Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

International Journal of Infectious Diseases (2004) **8**, 217—222

ELSEVIER

INTERNATIONAL SOCIETY FOR INFECTIOUS DISEASES

http://intl.elsevierhealth.com/journals/ijid

Identification of six new polymorphisms in the human coronavirus 229E receptor gene (aminopeptidase N/CD13)☆

Leen Vijgen, Els Keyaerts, Kalina Zlateva, Marc Van Ranst*

Laboratory of Clinical and Epidemiological Virology, Department of Microbiology and Immunology, Rega Institute for Medical Research, University of Leuven, Minderbroedersstraat 10, BE-3000, Belgium

Received 4 March 2004; accepted 18 March 2004
Corresponding Editor: Jonathan Cohen, Brighton, UK

KEYWORDS
Aminopeptidase N;
CD13;
Human coronavirus

Summary *Objective:* Human aminopeptidase N (APN/CD13/ANPEP) has been identified as the receptor for human coronavirus (HCoV) 229E. In this study, we analyzed the region of the APN gene that encodes a stretch of amino acid residues, essential for its HCoV-229E receptor function (amino acids 260—353).

Methods: Full-length APN exon 3, intron 3 and exon 4, was PCR-amplified and sequenced in DNA samples from 100 unrelated Caucasian Belgian healthy volunteers.

Results: We identified seven polymorphisms, including four intron 3 and three exon 4 variations. Apart from the already known C956T exon 4 mutation, the six other polymorphisms have not yet been described. The most prevalent APN variations in this population (C956T leading to an alanine to valine substitution, G978T, G987A and intron3-C389T) always occurred together at an allele frequency of 8.5%. Haploid DNA sequencing demonstrated the presence of these four variations on the same allele. Three polymorphisms in intron 3, intron3-G395C, intron3-C86T, and intron3-C429T, were identified with an allele frequency of 3.5%, 1% and 0.5% respectively. Five haplotypes were identified in the population of 100 individuals.

Conclusion: These results demonstrate that there is a relatively broad spectrum of variations in the APN domain critical for coronavirus binding.

The nucleotide sequence reported here has been submitted to the GenBank database with the following accession number: AF527789.

Introduction

Coronaviruses are large, enveloped, single-stranded, positive-sense RNA viruses, with a genome size of approximately 30 kb. Human coronaviruses (HCoV), with two known serogroups designated OC43 and 229E, are an important cause of upper respiratory tract illnesses, and are also implicated in diseases involving the digestive and the central nervous

☆ Paper received at the International Society for Infectious Diseases meeting in Cancun, March 2004 and fast-tracked through review to publication.
*Corresponding author. Tel.: +32-16-347908; fax: +32-16-347900.
E-mail address: Marc.VanRanst@uz.kuleuven.ac.be (M. Van Ranst).

1201-9712/$30.00
doi:10.1016/j.ijid.2004.03.004

system.[1] After the recent SARS outbreak and the discovery of a new coronavirus (SARS-CoV) as the causative agent, the importance of coronaviruses in human medicine can no longer be neglected.[2]

As a first step in viral entry into the cell, human coronavirus 229E (HCoV-229E) attaches, in a species-specific manner, to aminopeptidase N (APN/CD13/ANPEP).[3] In other species APN is also used as a receptor by the HCoV-229E related group I coronaviruses (porcine transmissible gastroenteritis virus, porcine epidemic diarrhea virus, feline infectious peritonitis virus and canine coronavirus).[4,5] Although human APN had also been suggested as a putative receptor for the spike glycoprotein of the SARS-associated coronavirus (SARS-CoV), it has now been proven that not APN but angiotensin-converting enzyme 2 is a functional receptor for SARS-CoV.[6,7] The receptor for human coronavirus serotype OC43 (HCoV-OC43) is still unknown, although evidence has been presented that 9-O-acetylated sialic acid is used as a receptor determinant for infection of human cells.[8]

APN is a type II glycoprotein that belongs to the family of membrane-bound metalloproteases. The 150 kDa APN glycoprotein (EC 3.4.11.2) is composed of 967 amino acids and is expressed in a variety of tissues including cells of the monocytic and granulocytic lineage, synaptic membranes of the central nervous system, and intestinal, lung and kidney epithelial cells.[9] A soluble form of APN is reported to be present in human plasma.[10] The physiological role of APN includes the conversion of oligopeptides in the small intestinal lumen into amino acids, and an involvement in the degradation of regulatory peptides in other tissues such as the kidney and the brain.[9] The human APN gene has been mapped to chromosome 15q25-26, and the coding part of the gene is divided in 20 exons.[11,12] Expression of APN is regulated by two different promoters, separated by an 8 kb intron, leading to the formation of two transcripts that only differ in their 5′ non-coding region. The epithelial promoter is located close to the coding part of the gene and is active in enterocytes and other epithelial cells, while the more upstream myeloid promoter is mainly active in myeloid cells.[13] Two splice variants in which exon 3 and exon 14 were lost, have recently been identified, resulting in both cases in a carboxy-terminal truncated protein.[14]

Several applications of APN in therapy and diagnosis have been described. APN/CD13 is frequently used in the characterization and typing of leukemia or lymphoma cells, as APN/CD13 is expressed on stem cells and during most developmental stages of myeloid cells.[15] It has recently been demonstrated that APN/CD13 is both a receptor for tumor homing peptides and a functional regulator of angiogenesis, implicating a role for APN/CD13 as a target for anti-angiogenic therapy and tumor targeted drug delivery.[16,17] APN has also been shown to mediate human cytomegalovirus (HCMV) infection.[18] The inhibition or neutralization of HCMV infection by soluble APN implicates the possibility of using soluble APN as an antiviral agent.

The domain of human aminopeptidase N that is essential for its HCoV-229E receptor activity has been assigned to APN amino acids 260—353, with a critical stretch of eight residues (amino acids 288—295).[19,20] This region is encoded by the major part of APN exon 3, whole exon 4 and a small part of APN exon 5 (nt 802—1083, hAPN mRNA, GenBank accession number X13276). In this study we examined whether polymorphisms could be detected in the HCoV-229E binding domain of APN in a Caucasian population of 100 unrelated, healthy individuals, assuming that these mutations could be of importance in HCoV-229E attachment to human cells.

Materials and methods

Subjects

The study involved a Caucasian population of 100 healthy, unrelated individuals from the region of Flanders, the northern part of Belgium. Informed consent was obtained from all participants and the study had the approval, based upon the guidelines from the World Medical Association's Declaration of Helsinki, of the Ethics Board of the University of Leuven.[21]

DNA extraction

DNA samples were collected through a non-invasive 'swish-and-spit' technique. Genomic DNA was acquired from oral epithelial cells, by rinsing the oral cavity with a 0.9% saline solution, after which DNA was extracted using an alkaline lysis procedure.[22]

PCR amplification

An 871 bp fragment, encompassing exon 3, intron 3 and exon 4 of the aminopeptidase N gene, was amplified by polymerase chain reaction (PCR), using 5′-TGCTTCCCAAAGGTGAGTGG-3′ as the forward primer and 5′-CCATTGGCAGGATGAACTCC-3′ as the reverse primer (GenBank accession number AC018988). PCR amplification was performed in a volume of 50 μL reaction mix, with a concentration

of 0.2 μM of forward and reverse primer, 0.2 mM of nucleotides, 1.5 mM of $MgCl_2$ and 1 unit of Taq polymerase (Applied Biosystems/Roche Molecular Systems, Belgium) at pH 9. PCR conditions were composed as follows: an initial denaturation at 95 °C for five minutes, then 40 cycles of 30 seconds at 95 °C, 30 seconds at 57 °C and 35 seconds at 72 °C, and a final elongation at 72 °C for seven minutes. The amplification reaction was performed in a Geneamp® PCR System 9700 thermal cycler (Applied Biosystems, Foster City, CA, USA). The resulting PCR products were visualized after electrophoresis, on an ethidium bromide-stained polyacrylamide gel.

Sequence analysis

After purification of the PCR products with the QIAquick PCR purification kit (Qiagen, Westburg, The Netherlands), the purified products were cycle sequenced in forward and reverse direction using the ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction kit (Applied Biosystems, Foster City, CA, USA). Besides the PCR primer set, two additional primers, a forward: 5′-CTGCCCCAGGATCAAACAGG-3′ and a reverse: 5′-GCACAGGGATGAAGAGAACG-3′, both located in APN intron 3, were used to obtain the full-length sequence of the 871 bp fragment (GenBank accession number AJ421875). Electrophoretic separation and detection were performed on an ABI PRISM 3100 Genetic Analyzer (Perkin Elmer, Applied Biosystems, Foster City, CA, USA).

Haplotype analysis

PCR products from individuals heterozygous for more than one variation were cloned into the pDrive cloning vector (Qiagen PCR Cloning kit) using One Shot Max efficiency DH5α-T1 competent cells (Invitrogen, Merelbeke, Belgium). Plasmids were purified with the Qiaprep miniprep spin kit (Qiagen, Westburg, The Netherlands) and checked for inserts by *Eco*RI restriction analysis. Bidirectional cycle sequencing of plasmids with inserts was performed with the PCR primer set using the ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction kit (Applied Biosystems, Foster City, CA, USA). Electrophoretic separation and detection were performed on an ABI PRISM 3100 Genetic Analyzer (Perkin Elmer, Applied Biosystems, Foster City, CA, USA).

Results

A total of 100 healthy unrelated Belgian individuals were screened for polymorphisms in the human aminopeptidase N domain that is essential for its HCoV-229E receptor activity. An 871 bp fragment, encompassing APN exon 3, intron 3 and exon 4 was PCR-amplified and sequenced, and was submitted to GenBank under accession number AF527789. The resulting chromatograms were analyzed using the SeqMan multiple sequence alignment tool (LaserGene, DNAStar, Madison, WI). Consensus sequences were compared with a reference APN sequence in GenBank (accession number AJ421875) using BLAST (Basic Local Alignment Search Tool).[23]

Seven variations were identified, of which six have not yet been described. Table 1 shows the distribution of the allele frequencies. Intron variations are described referring to their relative position in the intron, and exon variation positions are referred to their mRNA nucleotide position (GenBank accession number X13276). The location of the polymorphisms in the APN gene is shown in Figure 1. While no variations could be found in APN exon 3, four of the identified polymorphisms were located in intron 3 and three in exon 4. In intron 3, a C to T variation was found at intron 3 position 86 (C86T) (GenBank AJ421875, contig position 327) in two individuals, as well as a G to C variation at intron position 395 (G395C) in seven individuals, and a C to T variation at intron position 429 (C429T) in one sample. All individuals were heterozygous

Table 1 Allele frequencies of the APN gene mutations.

Position	Variation	Frequency	
		n	%
Intron 3[a]			
86	C	198	99.0
	T	2	1.0
389	C	183	91.5
	T	17	8.5
395	G	193	96.5
	C	7	3.5
429	C	199	99.5
	T	1	0.5
Exon 4[b]			
956 (codon 311)	C	183	91.5
	T	17	8.5
978 (codon 318)	G	183	91.5
	T	17	8.5
987 (codon 321)	G	183	91.5
	A	17	8.5

[a] Variations in intron 3 are referred to their position in intron 3.
[b] Variations in exon 4 are referred to their mRNA position (GenBank accession number X13276).

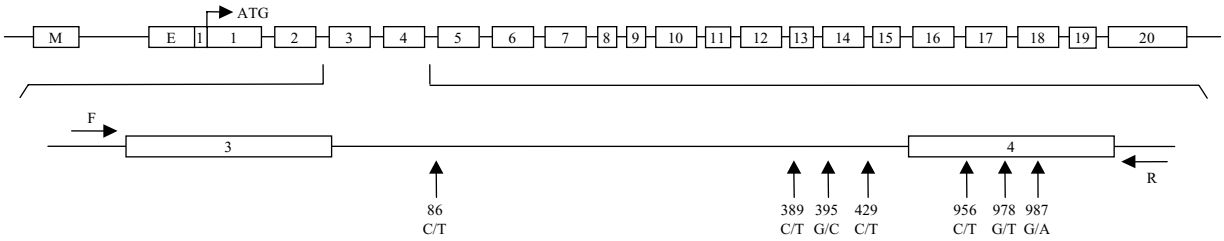

Figure 1 Schematic representation of the human aminopeptidase N (APN) gene. M represents the myeloid promoter, and E the epithelial promoter. The primers used in the PCR reaction are indicated by F (forward) and R (reverse). The locations of the identified variations are indicated (arrows).

for these polymorphisms, which have no apparent functional consequence, as they are located in a non-coding intron region of the APN gene. The fourth observed intron polymorphism, a C to T variation at intron position 389 (C389T) was found in association with three variations in exon 4: an already described C to T polymorphism in codon 311 (GenBank X13276, mRNA, nt position 956),[24] leading to an alanine to valine substitution, and two silent mutations, a G to T variation in codon 318 (nt 978), and a G to A variation in codon 321 (nt 987). Fifteen heterozygotes and one homozygote for these four polymorphisms were found in the group of 100 Belgian individuals. In one of these fifteen heterozygous individuals, the intron variation C429T was found. No deviations from expectations under Hardy-Weinberg assumptions were observed.

Haplotypes were determined by sequence analysis of cloned PCR-amplicons in samples of individuals heterozygous for more than one variation. Five haplotypes were identified, and their frequencies are shown in Table 2. The four linked variations, identified in samples of 16 individuals, were demonstrated to be present on the same allele, and this haplotype was detected with a frequency of 8.5%.

Discussion

In this study we analyzed an 871 bp region of the APN gene, encompassing the full-length APN exon 3, intron 3 and exon 4. In our search for polymorphisms in the APN domain that is essential for its HCoV-229E receptor function, we identified seven polymorphisms, of which four were located in the non-coding intron 3. In 25% of the Belgian individuals one or more of these seven variations were found. Haplotype analysis revealed the presence of five haplotypes in the population. Three polymorphisms in APN exon 4 (C956T, G978T and G987A) in association with an intron 3 variation (C389T), were identified at a relatively high allele frequency (8.5%) in our Belgian population. Fifteen heterozygotes and one homozygote for these four variations were observed. One of the exon 4 single nucleotide polymorphisms is a C to T mutation, C956T, leading to an amino acid change in codon 311. This alanine to valine substitution, a conservative amino acid change, has already been described in a population of Italian coeliac disease patients, in which it was observed with an allele frequency of 7.9%.[24] In our population, this polymorphism was approximately equally frequent (8.5%). The C956T polymorphism was always found in combination with the

Table 2 Distribution of the APN haplotypes.

Intron 3[a]				Exon 4[b]			Frequency	
86	389	395	429	956	978	987	*n*	%
C	C	G	C	C	G	G	173	86.5
C	T	G	C	T	T	A	17	8.5
C	C	C	C	C	G	G	7	3.5
T	C	G	C	C	G	G	2	1.0
C	C	G	T	C	G	G	1	0.5

[a] Variations in intron 3 are referred to their position in intron 3.
[b] Variations in exon 4 are referred to their mRNA position (GenBank accession number X13276).

two other observed exon 4 variations (G978T and G987A) and one intron 3 variation (C389T). Haploid DNA sequencing demonstrated the presence of these four variations on the same allele. Here we report for the first time an association between the C956T polymorphism and other variations in the APN gene. Absolute linkage between the SNPs within this short physical distance is commonly observed. Apart from the codon 311 variation (C956T), none of the other associated polymorphisms has an apparent functional effect, as the codon 318 and 321 variations (G978T and G987A) are silent, and C389T is located in an intron.

Recent studies revealed that the human APN gene is subject to alternative splicing, and two splice variants, in which exon 3 and exon 14 were lost, have been identified.[14] Alternative splicing is a mechanism that allows different protein isoforms to be created from a single gene. The splicing process is regulated by both cis- and trans-acting factors, which control the choice of 5′ or 3′ splice sites either positively or negatively. Cis-acting regulatory elements include exonic or intronic splice enhancers or silencers, which can be changed by mutations with a possible effect on mRNA splicing.[25,26] In this perspective, the importance of silent mutations or intron variations, which seem to have no apparent functional effect, should be reconsidered. Supporting this assumption, a silent mutation (C77G) in exon 4 of the human protein-tyrosine phosphatase CD45 gene has been shown to increase the expression of an aberrant exon 4-included CD45 by disrupting the activity of an exonic splice silencer. An association of this silent mutation with multiple sclerosis has been suggested.[27] In this way, it might be possible that the silent mutations in APN exon 4 and the APN intron 3 variations that we observed in our population could be of functional importance.

The alanine to valine substitution in codon 311, caused by the single nucleotide polymorphism C956T, is an amino acid change occurring in the APN domain that is essential for its HCoV-229E receptor function. Although this amino acid variation is rather conservative, an effect on the HCoV-229E binding capacity of aminopeptidase N could be possible. Since the three-dimensional structure of APN has not yet been described, molecular modeling studies of the whole protein could not be performed.

Further research is needed to elucidate the biological relevance of the APN variations described here. In vitro studies should be carried out including in vitro transcription, translation and viral binding assays of the described APN gene variants to detect a possible functional effect. Furthermore, a virus susceptibility test can be performed involving healthy volunteers homozygous or heterozygous for these variations. HCoV-229E is a relatively innocuous agent and it would be possible to determine whether these polymorphisms could affect susceptibility to infection.

Given the substantial role of aminopeptidase N in several domains, including disease therapy and diagnosis, studies analyzing the APN gene and the functional consequences of possible gene variations will be of importance in these and future applications.

Acknowledgements

We would like to thank the colleagues of the Laboratory of Clinical and Epidemiological Virology, Department of Microbiology and Immunology, Rega Institute for Medical Research, University of Leuven, Belgium, for helpful comments and discussion. This work was supported by a fellowship of the Fund for Scientific Research (FWO), Brussels, Belgium.

Conflict of Interest: No conflicting interest declared.

References

1. Arbour N, Day R, Newcombe J, Talbot PJ. Neuroinvasion by human respiratory coronaviruses. *J Virol* 2000;**74**:8913—21.
2. Drosten C, Gunther S, Preiser W, van der Werf S, Brodt HR, Becker S, et al. Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *N Engl J Med* 2003;**348**:1967—76.
3. Yeager CL, Ashmun RA, Williams RK, Cardellichio CB, Shapiro LH, Look AT, Holmes KV. Human aminopeptidase N is a receptor for human coronavirus 229E. *Nature* 1992;**357**:420—2.
4. Delmas B, Gelfi J, L'Haridon R, Vogel LK, Sjöström H, Noren O, Laude H. Aminopeptidase N is a major receptor for the entero-pathogenic coronavirus TGEV. *Nature* 1992;**357**:417—20.
5. Tresnan DB, Levis R, Holmes KV. Feline aminopeptidase N serves as a receptor for feline, canine, porcine, and human coronaviruses in serogroup I. *J Virol* 1996;**70**:8669—74.
6. Yu XJ, Luo C, Lin JC, Hao P, He YY, Guo ZM, et al. Putative hAPN receptor binding sites in SARS-CoV spike protein. *Acta Pharmacol Sin* 2003;**24**:481—8.
7. Li W, Moore MJ, Vasilieva N, Sui J, Wong SK, Berne MA, et al. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature* 2003;**426**:450—4.
8. Krempl C, Schulze B, Herrler G. Analysis of cellular receptors for human coronavirus OC43. *Adv Exp Med Biol* 1995;**380**:371—4.
9. Olsen J, Kokholm K, Noren O, Sjöström H. Structure and expression of aminopeptidase N. *Adv Exp Med Biol* 1997;**421**:47—57.
10. Favaloro EJ, Browning T, Facey D. CD13 (GP150; aminopeptidase-N): predominant functional activity in blood is localized to plasma and is not cell-surface associated. *Exp Hematol* 1993;**21**:1695—701.

11. Look AT, Peiper SC, Rebentisch MB, Ashmun RA, Roussel MF, Lemons RS, et al. Molecular cloning, expression, and chromosomal localization of the gene encoding a human myeloid membrane antigen (gp150). *J Clin Invest* 1986;**78**:914—21.
12. Lerche C, Vogel LK, Shapiro LH, Noren O, Sjöström H. Human aminopeptidase N is encoded by 20 exons. *Mamm Genome* 1996;**7**:712—3.
13. Shapiro LH, Ashmun RA, Roberts WM, Look AT. Separate promoters control transcription of the human aminopeptidase N gene in myeloid and intestinal epithelial cells. *J Biol Chem* 1991;**266**:11999—2007.
14. Dybkaer K, Kristensen JS, Pedersen FS. Single site polymorphisms and alternative splicing of the human CD13 gene — different splicing frequencies among patients with acute myeloid leukaemia and healthy individuals. *Br J Haematol* 2001;**112**:691—6.
15. Lendeckel U, Arndt M, Frank K, Wex T, Ansorge S. Role of alanyl aminopeptidase in growth and function of human T cells (review). *Int J Mol Med* 1999;**4**:17—27.
16. Pasqualini R, Koivunen E, Kain R, Lahdenranta J, Sakamoto M, Stryhn A, et al. Aminopeptidase N is a receptor for tumor-homing peptides and a target for inhibiting angiogenesis. *Cancer Res* 2000;**60**:722—7.
17. Curnis F, Arrigoni G, Sacchi A, Fischetti L, Arap W, Pasqualini R, et al. Differential binding of drugs containing the NGR motif to CD13 isoforms in tumor vessels, epithelia, and myeloid cells. *Cancer Res* 2002;**62**:867—74.
18. Soderberg C, Giugni TD, Zaia JA, Larsson S, Wahlberg JM, Moller E. CD13 (human aminopeptidase N) mediates human cytomegalovirus infection. *J Virol* 1993;**67**:6576—85.
19. Kolb AF, Maile J, Heister A, Siddell SG. Characterization of functional domains in the human coronavirus HCV 229E receptor. *J Gen Virol* 1996;**77**:2515—21.
20. Kolb AF, Hegyi A, Siddell SG. Identification of residues critical for the human coronavirus 229E receptor function of human aminopeptidase N. *J Gen Virol* 1997;**78**:2795—802.
21. World Medical Association. World Medical Association's Declaration of Helsinki 1964. World Medical Association 1996, Fernay-Voltaire. http://www.wma.net.
22. Lench N, Stanier P, Williamson R. Simple non-invasive method to obtain DNA for gene analysis. *Lancet* 1988;**1**:1356—8.
23. Altschul SF, Gish W, Miller W, Myers E, Lipman DJ. Basic local alignment search tool. *J Mol Biol* 1990;**215**:403—10.
24. Giordano M, Bolognesi E, D'Alfonso S, Lessi M, Zavattari P, Oderda G, et al. Linkage disequilibrium between intra-locus variants in the aminopeptidase n gene and test of their association with coeliac disease. *Ann Hum Genet* 1999;**63**:207—15.
25. Stoss O, Stoilov P, Daoud R, Hartmann AM, Olbrich M, Stamm S. Misregulation of pre-mRNA splicing that causes human diseases. Concepts and therapeutic strategies. *Gene Ther Mol Biol* 2000;**5**:9—30.
26. Maquat LE. The power of point mutations. *Nat Genet* 2001;**27**:5—6.
27. Lynch KW, Weiss A. A CD45 polymorphism associated with multiple sclerosis disrupts an exonic splicing silencer. *J Biol Chem* 2001;**276**:24341—7.