STRUCTURE AND FUNCTION OF AMINOPEPTIDASE N

Hans Sjöström, Ove Norén and Jørgen Olsen
Department of Medical Biochemistry and Genetics, Biochemistry Laboratory C, The Panum Institute, University of Copenhagen, Copenhagen, Denmark

1. INTRODUCTION

Aminopeptidase N (EC 3.4.11.2) is an enzyme preferentially releasing neutral amino acids from the N- terminal end of oligopeptides. It is a type II membrane glycoprotein with a zinc-dependent catalytic activity and is expressed in many tissues, being most abundant in the enterocytes of the small intestine and in the epithelium of kidney proximal tubules. For a recent review see Norén, O. *et al* (1997). The enzyme has over the years been known under several different names e.g. microsomal aminopeptidase, aminopeptidase M, particle-bound aminopeptidase, p 146, p161 and gp 150. The nomenclature committee of IUBMB (http://www.chem.qmw.ac.uk/iubmb/enzyme/) recommends the name membrane alanyl aminopeptidase. Knowledge of the primary structure of aminopeptidase N (Olsen *et al* 1988) led Look *et al* (1989) to identify CD 13 as the same protein. Furthermore aminopeptidase N serves as a receptor for coronaviruses belonging to a certain genetic subset (Delmas *et al* 1992, Yeager *et al* 1992). It is the aim of the present paper to summarise our present knowledge on the structure of aminopeptidase N and to integrate it into functional data giving a view of structure-function relationships of this ubiquitous enzyme.

2. CLASSIFICATION

Aminopeptidase N belongs to the M1 family of the MA clan of peptidases (Rawlings and Barrett 1999) also called gluzincins (Hooper 1994). The M1 family has 20 different entries. Four of them (aminopeptidase N, cystinyl aminopeptidase, EC 3.4.11.3, glutamyl aminopeptidase, EC

Cellular Peptidases in Immune Functions and Diseases 2
Edited by Langner and Ansorge, Kluwer Academic/Plenum Publishers, 2000

3.4.11.7 and pyroglutamyl-peptidase II, EC 3.4.19.6) are membrane-bound type II glycoproteins. Aminopeptidase N has been cloned and sequenced from six different mammalian species (table 1) and from chicken egg yolk (called aminopeptidase Ey). Interestingly an insect aminopeptidase N, which has been cloned and characterised is inserted in the membrane via a C-terminal GPI anchor. There is no known three-dimensional structure in the M1 family but it has been described for thermolysin in the M4 family of the same clan.

Table 1. Cloned and sequenced aminopeptidase N from different species and tissues

Species	Tissue/cell type	Reference
Human	Caco-2 cells	Olsen *et al*, 1988
	Myeloid cells	Look *et al* ,1989
Rat	Kidney	Malfroy *et al*, 1989
		Watt and Yip, 1989
Rabbit	Kidney/intestine	Yang *et al*, 1993
Pig	Small intestine	Delmas *et al*, 1994
Cat	Fcwf-cells	Tresnan *et al*, 1996
	CRFK-cells	Kolb *et al*, 1997
Mouse	CFTL-cells	Chen *et al*, 1996
Chicken	Liver	Midorikawa *et al*, 1998

3. STRUCTURAL INFORMATIONS

The following structural discussion is based on electron microscopic studies on the purified enzyme incorporated into liposomes, knowledge of the exon-intron organisation of the gene, computer-aided predictions of the secondary structure and comparison of the primary structure with the same enzyme of different species or similar enzymes from the same or different species (see Fig 1).

3.1 Electron microscopy

The overall structure of the purified pig intestinal aminopeptidase N was studied by Hussain *et al* (1981) after incorporation into lipid membranes. Electron microscopic observations demonstrated a dimeric symmetrical structure of aminopeptidase N with the dimensions 13,5 x 5,5 nm separated by a 5 nm gap from the membrane. Biochemical studies showed that the catalytically active part was located outside the liposomes and the anchoring peptide associated with the membrane.

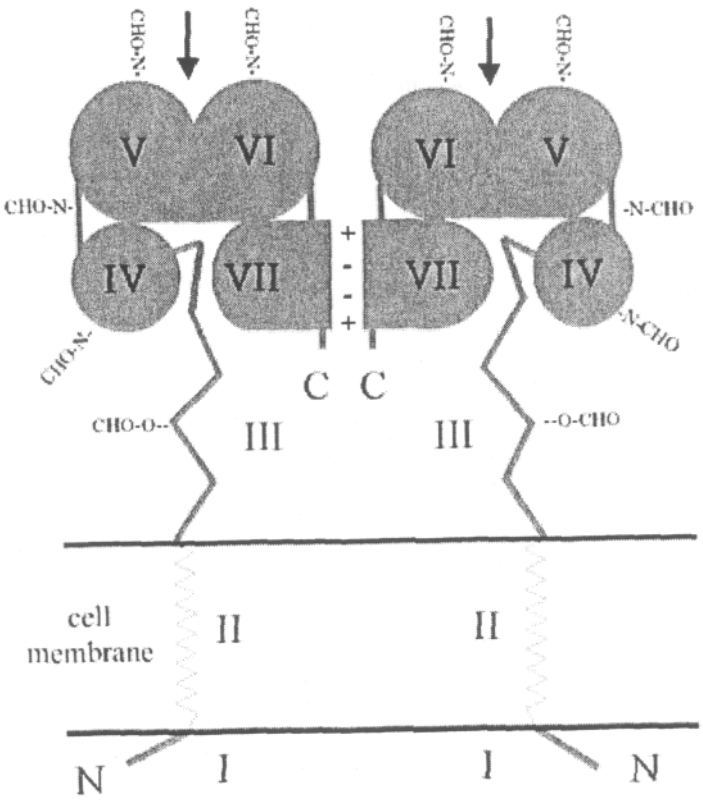

Figure 1. Hypothetical model of aminopeptidase N with different domains (I – VII), N- and C-terminus (N, C), active site (arrow), O-glycosylation (CHO-O--), N-glycosylation (CHO-N-) and dimerisation area (+--+) indicated.

3.2 Exon-intron structure

The exon-intron structure of the aminopeptidase N gene was mapped by Lerche *et al* (1996). It was originally suggested that exons define different structural domains (Gilbert 1987). This suggestion may not be true (Stoltzfus *et al* 1994), as introns instead may have spread in the genomes rather late in the evolution. However, there is evidence that the occurrence of introns has allowed exons to be shuffled between genes later in the evolution and thereby contribute to the generation of new genes. The aminopeptidase N gene consists of 20 exons encoding segments differing in size from 18 to 205 amino acids.

3.3 Secondary structure predictions

For secondary structure predictions the PredictProtein service (Rost *et al* Internet server was used. This program uses a neural network protocol with a stated accuracy of more than 70 %. The algorithm uses comparison of the target sequence with evolutionary related sequences. Proteins from the SwissProtein database were used and some proteins not occurring here were added manually. In all structure comparisons the numbering of human aminopeptidase N (Olsen *et al* 1988) is used.

4. DOMAIN STRUCTURE

The considerations mentioned above and further detailed in the following sections led us to suggest a domain structure of seven different domains. Part of this domain model fits with the domain structure of porcine aminopeptidase A, including the generation of a 45 kDa C-terminal fragment, suggested by Hesp and Hooper (1997). Bauer (1994) also demonstrated that trypsin released a fragment from pyroglutamyl-peptidase II, indicating that there is a similar model for all type II membrane bound aminopeptidases.

4.1 Domain I

Domain I is the cytosolic part of aminopeptidase N. The size of this domain varies a lot between the four mammalian type II aminopeptidases. Aminopeptidase N contains seven residues, glutamyl aminopeptidase has 17 residues, the rat pyroglutamyl-peptidase II 38 residues, whereas this domain of cystinyl aminopeptidase comprises 137 residues. Domain I of cystinyl aminopeptidase has been suggested to contain a signal for correct intracellular transport in the form of two di-leucine motives (Keller *et al* 1995). It has not been possible to ascribe any function to this domain of aminopeptidase N, although the enzyme is known to cluster in coated pits in connection with TGEV infection. As this clustering also occurs using an aminopeptidase N devoid of domain I (Hansen *et al* 1998), it is tempting to suggest that aminopeptidase N signals over the membrane via an auxiliary membrane protein.

4.2 Domain II

Domain II is the membrane-spanning domain probably existing as an α-helix. The domain is well conserved within the aminopeptidase Ns, but

differs a lot from the other three type II aminopeptidases. It may be that the amino acid sequence of this domain determines the Triton X-100 insolubility of aminopeptidase N at 4 °C (Danielsen 1995). This property has been suggested to form the basis for selective apical transport of newly synthesised aminopeptidase N from the trans-Golgi network.

4.3 Domain III

Comparisons between type II membrane aminopeptidases on one side and cytosolic and GPI anchored aminopeptidases of the Ml type on the other side allow a suggestion of a stalk region between residue 40 and residue 70. After this residue there in aminopeptidase N is a sequence WNXXRLP, with homologies also to non-membrane bound aminopeptidases. Aminopeptidase N from human seminal plasma, and normal and pregnancy serum forms have residues 48 and 59 and 69 as N-terminals (Huang *et al* 1997, Watanabe *et al* 1995, 1998). As exon 1 encodes the first 205 amino acids it can be excluded that these forms are generated by an alternative splicing mechanism. Studies on the release of other membrane proteins have suggested that there are several membrane-bound enzymes, which release ectoenzymes. Cleavage between position 68 and 69 demonstrates that this residue is localised within an exposed part of the stalk. The divergence of amino acid composition even within aminopeptidase Ns from different species is striking. A common feature, however, is the frequent occurrence of prolines. It seems reasonable to suggest that a stalk has an ordered structure, even if the secondary structure predictions suggest an unordered structure. The presence of prolines excludes an α-helix but is suggestive of a polyproline II-helix structure, having 3 amino acids per turn.

4.4 Domain IV

Domain IV comprises amino acid 70 to 252. The domain corresponds to most of exon 1 and all of exon 2. The end is defined by the start of a domain corresponding to thermolysin. The secondary structure prediction suggests an unordered structure with minor stretches of β-strands. Residues 216-227 represent a conserved region with similarities also to soluble mammalian and bacterial aminopeptidases. This could indicate a function of this region connected to the aminopeptidase enzymatic activity. However mutagenesis experiments with a conserved aspartic acid (residue 225) has shown that this residue is not critically involved in the hydrolytic mechanism (Luciani *et al* 1998). One of the two conserved N-glycosylation sites occurs in this domain (see later).

4.5 Domain V and VI

Domain V and VI are distinguished by a thermolysin comparison using the three zinc ligands as fixpoints and correspond to the two spherical domains of thermolysin connected by a loop between helix 2 and 3 of thermolysin. It includes amino acids 253 to 580 corresponding to exons 3 to 10.

Residue 576 is an arginine followed by proline in all examined aminopeptidase N structures except the pig where reside 577 is serine. This allows the bond to be cleaved by trypsin, a proteolysis that has been shown to occur *in vivo* in the pig, providing further evidence that this region is a linkage between two domains (VI and VII) (Sjöström *et al* 1978).

It might be reasonable to substantiate the possible structure similarity between the structure of thermolysin and the secondary structure prediction of APN in domain V and VI. First a comparison between the known three-dimensional structure of thermolysin, having seven α-helices, with the predicted one showed a rather high concordance with some evident mistakes. The first two α-helices are not predicted and three short predicted α-structures (129-132, 155-159 and 230-233) are not found in thermolysin. The four C-terminal α-helices are predicted with high reliability.

In aminopeptidase N the region around the HEXXH motif is not predicted as an α-helix, but this is also the case for thermolysin. It seems therefore reasonable to suggest that the region from positions 382 to 400 is an α-helix, which corresponds to helix 2 in thermolysin. The third zinc ligand (residue 411) is contained in a region, which for both enzymes is suggested to be α-helical. On the N-terminal side residues 315 to 332 in aminopeptidase N are located in a position similar to helix 1 in thermolysin. Two predicted β-strands in aminopeptidase N have counterparts in thermolysin, whereas the predicted α-helix in position 273-290 cannot be found in thermolysin. C-terminally to the active site there are four α-helices in thermolysin and four predicted α-helices in aminopeptidase N (481-490, 493-507, 515-524 and 536-546). Whereas the position of the first α-helix correlates well with that of thermolysin, the other three helices are positioned 12 residues closer to the N-terminus than they are in thermolysin. It is suggested that 12 amino acids are replaced by a short loop between helix 4 and helix 5 as the end of helix 4 and the start of helix 5 are located very close to each other in thermolysin.

Mutation of glutamic acid 355 in an aminopeptidase conserved region (the GAMEN motif) led to an almost completely inactive enzyme (Luciani *et al* 1998) Further characterisation indicated that this glutamic acid belongs to the anionic binding site in aminopeptidase N thereby specifying its aminopeptidase activity. Interestingly, using the above structural

comparison, this glutamic acid, (not occurring in thermolysin, which has no aminopeptidase activity) is located in the opening of the active site cleft.

4.6 Domain VII

Domain VII constitutes the remaining C-terminal part of the enzyme and includes amino acids 581 to 967 (exons 11-20). This domain has a very high content of predicted α-helices. In the pig enzyme the domain can be dissociated from the rest of the enzyme under non-reducing conditions (Sjöström and Norén 1982). Thus there are no disulfide bridges between domain VII and the rest of the enzyme, in spite of the fact that this domain harbours four of the seven cysteine residues. Of these four residues only residue 761 is conserved in all four type II aminopeptidases. It is therefore difficult to predict which (if any) of the four cysteins that are connected via disulfide linkages.

The final part of aminopeptidase N (residue 936 to 964) is predicted to constitute an extraordinarily long α-helix. It may be hard to accommodate an α-helix of this size within the domain VII structure and it is therefore suggested to be located outside the main part. This segment is furthermore characterised by a high amount of charged amino acids in a regular pattern with most of them on the same side of the putative helix. These properties, together with the high conservation within aminopeptidase N from different species suggest a distinct biological function.

5. POST-TRANSLATIONAL MODIFICATIONS

5.1 N-glycosylation

Pig aminopeptidase N is heavily N-glycosylated corresponding to a molecular weight of about 25 kDa (Danielsen *et al* 1983). As judged from comparisons of apparent molecular weights in SDS gel electrophoresis with number of amino acids, most of the carbohydrates are bound to residues between number 1 and 580. Only two of the eleven potential N-glycosylation sites (128 and 625) of human aminopeptidase N are conserved within the seven aminopeptidase N molecules compared. Interestingly one of them is located within domain III. Provided this domain is facing the membrane, it may be of significance for the interaction with another membrane protein of importance for signalling/membrane localisation of aminopeptidase N. Alternatively it can be suggested that the conserved sites

mark regions which are particularly well suited to monitor the progress of folding during the biosynthesis (Tatu and Helenius 1997).

5.2 O-glycosylation

The occurrence and importance of O-glycosylation were studied in a cell-line by transfecting different forms of human and pig aminopeptidase N into a mutant CHO cell line ldl(D), which has a UDP-Gal/UDP-GalNAc 4-epimerase deficiency making it unable to convert UTP-glucose and UTP-N-acetylglucoseamine to the corresponding galactose derivatives. This made it possible to selectively block the O-glycosylation (Norén, K. *et al* 1997). Using ^{3}H-labelled GalNAc and either wild type aminopeptidase N or aminopeptidase N devoid of the anchor and the stalk it could be demonstrated that the enzyme is O-glycosylated and that this glycosylation mainly or exclusively is confined to the stalk. These results corroborated the outcome of a computer analysis of human aminopeptidase N for the presence of O-glycosylation. This analysis indicated a high probability for O-glycosylation in the stalk region and in the link region between domain VI and domain VII. The experiments with the particular cell line furthermore showed that the absence of O-glycosylation does not affect the intracellular transport to cell surface nor the enzymatic activity.

5.3 Dimerisation

The dimerisation of the pig enzyme is due to non-covalent bonds as monomers can be generated by treatment with SDS only (Sjöström and Norén 1982), This is in contrast to glutamyl aminopeptidase which is suggested to be dimerised by an S-S-linkage (Hesp and Hooper 1997). The interaction is in the pig enzyme independent of the cytosolic and the membrane spanning domains, as proteolytically solubilised forms behave as dimers in gel filtration experiments (Sjöström *et al* 1983). An interacting structure - besides being specific for aminopeptidase N - is therefore expected to be of non-covalent character and localised within domains V to VII. The earlier described charged C-terminal helix is a candidate.

ACKNOWLEDGMENTS

Work performed in the authors′ laboratory has been supported by the Danish Research Council of Health Sciences, the Danish Cancer Society, the Lundbeck foundation, the Novo-Nordic Foundation and the Benzon

Foundation. The authors are members of the Biomembrane Research Center, Aarhus University, Denmark.

REFERENCES

Bauer, K., 1994, Purification and characterisation of the thyrotropin-releasing-hormone degrading ectoenzyme. *Eur. J. Biochem.* **224:** 387-396.

Chen, H., Kinzer, C.A., and Paul, W.E., 1996, p161, a murine membrane protein expressed on mast cells and some macrophages, is mouse CD13/aminopeptidase N. *J. Immunol.* **157:** 2593-2600.

Danielsen, E.M., Cowell, G.M., Norén, O., Sjöström, H., and Dorling, P.R., 1983, Biosynthesis of intestinal microvillar proteins. The effect of swainsonine on post-translational processing of aminopeptidase N. *Biochem. J.* **216:** 325-331.

Danielsen, E.M., 1995, Involvement of detergent-insoluble complexes in the intracellular transport of intestinal brush border enzymes. *Biochem.* **34:** 1596-1605.

Delmas, B., Gelfi, I, L'Haridon, R., Vogel, L.K., Sjöström, H., Norén, O., and Laude, H., 1992, Aminopeptidase N is a major receptor for the enteropathogenic coronavirus TGEV. *Nature* **357:** 411-419.

Delmas, B., Gelfi, J., Kut, E., Sjöström, H., Norén, O., and Laude, H., 1994, Determinants essential for the transmissible gastroenteritis virus-receptor interaction reside within a domain of aminopeptidase-N that is distinct from the enzymic site. *J. Virol.* **68:** 5216-5224.

Gilbert, W., 1987, The exon theory of genes. *Cold Spring Harbor Symposia Quant. Biol.* **52:** 901-905.

Hansen, G.H., Delmas, B., Besnardeau, L., Vogel, L.K., Laude, H., Sjöström, H., and Norén, O., 1998, The coronavirus transmissible gastroenteritis virus causes infection after receptor-mediated endocytosis and acid-dependent fusion with an intracellular compartment *J. Virol.* **72:** 527-534.

Hesp, J.R., and Hooper, N.M., 1997, Proteolytic fragmentation reveals the oligomeric and domain structure of porcine aminopeptidase A. *Biochem.* **36:** 3000-3007.

Hooper, N.M., 1994, Families of metalloproteases. *FEBS Lett.* **354:** 1-6.

Huang, K., Takahara, S., Kinouchi, T., Takeyama, M., Ishida, T., Ueyama, H., Nishi, K., and Ohkubo, I., 1997, Alanyl aminopeptidase from human seminal plasma: purification, characterization, and immunohistochemical localization in the male genital tract. *J. Biochem.* **122:** 779-787.

Hussain, M.M., Tranum-Jensen, J., Norén, O., Sjöström, H., and Christiansen, K., 1981, Reconstitution of purified amphiphilic pig intestinal aminopeptidase. Mode of membrane insertion and morphology. *Biochem. J.* **199:** 179-186.

Keller, S.R., Scott, H.M., Mastick, C.C., Aebersold, R., and Lienhard, G.E., 1995, Cloning and characterization of a novel insulin-regulated membrane aminopeptidase from Glut4 vesicles. *J. Biol. Chem.* **270:** 23612-23618.

Kolb, A.F., Hegyi, A., and Siddell, S.G., 1997, Identification of residues critical for the human coronavirus 229E receptor function of human aminopeptidase N. *J. Gen. Virol.* **78:** 2795-2802.

Lerche, C., Vogel, L.K., Shapiro, L.H., Norén, O., and Sjöström, H., 1996, Human aminopeptidase N is encoded by 20 exons. *Mammalian Genome* **7:** 712-713.

Look, A.T., Ashmun, R.A., Shapiro, L.H., and Peiper, S.C., 1989, Human myeloid plasma membrane glycoprotein CD13 (gp150) is identical to aminopeptidase N. *J. Clin. Invest.* **83:** 1299-1307.

Luciani, N., Marie-Claire, C., Ruffet, E., Beaumont, A., Roques, B.P., and Fournié-Zaluski, M.-C., 1998, Characterisation of Glu^{350} as a critical residue involved in the N-terminal

amine binding site of aminopeptidase N (EC 3.4.11.2): Insights into its mechanism of action. *Biochem.* **37:** 686-692.

Malfroy, B., Kado-Fong, H., Gros, C., Giros, B., Schwartz J.-C., and Helmiss, R., 1989, Molecular cloning and amino acid sequence of rat kidney aminopeptidase M: a member of a super family of zinc-metallohydrolases. *Biochem. Biophys. Res. Comm.* **161:** 236-241.

Midorikawa, T., Abe, R., Yamagata, Y., Nakajima, E., and Ichishima, E., 1998, Isolation and characterization of cDNA encoding chicken egg yolk aminopeptidase Ey. *Comp. Biochem. Phys. Part B* **119:** 513-520.

Norén, K., Hansen, G.H., Clausen, H., Norén, O., Sjöström, H., and Vogel, L.K., 1997, Defectively N-glycosylated and non-O-glycosylated aminopeptidase N (CD13) is normally expressed at the cell surface and has full enzymatic activity. *Exp. Cell Res.* **231:** 112-118.

Norén. O., Sjöström, H., and Olsen, J., 1997, Aminopeptidase N. *In Cell-surface peptidases in health and disease* (A.J. Kenny, and C.M. Boustead, eds.), BIOS Scientific Publishers, Oxford, pp. 175-191.

Olsen, J., Cowell, G.M., Kønigshøfer, E., Danielsen, E.M., Møller, J., Laustsen, L., Hansen, O.C., Welinder, K.G., Engberg, J., Hunziker, W., Spiess, M., Sjöström, H., and Norén, O., 1988, Complete amino acid sequence of human intestinal aminpeptidase N as deduced from cloned cDNA. *FEBS Lett.* **238:** 307-314.

Rawlings, N.D., and Barrett, A.J., 1999, MEROPS: the peptidase database. *Nucl. Acids Res.* **27:** 325-331.

Rost, N., Sander, C., and Schneider, R., 1994,. PHD - an automatic mail server for protein secondary prediction. *Comput. Appl. Biosci.* **10:** 53-60.

Sjöström, H., Norén, O., Jeppesen, L., Staun, M., Svensson, B., and Christiansen, L., 1978, Purification of different amphiphilic forms of a microvillus aminopeptidase from pig small intestine using immunoadsorbent chromatography. *Eur. J. Biochem.* **88:** 503-511.

Sjöström, H., and Norén, O., 1982, Changes of the quaternary structure of microvillus aminopeptidase in the membrane. *Eur. J. Biochem.* **122:** 245-250.

Sjöström, H., Norén, O., Danielsen, E.M., and Skovbjerg, H., 1983, Structure of microvillar enzymes in different phases of their life cycles. In *Brush border membranes,* Ciba Foundation Symposium 95, Pitman, London, pp 50-69.

Stoltzfus, A., Spencer, D.F., Zuker, M., Logsdon, J.M., and Doolittle, W.F., 1994, Testing the exon theory of genes: the evidence from protein structure. *Science* **265:** 202-207.

Tatu, U., and Helenius, A., 1997, Interactions between newly synthesized glycoproteins, calnexin and a network of resident chaperons in the endoplasmic reticulum. *J. Cell Biol.* **136:** 555-565.

Tresnan, D.B., Levis, R., and Holmes, K.V., 1996, Feline aminopeptidase N serves as a receptor for feline, canine, porcine and human coronaviruses in serogroup I. *J. Virol.* **70:** 8669-8674.

Watanabe, Y., Iwaki-Egawa, S., Mizukoshi, H., and Fujimoto, Y., 1995, Identification of an alanine aminopeptidase in human maternal serum as a membrane-bound aminopeptidase N. *Biol. Chem. Hoppe Seyler* **376:** 397-400.

Watanabe, Y., Ito, K., Iwaki-Egawa, S., Yamaguchi, R., and Fujimoto, Y., 1998, Aminopeptidase N in sera of healthy subjects is a different N-terminal processed derivative from the one obtained from maternal serum. *Mol. Gen. Metab.* **63:** 289-294.

Watt, V.M., and Yip, C.C., 1989, Amino acid sequence deduced from a rat kidney cDNA suggests it encodes Zn-peptidase aminopeptidase N. *J. Biol. Chem.* **264:** 5480-5487.

Yang, X.F., Milhiet, P.E., Gaudoux, F., Crine, P., and Boileau, G., 1993, Complete sequence of rabbit kidney aminopeptidase N and mRNA localization in rabbit kidney by *in situ* hybridization. Biochem. *Cell. Biol.* **71:** 278-287.

Yeager, C.L., Ashmun, R.A., Williams, R.K., Cardellichio, C.B., Shapiro, L.H., Look, A.T., and Holmes, K.V., 1992, Human aminopeptidase N is a receptor for human coronavinis 229E. *Nature* **357:** 420-422.