

# BIOINFORMATICS ANALYSIS USING HOMOLOGY MODELING TO PREDICT THE THREE-DIMENSIONAL STRUCTURE OF *Spodoptera littoralis* (Lepidoptera: Noctuidae) AMINOPEPTIDASE N RECEPTOR

A. M. SHOKRY<sup>1,2</sup>, M. A. ISMAIL<sup>3</sup>, HEBA M. YASSIN<sup>3</sup>, S. A. MOSTAFA<sup>1</sup>, M.  
S. SALAMA<sup>4</sup> AND M. A. SHAHIN<sup>3</sup>

1. Bioinformatics Dept., Agriculture Genetic Engineering Research Institute (AGERI), ARC, Giza, Egypt
2. Genomic and Biotechnology Section, Biology Dept., Fac. of Sciences, King Adulaziz Univ., Jeddah, Saudi Arabia
3. Faculty of Education - Ain Shams University, Cairo, Egypt
4. Faculty of Science- Ain Shams University, Cairo, Egypt

**D**uring the last years, the number of protein sequences of has increased rapidly. In spite of the fact that the X-ray crystallography is the main method for the determination of protein structure, it is consuming time and succeeds only if suitable conditions for growing crystals are possible (Wieman *et al.*, 2004). In this concern, three main methods of computational prediction for protein structure are used to determine three dimensional (3-D) structure of a protein from its sequence which are homology modeling, threading and *de novo* methods (Polanski and Kimmel, 2007). Threading and *de novo* methods are used when no homologous structure is available, but these methods are not yet very accurate. At the same time, the homology modeling is an improved method based on the fact that homologous proteins have similar 3-D structures. Therefore, it is highly desired to model the 3-D structure of protein by using structural bioinformatics approach with homology

modeling as a computational biological tool (Sateesh *et al.*, 2010).

The lepidopteran worm, *Spodoptera littoralis* is a polyphagous pest affecting various economically important crops. The homology modeling approach was utilized to predict the 3-D structure of aminopeptidase-N (APN) in *Spodoptera littoralis*, *in silico* to identify its gene function (Bravo *et al.*, 2007; Choi *et al.*, 2009; Singh *et al.*, 2010). Aminopeptidase-N (APN) is one of the four different kinds of insect receptors, which binds to *Bacillus thuringiensis* (*Bt*) toxins *via* oligomerization process. The other three are cadherins, glycoproteins and alkaline phosphatase (Knight *et al.*, 1994; Vadlamudi *et al.*, 1995; Valaitis *et al.*, 2001; Jurat-Fuentes and Adang, 2004).

Aminopeptidases N (APNs) are a class of endoproteases that cleave the N-terminus of the polypeptides to release

single amino acids (Piggot and Ellar, 2007; Crava *et al.*, 2010). They are members of the zinc dependant metalloprotease M1 type-I, that need the divalent cation zinc to activate a molecule of water, and belong to a subfamily named gluzincins (Albiston *et al.*, 2004; Crava *et al.*, 2010). Furthermore, Luan and Xu (2007) and Crava *et al.* (2010) stated that a single zinc ion is joined with a highly conserved HEX<sub>2</sub>HX<sub>18</sub>E amino acid motif, which playing a major role with Cry toxin interactions (Yang *et al.*, 2010). In more depth, the APN receptor plays a major role with Cry toxin interactions located in the apical membrane of midgut epithelial cells of insect's microvilli (Bravo *et al.*, 2004 & 2005; Crava *et al.*, 2010). Luo and Adang (1996); Denolf *et al.* (1997); Agrawal *et al.* (2002) and Crava *et al.* (2010) reported that all the different APNs genes encoded proteins of approximately 1000 amino acids that undergo various forms of post-translational modification *via* glycosylphosphatidylinositol (GPI) membrane anchored and N-glycosylation process to produce mature proteins ranging 90-170 kDa in size.

During sporulation, of the Gram-positive bacteria, *Bacillus thuringiensis* (*Bt*) forms crystalline protein inclusions, which possess insecticidal activity (Bravo *et al.*, 2005). The mode of action of *Bt* toxins to kill insects is not fully understood. However, Bravo *et al.* (2007) discussed the principal characterization of *Bt* toxins effects in lepidopteran insects. Whereas, the crystals of *Bt* toxins are solubilized in the insect midgut lumen due to

its characteristic pH and reducing conditions. The soluble protoxins are then activated by midgut proteases to release the toxin fragment (Bravo *et al.*, 2005 & 2007; Piggot and Ellar, 2007).

*Spodoptera litura* and *S. littoralis* are susceptible to *Bt* Cry1C toxin (Agrawal *et al.*, 2002; Yassin *et al.*, 2010). Moreover, Yassin *et al.* (2010) added that such toxin (Cry1C) generated susceptibility to 2882 bp gene encoding a 109 kDa APN receptor protein, which has been isolated from *S. littoralis* and cloned into convenient system pGEM-T easy vector.

As Known the 3-D structure of any protein is essential to understand how protein performs its function. So, the protein structure could be determined at high resolution by either experimental (i.e., X-ray crystallography and nuclear magnetic resonance; NMR) or computational methods using bioinformatics tools as described by Sasin and Bujnicki (2004). A variety of advanced homology modeling methods have been developed to provide reliable models of a protein that sharing in even 30% or more sequence identity with a known protein structure (Burley, 2000). The author reported also that the homology modeling could offer a possibility for the identification of target amino acid residues for protein engineering. The comparative homology modeling takes advantages of the structural similarities within the same family to construct an atomic resolution model of a protein from its amino acid sequence (Sali and Blundell, 1999). Such view depended on that the

proteins in the same family share the same basic folding in spite of low level of sequence identity (Choi *et al.*, 2009).

The objective of this study was to predict the complete 3-D structure of *Spodoptera littoralis* aminopeptidase N (SIAPN) protein based on comparative homology modeling. This 3-D model explores the molecular basis of a potential reaction mechanism between such APN protein receptor and *Bt* Cry toxin.

## MATERIALS AND METHODS

### *Sequence identification*

The APN sequence of *Spodoptera littoralis* (SIAPN) having length of 2882 bp was obtained from Yassin *et al.* (2010). Such sequence was translated using Vector NTI® Suite software version 11 available from Informax, Inc., Bethesda, Md.

### *Sequence retrieval alignment and homology modeling*

Protein sequence homology analysis and homology modeling of *Spodoptera littoralis* APN was compared with the predicted model of *Spodoptera litura* APN that has been deposited in the Model Protein Database at <http://www.caspur.it/PMDB/> with ID: PM0074654. A homology model for *Spodoptera litura* was constructed based on X-ray crystallographic structure of both proteins, tricorn interacting factor F3 of *Thermoplasma acidophilum* (PDB: 1z1w) and leukotrien A4 hydrolase D375n mutant of human (PDB: 1gw6).

### *In silico: the 3-D homology modeling software*

Homology modeling was performed using SWISS-MODEL, protein modeling server (Guex and Peitsch, 1997; Schwede *et al.*, 2003; Arnold *et al.*, 2006), accessible via the EXPASY at <http://www.expasy.org/>. Superimposition of *Spodoptera littoralis* APN model on that of *Spodoptera litura* APN was constructed by using RasMol (<http://www.umass.edu/microbio/rasmol/>) and Deep-View program (<http://spdbv.vital-it.ch/>).

The functional domains were identified from the National Center for Biotechnology Information (NCBI) conserved domain database (CDD) at <http://www.ncbi.nlm.nih.gov/blast/blast.cgi?db=cdd>.

## RESULTS AND DISCUSSION

In the current study, homology modeling method and computer programs were performed to predict the 3-D structure of SIAPN protein. The present study recorded the schematic representation of the predicted amino acid sequence of SIAPN protein of previously nucleotide sequenced by Yassin *et al.* (2010) as illustrated in Fig. (1). The results also showed that the isolated SIAPN gene encoded a putative 952 amino acid residues. In turn, the molecular weight was calculated to be 108.58 kDa. In this concern, Khan and Ranganathan (2009), Sahay and Shakya (2010) and Tajne *et al.* (2012) used the

same idea in their studies on a-D-Mannosidase in mammalian, antioxidant proteins of spinach and on *Manduca sexta*, respectively.

*In silico* analysis of the present observations showed the primary amino acid sequence that identified *SIAPN* as a member of the aminopeptidase family. Such results depended on the presence of common motifs of aminopeptidase family which included the following conservative regions: zinc-binding/gluzincin motif, gluzincin aminopeptidase motif, GPI anchor aminopeptidase. These results are in agreement with Crava *et al.* (2010) and Tajne *et al.* (2012) who found such observations on aminopeptidase N gene family in the lepidopterans.

In the present study, to analyze 3-D model of the deduced amino acid sequence, both signal peptide and GPI regions on *SIAPN* gene were excluded. Whereas, the signal peptide is a sequence of 15-30 amino acids at the N terminus of a secreted protein requiring for transport through a membrane and cleaved off after secretion (Lodge *et al.*, 2007). Whilst, GPI anchor tethers C terminus end of APN protein on the cell membrane by covalent linkage, as reported by Pierleoni *et al.* (2008). So, these regions are not involved in the *Bt* toxicity interaction at this investigation. However, this step was completed after using both Signal P and GPI Prediction GPI-SOM programs at <http://www.cbs.dtu.dk/services/SignalP/> and <http://gpi.unibe.ch/>, respectively.

After using the previous Signal P program, the obtained observations revealed that N-terminal cleavable signal peptide for retention in the endoplasmic reticulum was at residues 1 to 20. This result was also defined by Lodge *et al.* (2007). Besides, the study also showed that the cleavage site between the signal and the mature protein was between amino acid residues 20 and 21. Moreover, in the present investigation, GPI prediction program exhibited the presence of GPI anchor signal sequence at the C-terminus. Such sequence consists of three small amino acids SNS and followed by a stretch of 20 hydrophobic amino acid residues PTIFASSFLILAAMLIQLYR. There is also a signal sequence of three amino acids DSA to attach with GPI anchor.

The APN analysis in the present study showed the presence of three conserved regions. Firstly, zinc-binding/gluzincin motif HEX<sub>2</sub>HX<sub>18</sub>E (residues 355 to 378), which is a part of a typical catalytic active site for the majority zinc-dependent metallopeptidases and is required for enzymatic function. Secondly, the third zinc-binding ligand which is conserved in the sequence motif NEXFA (residues 377-381). Finally, gluzincin aminooeptidase motif GAMEN (residues 319 to 323), which is believed to form part of the active site and it also involves in aminopeptidase activity (Herrero *et al.*, 2005; Kyrieleis *et al.*, 2005).

Furthermore, the present results showed that the potential N-linked glyco-

sylation sites (NXS/T) are observed at residues 103, 377, 430, 574, 711 and 782. Besides, the amino acid sequence also contained four Cys residues. These residues are highly conserved among APN molecules of higher mammals (rat, rabbit, pig and human), as detected by Agrawal *et al.* (2002).

Finally, the present analysis revealed a highly conserved 64 amino acid residues from Leu 129 to Pro 193 in *S. littoralis* APN. These residues are in common with 11 APN family proteins. However, these 64 amino acid residues are believed to be important for Cry1Aa toxin binding (Nakanishi *et al.*, 1999 and 2002) and after binding to a receptor in the insect midgut, the toxin undergoes a conformational change leads to form pores (Sanjay *et al.*, 2001).

Based on computational alignment, a theoretical 3-D model structure of *S. littoralis* APN was obtained in this study, whereas, about 814 amino acids (aa) residues of whole suggested 3-D model were corresponded to residues 58-871 of the primary structure. Considering that the N-terminal cleavable signal peptide (residues 1-20) and the C-terminal GPI modification site (residues 930-952) were removed from the predicted 3-D model as previously discussed.

The present *SlAPN* model contained four structural domains, which spread from N-terminal domain I to C-terminal domain IV over the regions Asn<sup>58</sup>-Ile<sup>266</sup>; Ser<sup>267</sup>-Gly<sup>506</sup>; Asn<sup>507</sup>-Leu<sup>581</sup>

and Ser<sup>582</sup>-Ala<sup>871</sup>, respectively (Fig. 2). The recorded overall dimensions of this model were 91Å x 55Å x 65Å forming together a hook like structure.

Nevertheless, the former recorded N-terminal domain I contained highly conserved *Bt* Cry1 toxin-binding region (64 amino acid residues from Leu<sup>129</sup> to Pro<sup>193</sup>). The second domain was the most important one contained highly conserved regions of aminopeptidase family; including zinc-binding/gluzincin motif, the third zinc-binding ligand and the gluzincin aminopeptidase motif. The N-linked glycosylation sites are located in all the four domains starting from domain I (residues Asn<sup>103</sup>-Thr<sup>105</sup>), domain II (residues Asn<sup>377</sup>-Ser<sup>379</sup>, Asn<sup>430</sup>-Thr<sup>432</sup>), domain III (residues Asn<sup>574</sup>-Thr<sup>576</sup>, Asn<sup>711</sup>-Ser<sup>713</sup>) and finally in domain IV (residues Asn<sup>782</sup>-Ser<sup>784</sup>). The four highly conserved Cys, among APN molecules of higher eukaryotes, (residues 728, 735, 763 and 799) reported by Agrawal *et al.* (2002) are located in domain IV. In the 814 aa of *SlAPN*, the contents of the  $\alpha$ -helix,  $\beta$ -sheet, turn, and random coil were 312 aa, 124 aa, 378 aa, and 0 aa, respectively.

The structure of the N-terminal domain I (Fig. 3A) consists of a short  $\alpha$ -helix beside three clusters of  $\beta$ -sheets. The first cluster contains five non consecutive  $\beta$ -strands which are: S1 (Leu<sup>59</sup>-Thr<sup>70</sup>), S2 (Ile<sup>78</sup>-Met<sup>89</sup>), S3 (Asn<sup>109</sup>-Gln<sup>111</sup>), S4 (Thr<sup>143</sup>-Asn<sup>153</sup>) and S7 (Ile<sup>201</sup>-Ile<sup>203</sup>). The first four strands are in anti-parallel form while the last strand (S7) runs parallel to the first one (S1). Each of the second

(55&56) and the third (58&59) clusters of  $\beta$ -sheet are composed of two anti-parallel strands which are S5 (Arg<sup>161</sup>-Tyr<sup>163</sup>) and S6 (Glu<sup>167</sup>-Thr<sup>169</sup>) and S8 (Glu<sup>222</sup>- Arg<sup>229</sup>) and S9 (Thr<sup>231</sup>-Ile<sup>237</sup>), respectively. The  $\alpha$ -helix of domain I (H1; Phe<sup>187</sup>- Cys<sup>189</sup>) is located between strands S6 and S7.

Concerning the description of deduced amino acids that are present in domain I of *SlAPN* protein, a highly conserved *Bt* Cry1 toxin-binding region (64 amino acid residues from Leu<sup>129</sup> to Pro<sup>193</sup>) is found. So, this region began a little before S4 and continuing until reaching the Pro<sup>193</sup> residue, which located at the loop after H1 helix. Since the Cry1 toxin-binding regions have many conserved amino acid residues that may recognize and bind to a common structure in Cry1 toxin regions. Most of those amino acids are also conserved in lepidopteran APNs, suggesting that this toxin might bind to this region of APN in insects (Nakanishi *et al.*, 2002). This suggestion is confirmed with the structure of that region which is formed by various  $\beta$ -strands and  $\alpha$ -helix, as well as the large loops providing the flexibility for binding to Cry1 protein; also confirmed with the position of that part in the most distal part of membrane anchor segment and localization in the bottom of the saddle-like structure. All these findings providing the stability of that conserved region which is very necessary for toxin binding.

In closer look, such domain reveals also the oligosaccharide-binding site 103 is located within the *Bt* Cry1 toxin-

binding region that possibly allowing initial interaction between toxin and receptor and then an irreversible binding on the APN recognition site. As, the binding of toxin to the receptor is mediated by a two-step mechanism involving initial reversible binding followed by irreversible binding and membrane insertion in the midgut (Saraswathy and Kumar, 2004). In more depth, Chen *et al.* (2005) proposed that Cry1-APN interaction have two steps: carbohydrate recognition and irreversible protein-protein interaction.

Moreover, Pigott and Ellar (2007) regarded glycosylation as an important determinant of Cry1A binding. Knight *et al.* (2004) also considered the carbohydrates that attached to 120 kDa APN of *Manduca sexta* (tobacco hornworm) were epitope sites of Cry1Ac toxin. Also, Ning *et al.* (2010) found that binding between Cry1Ac and *HaALPs* was depended on the presence of N-linked oligosaccharides of these proteins, since digestion with N-glycosidase F eliminated toxin binding.

Domain II (Fig. 3B) was the membrane-spanning domain and forms the catalytic region of class proteases. The structure of this domain appeared to be consisting of two clusters of  $\beta$ -sheets and nine  $\alpha$ -helixs, presenting the structure of this domain to be highly conserved (Luan and Xu, 2007). Each one of both  $\beta$ -sheet clusters is composed of two  $\beta$ -strands. The two  $\beta$ -strands were in parallel at the first cluster, involving S10 (Ser<sup>267</sup>- Ala<sup>271</sup>) and S11 (Asn<sup>308</sup>-Ala<sup>312</sup>). Whilst, the two  $\beta$ -strands of the second cluster were in anti-

parallel configuration, including S12 (Ala<sup>320</sup>-Glu<sup>322</sup>) and S13 (Tyr<sup>329</sup>-Glu<sup>331</sup>). The rest of domain II was nine  $\alpha$ -helices containing H2 (Ala<sup>276</sup>-Tyr<sup>277</sup>), H3 (Gln<sup>288</sup>-His<sup>298</sup>), H4 (Val<sup>344</sup>-Trp<sup>361</sup>), H5 (Trp<sup>375</sup>-Ala<sup>392</sup>), H6 (Leu<sup>398</sup>-Asp<sup>414</sup>), H7 (Thr<sup>439</sup>-Phe<sup>455</sup>), beside H8 (Arg<sup>459</sup>-Asp<sup>471</sup>), H9 (Pro<sup>479</sup>-Lys<sup>488</sup>) and H10 (Thr<sup>495</sup>-Ile<sup>503</sup>) helices. The first two  $\alpha$ -helices were located in between  $\beta$ -strands S10 and S11, while the remains located after S13. The recorded loop between H6 and H7 (Ala<sup>415</sup>-Ser<sup>438</sup>) was a highly flexible and covered the front of this domain.

On the avenue of that domain (II), many conserved structures were observed in this investigation, such as the conserved zinc-binding/gluzincin motif (HEXXHX18E) in helix H4 and H5; the third zinc-binding ligand (NEXFA) in helix H5, and the gluzincin aminopeptidase (GAMEN) motif in  $\beta$ -strand S12. Hence, Cry1 proteins can develop toxic activity in a broad spectrum of pest insects by recognizing the conserved structures (Nakanishi *et al.*, 1999; Agrwal *et al.*, 2002; Herrero *et al.*, 2005). This also bearing in mind the appearance of conserved residues in domains II and III of Cry toxin; those domains are implicated in receptor recognition and pore formation (Shinkawa *et al.*, 1999). Thus, the study of the conserved regions in both toxin and receptor could be helpful to know the interaction of that enzyme with *Bt* toxins. However, the main mechanism of resistance to Cry toxins are due to the mutations affect toxin receptor interaction (Fernández *et al.*, 2008).

Comparison of APNs sequence of experimented insect as well as vertebrate and fungi aminopeptidases showed that the most striking similarity was around the zinc-binding motif, suggesting the role of zinc metal in enzyme catalysis. The catalytic role of domain II in binding to different substrates has been demonstrated in several zinc-metallopeptidases including aminopeptidase of the beetle *Tenebrio molitor* (Cristofaletti and Terra, 2000).

Domain III (Fig. 3C) contains seven  $\beta$ -strands; S14 (Asn<sup>507</sup>-Asn<sup>515</sup>), S15 (Gly<sup>517</sup>-Asp<sup>526</sup>), S16 (Arg<sup>538</sup>-Glu<sup>544</sup>), S17 (Thr<sup>547</sup>-Pro<sup>553</sup>), S18 (Thr<sup>557</sup>-His<sup>561</sup>), S19 (Ser<sup>566</sup>-Lys<sup>568</sup>) and S20 (Thr<sup>576</sup>-Leu<sup>580</sup>). They are arranged in a  $\beta$ -sandwich forming the base of the saddle-like APN structure.

The C-terminal superhelix domain IV (Fig. 3D) is formed of 18  $\alpha$ -helices and is divided into two helical modules. The helices are organized in the following arrangement: H11 (Ser<sup>582</sup>-Val<sup>589</sup>), H12 (Ser<sup>596</sup>-Ala<sup>612</sup>), H13 (Ser<sup>617</sup>-Arg<sup>628</sup>), H14 (Asn<sup>633</sup>-Thr<sup>647</sup>), H15 (Ile<sup>653</sup>-Ala<sup>668</sup>), H16 (Trp<sup>674</sup>-Phe<sup>689</sup>), H17 (Leu<sup>693</sup>-Asn<sup>699</sup>), H18 (Ser<sup>713</sup>-Ile<sup>723</sup>), H19 (Ala<sup>727</sup>-Cys<sup>735</sup>), H20 (Leu<sup>740</sup>-Lys<sup>747</sup>), H21 (Leu<sup>757</sup>-Cys<sup>763</sup>), H22 (Leu<sup>777</sup>-Asn<sup>782</sup>), H23 (Ala<sup>787</sup>-Ala<sup>798</sup>), H24 (Asp<sup>802</sup>-Leu<sup>806</sup>), H25 (Ser<sup>813</sup>-Met<sup>814</sup>), H26 (Lys<sup>818</sup>-Phe<sup>828</sup>), H27 (Leu<sup>832</sup>-Glu<sup>837</sup>) and H28 (Val<sup>842</sup>-Val<sup>867</sup>).

The same Figure (3D) exhibited that each of the two helical modules is arranged into two layers. Whereas, the first module is composed of eight parallel helices; four of them are located in the

outer layer (H11, H13, H15 and H17) and the other four helices are found in the inner layer (H12, H14, H16 and H18). The second module contains ten helices, seven of them (H19, H21, H23, H24, H25, H27 and H28) are in the outer layer and the rest three helices (H20, H22 and H26) are in the inner layer.

The evidences by several authors who suggested the role of APNs to function as Cry1 receptors are increased. Most of these evidences are based on *in vitro* binding and membrane reconstitution experiments, and on *in vivo* expression in transgenic *Drosophila* (Banks *et al.*, 2003). From *in vivo* evidences, silencing of the *APN* gene resulted in the elevated resistance of *Spodoptera litura* larvae to Cry1C protein, thereby demonstrating a functional role for this protein in Cry protein-mediated toxicity (Rajagopal *et al.*, 2002). Similarly, Sf21 cells expressing *HaAPN1* obtained from cotton bollworm (*Helicoverpa armigera*) showed increase-toxin sensitivity (Sivakumar *et al.*, 2007). The same authors also found that the silencing of *HaAPN1* in *H. armigera* by dsRNA resulted in a decreased larval susceptibility to Cry1Ac toxins. Moreover, findings reported by Herrero *et al.* (2005) suggested that the lack of APN production in laboratory-selected beet armyworm (*Spodoptera exigua*) were correlated with resistance to Cry1C toxin. Recently, the strongest evidence which support the assumption that APN proteins play an important role in Cry toxicity comes from the recent findings of Zhang *et al.* (2009) who postulated that toxin resistant of *H.*

*armigera* had a mutation in the *APN* gene. These findings are in accordance with those of Yang *et al.* (2010) who stated that knocking down of any one of the three APNs in sugarcane borer (*Diatraea saccharalis*) resulted in a decrease in Cry1Ab susceptibility.

As explained previously, Cry1 toxins showed toxicity for several lepidopteran insects. In addition, APNs from several insect species have been identified as Cry1 toxin receptors. It is important to identify the binding sites on both Cry toxin and receptor to understand the interaction between them. In this concern, Atsumi *et al.* (2005) hypothesized that the receptor binding sites on Cry1 toxins have two basic features. The former, as described by the previous authors, was a highly conserved structure, due to that the Cry1 toxins have similar primary sequences, three-dimensional structures and can recognize similar APNs in the midguts of several lepidopteran insects. Besides, a nonconserved structure, because the Cry1 toxins also exhibited a highly specific insecticidal activity and could distinguish host species in the lepidopteran spectrum.

From all the foregoing results, it can be concluded that the three dimensional structure of APNs, of economically important pest insect, must be identified and investment in biological control using Cry toxins scope. The results will provide insights on the functional properties of APN towards the understanding of the receptor-toxin interactions which will be



valuable for the production of Cry toxin proteins with a greater activity.

### SUMMARY

Insect pests are the major cause of damage to commercially important agricultural crops. The continuous application of synthetic pesticides resulted in developing severe insect resistance in addition to induce irreversible damage to the environment. *Bacillus thuringiensis* (*Bt*) emerged as a valuable biological alternative in pest control. The midgut aminopeptidase N (APN) of pest insect is a receptor for *Bacillus thuringiensis* Cry1 toxin. A 108.58 kDa APN has been characterized in *Spodoptera littoralis*. In the present *in silico* study, a homology model of SIAPN was constructed using Swiss-Model, Protein Modeling Server. The study detected that SIAPN three-dimensional structure has 4 structural domains. Domain I of the receptor is the region that recognizes Cry1 toxins, a part of this section might be very important in this role. Domain II has functions in Cry1 protein-APN interaction. Domain III has a sandwich topology and domain IV is a superhelix. The present data help in the development of a roadmap for the design and synthesis of novel Cry toxins and improve toxic activities depending on the APN's conserved structures which will contribute to the management of insect resistance in the field.

### REFERENCES

- Agrawal, N., P. Malhotra and R. K. Bhatnagar (2002). Interaction of

gene cloned and insect cell expressed aminopeptidase N of *Spodoptera litura* with insecticidal crystal protein Cry1C. Appl. Environ. Microbiol., 68: 4583-4592.

Albiston, A. L., S. Ye and S. Y. Chai (2004). Membrane bound members of the M1 family: more than aminopeptidases. Protein Pept. Lett., 11: 491-500.

Arnold, K., L. Bordoli, J. Kopp and T. Schwede (2006). The SWISS-MODEL Workspace: A web-based environment for protein structure homology modeling. Bioinformatics, 22: 195-201.

Atsumi, S., E. Mizuno, H. Hara, K. Nakanishi, M. Kitami, N. Miura, H. Tabunoki, A. Watanabe and R. Sato (2005). Location of the *Bombyx mori* aminopeptidase N type 1 binding site on *Bacillus thuringiensis* Cry1Aa toxin. Appl Environ. Microbiol., 71: 3966-3977.

Banks, D. J., G. Hua and M. J. Adang (2003). Cloning of a *Heliothis virescens* 110 kDa amino-peptidase N and expression in *Drosophila* S2 cells, Insect Biochem. Mol. Biol., 33: 499-508.

Bravo, A., S. S. Gill and M. Soberón (2005). *Bacillus thuringiensis* mechanisms and use. In: Comprehensive Molecular Insect Science. Elsevier ed., 175-206.

- Bravo, A., S. S. Gill and M. Soberón (2007): Mode of action of *Bacillus thuringiensis* Cry and Cyt toxins and their potential for insect control. *Toxicon.*, 49: 423-435.
- Bravo, A., I. Gómez, J. Conde, C. Muñoz-Garay, J. Sánchez and M. Zhuang (2004). Oligomerization triggers differential binding of a pore-forming toxin to a different receptor leading to efficient interaction with membrane microdomains. *Biochem. Biophys. Acta*, 1667: 38-46.
- Burley, S. K. (2000). An overview of structural genomics. *Nat. Struct. Biol.*, 7: 932-938.
- Chen, J., M. R. Brown, G. Hua and M. J. Adang (2005). Comparison of the localization of *Bacillus thuringiensis* Cry1A delta-endotoxins and their binding proteins in larval midgut of tobacco hornworm, *Manduca sexta*. *Cell Tissue Res.*, 321: 123-129.
- Choi, S. H., H. S. Kim and E. Y. Lee (2009). Comparative homology modeling-inspired protein engineering for improvement of catalytic activity of *Mugil cephalus* epoxide hydrolase. *Biotechnol. Lett.*, 31: 1617-1624.
- Crava, C. M., Y. Bel, S. F. Lee, B. Manachini, D. G. Heckel and B. Escriche (2010). Study of the aminopeptidase N gene family in the lepidopterans *Ostrinia nubilalis* (Hübner) and *Bombyx mori* (L.): Sequences, mapping and expression. *Insect Biochemistry and Molecular Biology*, 40: 506-515
- Cristofaletti, P. T. and W. R. Terra (2000). The role of amino acid residues in the active site of a midgut microvillar aminopeptidase from the beetle *Tenebrio molitor*. *Biochim. Biophys. Acta*, 1479: 185-195.
- Denolf, P., K. Hendrickx, J. Van Damme, S. Jansens, M. Peferoen, D. Degheele and J. Van Rie (1997). Cloning and characterization of *Manduca sexta* and *Plutella xylostella* midgut aminopeptidase N enzymes related to *Bacillus thuringiensis* toxin-binding proteins. *Eur. J. Biochem.*, 248: 748-761.
- Fernández, L. E., I. Gómez, S. Pacheco, I. Arenas, S. S. Gill, A. Bravo and M. Soberón (2008). Employing phage display to study the mode of action of *Bacillus thuringiensis* Cry toxins. *Peptides*, 29: 324-329.
- Guex, N. and M. C. Peitsch (1997). SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis*, 18: 2714-2723.
- Herrero, S., T. Gechev, P. L. Bakker, W. J. Moar and R. A. de Maagd (2005). *Bacillus thuringiensis*

- Cry1Ca-resistant *Spodoptera exigua* lacks expression of one of four aminopeptidase N genes. *BMC Genomics*, 6: 96
- Jurat-Fuentes, J. L. and M. J. Adang (2004). Characterization of a Cry1Ac receptor alkaline phosphatase in susceptible and resistant *Heliothis virescens* larvae. *Eur. J. Biochem.*, 271: 3127-3135.
- Khan, J. M. and S. Ranganathan (2009). A multi-species comparative structural bioinformatics analysis of inherited mutations in a-D-Mannosidase reveals strong genotype-phenotype correlation. *BMC Genomics*, 10(Suppl 3):S33
- Knight, P. J., J. Carroll and D. J. Ellar (2004). 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: 101-12.
- Knight, P. J., N. Crickmore and D. J. Ellar (1994). The receptor for *Bacillus thuringiensis* Cry1A (c)  $\delta$  endotoxin in the brush border membrane of the lepidopteran *Manduca sexta* is aminopeptidase N. *Mol. Microbiol.*, 11: 429-436.
- Kyrieleis, O. J. P., P. Goettig, R. Kiefersauer, R. Huber and H. Brandstetter (2005). Crystal structures of the tricorner interacting factor F3 from *Thermoplasma acidophilum*, a zinc aminopeptidase in three different conformations. *J. Mol. Biology*, 349: 787-800.
- Lodge, J., P. Lund and S. Minchin (2007). *Gene Cloning: principles and applications*. Taylor & Francis Group, ISBN 0-203-96728-3 Master e-book ISBN, UK, pp 451.
- Luan, Y. and W. Xu (2007). The structure and main functions of aminopeptidase N. *Current Medicinal Chemistry*, 14: 639-647.
- Luo, K., Y. LU and M. J. Adang (1996). A 106 kDa form of aminopeptidase is a receptor for *Bacillus thuringiensis* Cry1C  $\delta$ -endotoxin in the brush border membrane of *Manduca sexta*. *Insect Biochem. Mol. Biol.*, 26: 783-791.
- Nakanishi, K., K. Yaoi, N. Shimada, T. Kadotani and T. Sato (1999). *Bacillus thuringiensis* insecticidal Cry1Aa toxin binds to a highly conserved region of aminopeptidase N in the host insect leading to its evolutionary success. *Biochim. et Biophysica Acta*, 1432: 57-63.
- Nakanishia, K., K. Yaoib, Y. Naginoa, H. Haraa, M. Kitamia, S. Atsumia, N. Miuraa and R. Satoa (2002). Aminopeptidase N isoforms from the midgut of *Bombyx mori* and *Plutella xylostella*, their classifica-

- tion and the factors that determine their binding specificity to *Bacillus thuringiensis* Cry1A toxin. *FEBS Lett.*, 519: 215-220.
- Ning, C., K. Wu, C. Liu, Y. Gao, L. Jurat-Fuentes and X. Gao (2010). Characterization of a Cry1Ac toxin-binding alkaline phosphatase in the midgut from *Helicoverpa armigera* (Hübner) larvae. *Journal of Insect Physiology*, 56: 666-672.
- Pierleoni, A., P. L. Martelli and R. Casadio (2008). PredGPI: a GPI-anchor predictor. *BMC Bioinformatics*, 9: 392-403.
- Pigott, C. R. and D. J. Ellar (2007). Role of receptors in *Bacillus thuringiensis* crystal toxin activity. *Microbiology and Molecular Biology Reviews*, 71: 255-281.
- Rajagopal, R., S. Sivakumar, N. Agrawal, P. Malhotra and R. K. Bhatnagar (2002). Silencing of midgut aminopeptidase N of *Spodoptera litura* by doublestranded RNA establishes its role as *Bacillus thuringiensis* toxin receptor. *J. Biol. Chem.*, 277: 46849-46851.
- Sahay, A. and M Shakya. (2010). *In silico* analysis and homology modeling of antioxidant proteins of Spinach. *J. Proteomics Bioinformatics*, 3: 148-154.
- Sali, A. and T. L. Blundell (1993). Comparative protein modeling by satisfaction of spatial restraints. *J. Mol. Biol.*, 234: 779-815.
- Sanjay, S., I. N. Trivedi, R. Prasad, J. Kuruvilla, K. K. Rao and H. S. Chhatpar (2001). Aminopeptidase-N from the *Helicoverpa armigera* (Hubner) Brush Border Membrane Vesicles as a Receptor of *Bacillus thuringiensis* Cry1Ac  $\delta$ -Endotoxin. *Current Microbiology*, 43: 255-259.
- Saraswathy, N. and P. A. Kumar (2004). Protein engineering of  $\delta$ -endotoxins of *Bacillus thuringiensis*. *Electronic Journal of Biotechnology*, 7: 178-188.
- Sasin, J. M. and J. M. Bujnicki (2004). *In silico* structural analysis of parasporin 2 protein sequences of non-toxic *Bacillus thuringiensis*. *Nucleic Acids Res.*, 32: 586-589.
- Sateesh, P., A. Rao, S. K. Sangeeta, M. N. Babu and R. S. Grandhi (2010). Homology modeling and sequence analysis of anxC3.1. *International Journal of Engineering Science and Technology*, 2: 1125-1130.
- Schwede, T., J. Kopp, N. Guex and M. C. Peitsch (2003). SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Res.*, 31: 3381-3385.
- Shinkawa, A., K. Yaoi, T. Kadotani, M. Imamura, N. Koizumi, H. Iwahana and R. Sato (1999). Binding of

- phylogenetically distant *Bacillus thuringiensis* Cry toxins to a *Bombyx mori* aminopeptidase N suggests importance of Cry toxin's conserved structure in receptor binding. *Current Microbiology*, 39: 14-20.
- Singh, S., A. Kumar, A. Patel, A. Tripathi, D. Kumar and M. K. Verma (2010). *In Silico* 3D Structure Prediction and Comparison of Nucleocapsid Protein of H1N1. *Journal of Modeling and Simulation of Systems*, 1: 108-111
- Sivakumar, S., R. Rajagopal, G. R. Venkatesh, A. Srivastava and R. K. Bhatnagar. (2007). Knockdown of aminopeptidase-N from *Helicoverpa armigera* larvae and in transfected Sf21 cells by RNA interference reveals its functional interaction with *Bacillus thuringiensis* insecticidal protein Cry1Ac. *J. Biol. Chem.*, 282: 7312-7319.
- Tajne, S., R. Sanam, R. Gundla, N. S. Gandhi, R. L. Mancera, D. Boddupally, D. R. Vudem and V. R. Khareedu (2012). Molecular modeling of *Bt* Cry1Ac (DI-DII)-ASAL (*Allium sativum* lectin)-fusion protein and its interaction with aminopeptidase N (APN) receptor of *Manduca sexta*. *Journal of Molecular Graphics and Modeling*, 33: 61-76.
- Vadlamudi, R. K., E. Weber, I. Ji, T. H. Ji and L. A. Bulla (1995). Cloning and expression of a receptor for an insecticidal toxin of *Bacillus thuringiensis*. *J. Biol. Chem.*, 270: 5490-5494.
- Valaitis, A. P., J. L. Jenkins, M. K. Lee, D. H. Dean and K. J. Garner (2001). Isolation and partial characterization of gypsy moth BTR-270, an anionic brush border membrane glycoconjugate that binds *Bacillus thuringiensis* Cry1A toxins with high affinity. *Arch. Insect Biochem. Physiol.*, 46: 186-200.
- Yang, Y., Y. C. Zhu, J. Ottea, C. Husseneder, B. R. Leonard, C. Abel and F. Huang (2010). Molecular characterization and RNA interference of three midgut aminopeptidase N isozymes from *Bacillus thuringiensis*-susceptible and resistant strains of sugarcane borer, *Diatraea saccharalis*. *Insect Biochemistry and Molecular Biology* 40: 592-603.
- Yassin, H. M., A. M. Shokry, S. A. Mostafa, M. S. Salama and M. A. Shahin (2010). Identification, isolation and cloning of cDNA encoding aminopeptidase from the midgut of the Egyptian cotton leaf worm that serves as a receptor for Cry toxin. *Egypt. J. Genet. Cytol.*, 39: 179-192.
- Wieman, H., K. Tondel, E. Anderssen and F. Drablos (2004). Homology-Based Modeling of Targets for Ra-

tional Drug Design. Mini-Reviews in Medicinal Chemistry, 4: 793-804.

Zhang, S., H. Cheng, Y. Gao, G. Wang, G. Liang and K. Wu (2009). Mu-

tation of an aminopeptidase N gene is associated with *Helicoverpa armigera* resistance to *Bacillus thuringiensis* Cry1Ac toxin. Insect Biochem. Mol. Biol., 39: 421-430.

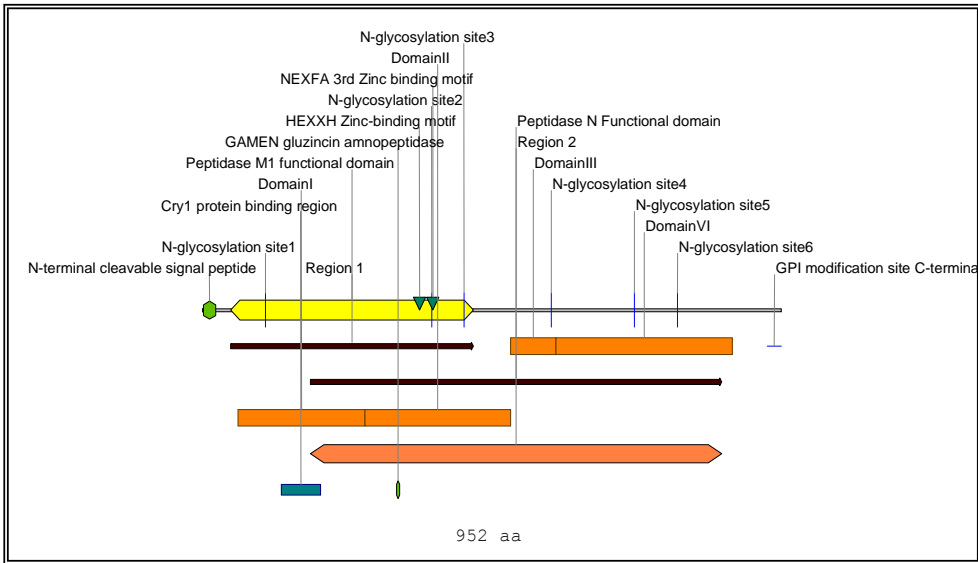


Fig. (1): Schematic representation and putative signature domains of *Spodoptera littoralis* APN.

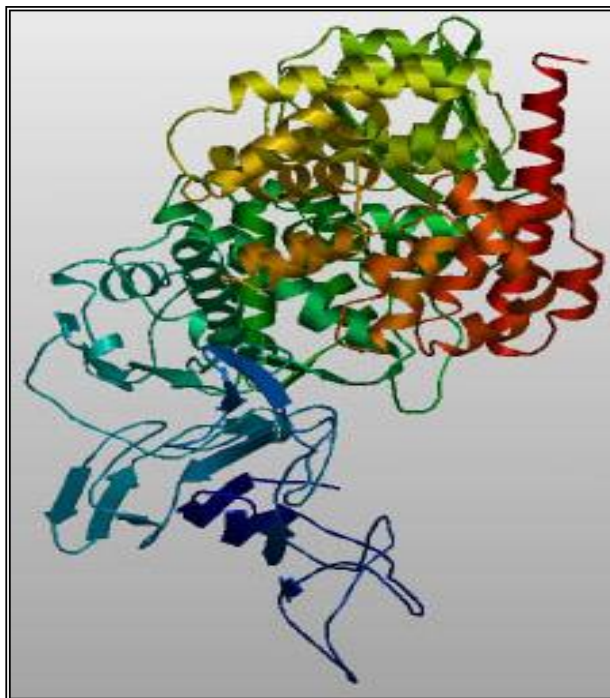


Fig. (2): *Spodoptera littoralis* APN 3-D structure of homology modeling. Domain I is in turquoise, domain II in green, domain III in blue and domain IV in orange.

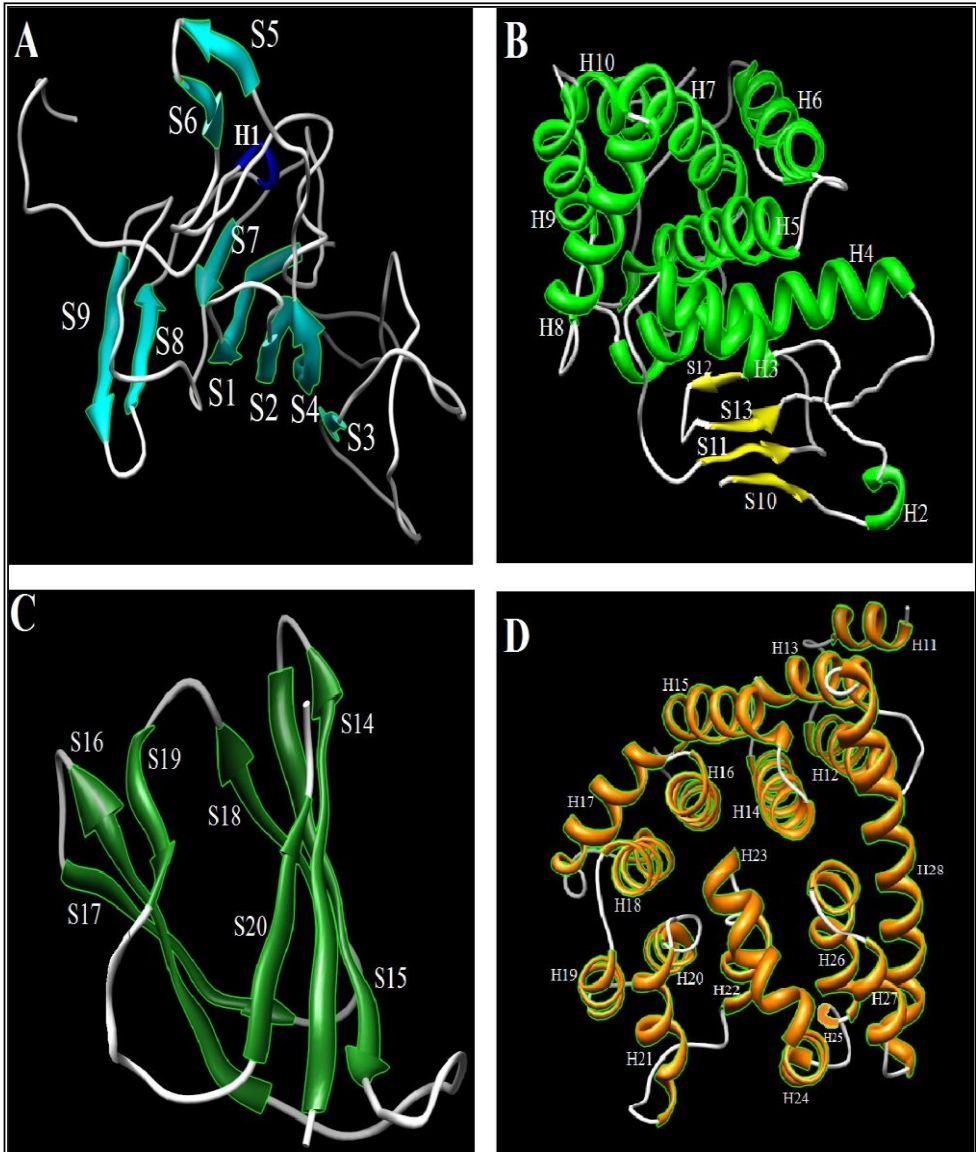


Fig. (3): Structural domains of *Spodoptera littoralis* APN; A- Domain I, B-Domain II, C-Domain III and D-Domain IV. Letter (S) Stands for  $\beta$  Strands and letter (H) stands for  $\alpha$  helix.