Volume : 2 | Issue : 12 | December 2013 • ISSN No 2277 - 8179

Research Paper

Insilico Comparative Analysis of Functional Residues in Aminopeptidase-N

International Journal of Scientific Research

Bioinformatics

KEYWORDS : Aminopeptidase-N, Protein topology, Cry toxins, Multiple sequence alingnment, active site.

Anamika Singh — Bioinformatics and Applied science division., Indian Institute of Information Technology, Allahabad.211012.

C. V. S. Siva Prasad — Bioinformatics and Applied science division., Indian Institute of Information Technology, Allahabad.211012.

ABSTRACT

Aminopeptidase-N receptor is one of the main receptor for Cry toxin binding in insects. Except insects it is also playing important role in humans, bacteria, fungi and many other lower organisms. Due to lack of crystal structure and till now we don't have sufficient data about it binding site and catalytic sites. Here we are trying to show the evolutionary significance and conservation strategies of aminopeptidase-N receptor. For this analysis we have taken 33 aminopeptidase from different organism source. We are trying to predict 3-D structure of insect APN and tracing its active sites, which will help further for drug designing.

Methodology:

Sequence retrieval:

We retrieved different Aminopeptidase-N from many species on the basis of non redundant BLAST and sequences are taken from NCBI and swissprot. *M.sexta* (Q11001), *Bombyx mori* (NP_001037013.1), *Spodoptera exigua* (AAP44964.1), *Trichoplusia ni* (AAX39863.1), *Lymantria dispar* (AAL26895.1), *Plutella xylostella* (AAB70755)1, *Chilo suppressalis* (BC69855.3), *Helicoverpa armigera* (AF521659), *Epiphyas postvittana* (AAF99701.1), *Achaea janata* (ABH07377.2), *Anopheles gambiae* (XP_313486.3), *Aedes aegypti* (EAT42698.1), *Drosophila melanogaster* (CG31198) *P*SA_HUMAN Puromycin-sensitive aminopeptidase (P55786), PSA_MOUSE Puromycin-sensitive aminopeptidase (Q11011), *Nitrosomonas europaea* (NP_840710.1) and *Aplysia californica* (U42380_1, *Felis silvestris catus* (P79171), *Homo sapiens* CD13(P15144), *Mus musculus* (P97449), *Oryctolagus cuniculus* (P15541), *Lactobacillus delbrueckii subsp. Lactis* (P37896), *Escherichia coli* K12 (P04825), *Bos taurus* (P79098), *Sus scrofa* (P15145), *Lactobacillus helveticus* (Q10730), *Caulobacter crescentus* (P37893), *Streptomyces lividans* (Q11010), *Haemophilus influenzae* (P45274), *Lactococcus lactis subsp. Lactis* (Q9CIQ1), *Lactobacillus delbrueckii subsp. Lactis* (P37896), *Haemonchus contortus* (Q10737)and *Lactococcus lactis subsp. cremoris* (P0C2T8).

Multiple sequence Alignment:

Multiple sequence alignment of thirty four sequences was done using CLUSTAL-W (16). Different parameters were tested and manual editing was performed wherever required to get significant alignment.

Motif Identification:

The conserved motifs identified by multiple sequence alignment are submitted to PROSITE web server to scan against existing signatures and identify motifs unique to APN receptor (http://www.expasy.ch/PROSITE/).

Accessible Surface Area Prediction:

Surface Accessible area (SASA values) for APN and Cry toxins were calculated by using POPS Server (Parameter Optimized Surface) (http://ibivu.cs.vu.nl/programs/popswww/). It is a fast algorithm to calculate accessible surfaces areas at atomic (default) and residue (coarse grained) level. For the amino acid conservation score for APN was calculated by using Consurf program (http://consurf.tau.ac.il/). Whereas literature survey shows Cadherin and Cry toxins active sites, hence there is no need to predict active sites further. Identification of functional and structural residues of APN was done by using Conseq program (http://conseq.tau.ac.il/).

Results and Discussions:

Sequence analysis of APNs indicates that most of them were evolved through common ancestors. In this work we are more focused on HaAPN2. HaAPN2 is having differential sequence similarity with different insect species of Lepidopteran, Colepterana and Hymenopterans. APN2 shows similarity of 63% *Sprodoptera exigua,* 57% similarity with *Bombyx mori,* 63.4% with *Manduca sexta,* 61.7% with *Trichiplusia ni*, 59% with *Plutella xylostella* , 56.7% *with Ostrinia nubilalis,* 58% with *Chilo suppressalis* and 57% with *Lymantaria dispar.* In contrast to this less then 50% similarity is found 45.8% *Epiphyas postvittana,* 37.9% with *Aedes gambia,* 37.2% with *Aedes aegypti,* 33.8% with *Achaea janata,* 29.2% with *Mus musculusand,* 29.2% with *Homo sapience. M. sexta* also shows 31% identity with human, rabbit and rat. *M. Sexta* APN is identical to *Saccharomyces* and Mouse APN, 27% and 29% respectively. Alignment results indicates that the maximum sequence length is 1074 amino acid minimum is 810 and average sequence length is 935 (figure 1). Percentage of conservation and the major substitutions present at the different position (Table1). Manually we have selected those amino acids which are having conservation score more than 75%. Table 1 shows amino acid G310, R338, P347, Y402, F405, F484, G487, A488, M489, E490, N491, G493, H523, E524, H527, W529, G531, T535, W539, E546, Y618, G621, Y644, G694 and P743 are 100% conserved. Few amino acids are having conservation 90% or more than 90% are Y312, T330, Q331, D345, N371, L403, Y465, D483, L503, Y534, L542, D585, M627, W689, P689 and Y810. From the multiple sequence alingnment it is very clear that Glysine are highly conserved in all APN. Experimental data shows that HEXXH residues are highly conserved and are Gluzine binding motif. Residues from 487 to 491 are highly conserved are they can be considerd as main active site for aminopeptiudase N. In human it is bestatin binding site.

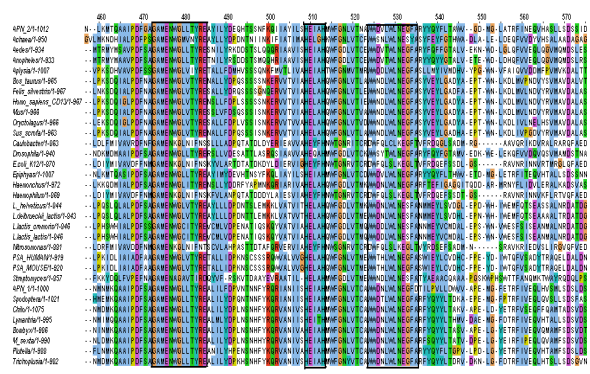

Fig 1: Showing multiple sequence analysis of different Aminopeptidase-N.

Alignment Position	Residue	Percentage	Alignment Position	Residue	Percentage
310	G	100	527	H	100
312	Y H	96 4	529	W	100
330	T S	93 7	530	F T S	

Volume : 2 | Issue : 12 | December 2013 • ISSN No 2277 - 8179

331	Q	96	531	G	100
337	A F P	84 13 3	533	L R Y	81 14 5
338	R	100	534	V I	96 4
341	F T V	84 12 4	535	T	100
342	P Y A	84 12 4	538	W D	87 13
343	C S Y F	75 18 5 2	539	W	100
345	D E	96 4	542	L T	90 10
346	E R	84 16	543	W S	87
347	P	100	544	L	90
350	K L M	84 12 4	545	N K	87 13
351	A S	87 13	546	E	100
353	F Y W	87 9 4	547	G S	78 22
370	S A	87 13	548	F L	87 13
371	N C	96 4	549	A T D	84 12 4
393	F W Y	81 13 6	554	Y Q S V	75 15 5 5
396	T P	87 13	585	D Q	96 4
400	S P	84 16	591	H Q N G	75 20 3 2
402	Y	100	618	Y	100
403	L V	96 4	620	K R	87 13
405	A	100	621	G	100
447	L M Y	78 18 4	622	A S	81 19
465	Y F	90 10	626	R V T	78 20 2
480	A F G S	87 5 4 4	627	N Q	96 4
482	P D A R	81 10 6 3	632	L I M V	75 15 5 5
483	D F E	90 8 2	640	G A S	75 20 5
484	F	100	644	Y	100
486	A M S F	78 12 10 10	658	L F	84 16
487	G	100	662	L M W	75 20 5
488	A	100	689	W F	96 4
489	M	100	692	Q K T	75 20 2
490	E	100	694	G	100
491	N	100	696	P N	96 4
492	W K	84 16	715	Q V L I	84 10 3 3
493	G	100	740	W L Y	78 18 4
494	L C A	87 8 5	743	P	100
496	T I	81 19	799	N L	87 13
497	Y F I	84 10 6	810	Y T L H	90 4 3 3
498	R N	87 13	811	D S K	75 8 7
499	E S T	84 10 6	837	R Q A	75 18 7
503	L M F	90 5 5	980	F Y E D	75 10 10 5
505	D H R N	78 8 8 6	1009	V Q S	75 13 12
523	H	100	1116	Q A	81 19
524	E	100	1159	G Q A S	84 10 3 3
526	A T F	81 16 3			

Table 1. A list of conserved amino acids is presented

Conserved cysteines residues:

Cysteine amino acid are one of the important residue for marinating the structural integrity of any protein as it forms disulfide bridges that help to keep the molecule intact and to maintain the confirmation of elements of the active site. All analyzed 33 sequences Cysteine residues are aligning in 9 places but they are conserved only in six places all other alignment positions they are highly variable. Manually we have selected those cysteines which are having conservation more than 50% (Table: 2). In multiple sequence alignment Cystein residues are 59 comes in alignment and conservation percentage lies between 2 to 50% and at six positions it les more than 50%. At alignment position 343 cysteine is highly conserved *i.e.* 75% while in alignment position 1084 conservation of cystein is 60% and this position is highly variable.

S.No	Alignment Position	Percentage	Substituents
1.	343	75%	F, S, Y
2.	536	51%	M, L, I, V
3.	965	63%	L, Q, G, N, A
4.	972	60%	S, A, T, H
5.	1011	60%	I, N, L
6.	1084	60%	H, N, A, D, S, K, L, W, A, R

Table 2: Showing Conservation percentage of Cystein residues.

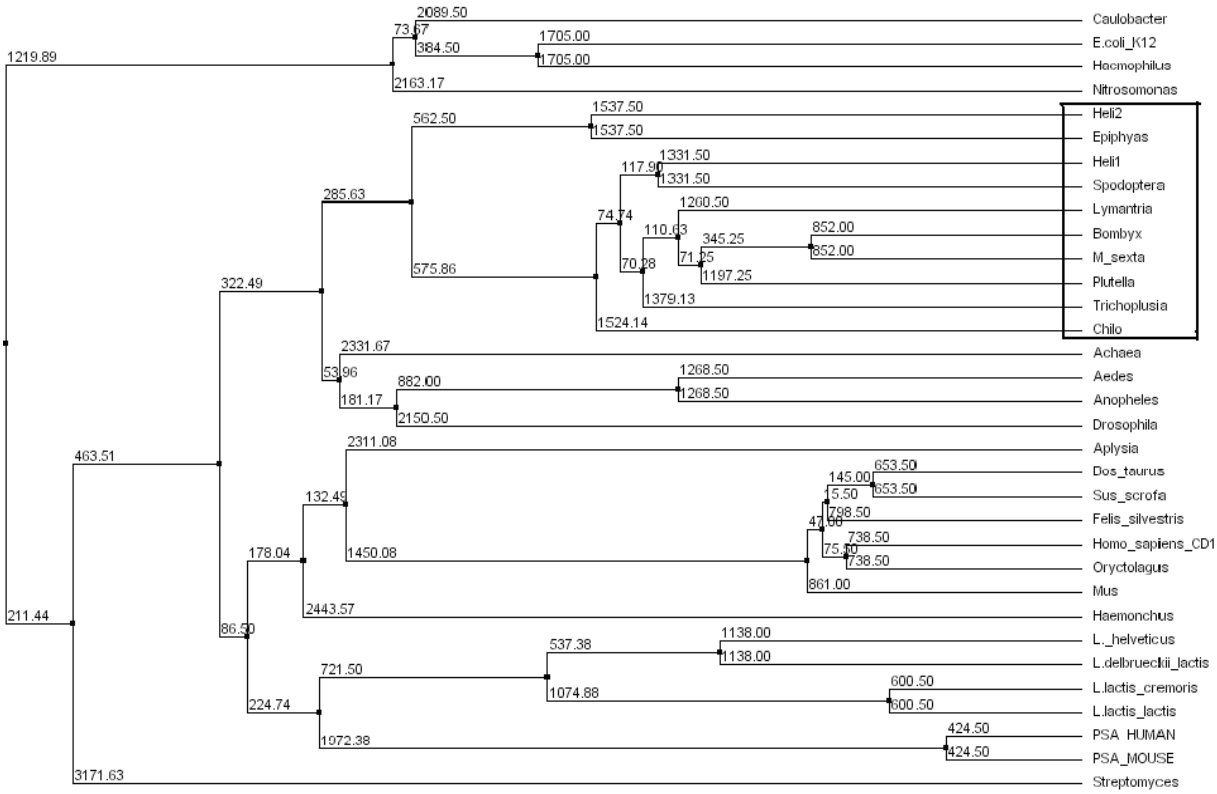

Fig 1: Thirty three Aminopeptidase-N sequensces were used to generte Average distance tree (BLOSUM62) based on phylogeny and it was found that insect APNs are present in two major clusters, which are having Lepidopterans and Dipterans

Conserved Motif: From the MSA it is very clear that residue GAMENWGLXTYRE (487-491) and HEXXH (523-527) are highly conserved and they may act as active site or binding site. From the MSA it is very clear that residue GAMENWGLXTYRE (487-491), HEXXH (523-527) and AWWDNLWLNEGF are highly conserved and they may act as active or binding sites. In both regions amino acids are almost 100% conserved. Amino acid sequence analysis shows that APN is having two most conserved regions one is HEXXH in which H, E, H residues are 100% conserved while other residue shows 85-100% conservation. Second site is GAMENWGLRE in which all are 100% conserved. In MSA of APNs residues from amino acid WWDNLWLNEGFA motif is highly conserved and may act as receptor binding site.

Discussion: Multiple alignment results of different APN sequence showed highest similarity between the Lepidopterus APN sequences. Phylogenetic tree studies also supported the above said facts by grouping Lepidopterus APN sequences into a single cluster (Raj Bhatnagar *et al.,* 2004). In this study multiple alignments showed three major conserved sites in almost all the aligned sequences (Dilip Chandu *et al.,* 2003, Wetterholm, A., *et al.,* 1992). Where HEXXH motif is also predicted, which is responsible for peptidase activity (Wetterholm, A., *et al.,* 1992, Blomster, M., *et al.,* 1995, Rudberg, P. C., *et al.,* 2004). Second most conserved site was predicted as GAMEN motif, previous studies supports that Glutamine is responsible for substrate binding. Mutational studies indicates that this motif is actually responsible for metallopeptidase activity. Many workers have determined hydrolysis and Catalytic activities of APNs. Above said site GAMEN found as Bastatin binding site in APN of *E. coli* (Ito Kiyoshi, 2006).

Third motif WWDNL found as a peptide binding site in human APN and it was found that this motif is also important in tumor angiogenesis and binding with peptides (R. Pasqualini *et al.,* 1995). Hence, we have accepted this site for Cry toxin protein binding region on APN. All the above said three motif are found to be conserved in APN of *M.sexta* and *H.armigera* the similar sites were predicted by Conseq and Consurf servers. Highly conserved residues were represented in the table with level of conservation (Table 1) Cysteine residues are important in the structural integrity of proteins by forming disulphide bridges, which in turn helps to keep the molecules intact and to maintain the conformation of the active sites. Number of Cysteine residues were also mentioned with percentage in Table 2 , which helps in enhancing the stability of the proteins (**Pier Luigi Martelli** *et al.,* **2002 & Rebecca S. Myers** *et al.,* 2005).

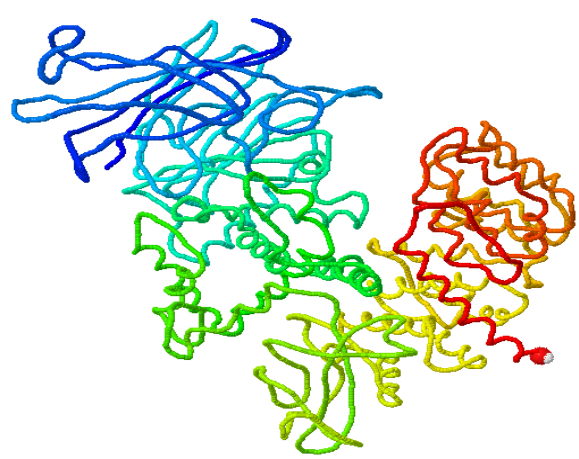

Fig 3: APN 3D structure

There are four domain present in APN 3D structure (Anamika Singh and C. V. S. Siva Prasad, 2009). As per literature and from our predictions it is clear that conserved functional sites are almost present in domain II region, highly conserved GAMEN motif is present on the sheet-16 (S16) which may have functions as ligand binding site in APN of *E.coli*. Zinc metal binding site HEXXH is found on helix seven (H7), whereas third motif WWDNL is considered as peptide binding site, hence we con-

sidered this site as one of the Cry1A toxin binding site, which is present on helix eight (H8). The above said three sites are highly conserved in APN of insects, mammals and also in some bacteria. These sites were also observed in APN of *M. sexta* and *S. littura* (Lopez Pazos, S. A. & J. A. 2008). In case of Lepidopterous insects not only the above said functional sites, there are few more oligosaccharide binding conserved sites are present. In our predictions these sites were found in the multiple sequence alignments with other APN sequences and also experimentally these four oligo binding motifs were determined in *M. sexta* by using NMR technique (Stephen *et al.,* 2004). In fact the similar conserved sites were predicted on APN of *H. armigera* but we do not have any experimental evidence till now which oligosaccharides are binding with APN of *H. armigera*. Essentially the above said information clearly points out that there are chances that in *H. armigera* APN binding with oligosaccharides but they may be explored in due course of time. Post translational modifications with oligosaccarides is an essential phenomenon in many APN receptors of mammals, insects and also in plants (Anamika Singh and C. V. S. Siva Prasad, 2009). Which helps to attain the final protein folding (Richard M. Twyman *et al.,* 2003).

Conclusion:

Insect pests are major cause of damage to the commercially important agricultural crops of the world. Indiscriminate use of synthetic pesticides resulted in sever insect resistance and causing irreversible damage to the environment. *Bacillus thuringiensis* (Bt) emerged as a valuable biological alternative in pest control. However, insect resistance against Bt has been frequently reported. One of the insects resistance development mechanisms is due to reduction in the binding of toxins to gut receptors. Among different Bt toxins, Cry toxins were extensively used in Bt transgenic plant and in many cases insect showed resistance against Cry toxins. Therefore, it is very important to study the mechanism of Cry toxins binding to the Aminopeptidase-N (APN) and Cadherin like insect gut receptors. In this process we have studies the detail seqential information of APN by which new mutations can be introduced in the APN. Based on this study we can proposed some novel mutations and developed important genes experimentally based on our proposed mutations.

The sequence and 3D structure studies were used to predict the active sites, which will give a new insight for the interactions.

REFERENCE

Valaitis A.P, Jenkins J.L, Lee M.K, Dean D.H, Garner K.J. 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. | | Milan Chromek, Zuzana Slamová, Peter Bergman, László Kovács, L'udmila Podracká Ingrid Ehrén, Tomas Hökfelt, Gudmundur H Gudmundsson, Richard L Gallo, Birgitta Agerberth & Annelie Brauner. 2006. The antimicrobial peptide cathelicidin protects the urinary tract against invasive bacterial infection. Nature Medicine - 12, 636 – 641. | | Blomster, M., Wetterholm, A., Mueller, M. J., and Haeggström, J. Z., 1995. Evidence for a catalytic role of tyrosine 383 in the peptidase reaction of leukotriene A4 hydrolase. Eur. J. Biochem. 231, 528–534. | | Reguera J, Santiago C, Mudgal G, Ordoño D, Enjuanes L, et al. (2012) Structural Bases of Coronavirus Attachment to Host Aminopeptidase N and Its Inhibition by Neutralizing Antibodies. PLoS Pathogens 8(8) | | | Dilip Chandu and Dipankar Nandi. PepN is the major aminopeptidase inEscherichia coli: insights on substrate specificity and role during sodium-salicylate-induced stress Microbiology December 2003 vol. 149 no. 12 3437-3447 | | | Dilip Chandu, Anujith Kumar and Dipankar Nandi. PepN, the Major Suc-LLVY-AMC-hydrolyzing Enzyme inEscherichia coli, Displays Functional Similarity with Downstream Processing Enzymes in Archaea and Eukarya. IMPLICATIONS IN CYTOSOLIC PROTEIN DEGRADATION. 2003 The Journal of Biological Chemistry, 278, 5548-5556. | , | | Andreas F. Kolb, Annette Hegyi and Stuart G. Siddell. Identification of residues critical for the human coronavirus 229E receptor function of human aminopeptidase N. Journal of General Virology (1997), 78, 2795–2802 | | | Takesue Y, Yokota K, Miyajima S, Taguchi R, Ikezawa H., 1989. Membrane anchors of alkaline phosphatase and trehalase associated with the plasma membrane of larval midgut epithelial cells of the silkworm, Bombyx mori.J Biochem. 105(6):998-1001. | | Garczynski, S. F. and M. J. Adang. 1995. Bacillus thuringiensis CryIA(c) -endotoxin binding aminopeptidase in the Manduca sexta midgut has a glycosyl-phosphatidylinositol anchor. Insect Biochem. Mol. Biol. 25:409-415. | | Agarwal, N., Malhotra, P. and Bhatanagar, R. K., 2002. Introduction of Gene-Cloned and Insect Cell-Expressed Aminopeptidase N of Spodoptera litura with Insecticidal Crystal Protein Cry1C. Applied and Environmental Microbiology, 68(9): 4583-4592. | | Luo, K., Y. J. Lu, and M. J. Adang. 1996. A 106-kDa form of aminopeptidase is a receptor for Bacillus thuringiensis Cry1C -endotoxin in the brush border membrane of Manduca sexta. Insect Biochem. Mol. Biol. 26:783-791. | | http://www.expasy.ch/PROSITE/ | | http://ibivu.cs.vu.nl/programs/popswww/ | | http://consurf.tau.ac.il/ | | http://conseq.tau.ac.il/ | | Rajagopal., Neema Agarwal, Angamuthu Selvapanditani, S. Sivakumar, Suhail Ahmad and Raj K. Bhatnagar, 2003. Recombinantly expressed isoenzymic aminopeptidases from Helicoverpa armigera (American cotton bollworm) midgut display differential interaction with closely related Bacillus thuringiensis insecticidal proteins., Biochem. J.370, 971±978. | | Wetterholm, A., Medina, J. F., Radmark, O., Shapiro, R., Haeggström, J. Z., Vallee, B. L., and Samuelsson, B., 1992. Leukotriene A4 hydrolase: abrogation of the peptidase activity by mutation of glutamic acid-296. Proc. Natl. Acad. Sci. U. S. A. 89, 9141–9145. | | Rudberg, P. C., Tholander, F., Andberg, M., Thunnissen, M. M., and Haeggström, J. Z. 2004. Leukotriene A4 Hydrolase. Identification of a comman carbohxylate recognition site for epoxide hydrolase and Aminopeptidase substrates. J. Biol. Chem. 279, 27376–27382. | | Kiyoshi Ito, Yoshitaka Nakajima, Yuko Onohara, Masahide Takeo, Kanako Nakashima, Futoshi Matsubara, Takashi Ito, and Tadashi Yoshimoto, 2006. Crystal Structure of Aminopeptidase N (Proteobacteria Alanyl Aminopeptidase) from Escherichia coli and Conformational Change of Methionine 260 Involved in Substrate Recognition. J. Biol. Chem., 281(44): 33664-33676. | | Pasqualini, R, E Koivunen and E Ruoslahti, 1995. A peptide isolated from phage display libraries is a structural and functional mimic of an RGD-binding site on integrins. The Journal of Cell Biology, 130:1189-1196. | | Pier Luigi Martelli, Piero Fariselli, Luca Malaguti and Rita Casadio, 2002. Prediction of the disulfide bonding state of cysteines in proteins with hidden neural networks. Protein Engineering, 15(12):951-953. | | Rebecca S. Myers, Rommie E. Amaro, V. Jo Davisson and Zaida A. 2005. Luthey-Schulten. Structural Elements in IGP Synthase Exclude Water to Optimize Ammonia Transfer. Biophysical Journal, 89(1):475-487. | | | Anamika Singh And C. V. S. Siva Prasad, 2009. Functional interpretation of APN receptor from M.sexta using a molecular model, Bioinformation 8, 321. | | López Pazos, S. A. & J. A., 2008. Cerón Salamanca: Minireview and hypothesis: homology modelling of Spodoptera litura (Lepidoptera: Noctuidae) aminopeptidase N receptor . Rev. Acad. Colomb. Cienc., 32(123): 139-144. | | | Richard M. Twyman, Stefan Schillberg and Rainer FischerMolecular farming of antibodies in plants Improvement of Crop Plants for Industrial End Uses. 435-469 978-1-4020-5485-3 | | Elaine Stephens, Jane Sugars, Sarah L. Maslen, Dudley H. Williams, Len C. Packman and David J. Ellar, 2004. The N-linked oligosaccharides of aminopeptidase N from Manduca sexta Site localization and identification of novel N-glycan structures. Eur. J. Biochem. 271, 4241–4258. |