpe, J. Wolff, M. Genest, A.T. Hagler, Structure and energetics of ligand binding to proteins: *Escherichia coli* dihydrofolate reductase-trimethoprim, a drug–receptor system, Proteins 4 (1988) 31–47.
[46] R.A. Laskowski, J.A. Rullmannn, M.W. MacArthur, R. Kaptein, J.M. Thornton, AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR, J. Biomol. NMR 8 (1996) 477–486.
[47] J.A. Cuff, M.E. Clamp, A.S. Siddiqui, M. Finlay, G.J. Barton, JPred: a consensus secondary structure prediction server, Bioinformatics 14 (1998) 892–893.
[48] G. Ramachandraiah, N.R. Chandra, A. Surolia, M. Vijayan, Re-refinement using reprocessed data to improve the quality of the structure: a case study involving garlic lectin, Acta Crystallogr. D: Biol. Crystallogr. 58 (2002) 414–420.
[49] P. Benkert, M. Kunzli, T. Schwede, QMEAN server for protein model quality estimation, Nucleic Acids Res. 37 (2009) W510–W514.
[50] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, C. Simmerling, Comparison of multiple Amber force fields and development of improved protein backbone parameters, Proteins 65 (2006) 712–725.
[51] D.A. Case, T.A. Darden, T.E. Cheatham, C.L. Simmerling, J. Wang, R.E. Duke, R. Luo, R.C. Walker, W. Zhang, K.M. Merz, B. Wang, S. Hayik, A. Roitberg, G. Seabra, I. Kolossvary, K.F. Wong, F. Paesani, Jiri Vanicek, Jian Liu, X. Wu, S.R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, V. Hornak, G. Cui, D.R. Roe, D.H. Mathews, M.G. Seetin, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, P.A. Kollman, B.P. Roberts, AMBER 11, University of California, San Francisco, 2010.
[52] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, Comparison of simple potential functions for simulating liquid water, AIP, 1983.
[53] H.J.C. Berendsen, J.P.M. Postma, W.F.v. Gunsteren, A. DiNola, J.R. Haak, Molecular dynamics with coupling to an external bath, AIP, 1984.
[54] J.-P. Ryckaert, G. Ciccotti, H.J.C. Berendsen, Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes, J. Comput. Phys. 23 (1977) 327–341.
[55] R.W. Pastor, B.R. Brooks, A. Szabo, An analysis of the accuracy of Langevin and molecular dynamics algorithms, Mol. Phys. 65 (1988) 1409–1419.
[56] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: an N-log(N) method for Ewald sums in large systems, AIP, 1993.
[57] J. Srinivasan, T.E. Cheatham, P. Cieplak, P.A. Kollman, D.A. Case, Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate–DNA helices, J. Am. Chem. Soc. 120 (1998) 9401–9409.
[58] P.A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, et al., Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models, Acc. Chem. Res. 33 (2000) 889–897.
[59] W.C. Still, A. Tempczyk, R.C. Hawley, T. Hendrickson, Semianalytical treatment of solvation for molecular mechanics and dynamics, J. Am. Chem. Soc. 112 (1990) 6127–6129.
[60] D. Sitkoff, K.A. Sharp, B. Honig, Accurate calculation of hydration free energies using macroscopic solvent models, J. Phys. Chem. 98 (1994) 1978–1988.
[61] M.F. Sanner, A.J. Olson, J.-C. Spehner, Reduced surface: an efficient way to compute molecular surfaces, Biopolymers 38 (1996) 305–320.
[62] J. Schlitter, Estimation of absolute and relative entropies of macromolecules using the covariance matrix, Chem. Phys. Lett. 215 (1993) 617–621.
[63] M.M. Teeter, D.A. Case, Harmonic and quasiharmonic descriptions of crambin, J. Phys. Chem. 94 (1990) 8091–8097.
[64] D. Case, T. Darden, T. Cheatham III, C. Simmerling, J. Wang, R. Duke, R. Luo, M. Crowley, R.C. Walker, W. Zhang, K. Merz, B. Wang, S. Hayik, A. Roitberg, G. Seabra, I. Kolossváry, K.F. Wong, F. Paesani, J. Vanicek, X. Wu, S. Brozell, T. Steinbrecher, H. Gohlke, L. Yang, C. Tan, J. Mongan, V. Hornak, G. Cui, D. Mathews, M. Seetin, C. Sagui, V. Babin, P. Kollman, Amber 11, University of California, San Francisco, 2011.
[65] V. Tsui, D.A. Case, Theory and applications of the generalized Born solvation model in macromolecular simulations, Biopolymers 56 (2001) 257–291.
[66] A. Onufriev, D. Bashford, D.A. Case, Exploring protein native states and large-scale conformational changes with a modified generalized born model, Proteins Struct. Funct. Bioinform. 55 (2004) 383–394.
[67] S. Pacheco, I. Gomez, I. Arenas, G. Saab-Rincon, C. Rodriguez-Almazan, S.S. Gill, et al., Domain II loop 3 of *Bacillus thuringiensis* Cry1Ab toxin is involved in a ping pong binding mechanism with *Manduca sexta* aminopeptidase-N and cadherin receptors, J. Biol. Chem. 284 (2009) 32750–32757.

[68] K.G. Tina, R. Bhadra, N. Srinivasan, PIC. protein interactions calculator, Nucleic Acids Res. 35 (2007) W473–W476.
[69] K. Nakanishi, K. Yaoi, N. Shimada, T. Kadotani, R. Sato, *Bacillus thuringiensis* insecticidal Cry1Aa toxin binds to a highly conserved region of aminopeptidase N in the host insect leading to its evolutionary success, Biochim. Biophys. Acta 1432 (1999) 57–63.
[70] S. Liu, C. Zhang, H. Zhou, Y. Zhou, A physical reference state unifies the structure-derived potential of mean force for protein folding and binding, Proteins 56 (2004) 93–101.
[71] S.C. MacIntosh, T.B. Stone, S.R. Sims, P.L. Hunst, J.T. Greenplate, P.G. Marrone, et al., Specificity and efficacy of purified *Bacillus thuringiensis* proteins against agronomically important insects, J. Invertebr. Pathol. 56 (1990) 258–266.
[72] J. Janin, The kinetics of protein–protein recognition, Proteins 28 (1997) 153–161.
[73] D. Xu, C.J. Tsai, R. Nussinov, Hydrogen bonds and salt bridges across protein–rotein interfaces, Protein Eng. 10 (1997) 999–1012.
[74] S. Atsumi, E. Mizuno, H. Hara, K. Nakanishi, M. Kitami, N. Miura, et al., Location of the *Bombyx mori* aminopeptidase N type 1 binding site on *Bacillus thuringiensis* Cry1Aa toxin, Appl. Environ. Microbiol. 71 (2005) 3966–3977.
[75] B.J. Carrington, R.L. Mancera, Comparative estimation of vibrational entropy changes in proteins through normal modes analysis, J. Mol. Graph. Model. 23 (2004) 167–174.
[76] B. Xu, H. Shen, X. Zhu, G. Li, Fast and accurate computation schemes for evaluating vibrational entropy of proteins, J. Comput. Chem. 32 (2011) 3188–3193.
[77] L. Masson, A. Mazza, L. Gringorten, D. Baines, V. Aneliunas, R. Brousseau, Specificity domain localization of *Bacillus thuringiensis* insecticidal toxins is highly dependent on the bioassay system, Mol. Microbiol. 14 (1994) 851–860.
[78] D. Bosch, B. Schipper, H. van der Kleij, R.A. de Maagd, W.J. Stiekema, Recombinant *Bacillus thuringiensis* crystal proteins with new properties: possibilities for resistance management, Biotechnology (NY) 12 (1994) 915–918.
[79] L. Mehlo, D. Gahakwa, P.T. Nghia, N.T. Loc, T. Capell, J.A. Gatehouse, et al., An alternative strategy for sustainable pest resistance in genetically enhanced crops, Proc. Natl. Acad. Sci. U. S. A. 102 (2005) 7812–7816.