MOLECULAR ANALYSIS OF THE CORONAVIRUS-RECEPTOR FUNCTION OF AMINOPEPTIDASE N

Andreas F. Kolb, Annette Hegyi, Julia Maile, Angelien Heister, Margitta Hagemann, and Stuart G. Siddell

Institute of Virology and Immunology
University of Würzburg, Versbacherstr. 7
D-97078 Würzburg, Germany

1. ABSTRACT

Aminopeptidase N (APN) is a major cell surface for coronaviruses of the serogroup I. By using chimeric APN proteins assembled from human, porcine and feline APN we have identified determinants which are critically involved in the coronavirus-APN interaction. Our results indicate that human coronavirus 229E (HCV 229E) is distinct from the other serogroup I coronaviruses in that determinants located within the N-terminal parts of the human and feline APN proteins mediate the infection of HCV 229E, whereas determinants located within the C-terminal parts of porcine, feline and canine APN mediate the infection of transmissible gastro-enteritis virus (TGEV), feline infectious peritonitis virus (FIPV) and canine coronavirus (CCV), respectively. A further analysis of the mapped amino acid segments by site directed mutagenesis revealed that a short stretch of 8 amino acids in the hAPN protein plays a decisive role in mediating HCV 229E reception.

2. INTRODUCTION

On the basis of serological and genetical data, coronaviruses are classified into three different serogroups (Siddell, 1995). The members of serogroup I utilise the cell surface molecule aminopeptidase N (APN) as a receptor (Delmas et al., 1992; Yeager et al., 1992; Tresnan et al., 1996). Human APN and porcine APN act species specifically in that they only mediate the infection of the human coronavirus 229E and the porcine coronavirus TGEV, respectively (Delmas et al., 1994; Kolb et al., 1996). In contrast, the feline APN protein is able to mediate infection of FIPV, TGEV, HCV 229E and CCV (Tresnan et al.,

1996). However, the feline coronavirus FIPV can only utilise fAPN as a receptor. By using chimeric molecules assembled from hAPN and pAPN the determinants which play an essential role in the infection of HCV 229E and TGEV have been identified (Delmas et al., 1994; Kolb et al., 1996). The region encompassing pAPN amino acids 717–813 has been delimited as the protein segment which is critically involved in TGEV infection via APN (Delmas et al., 1994). In contrast, a region located in the amino-terminal part of hAPN (hAPN amino acids 260–353) has been defined as being essential for the infection of HCV 229E via hAPN (Kolb et al., 1996). Crandell feline kidney (CRFK) cells are susceptible to FIPV, TGEV and HCV 229E. We anticipated that infection of HCV 229E into these cells would be mediated by a feline homologue of hAPN. We have isolated a fAPN cDNA from CRFK cells and by using chimeric molecules assembled from human, porcine and feline APN we have delimited the fAPN amino acid regions, which are involved in the HCV 229E, TGEV and FIPV receptor function. Within these regions we have analysed the amino acid residues which play an essential role in mediating coronavirus infections.

3. MATERIALS AND METHODS

3.1. Cells and Viruses

The HCV 229E isolate used in these studies was described previously. TGEV strain Purdue 46 was obtained from G. Herrler, University of Marburg, Germany. FIPV strain 79–1146 was obtained from R. de Groot, University of Utrecht, The Netherlands. CRFK cells (ECACC 86093002) and 293 HEK cells (ECACC 85120602) were cultivated according to standard techniques. Transfections were performed as described (Kolb et al., 1996).

3.2. Recombinant DNA

The construction of cDNA clones and chimeras used during these studies are described elsewhere (Delmas et al., 1994; Kolb et al., 1996; Kolb et al., submitted; Hegyi et al., in preparation).

3.3. Protein Detection

Detection of de novo synthesised HCV 229E or TGEV antigens was performed by Western blotting (Kolb et al., 1996). Detection of de novo synthesised FIPV antigens was done by immune-precipitation of viral proteins from infected and metabolically labelled cells as described (Grosse and Siddell, 1994). The enzymatic activity of APN proteins was analysed as described (Kolb et al., 1996).

4. RESULTS

A feline APN cDNA was isolated from CRFK cells by reverse transcription-PCR. The primers used for the amplification were derived from nucleotide sequences which are conserved between the known mammalian APN genes. Two PCR products encompassing the complete fAPN cDNA were cloned into an expression vector and transfected into 293 HEK cells. Cells transfected with the fAPN expression construct became susceptible to infection with FIPV, TGEV and HCV 229E (Table 1). In accordance with published data

Table 1. Receptor activity of mammalian APN proteins and chimeric proteins derived thereof[1]

Chimera	Backbone	Inserted residues	Susceptibility to HCV229E	Susceptibility to TGEV	Susceptibility to FIPV
pAPN			−	+	−
(Delmas et al., 1992)					
hAPN			+	−	−
(Yeager et al., 1992)					
fAPN [Fcwf]			+	+	+
(Tresnan et al., 1996)					
fAPN [CRFK]			+	+	+
(Kolb et al., subm.)					
AP24	hAPN	pAPN 717-813	+	+	−
(Delmas et al., 1994)					
AP18-hbh	hAPN	bAPN 643-841	nd	+	−
(Benbacer et al., 1997)					
AP18-hch	hAPN	cAPN 643-841	nd	+	+
(Benbacer et al., 1997)					
AP34	pAPN	fAPN 135-297	+	+	−
(Kolb et al., subm.)					
AP33	hAPN	fAPN 670-840	+	+	+
(Hegyi et al., in prep.)					

nd: not determined

[1]The susceptibility of transfected cells to the infection with HCV 229E, TGEV and FIPV is indicated. Susceptibility was analysed by detection of viral antigens in infected cells.

(Delmas et al., 1993), transfection of pAPN or hAPN expression vectors failed to induce susceptibility to FIPV infection (Table 1). The sequence of the fAPN cDNA derived from CRFK cells is highly homologous to the fAPN cDNA recently isolated from Fcwf cells (Tresnan et al., 1996). However, the predicted CRFK fAPN protein differs from the Fcwf fAPN protein by 16 amino acids (Kolb et al., submitted).

In order to define the determinants in the fAPN protein which mediate the infection of HCV 229E a chimeric protein was generated (AP34) which carries fAPN amino acids 135–297 incorporated into a pAPN backbone (Table 1; Fig. 1). This chimera was able to act as a receptor for HCV 229E, indicating that the critical determinants required for the infection of HCV 229E via feline APN are present within fAPN amino acids 135–297 (Kolb et al., submitted). The sequences of human, feline and porcine APN were aligned by using the CLUSTAL method of the MegAlign program (DNAStar, Madison, WI, USA). This alignment reveals that the chimeras AP32, which carries hAPN amino acids 260–353 inserted into a pAPN backbone (Kolb et al., 1996), and AP34 overlap by as little as 39 amino acids (Fig. 1). These 39 amino acids contain a sequence which is highly divergent between the 3 APN proteins (Fig. 1). Within this short sequence the hAPN and fAPN are more related to each other that to the pAPN protein. A number of site directed mutants was therefore generated to more closely characterise this amino acid segment. In one of these mutants the pAPN amino acids 283–290 (QSVNETAQ) were replaced by the hAPN amino acids 288–295 (DYVEKQAS) (mutant AP35, Table 2). This replacement of 6 amino acids of pAPN by the corresponding hAPN residues was sufficient to convert pAPN into a functional receptor for HCV 229E (Kolb et al., submitted). At the same time this mutant was still able to mediate the infection of TGEV, suggesting that the exchanged amino acid residues do not play an essential role for the TGEV-pAPN interaction (Table 2).
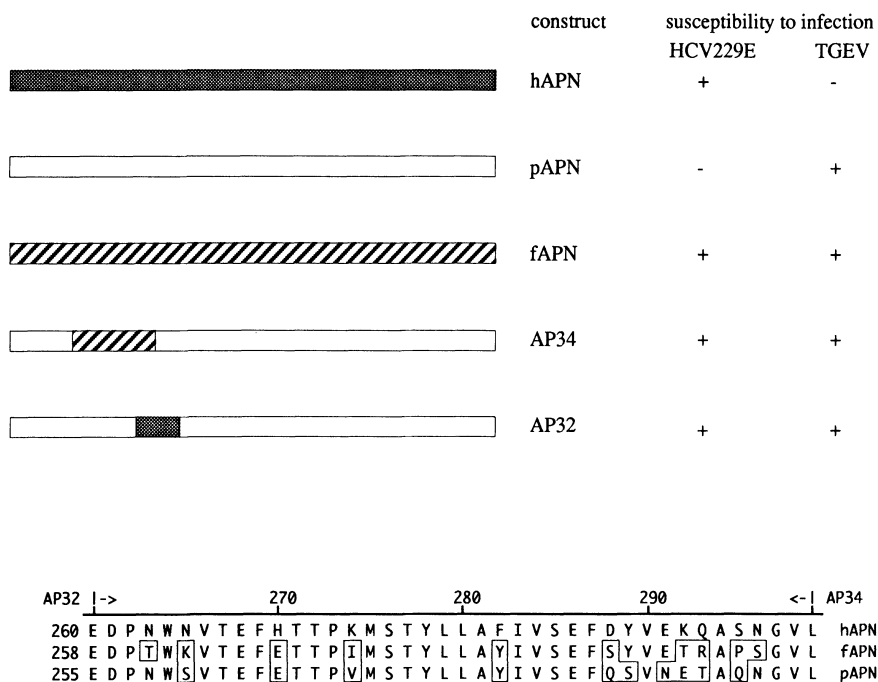
A. F. Kolb *et al.*

	construct	susceptibility to infection	
		HCV229E	TGEV
	hAPN	+	-
	pAPN	-	+
	fAPN	+	+
	AP34	+	+
	AP32	+	+

```
AP32 |->                        270            280           290        <-| AP34
  260 E D P N W N V T E F H T T P K M S T Y L L A F I V S E F D Y V E K Q A S N G V L   hAPN
  258 E D P[T]W[K]V T E F[E]T T P[I]M S T Y L L A[Y]I V S E F[S]Y V E[T R]A[P S]G V L   fAPN
  255 E D P N W[S]V T E F[E]T T P[V]M S T Y L L A[Y]I V S E F[Q S]V[N E T]A[Q]N G V L   pAPN
```

Figure 1. Schematic representation of the proteins hAPN, pAPN, fAPN and the chimeric proteins AP34 and AP32. The susceptibility of 293 HEK cells transfected with the corresponding expression constructs to infection with HCV229E and TGEV is indicated. Sequence alignment of hAPN amino acids 260 to 299 with the corresponding regions of the fAPN and pAPN proteins. The alignment was done using the CLUSTAL method of the MegAlign program (DNAStar, Madison, WI, USA). Residues diverging from the hAPN sequence are boxed.

The hAPN amino acid segment 288–295 was analysed for features that are similar in the fAPN sequence and different in the pAPN sequence to establish a molecular basis for the ability of hAPN and fAPN to act as receptors for HCV 229E. A first obvious difference between the pAPN sequence on the one side and the hAPN and fAPN sequences on the other is the presence of a N-linked glycosylation signal in the pAPN protein (residues: 283–290 QSV*NET*AQ) which is absent from hAPN and fAPN. The glycosylation signal was removed by site directed mutagenesis (converting the pAPN amino acids 283–290 to

Table 2. Receptor activity of mutated pAPN proteins. The susceptibility of transfected cells to the infection with HCV 229E and TGEV is indicated

Mutant	Mutated sequence: pAPN 283-290: QSVNETAQ	Susceptibility to HCV 229E	Susceptibility to TGEV
AP35	DYVEKQAS	+	+
AP36	QSVETRAQ	–	+
AP37	QYVNERAQ	–	+
AP38	QYVNKTAQ	–	+

QSVETRAQ). A mutant carrying this sequence alteration (AP36), however, failed to mediate HCV 229E infection (Table 2), suggesting that the absence of the glycosylation signal within this protein domain is not sufficient to convert pAPN into a receptor for HCV 229E. The mutant AP36 was still able to act as a receptor for TGEV, again indicating that the mutated sequence does not play an essential role in the pAPN-TGEV interaction.

Another prominent property which distinguishes hAPN and fAPN from pAPN is the presence of a tyrosine residue (hAPN: Y-289; fAPN Y-287) and a positively charged residue (hAPN: K-292; fAPN: R-291) within the critical amino acid segment defined by mutant AP35 (Fig. 1). A computer prediction using the predict-protein program (Rost, 1996) also indicates that this amino acid segment is located within an alpha-helical domain of the APN protein. The tyrosine residue and the positively charged residue are predicted to lie on the same surface of the helix. These two residues may therefore play a critical role in the HCV 229E receptor function of hAPN and fAPN. Two further mutants were thus generated in which the pAPN sequence QSVNETAQ (residues 283–290) was converted to QYVNERAQ (mutant AP37) or QYVNKTAQ (mutant AP38). Both of these mutants, however, failed to convert pAPN into a receptor for HCV 229E (Table 2). Again, both mutants could still act as receptors for TGEV (Table 2).

The amino acid residues which are critically involved in mediating the HCV 229E receptor function of both hAPN and fAPN are located within the amino-terminal part of the protein (Fig. 1). The mutant AP34, however, failed to mediate FIPV infection, suggesting that the determinants which are essential for the FIPV receptor function of fAPN are not present within this chimera (Kolb et al., submitted). The determinants mediating TGEV and CCV infection have been mapped to the carboxy-terminal part of porcine and canine APN, respectively (Delmas et al., 1994; Benbacer et al., 1997). To investigate whether the determinants mediating the FIPV receptor function are also located within the C-terminal part of fAPN, we generated a chimera (AP33) in which fAPN amino acids 670–840 were integrated into a hAPN backbone (Table 1, Fig. 2). Transfection/infection experiments revealed that AP33 can mediate infection of FIPV and TGEV (Hegyi et al., in prep.). This indicates that a) the determinants which are essential for the fAPN-FIPV interaction are indeed located within the C-terminal segment of fAPN, b) TGEV and FIPV interact with similar determinants within their respective receptor proteins and c) as in hAPN and pAPN, HCV 229E and TGEV interact with different amino acid segments of the fAPN protein.

In order to define critical amino acid residues within the domain involved in TGEV, FIPV and CCV reception we aligned the amino acid sequences of APN proteins which have been analysed for their coronavirus receptor function (Fig. 2). Chimeric proteins carrying fragments of the C-terminal part of porcine, canine, bovine and feline APN (encoding the amino acids indicated in Table 1) within an hAPN backbone can all act as functional receptors for TGEV (Delmas et al., 1994; Benbacer et al., 1997). Only the corresponding hAPN sequence is unable to mediate the TGEV-receptor function, suggesting that residues critically involved in the TGEV-APN interaction are not present within the hAPN sequence. As few as 9 amino acids in the 717–813 (pAPN) region are specific for the human APN sequence (Fig. 2).

By using site directed mutagenesis we converted two stretches of amino acids within the C-terminal part of the hAPN protein (alone and in combination) to the corresponding porcine sequence (Table 3). None of the mutations, however, converted hAPN into a functional receptor for TGEV. The EEE->QDQ mutation in the mutant AP42 rather rendered the APN protein non-functional in that the expressed protein neither showed any HCV 229E receptor activity nor retained its aminopeptidase function (Table 3). The catalytic centre of the hAPN protein as well as the stretch of amino acids which plays an essential
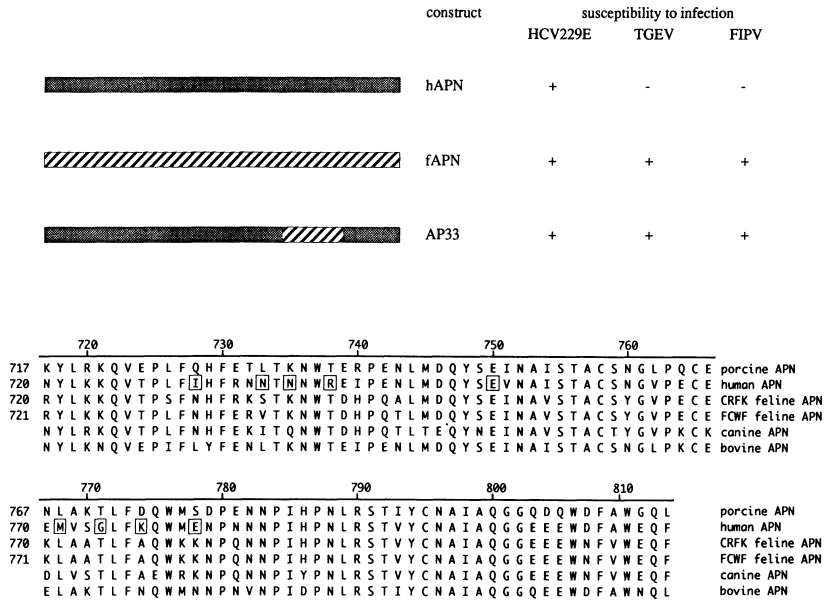
Figure 2. Schematic representation of the proteins hAPN, fAPN and the chimeric protein AP33. The susceptibility of cells transfected with the corresponding expression constructs to infection with HCV229E, TGEV and FIPV is indicated. Sequence alignment of pAPN amino acids 717 to 813 with the corresponding regions of the human, feline (from FCWF and CRFK cells) canine and bovine APN proteins. The alignment was done using the CLUSTAL method of the MegAlign program (DNAStar, Madison, WI, USA). Residues unique to the hAPN sequence are boxed.

role in the HCV 229E receptor function of hAPN is distant to the site of the EEE->QDQ mutation in the primary structure. This result may thus suggest that the determinants identified as being essential for the HCV 229E- and TGEV-receptor function of human, porcine and feline APN are not independent protein domains. Most of the sequenced mammalian APN proteins carry the EEE sequence at the position corresponding to pAPN amino acids 806–808, whereas the QDQ sequence is unique to the porcine APN protein (Table 3). As the introduction of the QDQ sequence into hAPN destroys the HCV 229E receptor function and the enzymatic activity, a second site in the pAPN protein must have a compensatory effect to ensure the proper biological function of pAPN.

Table 3. Receptor activity of mutated hAPN proteins. The susceptibility of transfected cells to the infection with HCV 229E and TGEV and the enzymatic activity of the mutated proteins are indicated

Mutant	Original hAPN sequence (residues)	Mutated sequence (residues)	Susceptibility to HCV 229E	Susceptibility to TGEV	Enzymatic activity
AP40	IHFRNN (731-736) + EEE (806-808)	QHFETL (731-736) + QDQ (806-808)	–	–	–
AP41	IHFRNN (731-736)	QHFETL (731-736)	+	–	+
AP42	EEE (806-808)	QDQ (806-808)	–	–	–

5. DISCUSSION

We have identified, by using chimeric APN proteins and site directed mutagenesis, a short amino acid segment in the hAPN protein which is critically involved in mediating the HCV 229E receptor function (Kolb et al., submitted). Additionally, we have demonstrated that the determinants in the fAPN protein which mediate reception of HCV 229E and TGEV are located in similar parts of the protein as the corresponding determinants in hAPN and pAPN (Kolb et al., 1996; Delmas et al., 1994). The utilisation of different determinants of the receptor protein may also reflect the serological and genetical differences between HCV 229E and the other group I coronaviruses like TGEV, FIPV and CCV (Siddell, 1995). In the absence of an APN crystal structure, however, it is difficult to predict whether these determinants are really independent protein domains or rather assemble into a single epitope which mediates coronavirus infections.

REFERENCES

Benbacer, L., Kut, E., Besnardeau, L., Laude, H., and Delmas, B., 1997, Interspecies aminopeptidase-N chimeras reveal species-specific receptor recognition by canine coronavirus, feline infectious peritonitis virus, and transmissible gastroenteritis virus, *J. Virol.* **71**: 734–737.

Delmas, B., Gelfi, J., L'Haridon, R., Vogel, L.K., Sjostrom, H., Noren, O., and Laude, H., 1992, Aminopeptidase N is a major receptor for the entero-pathogenic coronavirus TGEV, *Nature* **357**: 417–20.

Delmas, B., Gelfi, J., Sjostrom, H., Noren, O., and Laude, H., 1993, Further characterization of aminopeptidase-N as a receptor for coronaviruses, *Adv. Exp. Med. Biol.* **342**: 293–298.

Delmas, B., Gelfi, J., Kut, E., Sjostrom, H., Noren, O., and Laude, H., 1994, Determinants essential for the transmissible virus-receptor interaction reside within a domain of aminopeptidase-N that is distinct from the enzymatic site, *J. Virol.* **68**: 5216–24.

Grosse, B. and Siddell, S.G., 1994, Single amino acid changes in the S2 subunit of the MHV surface glycoprotein confer resistance to neutralization by S1 subunit-specific monoclonal antibody, *Virology* **202**: 814–824.

Kolb, A.F., Maile, J., Heister, A., and Siddell, S.G., 1996, Characterization of functional domains in the human coronavirus HCV 229E receptor, *J. Gen. Virol.* 2515–2521.

Rost, B., 1996, PHD: predicting one dimensional protein structure by profile based neural networks., *Meth. Enzymol.* **266**: 525–539.

Siddell, S.G., 1995, The Coronaviridae: an introduction, in: *The Coronaviridae* (S. G. Siddell), Plenum Press, New York, pp. 1–10.

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.

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 coronavirus 229E, *Nature* **357**: 420–2.