Study and comparison of mosquito (Diptera) aminopeptidase N protein with other order of insects

Renu Jakhar and SK Gakhar

Abstract
The Aminopeptidase N group of exopeptidases are abundant proteins on the midgut brush border of mosquitoes. It belongs to a group of membrane-bound zinc metalloproteases. It is a ubiquitous enzyme that is found in a wide range of organisms from insects to mammals. Studies carried out on mosquitoes APNs are proteolytic enzymes involved in a wide range of functions, including digestion, involved in defense responses receptors for Bacillus thuringiensis (Bt) Cry1 insecticidal toxins and very important to play a role in parasite-vector interactions. Mosquito based Malaria Transmission Blocking Vaccines (TBV) target midgut surface antigens of Anopheles mosquitoes, the obligate vector of malaria parasite. Aminopeptidase N 1 is one of the most potent immunogenic protein involved in parasite infection and development in mosquitoes.

Keywords: Aminopeptidase N, mosquitoes, zinc metalloproteases, transmission blocking vaccines

Introduction
The insect digestive system comprises of a variety of digestive enzymes which are expressed over a period of time after ingestion of food. The majority of them are proteolytic enzymes those are secreted during the course of digestion. They are categorized as endopeptidases (Trypsin, Chymotrypsin) and exopeptidases (Aminopeptidase and Carboxypeptidase) based on they either cut the middle of a peptide or at the ends respectively [1]. Exopeptidases play an important role in digestion as they convert proteins in di and tri peptides. Although these exopeptidases are well described in mammals yet their knowledge in insects is limiting.

Exopeptidases are further classified into aminopeptidase and carboxypeptidase for their preference for N and C terminal residues respectively. The Exopeptidases activity is performed by metalloproteases which possess Zn divalent ion for the catalytic activity. The class metalloprotease of exopeptidase family is characterized by the presence of metal ion in the active centre of the enzyme [2, 3], and they are also in the largest number among proteolytic enzymes [4].

Activity and Specificity
Aminopeptidase N (APN) in the insect midgut divides N-terminal amino acids from peptides, an important step for co-transportation of amino acid into epithelial cells. The APNs varies in specificity for amino acids on the basis of digestive adaptability according to blood meal. The enzyme activity of recombinant near full-length S2-expressed AnAPN1 (amino acid 22-942) is catalytically active due to cleavage of the substrate L-leucine-p-nitroanilide and release of 4-nitroaniline, observed a optimum pH of 6.0-7.5. The metalloprotease inhibitor 1, 10-phenanthroline inactivated AnAPN1, but get substantially activated in the presence of Co2+, Mn2+, and Zn2+ cations. Using synthetic substrates, midgut aminopeptidase shows highest affinity for leucine (100%), medium affinity for alanine (76%), arginine (67%), methionine (55%) and lysine (49%), and a very low affinity (10%) for the remaining amino acids tested (tyrosine, serine, glutamic acid and proline) as reported by Graf and Briegel (1982) [5].

Horler and Briegel (1995) [6] reported an integrated aminopeptidase activity in midguts of sugar-fed Anopheles albimanus females. In blood-fed females, maximal aminopeptidase activity was reached at 24 h. similar constitutive and induced aminopeptidase activities were
reported in midguts of An. stephensi \[7, 8\], An. quadrinaculatus \[6\] and An. gambiae \[9\]. Although, after a blood meal activity increased by a factor of six within 24 h getting a maximum coincident with trypsin activity in the midgut of Aedes aegypti \[10\]. Amastatin, leuhiustin, and bestatin are inhibitors of various aminopeptidases in mammals. These inhibitors showed specificity for the aminopeptidase activity present in Ae. aegypti midgut extracts. Graf and Briegel (1989) \[11\], reported that almost 40% of the aminopeptidase activity was detected within the midgut epithelial cells of Ae. aegypti. The APNs differ in specificity for amino acids enabling digestive adaptability to host plants. The 120-kDa APN from Bombbyx mori (BmAPN1) has been reported to have highest activity for non-polar amino acid conjugates \[12\]. It shows 2.7 times higher activity with methionine p-nitroanilide as substrate than with leucine and alanine p-nitroanilide and hydrolyzes arginine and lysine p-nitroanilides at much slower rate. Hua et al., 1998 \[12\], has further added that BmAPN1 is less efficient in hydrolyzing proline, glycine and valine derivatives. The 120-kDa APN from Epiphyas postvittana (EpAPN1) prefers leucine amides, but also uses isoleucine, phenylalanine, tryptophan, histidine and valine amides. It does not show significant hydrolysis of serine, alanine, glycine, proline and arginine \[13\].

Many coleopteran midgut APNs have also been enzymatically characterized. Purified a 90-kDa APN from Tenebrio molitor has been determined for its substrate specificity and inhibition report \[14\]. The APN shows a broad specificity for N-terminal amino acids, but could not hydrolyze the acidic amino acid substrate, Asp-β-naphthylamide. T. molitor APN shows preferences towards tri- and tetrapeptides than towards di- and pentapeptides \[14\]. Amastatin being a strong inhibitor of APN than bestatin, supports enzyme’s preference for longer substrates. The coleopteran Morimus funereus has a 67-kDa membrane-bound APN that is solubilized from membrane as a 400-kDa oligomer \[15\]. They have also reported that oligomeric APN preferentially hydrolyzes hydrophobic amino acid para-nitroanilides.

The T. molitor 90-kDa APN is a glycoprotein with a C-terminal anchor. The N-terminus has a Trp-Arg-Leu-Pro motif typical of insect APNs. The enzyme has a pI of 7.4 and an optimum pH of 8.0. T. molitor APN shares the basic features of catalysis and substrate binding with mammalian APNs, yet differs in several amino acid residues participating in those events \[16\].

A 130-kD hemipteran Acycthisophon pisum APN is found in the midgut membrane \[17\]. The enzyme is more abundant being about 16% of total gut protein, than in lepidopteran APNs. It also has a broad specificity for N-terminal amino acid residues; but is more unique in having inability to hydrolyze acidic aminoacyl substrates \[17\]. Many Aminopeptidase inhibitors active against insect APNs are commercially available. Amastatin, protesbin and actinonin with IC50 values of 76, 74, and 148 μM respectively are most competent APN inhibitors against the 110-kDa APN of B. mori (also called BmAPN4) \[12\]. Manduca sexta aminopeptidase N1 (MsAPN1) activity is inhibited by actinonin, amastatin and bestatin, as by Zn21 chelating agents 2, 20-dipiridyl and 1, 10-phenanthrolime \[18\]. Aminopeptidase inhibitors being toxic to various insect species are considered potential insect control agents \[19\]. A novel synthetic APN inhibitors is synthesized in vitro with H. armigera gut extracts and in vivo via larval growth inhibition assays by Duncan et al. (2006) \[20\], on the basis of effectiveness of actiontin when given to Helicoverpa armigera.

### Structural Chemistry

Various APNs are present on the midgut brush border surface range in size from about 60- to 200-kDa. The pl of mosquitoes APNs is about 5.0. Mosquitoes brush border APNs are glycoproteins attached to the membrane by glycosphosphatidylinositol (GPI) anchors found on the midgut brush border vary in size from ~60- to 200-kDa. The pl of lepidopteran APNs is ~5.5 and their optimal pH varies from 7.5-8.5. Insect brush border APNs are glycoproteins anchored to the membrane by glycosphosphatidylinositol (GPI) anchors. The MsAPN1 protein has 990 amino acid residues pre-protein having four potential N-linked glycosylation sites (Asn-Xaa-Ser/Thr) \[21\]. The N-terminus contains a hydrophobic signal sequence. The C-terminus of APNs has the typical hydrophobic region and processing signal of a GPI-anchored protein. Purified 115-kDa MsAPN1 with a chopped GPI-anchor MsAPN1 is 4.1 mole % carbohydrate, has N-acetyl galactosamine, a moiety that is involved in Cry1Ac toxin binding and ~50% bound lipids \[22\]. Mass spectrometric analysis of MsAPN1 showed that three out of four predicted sites are N-glycosylated with highly fucosylated N-glycans and the fourth site has an attached paucimannosidic N-glycan \[23\]. MsAPN1 and other lepidopteran APNs near the mature C-terminus have abundant serine and threonine residues (potential O-glycosylation sites). MsAPN1 O-glycosylation sites (probable Bt Cry1Ac binding sites) have N-acetylgalactosamine (GalNAc)-peptide glycans \[24\]. A central Zn21 coordination site (His-Glu-Xaa-Xaa-His (18 residues)-Glu) \[25\] and the conserved glazucin aminopeptidase motif (Gly-Ala-Met-Glu-Asn-Try-Gly) are present in MsAPN1 and other lepidopteran APNs.

Immunoblot analysis of the supernatant and pellet fractions of phosphotidylinositol-specific phospholipase C (PI-PLC)-treated MMV solutions suggest the presence of a GPI anchor on An. gambiae APN1(AnAPN1) and that recognition of AnAPN1 is midgut-specific. AnAPN1 is comprised of an N-terminal signal peptide (residues 1-19) and C-terminal ectodomain (residues 22-993) that contains a putative mucin O-glycosylated region (resides 952-993). A glycosphosphatidylinositol (GPI)-anchor (residues 997-1020) resides at the C-terminus.

The tertiary structure of the AnAPN1 ectodomain (residues 57-942) exhibited the classical four-domain assembly of M1-family metalloendopeptidases, designated domains I-IV. The N-terminal domain I, visible from residue 57 (residues 57-270), comprised a 15-stranded β-barrel fold, acts as vaccine immunogen (residues 60-194). Domain II, or the catalytic domain (residues 271-523), is involved in substrate recognition and contains the G380AMEN motif common to many exopeptidases, in addition to the zinc-binding motifs HEXXH (H380EYAH390) and NEXFA (NE389GFA). This domain had a thermolysin-like fold and comprised a four-stranded β-sheet with connecting α-helix (helix α4) and an eight α-helical superhelix subdomain. Domain III (residues 524-613) adopted a β-sandwich architecture in which two separate anti-parallel β-sheets packed together. The C-terminal domain IV (residues 614-942) adopted the most variable conformation compared to other M1-family metalloendopeptidases and comprised a 17 α-helical super-helix. AnAPN1 was most similar to porcine APN (PDB ID
4FKE10) with 32% sequence identity and 48% sequence similarity. The first active site was flanked by domains II and IV and the second was located on domain IV. The surface representation of AnAPN1, with a large rift between domains II and IV, was most similar to porcine APN and the open conformation of ERAP.

In AnAPN1 protein the zinc ion was coordinated by Nε2 atoms of His566 and His570 from the H566EXXH570 motif, the carboxyl Oε of Glu569 of the NE569XFA motif and the N-terminus of the modeled peptide. The N-terminal residue of this peptide was anchored to the active site, coordinated to the highly conserved Glu333 of the GAMEN motif in addition to Glu156 and Glu339. This residue also formed hydrogen bonds with His566. The alanine at the second position of the peptide also formed contacts with residues that lined the electronegative cleft and formed hydrogen bonds with Gly330 and Ala331 from the GAMEN motif. The conformation of Tyr422 (equivalent to Tyr429 in ERAP1) is often used as a partial indicator of whether M1-family metalloenzymes are in an active or inactive state. Tyr422 pointed towards the active site in AnAPN1 and suggested an active enzyme conformation.

The quaternary structure (dimer) of AnAPN1 is present in aqueous solutions; two molecules from adjacent asymmetric units formed a putative AnAPN1 dimer, with contacts via domains III and IV. Several salt bridges and hydrogen bonds were formed between Asp339 (domain IV) and Arg551 (domain III), Lys555 (domain III) from each monomer formed hydrogen bonds with each other, as well as an extensive network of hydrophobic contacts. As opposed to the dimer arrangement in the asymmetric unit, this dimer allowed free movement of domain IV into an open and closed conformation and an unobstructed cleft for peptide entry and exit. AnAPN1 are rich in alanine and threonine residues [26].

Biological Aspects and tissue Expression time

APNs of Mosquitoes are attached to the midgut surface where they are preferentially localized in cholestere-rich membrane rafts [27]. Lipid rafts allow for the partitioning of multiple ookinete interacting proteins into discrete locations on the midgut surface, enhancing multivalent interactions between the ookinete and the midgut. Multivalent interactions have been suggested to be a conserved process for strengthening single, protein-protein, or protein-glycan interactions for several vector-borne pathogens, including Plasmodium. APNs of Lepidoptera are anchored to the midgut surface where they are preferentially confined in cholesteral-rich membrane rafts. A complex set of events that require lipid raft integrity, lateral mobility, and partitioning on midgut surfaces has been shown to be essential for Cry1A toxin insertion and pore formation in Heliothis virescens and Manduca sexta [28].

RT-PCR analysis of midgut mRNA suggests that AgAPN1-specific transcripts are constitutively expressed in midguts of sugar-fed and blood-fed mosquitoes. APN was significantly up-regulated following feeding with an infected compared to an uninfected but was quickly downregulated thereafter. AgAPN1 is present during the ookinete traversal of the midgut post blood-feeding (PBF), and antibodies against AgAPN1 clearly interfere with the ookinete-to-oocyst transition. In An. stephensi, aminopeptidase activity is higher in a strain selected for refractoriness to P. falciparum [29]. In a Plasmodium-resistant strain of An. gambiae (L35), a midgut aminopeptidase was up-regulated after a P. berghei-infected blood-meal, and no up-regulation was observed in the Plasmodium-susceptible G3 strain [30]. These results for Ae. aegypti and the cited reports for Anopheles spp. provide complementary evidence that APN induction is likely to be related to the mosquito immune response against malaria parasites.

Molecular characterization

While mosquito APNs from species of the genera Aedes, Culex and Anopheles are well-characterized at the sequence level due to the availability of genome sequence data on Vectorbase. Five APN gene APN1 [31], and APN2, APN3, APN4, and APNs [32] have been characterized in A. gambiae. The APN1 in An. gambiae was found located on chromosome 2R and contains four introns. The gene harbors an open reading frame of 3063 which codes for a protein of 1020 residues, with a predicted molecular mass of 113.2 kDa. Sequence analysis of the predicted protein suggests that it is first translated as a pre-pro-protein, with a signal sequence between amino acids 1–19. The presence of a putative signal sequence suggests that zymogen is secreted and activated after tryptic cleavage releasing the pro-domain. On Vectorbase 20 APNs are identified on different locations of chromosome. Angelluci et al., 2008 [33] have reported nine APNs in the midgut of lepidopteran larvae. Genomic organization analysis showed that four types of APN (APN2 to APN5) genes were localized on same chromosome, adjacent to each other in the genome [32]. Characterization of APN in An. stephensi revealed more or less similar results as previously described for An. gambiae. The An. stephensi APN1 was found 81% similar to An. gambiae APN1 [34]. The reported An. stephensi APN1 has 3102 ORF having four introns. The Ae. Aegypti APN1 contains three introns with an ORF of 2697 bp and a predicted protein of 112 kDa. Two isoforms of Aminopeptidase N was also identified in Ae. aegypti larval midgut [35].

Transmission blocking activity

The alanyl Aminopeptidase N is the leading malarial TBV immunogen [36]. A midgut specific protein, reported to play an important role in ookinete invasion of Plasmodium in the An. gambiae. The previous studies have shown that APN1 acts as a receptor. It is glycosyl phosphatidyl inositol anchored protein for attachment of Plasmodium parasite and later helps in sexual growth [37, 38, 31]. Rosenfeld and Vanderberg, 1999 [39] validated the potential use of APN for malaria transmission blocking in An. gambiae by recording a significantly reduced number of oocysts using rabbit polyclonal antibodies against APN in infected mice [39]. Identified potent epitope from AgAPN1 is highly immunogenic. Similar results were observed in An. stephensi [40], suggesting possible multispecies blocking strategies. Also, antibodies against APN1 responsible for 73% blockage in An. gambiae and 67% in An. stephensi [31]. Based on the crystal structure of the near-full length APN1, a structure-guided construct has been expressed in E. coli and revealed B cells epitopes as transmission-blocking antigens [26]. Studies on the APN1 has also been identified and characterized from An. stephensi as a candidate for Transmission Blocking Vaccine [34].

Cloning and Expression Systems

APN cDNAs have been expressed in E. coli and in insect cells, though the level of expression is variable. In general, it
has been observed that while truncated fragments of APN are readily expressed, full-length APN expression in E. coli is more difficult. Initial efforts to express APNs in cultured insect cells were challenging. However, since about 2000 there have been a number of reports of insect cultured cell expression of APNs using Sf21 cells and the baculovirus expression system. For example, Chen et al. 2009 [41], expressed Ae. aegypti APN1 yielding a protein that was catalytically active and suggests that Sf21 cells expressing AeAPN1 were not susceptible to Cry11Aa toxin. APN gene expression is developmentally regulated. In three out of four lepidopteran species, APN specific activity enhanced, as larvae grew from neonatal to the third instar stage [42]. Not all APNs are expressed at the same levels or during the same larval stages. In Helicoverpa punctigera, HpAPN1 mRNA expression increases from from the neonate stage to the third larval stage. HpAPN2 and HpAPN3 had less expression and expressed more evenly during larval development [43].

Differential expression of APN genes supports a model whereby each of the four APNs has a distinct role in digestion. The expression of a recombinant AnAPN1 (rAnAPN1) antigen (in E. coli), its formulation with Alhydrogel (Brenttag Biosector), its vaccine potency, immunogenicity and histopathology in a small-animal model was reported in AnAPN1 [39]. The transmission-blocking activity of antibodies isolated from rAnAPN1-immunized mice can generate functional transmission-blocking antibodies in the mice. Antibodies raised against AnAPN1 without adjuvant were found to have transmission-blocking activity, as indicated by an increased absence of oocysts (i.e., reduction in infected mosquito prevalence) or a reduction in oocyst intensity (i.e., the number of oocysts counted) in the midguts [40]. These studies have shown that when rabbit polyclonal antibodies were directed against the N-terminal portion of APN passively transferred to P. berghei (P. berghei) infected mice; they were able to significantly reduce the number of oocysts in both An. gambiae and An. stephensi [41].

Isolation and purification of APN

Insect midguts are dissected from larvae, frozen on dry ice and stored until needed. For smaller lepidopteran larvae such as Plutella xylostella, brush border membrane vesicles (BBMV) are prepared from whole larvae. BBMV are easily prepared by divalent cation precipitation methods resulting in 4 to 10-fold enrichment of brush border marker enzymes. Insect APNs are purified from BBMV. In Sangadala et al. (2001) [22], 200 g of M. sexta midguts yielded 300 mg of BBMV protein. BBMV are solubilised in 1% CHAPS. An endogenous phosphatidylinositol-specific phospholipase C (PIPLC) converts APN to the water-soluble form. Proteins are separated by anion exchange chromatography. Depending on the degree of purity desired, a second purification step is advised.

Sangadala et al. (2001) [22]. incorporated a second hydrophobic interaction chromatography step to yield 3 mg protein sufficiently pure for mass spectrometry analyses. Some APNs have unique sugar moieties allowing rapid purification by lectin-affinity chromatography. Gill et al. (1995) [44], separated solubilized BBMV proteins by anion-exchange chromatography followed by wheat germ agglutinin affinity chromatography to isolate the 120-kDa APN in Heliothis virescens. If the goal is to isolate APNs that bind Bt Cry toxins, then toxin affinity chromatography is a simple and direct approach. Cry toxins are coupled to cyanogen bromide-activated beads then incubated with CHAPS-solubilized BBMV. Following washes, bound proteins are eluted with sodium isothiocyanate [33, 45, 46].

Cry toxins receptor

Cry toxins in the form of Bt subsp. Israelensis formulations (VectoBac®, Valent BioScience Corp., Libertyville, USA) are used in several countries to control the An. gambiae, An. dirus, and Ae. aegypti mosquito population. As larvae of these insects are extremely susceptible to this toxin, it can be utilized to check the spread of mosquito-borne diseases. It was observed that digoxigenin-labeled Cry4A bind to the apical microvilli of the Culex pipens larval midgut suggesting the presence of receptor proteins [47]. Moreover, experiments of RNA-mediated interference [48], and transgenic technology [49], strongly indicated that receptor to the Cry toxins in other insect species is APN.

Several Cry toxin receptor APNs have been cloned from Aedes and Anopheles. 100-kDa APN in the An. quadriramaculatus was the first APN to be identified as a Bt Cry toxin receptor isolated from the BBMV [50]. This protein was used to screen different mosquitocidal Cry toxins binding discovered that only Cry11Ba was able to bind the APN. Another protein BLAST search was performed using the sequence of a known conserved region for aminopeptidases, which yielded 16 homologous proteins from the An. gambiae genomic database. This implicated that in mosquitoes, a huge number of aminopeptidase isomers are present. In An. gambiae a 106-kDa APN, called AgAPN2, identified as a putative midgut receptor for Cry11Ba toxin. AgAPN2 cDNA was cloned and predicted AgAPN2 protein analysis revealed a zinc-binding motif (HEIAH), three potential N-glycosylation sites and a predicted glycosylphosphatidylinositol (GPI) anchor site localized to the microvilli of the posterior midgut. This AgAPN2 found homologus to100-kDa APN in the An. quadriramaculatus and paralogous to three genes AgAPN3, AgAPN4 and AgAPN5 located adjacent AgAPN2 on 2L chromosome, shows 99% amino acid identity to AgAPN4, 95% to AgAPN3 and 56% to AgAPN5. Ae. aegypti 112 Kda AeAPN1 protein identified as a Cry11Aa toxin binding protein, show less toxicity than Cry11Ba. The well-studied Cry11Ba toxin binds APNs in An. quadriramaculatus via GalNAc moieties (GPI anchored) whereas AeAPN1-Cry11Aa interaction is glycan-independent and one of the Cry11Aa binding regions is localized to amino acids 525–778 of APN1. Nine distinct classes of APNs were identified by phylogenetic analysis of which AgAPN2 and its paralogues belong to the novel class 6 and Class 7 includes AgAPN1 and an AeAPN (AeAPNRc2) that shows 37% identity to AgAPN1. The ability to bind Bt Cry toxins is the novel feature of lepidopteran and dipteran APNs and serves as adventitious toxin receptors. The well-studied Cry1Ac toxin binds APNs in B. mori, M. sexta, H. virescens, H. armigera and probably other lepidopteran species via GalNAc moieties. Other Cry1A toxins bind amino acid regions on specific APNs. For example, in B. mori each APN isoform has a common Cry1Aa binding region [51] and Spodoptera litura with Insecticidal Crystal Protein Cry1C was performed [52].

References

1. Barrett AJ, Woessner JF, Rawlings ND. Handbook of


14. Simpson RM, Newcomb RD. Binding of Bacillus thuringiensis δ-endotoxins Cry1Ac and Cry1Ba to a 120-kDa aminopeptidase-N of Epiphyas postvittana purified from both brush border membrane vesicles and baculovirus-infected Sf9 cells. Insect biochemistry and molecular biology. 2000; 30(11):1069-78.

15. Cristofoletti PT, Terra WR. Specificity, anchoring, and subsites in the active center of a microvillar aminopeptidase purified from Tenebrio molitor (Coleoptera) midgut cells. Insect biochemistry and molecular biology. 1999; 29(9):807-19.


29. Zhuang M, Oltean DI, Gómez I, Pullikuth AK, Soberón


42. Chen J, Aimanova KG, Pan S, Gill SS. Identification and characterization of Aedes aegypti aminopeptidase N as a putative receptor of Bacillus thuringiensis Cry11A toxin.