# Structural and Molecular Basis for Multifunctional Roles of Aminopeptidase N

A DISSERTATION
SUBMITTED TO THE FACULTY OF
UNIVERSITY OF MINNESOTA
BY

**Chang Liu**

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

**Supervisor: Dr. Fang Li**

May 2016

## Acknowledgements

First, I want to express my sincere gratitude to my advisor, Dr. Fang Li, for his long lasting support and all the inspiring ideas for my thesis research. His excellent mentorship inspired me greatly to pursue academic excellence during my graduate study. I also really appreciate all my committee members, Dr. Sundaram Ramakrishnan, Dr. Cheuk Leung and Dr. Chengguo Xing, who took time and attention to serve on my committee and provided me a lot of insightful advices for improvement of my thesis work.

In addition, I am very thankful for the scientists from our collaborating labs for their critical comments and great work. The collaborators that have contributed to my thesis research include the labs of Dr. Jianrong Li (Ohio State University), Dr. Shibo Jiang and Dr. Lanying Du (New York Blood Center). I also want to thank Dr. Carrie Wilmot and Dr. Erik Yukl for discussion and comments, Dr. Jay Nix at Advance Light Source beamline 4.2.2 for assistance in X-ray data collection and personnel at the Consortium for Functional Glycomics for help in glycan screen arrays.

This work was supported in part by funding provided by the Graduate School Fellowship and Doctoral Dissertation Fellowship (University of Minnesota Graduate School), as well as NIH Grants R01AI089728 and R01AI110700 (to Dr. Fang Li). Computer resources were provided by the Basic Sciences Computing Laboratory of the University of Minnesota Supercomputing Institute.

## Abstract

Aminopeptidase N (APN, also known as CD13 or alanine aminopeptidase) is a zinc-dependent metallopeptidase widely expressed on the cell surfaces of many different tissues, including kidney, intestine, endothelium and central nervous system [1]. Aside from regulating peptide metabolism in diverse physiological pathways, APN also plays pivotal roles in tumor cell motility, tumor-homing therapy, and coronavirus infection, all of which are seemingly unrelated to its aminopeptidase activity. How APN functions in these processes is largely a mystery. Therefore, in my thesis research, I used a combination of structural and functional studies to delineate the underlying mechanisms of APN-mediated tumor cell motility, tumor-homing therapy and coronavirus infection.

APN is an important target for tumor-homing therapy because it serves as a receptor for tumor-homing peptide (peptides that bring anti-cancer drugs to tumor cells) Asn-Gly-Arg (NGR)[2]. To provide mechanistic insights into this interaction, I determined the crystal structure of APN in complex with a tumor-homing peptide containing a representative NGR motif. The tumor-homing peptide binds to the APN's enzymatic active site, but it resists APN degradation due to a distorted scissile peptide bond. APN mediates tumor cell motility via its interactions with extracellular matrix (ECM) proteins. Using a variety of biochemical assays. I found that APN binds to, but does not degrade, NGR motifs in ECM proteins that share similar conformations with the NGR motif in the APN-bound tumor-homing peptide. Therefore, APN-based tumor cell motility and tumor-homing therapy rely on a unified mechanism in which both functions are driven by the specific

and stable interactions between APN and the NGR motifs in ECM proteins and tumor-homing peptides. This study not only elucidates the molecular basis for APN-based tumor cell motility and tumor-homing therapy, but also facilitates the development of APN-targeting cancer therapies.

Besides its above functions, APN is also the receptor for several coronaviruses such as human respiratory coronavirus 229E (HCoV-229E), porcine transmissible gastroenteritis virus (TGEV), feline enteric coronavirus (FCoV)[3-5]. Previous studies have revealed the coronavirus-binding sites on APN and they are not located in APN's active site[6-8]. In 2013, a TGEV-related coronavirus, porcine epidemic diarrhea coronavirus (PEDV), emerged in the United States and has significantly damaged America's pork industry. I investigated the receptor usage and cell entry of PEDV, and found that PEDV used APN from both pig and human and a sugar coreceptor N-acetylneuraminic acid (Neu5Ac) for cell entry. Moreover, PEDV was able to infect cells from pig, human, monkey, and bat. These results provide insights in the transmission pathway of PEDV, implicate PEDV as a potential threat to other species, and suggest antiviral strategies to control its spread.

Coronavirus spike proteins mediate viral entry into cells, a process that requires the spike proteins to be proteolytically activated. It has been a conundrum what proteases activate PEDV entry. To further characterize the cell entry process of PEDV, I systematically examined the roles of different host cell proteases in PEDV entry using pseudovirus entry and biochemical assays. I discovered that PEDV spike is activated by lysosomal cysteine

proteases, but not proprotein convertases or cell-surface serine proteases. Extracellular trypsin activates PEDV entry when lysosomal cysteine proteases are inhibited. I further pinpointed cathepsin L and cathepsin B as the lysosomal cysteine proteases that activate PEDV spike. These results advance our understanding of the molecular mechanism for PEDV entry and identify potential antiviral targets for curbing the spread of PEDV.

Collectively, my thesis work provides critical structural and molecular insights in better understanding the diverse functions of APN involved in pathological conditions of tumor progression and coronavirus infection. In addition, this work also lays foundation for therapeutic development to treat these APN-related health-threatening diseases.

## Table of Contents

# List of Tables

# List of Figures

# Abbreviations

| | |
|---|---|
| APN | Aminopeptidase N |
| pAPN | Porcine APN |
| ECM | Extracellular matrix |
| MIRAS | Multiple isomorphous replacement and anomalous scattering |
| FBS | Fetal bovine serum |
| DMEM | Dulbecco's modified Eagle medium |
| PBS | Phosphate-buffered saline |
| PBST | Phosphate-buffered saline with Tween-20 |
| RLU | Relative luciferase units |
| RFU | Relative fluorescence units |
| SDS | Sodium dodecyl sulfate |
| PAGE | Polyacrylamide gel electrophoresis |
| HCoV-229E | Human coronavirus 229E |
| TGEV | Porcine transmissible gastroenteritis coronavirus |
| FCoV | Feline coronavirus |
| PEDV | Porcine epidemic diarrhea coronavirus |
| NTD | N-terminal domain |
| CTD | C-terminal domain |
| RBD | Receptor-binding domain |
| Neu5Ac | *N*-acetylneuraminic acid |
| DTT | Dithiothreitol |

| | |
|---|---|
| hPPC | Human proprotein convertases |
| hLCP | Human lysosomal cysteine proteases |
| r.m.s.d. | Root mean square deviation |
| FITC | Fluorescein isothiocyanate |
| MOI | Multiplicity of infection |

## Summary of published manuscripts resulting from thesis work

**Liu C***, Yang Y*, Zheng Y, Zhou Yu, Jiang S, Li J, Du L, Li F Entry of porcine epidemic diarrhea coronavirus is activated by lysosomal proteases. Manuscript in Preparation

**Liu C***, Tang J*, Ma Y*, Liang X, Yang Y, Peng G, Qi Q, Jiang S, Li J, Du L, Li F (2015) Receptor usage and cell entry of porcine epidemic diarrhea coronavirus. ***J Virol*** 89: 6121-5

**Liu C***, Yang Y*, Chen L, Lin YL, Li F (2014) A unified mechanism for aminopeptidase N-based tumor cell motility and tumor-homing therapy. ***J Biol Chem*** 289: 34520-34

# 1 Chapter One

## A unified mechanism for aminopeptidase N-based tumor cell motility and tumor-homing therapy

**Chang Liu*,** Yang Yang*, Chen Lang, Yi-Lun Lin, Fang Li

(*Co-first authors)

***J Biol Chem*** 289: 34520-34529.

## 1.1 Introduction

Mammalian aminopeptidase N (APN), also called CD13, is a tumor marker that is overly expressed on the cell surface of almost all major tumor forms, including skin, ovary, lung, stomach, colon, kidney, bone, prostate, renal, pancreatic, thyroid, and breast cancers[9]. As a zinc-dependent aminopeptidase, APN cleaves the N-terminal neutral residue off physiological peptides and functions ubiquitously in various peptide metabolism pathways[1]. Consequently, tumor cell surface APN is required for tumor growth and development by cleaving and activating angiogenic peptides that are essential for tumor angiogenesis[10]. However, tumor cell surface APN also mediates tumor cell motility and serves as a receptor for tumor-homing peptides that guide anti-cancer drugs to tumor cells; these functions appear to be unrelated to each other or to APN's enzymatic activity[2,10-14]. It is puzzling how APN, a zinc-dependent aminopeptidase in essence, can function as a cell motility molecule and as a tumor-homing receptor. Solving these puzzles can lead to a better understanding of cancer biology and more effective cancer treatment.

APN is the most extensively studied member of the large M1 family of zinc-dependent aminopeptidases[1]. APN is a cell surface-anchored ectoenzyme, with a small N-terminal intracellular domain, a single-pass transmembrane anchor, a small extracellular stalk, and a large C-terminal ectodomain (Fig. 1.1A). We recently determined the crystal structures of porcine APN (pAPN) ectodomain alone and in complex with a cleaved polyalanine peptide as well as the crystal structure of a catalytically inactive mutant of pAPN

ectodomain in complex with an uncleaved polyalanine peptide[6]. The pAPN ectodomain has a seahorse-like shape, with four distinct domains: head, side, body, and tail. pAPN forms a homodimer through interactions between the head domains. The zinc-dependent catalytic site is located in a cavity surrounded by the head, side, and body domains. The catalytic site contains a tightly chelated zinc, a zinc-activated catalytic water that attacks and breaks the scissile peptide bond of peptide substrates, and a number of pAPN residues that either participate directly in catalysis or anchor peptide substrates in position for catalysis. In addition to binding peptides, pAPN also binds the exposed N terminus of proteins by undergoing a closed-to-open conformational change, which opens up its active site cavity. These results have elucidated the enzymatic activity of APN, but the puzzling roles of APN in tumor cell motility and tumor-homing therapy are still unclear.

APN plays a critical role in tumor cell migration and metastasis. It was previously shown that increased expression of APN on tumor cell surfaces greatly enhanced the migratory capacity of these tumor cells[15,16]. Moreover, decreasing the expression of APN on tumor cell surfaces or use of anti-APN antibodies or APN inhibitors to treat tumor cells blocked tumor cell migration and metastasis[12,17,18]. It was suggested that APN degrades extracellular matrix (ECM) proteins for tumor cell migration and metastasis[19,20]. However, tumor cells overexpressing enzymatically inactive APN also demonstrate enhanced migration and metastasis[17,21]. Thus, APN-mediated tumor cell motility and metastasis reside on some unknown activity of APN that is independent from its zinc-dependent aminopeptidase activity. In addition to mediating tumor cell motility, APN

also functions broadly in other cell motility and adhesion processes such as immune cell chemotaxis, sperm motility, and monocytic cell adhesion[22-28]. These functions of APN are reminiscent of those of integrins, which mediate cell motility and adhesion via specifically interacting with a three-residue motif, arginine-glycine-aspartate (RGD), in ECM proteins or on the surface of other cells[29-33]. Whereas the structures and functions of integrins have been extensively studied and well characterized, little is known about how APN mediates tumor cell motility or other cell motility and adhesion processes.

Tumor-homing therapy, also called targeted drug therapy, has recently emerged as one of the most promising approaches for cancer treatment[11,34,35]. The concept of tumor-homing therapy is to link anti-tumor drugs to a tumor-homing peptide; the latter specifically recognizes a receptor that is uniquely or overly expressed on tumor cell surfaces and thereby actively guides the drugs to tumor cell surfaces. APN and integrins are two of the most promising tumor-homing receptors; APN- and integrin-based tumor-homing therapies are the only ones that are currently in clinical trials[35-40]. Consistent with serving as receptors for RGD motifs in ECM proteins, integrins are also receptors for tumor-homing peptides containing an RGD motif[11,34,35]. The detailed interactions between integrins and the tumor-homing RGD peptide have been delineated by structural studies[29,32,33]. However, APN is a zinc-dependent aminopeptidase that is not known to recognize any specific peptide motifs, and thus it was totally unexpected when phage display identified APN as the receptor for tumor-homing peptides containing an Asn-Gly-Arg (NGR) motif[2,11,13]. Nevertheless, the APN-based tumor-homing NGR peptides

have been rapidly developed and are now in phase I and II clinical trials to treat advanced solid tumors[37-39]. Despite the great promises that APN-based tumor-homing therapy holds, the basic knowledge about this therapy is still lacking. It is not known where the binding site in APN is for the tumor-homing NGR peptides or what the detailed interactions are between APN and these peptides. The above knowledge would be essential for rational design and development of tumor-homing peptides with improved affinity and specificity for APN.

Here, we investigated the interactions between APN and ECM proteins and between APN and a tumor-homing NGR peptide. Using x-ray crystallographic and biochemical methods, we have identified a unified mechanism governing both APN-based tumor cell motility and tumor-homing therapy. These results not only solve the puzzles surrounding these APN-related functions, but also lay the foundation for future development of better APN-based cancer treatments. Furthermore, our study establishes APN as an integrin-like cell motility and adhesion molecule that should be investigated in depth and exploited therapeutically.

## 1.2 Materials and Methods

### Protein Expression and Purification

Porcine APN (pAPN) (GenBankTM accession number CAA82641.1) ectodomain (residues 62–963) was expressed and purified as described previously[6]. Briefly, pAPN containing an N-terminal honeybee melittin signal peptide and a C-terminal $His_6$ tag was

cloned into pFastbac1 vector, expressed in Sf9 insect cells using Bac-to-Bac expression system (Life Technologies), and secreted to cell culture medium. The $His_6$-tagged pAPN was harvested and purified sequentially on HiTrap nickel-chelating HP column and Superdex 200 gel filtration column (GE Healthcare). Fc-tagged pAPN was obtained by fusion of the human $IgG_4$ Fc region to the C terminus of the pAPN ectodomain and was purified sequentially on HiTrap Protein G HP column and Superdex 200 gel filtration column. pAPN containing the E384Q mutation was constructed using site-directed mutagenesis, and it was expressed and purified in the same way as the wild type pAPN.

The human fibronectin gene that encodes residues 1–275 (including the free N terminus and domains 1–5) (GenBankTM accession number BAD52437.1) was synthesized commercially (Genscript). Constructs that encode different parts of the above fibronectin region along with a C-terminal His6 tag were cloned into vector pET42b(+) (Merck Millipore) and expressed in Escherichia coli BL21 cells (Life Technologies) through induction with isopropyl β-D-1-thiogalactopyranoside (Sigma-Aldrich). Different fibronectin domains with a C-terminal $His_6$ tag were purified subsequently on a HiTrap nickel-chelating HP column and Superdex 200 gel filtration column.

CNGRCG peptide and its mutants were fused to a C-terminal GST tag using a SGSGSGSG peptide linker. Constructs were cloned into vector pET42b(+) and expressed as described previously for fibronectin domains. Recombinant proteins were purified on GSTrap 4B column (GE Healthcare) and Superdex 200 gel filtration column sequentially.

All the above recombinant proteins used in this study were stored in buffer containing 20 mM Tris-HCl, pH 7.4, and 200 mM NaCl.

**Crystallization and Structural Determination**

Crystallization of $His_6$-tagged pAPN ectodomain was carried out as described previously[6]. Briefly, crystallization was set up in sitting drops at 4 °C by adding 2 μl of protein solution to 2 μl of well solution containing 18% (v/v) PEG3350, 200 mM $Li_2SO_4$, and 100 mM HEPES, pH 7.2. Crystals appeared in 2 days and were allowed to grow for another 2 weeks. The crystals were then transferred to buffer containing 5 mM CNGRCG (AnaSpec), 20% (v/v) ethylene glycol, 25% (v/v) PEG3350, 200 mM $Li_2SO_4$, and 100 mM HEPES, pH 7.2. After 2 days, the crystals were flash-frozen in liquid nitrogen. X-ray diffraction data were collected at ALS beamline 4.2.2 and processed using software HKL2000[41]. The structure of the pAPN-CNGRCG complex was determined by molecular replacement using the pAPN structure as the search template (PDB code 4FKE). The $F_o - F_c$ omit electron density map was calculated in the absence of CNGRCG and showed strong density for CNGRCG. Based on the $F_o - F_c$ omit electron density map, the model of CNGRCG was built. The structure of the complex was subsequently refined at 1.95 Å resolution using software CNS and Refmac[42,43].

**APN Catalysis Inhibition Assay**

The APN catalysis inhibition assay was carried out as described previously[44]. Briefly, 2 nM $His_6$-tagged pAPN and 0.5 mM L-alanine-*p*-nitroanilide (Sigma-Aldrich) were

incubated in 100 μl of 60 mM $KH_2PO_4$, pH 7.2, in the presence of gradient concentrations of CNGRCG or GNGRG peptide (Genscript). The reactions were incubated at 37 °C for 30 min. Formation of product *p*-nitroanilide was measured every 10 min using an absorbance plate reader (BioTek) at 405 nm. The $IC_{50}$ value was defined as the concentration of each peptide that led to 50% of maximal pAPN catalytic activity. $K_i$ values for each peptide were calculated from the $IC_{50}$ value using the Cheng-Prusoff equation, $K_i = IC_{50}/(1 + [S]/K_m)$, in which $K_m$ was determined previously[6].

**Wound Healing Assay**

Tumor cell wound healing assay was performed as described previously[45]. Briefly, $8 \times 10^4$ human fibrosarcoma HT-1080 cells were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (Invitrogen) and seeded onto 24-well plates. After serum starvation for 16 h, the scratch wounds were created on the confluent monolayers with a pipette tip. Each well was washed twice with serum-free media to remove cell debris. Then the cells in triplicates were treated with 10 μg/ml anti-APN antibody WM15 (BD Pharmingen) or 10 μg/ml anti-integrin αV/β3 antibody (Santa Cruz Biotechnology) for 8 h. Digital images at different time points were captured using an inverted contrasting microscope (Leica Microsystems). The wound healing effect was calculated as the distance of cells migrating into cell-free spaces compared with the initial wound. The relative migration was standardized against the control group without any antibody treatment.

**Transwell Migration Assay**

Transwell migration assay was conducted as described previously[46]. Briefly, Transwell inserts (Corning Glass) with 6.5-mm diameter and 8-μm pore size were coated with 10 μg/cm$^2$ fibronectin (Sigma-Aldrich) and air-dried. HT-1080 cells were seeded at a density of $8 \times 10^4$ and treated with 10 μg/ml anti-APN antibody WM15. 4 h after treatment, cells in the upper compartment were scraped off with a cotton swab. Cells passing through the membrane were fixed with 5% glutaraldehyde (Sigma-Aldrich) and stained with 0.5% crystal violet (Sigma-Aldrich). For quantification, cells that had migrated to the lower surface were counted under a microscope in three fields for triplicate experiments. The relative migration was standardized against the control group without antibody treatment.

**Dot-blot Hybridization Assay**

Dot-blot hybridization assay was carried out as described previously[47]. Briefly, 10 μg of fibronectin, laminin, or type IV collagen (Sigma-Aldrich) was each dotted onto a nitrocellulose membrane. The membranes were dried completely and blocked with BSA at 4 °C overnight. The membranes were then incubated at 37 °C for 2 h with 50 μg/ml $His_6$-tagged pAPN, which had been preincubated alone or with 20 μg/ml fibronectin domains 4–5, anti-APN antibody WM15, or anti-integrin αV/β3 antibody. The membranes were then washed five times with phosphate-buffered saline with Tween 20 (PBST), incubated with anti-$His_6$ mouse monoclonal IgG1 antibody (Santa Cruz Biotechnology) at 37 °C for 2 h, washed five times with buffer PBST again, incubated with HRP-conjugated goat anti-mouse IgG antibody (Santa Cruz Biotechnology) at 37 °C

for 1 h, and washed five times with buffer PBST. Finally, the bound proteins were detected using ECL Plus (GE Healthcare).

**ELISA**

Binding of pAPN to different ECM proteins was carried out using ELISA as described previously[48]. Briefly, ELISA plates were coated overnight at 4 °C with 10 μg/ml fibronectin, laminin, type IV collagen, or PBS. After blocking at 37 °C for 2 h, the plates were incubated at 37 °C for 2 h with 50 μg/ml $His_6$-tagged pAPN, which had been preincubated alone or with 20 μg/ml fibronectin domains 4–5 or anti-APN antibody WM15. The plate was then treated the same way as in the dot-blot hybridization assay. Finally, the bound proteins were detected using HRP substrate (G-Bioscience), and the color reaction was quantified using an absorbance plate reader (BioTek) at 630 nm.

**AlphaScreen Protein-Protein Binding Assay**

Binding of His6-tagged pAPN to GST-tagged proteins or peptides (e.g. fibronectin domains 4–5, NGR peptide, or its mutants) was carried out using AlphaScreen assay as described previously[48]. Briefly, each of the GST-tagged proteins or peptides at a final concentration of 30 nM was mixed with $His_6$-tagged pAPN also at a final concentration of 30 nM in ½ AreaPlate (PerkinElmer Life Sciences) for 1 h at room temperature. AlphaScreen anti-GST acceptor beads and nickel chelate donor beads (PerkinElmer Life Sciences) were added to the mixture at final concentrations of 10 μg/ml each. The mixture was incubated at room temperature for 1 h and protected from light. The assay

plates were read in an EnSpire plate reader (PerkinElmer Life Sciences).

Binding of Fc-tagged pAPN to $His_6$-tagged fibronectin domains was carried out in the same way as above, except that fibronectin domains had a final concentration of 9 nM and that AlphaScreen protein A acceptor beads (PerkinElmer Life Sciences) and AlphaScreen nickel chelate donor beads were added to the mixture. To block the binding interaction between pAPN and fibronectin domains 4–5, bestatin, methionine, or CNGRCG peptide at various concentrations was incubated with the mixture for 1 h before donor and acceptor beads were added.

**SDS-PAGE**

Fibronectin domains 4–5 at a final concentration of 0.5 μg/μl was incubated alone or with pAPN at a final concentration of 0.5 μg/μl at 37 °C for 2 h. Subsequently, the mixture was subjected to SDS-PAGE. The gel was stained by Brilliant Blue G (Sigma-Aldrich).

**Mass Spectrometry**

Fibronectin domains 4–5 at a final concentration of 100 μM was incubated alone or with pAPN at a final concentration of 10 μM at 37 °C for 2 h. Subsequently, the mixture was subjected to mass spectrometry at the Center for Mass Spectrometry and Proteomics at the University of Minnesota (Minneapolis, MN).

## 1.3 Results

### Crystal Structure of Porcine APN in Complex with a Tumor-homing NGR Peptide

To investigate the structural basis for the interactions between APN and tumor-homing NGR peptides, we determined the crystal structure of porcine APN (pAPN) in complex with peptide CNGRCG. CNGRCG was chosen for the study because it is the most commonly used tumor-homing NGR peptide and targets tumor cells more effectively than other NGR peptides such as GNGRG[11,49]. pAPN was chosen for this study because it was previously crystallized in a closed and catalytically active conformation under physiologically relevant pH, and the crystals diffracted to high resolution (~2.0 Å)[6]. In addition, pAPN and human APN share high sequence homology, with catalytic residues 100% conserved between the two proteins[6]. To crystallize the pAPN-CNGRCG complex, pAPN was expressed, purified, and crystallized as described previously, and CNGRCG was soaked into pAPN crystals as described previously for other APN-binding ligands[6]. The structure of the complex was determined by molecular replacement using the unliganded pAPN structure as the search template (Fig. 1.1A). The Fo − Fc omit electron density map calculated in the absence of CNGRCG showed strong density for this peptide, allowing the model to be built (Fig. 1.1B). Subsequently, the structure of the complex was refined at 1.95 Å resolution (Table 1.1).

The CNGRCG tumor-homing peptide binds to the catalytic site of pAPN (Fig. 1.1, C and D). The two cysteines in CNGRCG form a disulfide bond (Fig. 1.1B). Consequently, the NGR region forms a short loop with a sharp turn, facilitated by the presence of a flexible

glycine in the middle of the loop. The NGR loop makes sequence-specific interactions with the pAPN active site through the side chains of asparagine and arginine (Fig. 1.2, A and B). Specifically, the side chain of asparagine in the NGR loop is parallel to the side chain of pAPN residue His-383; it also forms five water-mediated hydrogen bonds with the side chains of pAPN residues Glu-413, Ser-410, and Glu-384, and one water-mediated interaction with the carbonyl oxygen of pAPN residue Val-380 (Fig. 1.2A). Moreover, the side chain of the arginine in the NGR loop forms a hydrophobic interaction with Cα of pAPN residue Ala-346, one hydrogen bond with the side chain of pAPN residue Asn-345, two hydrogen bonds with the carbonyl oxygen of pAPN residue Asn-345, and a water-mediated hydrogen bond with the side chain of pAPN residue Arg-358 (Fig. 1.2B). Although glycine in the NGR loop has no direct contact with pAPN, its conformational flexibility is critical for the formation of the NGR loop (Fig. 1.2C). CNGRCG also makes sequence nonspecific interactions with the pAPN active site through the main chain groups, as described previously for pAPN-bound polyalanine peptide (Fig. 1.2D)[6]. Because all of the APN residues involved in binding CNGRCG are completely conserved between human and porcine APNs (Fig. 1.2, A, B, and D)[6], the CNGRCG peptide is expected to interact with both APNs in the same way.

To further understand the APN/CNGRCG interactions, we investigated the roles of both the NGR motif and the disulfide bond-fortified loop conformation in APN binding using biochemical methods. To this end, we introduced mutations into CNGRCG, and fused CNGRCG and each of the mutant peptides to a C-terminal GST tag. We then measured

the binding affinity between APN and the GST-tagged peptides using AlphaScreen protein-protein binding assay. Mutations of each of the residues in the NGR motif to an aspartate significantly weakened APN binding and so did mutations of each of the two flanking cysteines in CNGRCG to a glycine (Fig. 1.2E). Furthermore, we evaluated the contribution of the disulfide bond in CNGRCG to APN binding using APN catalysis inhibition assay. CNGRCG inhibits APN catalysis more effectively than GNGRG (Fig. 1.2F), again suggesting that the loop formation in CNGRCG contributes critically to APN binding. Indeed, soaking the GNGRG peptide into pAPN crystals did not yield any clear electron density of the peptide, consistent with the weak APN binding by GNGRG. Taken together, both structural and biochemical data revealed that the interactions between APN and CNGRCG depend on the specific sequence of the NGR motif as well as the loop conformation of the peptide. These sequence-specific and conformation-dependent interactions between APN and CNGRCG explain the high affinity and specificity of APN binding to the tumor-homing peptide; they also define APN as a functional receptor for the NGR motif with a loop conformation in tumor-homing peptides or other biological settings.

The CNGRCG tumor-homing peptide resists enzymatic degradation by pAPN. Our previous study demonstrated that the crystallized pAPN is catalytically active, and when peptide substrate polyalanine was soaked into the pAPN crystals, electron density showed a broken scissile peptide bond and hence a degraded polyalanine (Fig. 1.3A). In contrast, when soaked into crystals of a catalytically inactive mutant of pAPN, polyalanine

remained uncleaved (Fig. 1.3B). In this study, CNGRCG was soaked into catalytically active pAPN crystals under the same soaking condition as for polyalanine. Electron density revealed that CNGRCG remained uncleaved in the crystals (Fig. 1.1B), suggesting that CNGRCG is a much poorer substrate than polyalanine for APN. To investigate why CNGRCG resists APN degradation, we compared the active site geometry of the pAPN-CNGRCG complex with that of the pAPN-polyalanine complex. At the active site of the pAPN-polyalanine complex (Fig. 1.3C), a zinc-activated catalytic water attacks and breaks the N-terminal scissile peptide bond of the peptide substrate; simultaneously, the catalytic water transfers a proton through the pAPN residue Glu-384 to the leaving nitrogen group of the peptide substrate, which subsequently becomes the N terminus of the newly formed peptide product[6]. At the active site of the pAPN-CNGRCG complex, however, the scissile peptide bond of CNGRCG deviates from the optimal geometry required for peptide bond hydrolysis due to the sharp turn of the NGR loop (Fig. 1.2C). The outcome is that the leaving nitrogen group of CNGRCG is too far away from Glu-384 to accept the transferring proton, leading to a disconnected proton transfer pathway and an intact CNGRCG (Fig. 1.2D). Consequently, CNGRCG can serve as an efficient and stable tumor-homing peptide by targeting tumor cell surface APN without being turned over immediately during the tumor-homing process.

**Interactions between APN and Extracellular Matrix Proteins**

The above structural study has established APN as a functional receptor for the NGR motif in tumor-homing peptides developed in vitro. Our structural finding raises the

possibility that the NGR motifs may exist in vivo and that APN may interact with in vivo NGR motifs to perform its physiological functions. Indeed, it was previously observed that fibronectin, one type of the ECM proteins, contains four NGR motifs[49]. In this study, we looked into the sequences of other types of ECM proteins and found that laminin contains three NGR motifs, whereas type IV collagen contains none. Then we investigated whether APN mediates tumor cell motility by interacting with ECM proteins containing NGR motifs. To this end, we examined the role of APN in tumor cell motility using both the tumor cell wound healing assay and transwell migration assay. The used cell line was HT-1080 (human fibrosarcoma). In the wound healing assay, tumor cells secrete a mixture of ECM proteins into the extracellular environment such that tumor cells can move by interacting with these secreted ECM proteins. The results showed that both anti-APN antibody and anti-integrin antibody significantly decreased tumor cell motility (Fig. 1.4, A and B). In the transwell migration assay, tumor cell motility was measured in a fibronectin-coated transwell chamber. The motility of HT-1080 was significantly reduced by anti-APN antibody (Fig. 1.4C). Both of these tumor cell motility assays demonstrated that APN plays an important role in tumor cell motility by interacting with ECM proteins (e.g. fibronectin) in the extracellular environment. Next, we analyzed the interactions between pAPN and individual ECM proteins using both dot-blot hybridization and ELISA assays (Fig. 1.5, A and B). pAPN specifically binds fibronectin and laminin, both of which contain NGR motifs, but not type IV collagen that does not contain any NGR motif. In addition, these interactions were significantly inhibited by anti-APN antibody (Fig. 1.5, A and B). These results suggest that APN

specifically interacts with ECM proteins containing NGR motifs, contributing to the APN-mediated cell motility.

We further investigated the detailed interactions between APN and ECM proteins containing NGR motifs. First, we examined the interactions between APN and fibronectin using AlphaScreen protein-protein binding assay. Because APN is capable of interacting with the exposed N terminus of proteins, we investigated whether APN binds fibronectin at its exposed N terminus or a specific domain containing the NGR motif (i.e. NGR domain) (Fig. 1.6A). To this end, we expressed different parts of fibronectin with or without the exposed N terminus or the NGR domain, and analyzed their binding interactions with pAPN (Fig. 1.6B). Although domain 5 contains the NGR motif, it needed to be expressed together with domain 4 due to the interactions between the two domains [50]. The results from the AlphaScreen assay showed that pAPN has significantly higher affinity for the NGR domain than for the exposed N terminus of fibronectin. Moreover, this NGR domain competitively blocked the interactions between pAPN and fibronectin or laminin (Fig. 1.5, A and B). In addition, when this NGR domain was fused to a C-terminal GST tag, the GST-tagged NGR domain has significantly higher pAPN-binding affinity than GST alone (Fig. 1.6C). Thus, APN primarily binds fibronectin in the NGR domain. Second, we analyzed the interactions between APN and the fibronectin NGR domain containing mutations in the NGR motif (Fig. 1.6D). The results from AlphaScreen assay showed that mutations in the NGR motif significantly decreased the binding affinity between pAPN and the fibronectin NGR domain. Thus, APN specifically

interacts with the NGR motif in the fibronectin NGR domain. Finally, we probed the binding site of the fibronectin NGR motif in APN using three active-site inhibitors of APN, methionine, bestatin, and the CNGRCG peptide (Fig. 1.6E). The results from the AlphaScreen assay showed that active-site inhibitors of APN inhibited the binding of the fibronectin NGR motif to pAPN, supporting direct binding of the fibronectin NGR motif to the APN active site. Moreover, a catalytically inactive mutant pAPN containing an E384Q mutation in the catalytic site bound the fibronectin NGR domain in the same way as the wild type pAPN did, implying that the catalytic activity of APN is not a prerequisite for binding the fibronectin NGR domain (Fig. 1.6E). Here, the E384Q mutation at the pAPN active site did not have a significant effect on the binding of pAPN to the fibronectin NGR domain or any of the active site inhibitors due to the numerous other interactions between pAPN and its ligands[6]. Collectively, these results suggest that APN specifically binds the NGR motif of ECM proteins and that the binding site for the NGR motif of ECM proteins is located at the APN active site.

The NGR motif of fibronectin shares similar structural conformations with the NGR motif in the APN-bound tumor-homing peptide. As revealed by the NMR solution structure of the fibronectin NGR domain (i.e. domain 5), the NGR motif is located on an exposed and extruding loop that is connected to a stem region (Fig. 1.7A)[50]. The NGR loop has a sharp turn, facilitated by the glycine residue in the middle of the NGR motif and two other glycine residues flanking the NGR motif. The conformation of the NGR loop depends on the tertiary structure of fibronectin, including the inter-domain

interactions between domain 5 and domain 4. As a result, efforts to prepare the NGR loop outside the context of the fibronectin domains have been unsuccessful in yielding the crystal structure of APN in complex with the fibronectin NGR loop. Nevertheless, comparison of the structural conformation of the fibronectin NGR loop to that of APN-bound CNGRCG shows that the two NGR loops share similar structural conformations, suggesting that APN interacts with the fibronectin NGR motif and the tumor-homing NGR peptide using similar structural mechanisms (Fig. 1.7A). Moreover, as demonstrated by SDS-PAGE and mass spectrometry, incubation of the fibronectin NGR domain with APN in solution did not lead to APN cleavage of the NGR motif of the fibronectin NGR domain, indicating that APN binding to the NGR motifs in ECM proteins is stable and nondamaging (Fig. 1.7, B and C).

## 1.4 Discussion

It is puzzling how APN, a zinc-dependent aminopeptidase in essence, can mediate tumor cell motility and serve as a receptor for tumor-homing peptides. Here, we systematically investigated the underlying molecular and structural mechanisms for these functions of APN that are seemingly unrelated to each other or to APN's enzymatic activity. Our study has identified a unified mechanism for APN-based tumor cell motility and tumor-homing therapy where APN carries out these functions by specifically interacting with the NGR motif in ECM proteins and in tumor-homing peptides, respectively (Fig. 1.8). The NGR motifs in ECM proteins and in tumor-homing peptides share similar structural conformations by both forming a short loop with a sharp turn. The structural

conformations of the NGR motifs are stabilized either by local tertiary structure as in ECM proteins or by a flanking disulfide bond as in tumor-homing peptides. The binding site of the NGR motifs in APN is located at the zinc-aminopeptidase active site. APN recognizes the NGR motifs in a sequence-specific and conformation-dependent manner. Despite binding to the APN active site, the NGR motifs resist APN degradation because the presumed scissile peptide bonds are in catalytically inactive conformations. Therefore, the interactions between APN and the NGR motifs are specific and stable, allowing APN to provide traction for tumor cell motility and to also serve as a receptor for tumor-homing peptides. The structural information provided in this study on the detailed interactions between APN and the NGR motif can guide the design and development of tumor-homing peptides that efficiently target tumor cell surface APN and inhibitors that effectively block APN-mediated tumor cell motility and metastasis.

The implications of our findings go beyond APN-based tumor cell motility and tumor-homing therapy; this study has established APN as an integrin-like cell motility and adhesion molecule. Indeed, despite sharing neither sequence similarity nor catalytic activity, APN and integrins resemble each other in several important ways (Fig. 1.8). First, integrins mediate cell motility and adhesion via binding to their signature RGD motifs in ECM proteins and on the surface of other cells. Like integrins, APN functions not only in tumor cell motility but also in other cell motility and adhesion processes such as immune cell chemotaxis, sperm motility, and monocytic cell adhesion[22-28]. Here, we propose that APN mediates these other cell motility and adhesion processes by binding to

its signature NGR motifs in ECM proteins and on the surface of other cells. Second, both APN and integrins relay signal transduction between cells and the extracellular environment[1]. The closed-to-open conformational changes of APN may contribute to APN-mediated signal transduction[6]. Third, both APN and integrins serve as functional receptors for tumor-homing peptides that contain NGR and RGD motifs, respectively. It is worth noting that although the above functions represent the main function of integrins, they are only secondary functions for APN, whose main role is to regulate the metabolism of peptides. Whereas integrins have been extensively studied and therapeutically targeted, our knowledge about APN has been rather limited. This study has laid the foundation for a better understanding of the physiological functions and therapeutic implications of APN.

**Table 1.1 Data collection and refinement statistics**

| | pAPN/CNGRCG complex |
|---|---|
| **Data Collection** | |
| Space group | C2 |
| Cell dimensions | |
| a, b, c (Å) | 260.3, 62.9, 82.0 |
| α, β, γ (°) | 90, 100.6, 90 |
| Resolution (Å) | 50-1.92 (1.96-1.92)[a] |
| $R_{sym}$ or $R_{merge}$ | 0.079 (0.615) |
| $I/\sigma I$ | 20.7 (1.7) |
| Completeness (%) | 97.9 (97.0) |
| Redundancy | 3.8 (3.9) |
| **Refinement** | |
| Resolution (Å) | 47.73-1.95 |
| No. reflections | 88780 |
| $R_{work}/R_{free}$ | 0.140 / 0.189 |
| No. atoms | 8498 |
| Protein | 7259 |
| CNGRCG | 40 |
| *B*-factors (Å$^2$) | 48.9 |
| Protein | 47.3 |
| CNGRCG | 44.0 |
| R.M.S. derivations | |
| Bond lengths (Å) | 0.012 |
| Bond angles (°) | 1.614 |
| Ramachandran plot | |
| Favored (%) | 97 |
| Allowed (%) | 2.6 |
| Disallowed (%) | 0.4 |

[a]Values in parentheses are for highest-resolution shell.

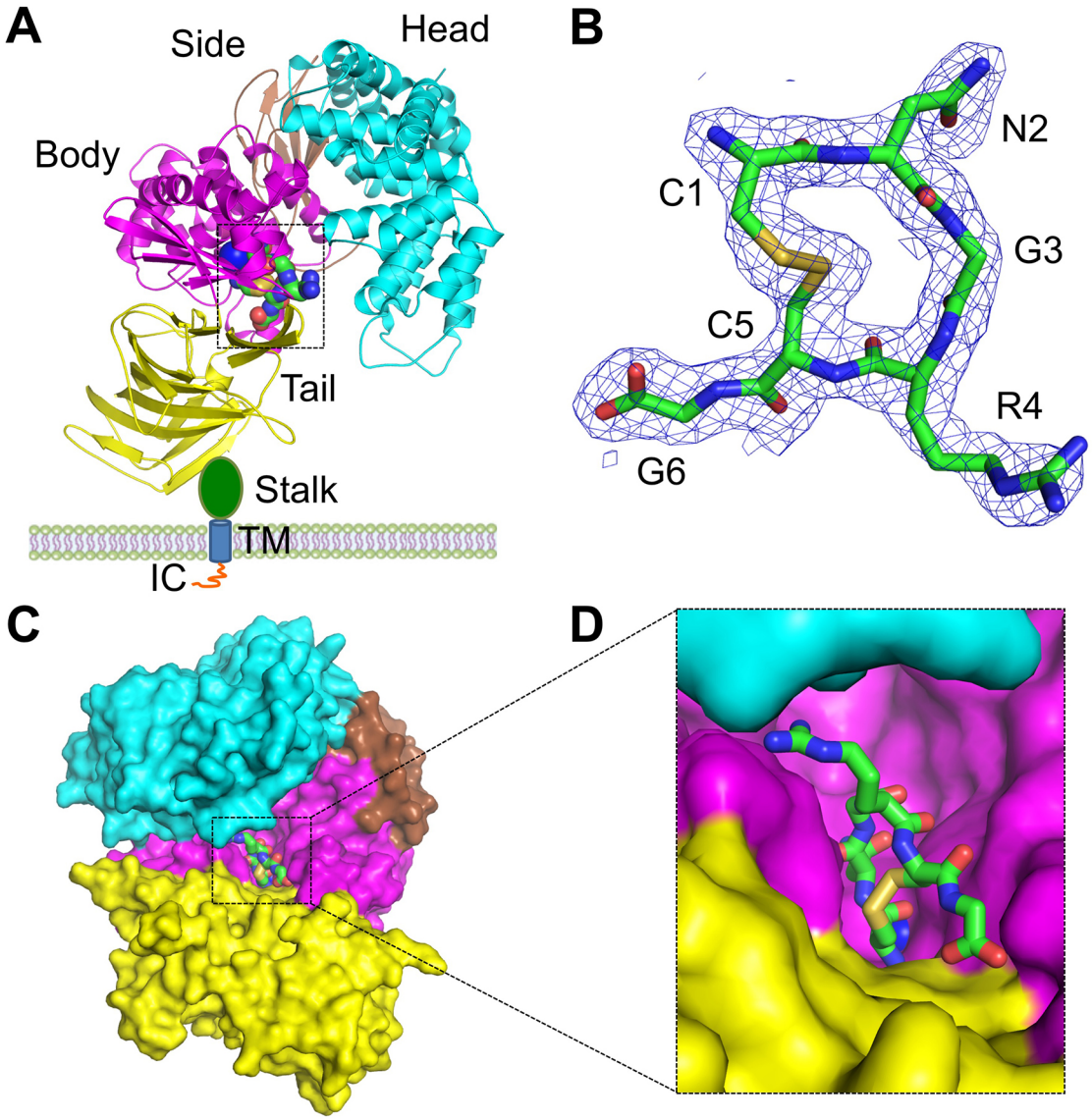

A
Side
Head
Body
Tail
Stalk
TM
IC
B
N2
C1
G3
C5
R4
G6
C
D

**Figure 1.1 Crystal structure of porcine APN in complex with a tumor-homing peptide CNGRCG.**

(A) Overall structure of the pAPN-CNGRCG complex. pAPN contains an ectodomain, a stalk, a transmembrane anchor (TM), and an intracellular tail (IC). The ectodomain contains four domains: head (cyan), side (brown), body (magenta), and tail (yellow). CNGRCG is shown in green as balls and sticks. Zinc is shown as a blue ball. Only one monomer of the dimeric pAPN is shown. (B) Electron density map of CNGRCG. The electron density map corresponds to $F_o - F_c$ omit map calculated in the absence of CNGRCG and contoured at 2.2 σ. (C) Another view of the pAPN-CNGRCG complex. The view of the complex is obtained by rotating the view in A first by $90^o$ along a vertical axis and then by $45^o$ along a horizontal axis, in such a way that the active site cavity of pAPN is facing the reader. (D) Enlarged view of CNGRCG in the active site of pAPN.

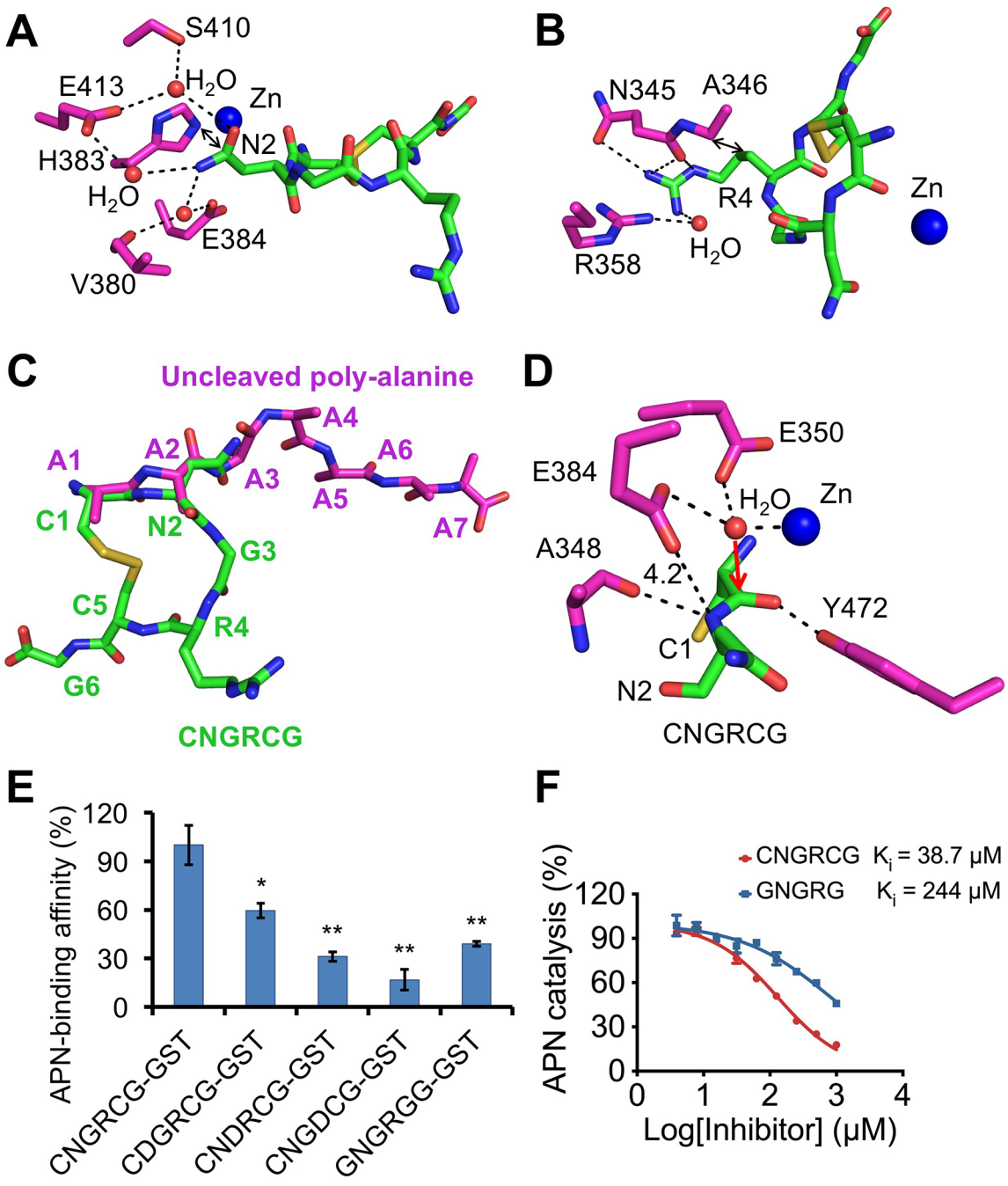
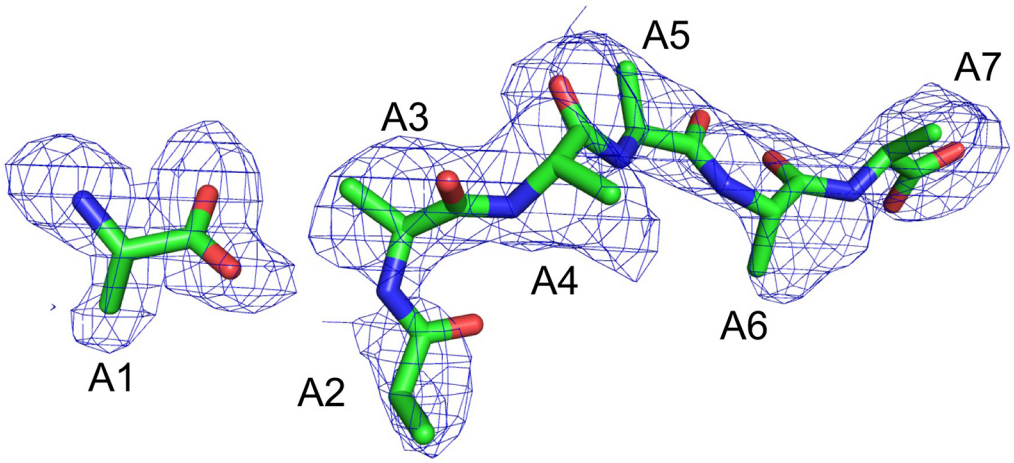

A
S410
E413
H2O
Zn
N2
H383
H2O
E384
V380
B
N345
A346
R4
Zn
H2O
R358
C
Uncleaved poly-alanine
A1
A2
A3
A4
A5
A6
A7
C1
N2
G3
C5
R4
G6
CNGRCG
D
E350
E384
H2O
Zn
A348
4.2
Y472
C1
N2
CNGRCG
E
APN-binding affinity (%)
120
90
60
30
0
*
**
**
**
CNGRCG-GST
CDGRCG-GST
CNDRCG-GST
CNGDCG-GST
GNGRGG-GST
F
CNGRCG Ki = 38.7 μM
GNGRG Ki = 244 μM
APN catalysis (%)
120
90
60
30
0
0
1
2
3
4
Log[Inhibitor] (μM)

**Figure 1.2 Detailed interactions between pAPN and CNGRCG.**

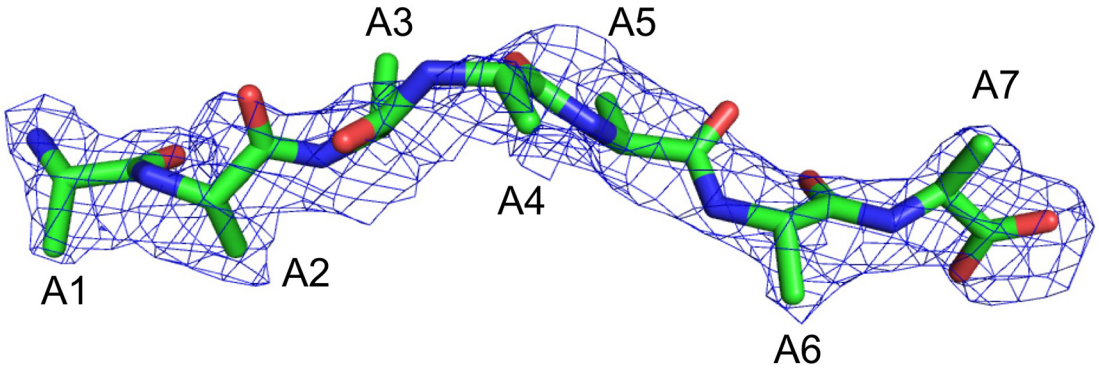
(A) Detailed interactions between pAPN and the side chain of the asparagine in CNGRCG. pAPN residues are in magenta, and CNGRCG is in green. (B) Detailed interactions between pAPN and the side chain of the arginine in CNGRCG. (C) Comparison of the conformations of CNGRCG in the crystal of the wild type pAPN-CNGRCG complex and the uncleaved polyalanine peptide in the crystal of mutant pAPN-polyalanine complex (PDB code 4NAQ). (D) Active site geometry of the pAPN-CNGRCG complex. The presumable scissile peptide bond of CNGRCG has a catalytically inactive conformation, resulting in its leaving nitrogen group being too far away from the proton-transferring pAPN resi- due Glu-384. Red arrow indicates the potential attack of the scissile peptide bond by the catalytic water at the pAPN active site. Unit of distance is in angstroms. (E) pAPN binding by CNGRCG and its mutants. CNGRCG and the mutant peptides were fused to a C-terminal GST tag. The binding affinities of these fusion proteins with pAPN were measured using AlphaScreen protein-protein binding assay. The binding affinity of GST-tagged CNGRCG with pAPN was used as the standard and taken as 100%. Error bars indicate S.E. (compared with the standard two-tailed t test; *, $p < 0.05$; **, $p < 0.01$; $n = 3$). (F) Inhibition of APN catalytic activity by CNGRCG and GNGRG peptides. The catalytic activity of pAPN on Ala-*p*-nitroanilide in the absence of any inhibitor was taken as 100%. Error bars indicate S.E. ($n = 3$).

**A**



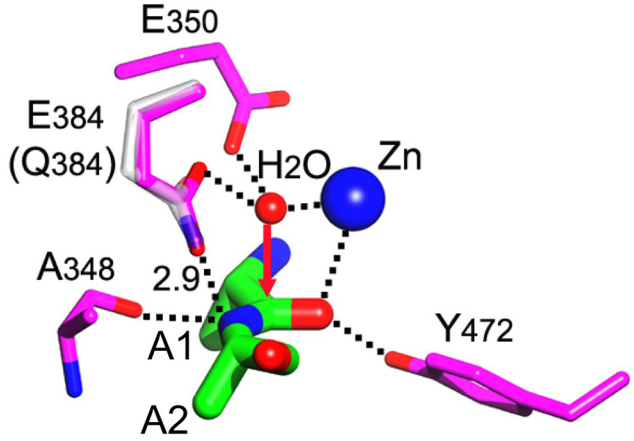

Cleaved poly-alanine in wild-type pAPN crystal

**B**




Uncleaved poly-alanine in E384Q mutant pAPN crystal

**C**




Uncleaved poly-alanine

**Figure 1.3 Previously published catalytic mechanism of APN.**

(A) Cleaved polyalanine peptide bound to wild type pAPN where the peptide was degraded in the crystal (PDB code 4NZ8). The electron density map corre- sponds to $F_o - F_c$ omit map calculated in the absence of the cleaved polyalanine and contoured at 1.5 σ. (B) Uncleaved polyalanine peptide bound to E384Q mutant pAPN, which is catalytically inactive (PDB code 4NAQ). The electron density map corresponds to the $F_o - F_c$ omit map calculated in the absence of the uncleaved polyalanine and contoured at 2.0 σ. (C) Interactions between catalytic residues of pAPN (magenta) and scissile peptide bond of polyalanine (green). Although Gln-384 (white) was introduced to pAPN to generate a catalytically incompetent enzyme for crystallographic studies (PDB code 4NAQ), Glu-384 (magenta) from wild type pAPN (PDB code 4FKE) was grafted here to illustrate the catalytic mechanism of pAPN. Red arrow indicates the potential attack of the scissile peptide bond by the catalytic water at the pAPN-active site. Zinc is shown as a blue ball and catalytic water as a red ball. Unit of distance is in angstroms.

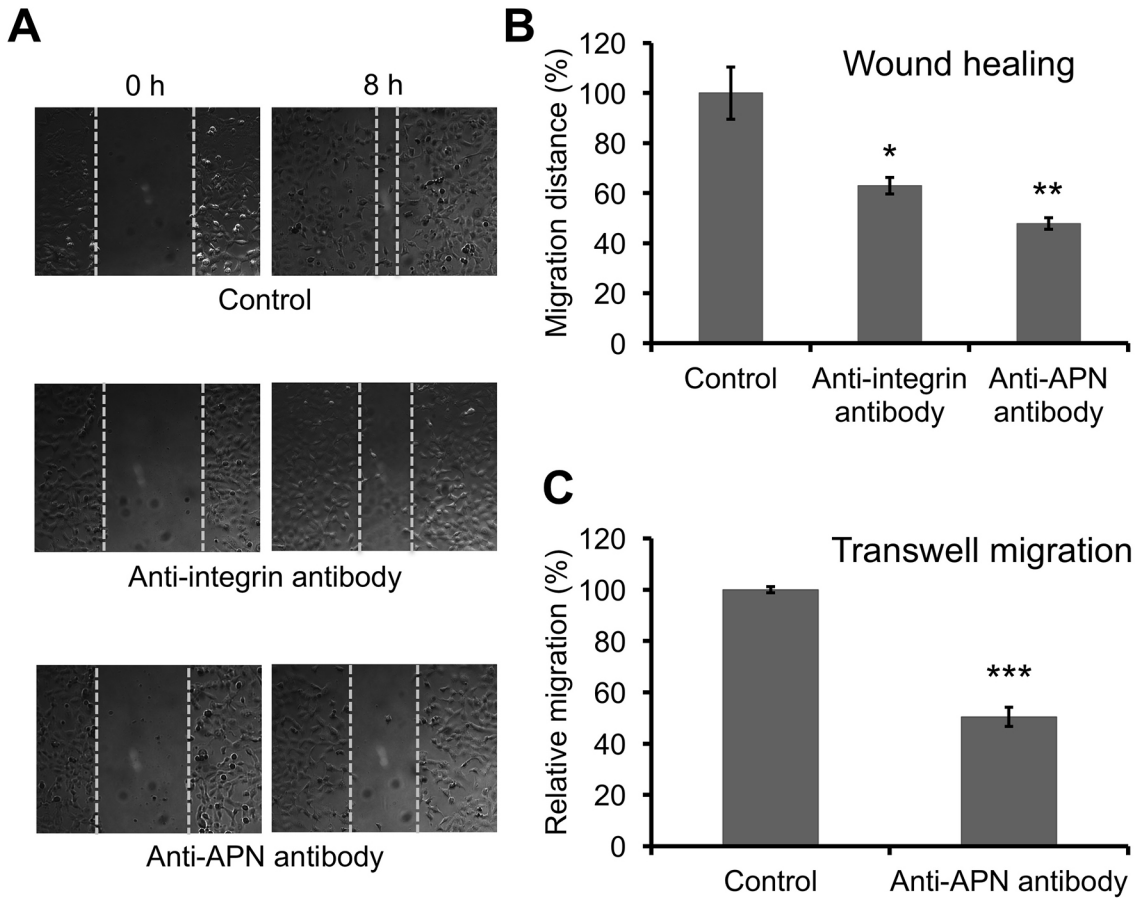



**Figure 1.4 Tumor cell motility assays.**

(A) Microscopic photos showing the inhibition of HT-1080 cell motility by anti-APN or anti-integrin antibody in tumor cell wound healing assay. (B) Quantification of HT-1080 cell motility in wound healing assay. The migration distance in the control group without any antibody treatment was taken as 100%. (C) Transwell migration assay showing the inhibition of HT-1080 cell motility by anti-APN antibody. The relative migration was standardized against the control group without anti- body treatment. Error bars indicate S.E. (two-tailed t test; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; $n = 3$).

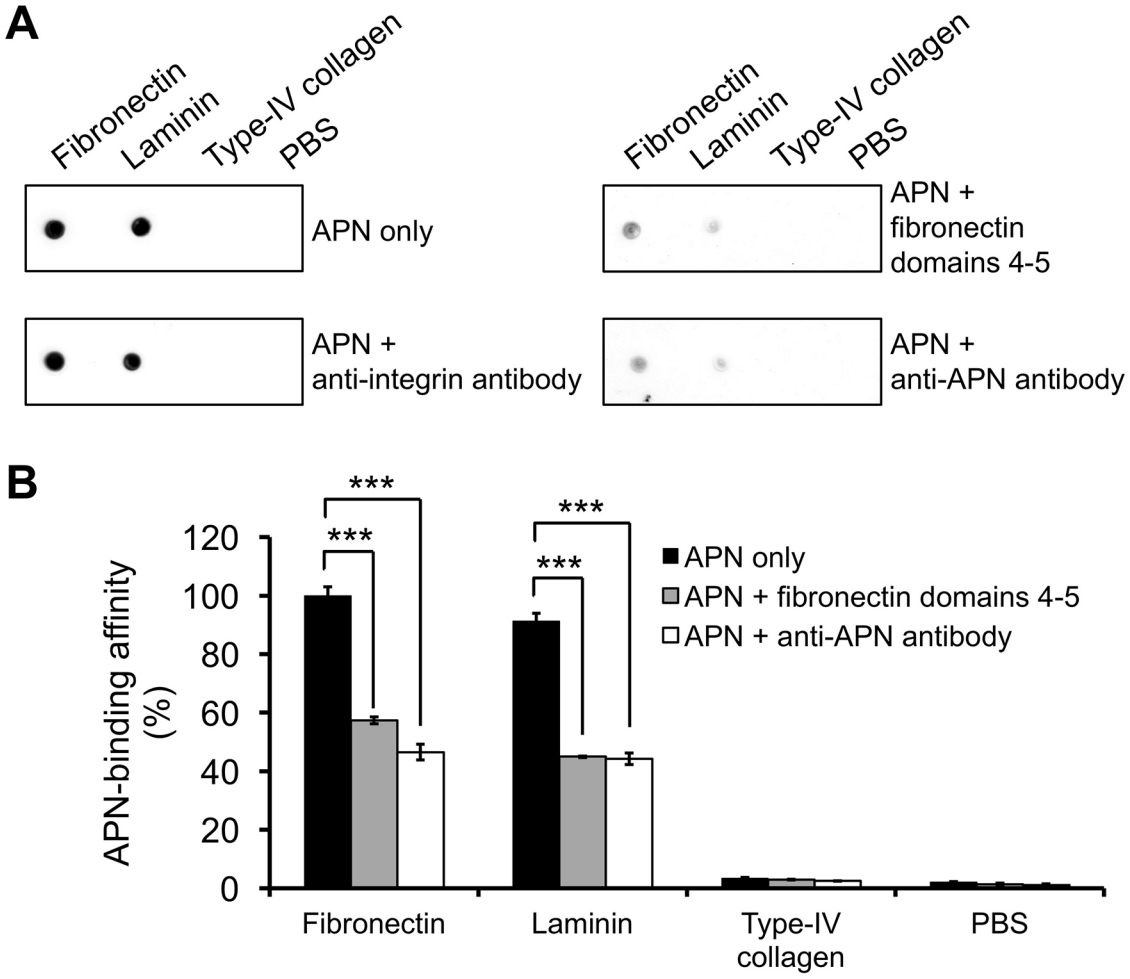



**Figure 1.5 Interactions between APN and extracellular matrix proteins.**

Both dot blot hybridization assay (A) and ELISA (B) were performed to detect the interactions between pAPN and individual ECM proteins with or without NGR motifs. These assays also measured competitive inhibition by anti-APN antibody or fibronectin NGR domains, with anti-integrin αV/β3 antibody as the negative control. $His_6$-tagged pAPN was detected by anti-$His_6$ antibody. Error bars indicate S.E. (two-tailed t test; ***, $p < 0.001$; $n = 3$).

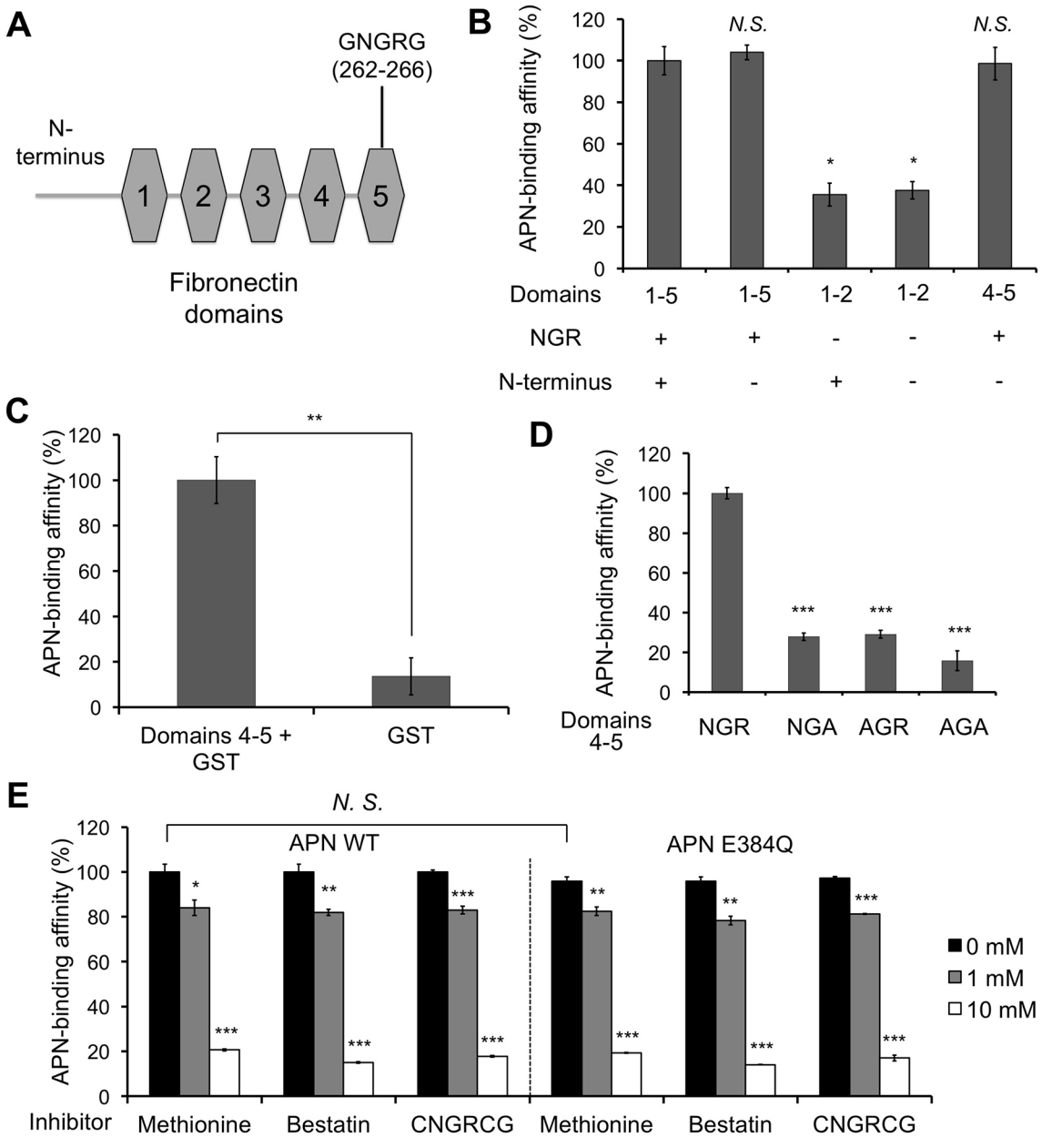

A
GNGRG
(262-266)
N-
terminus
1
2
3
4
5
Fibronectin
domains
B
APN-binding affinity (%)
120
100
80
60
40
20
0
N.S.
N.S.
*
*
Domains
1-5
1-5
1-2
1-2
4-5
NGR
+
+
-
-
+
N-terminus
+
-
+
-
-
C
APN-binding affinity (%)
120
100
80
60
40
20
0
**
Domains 4-5 +
GST
GST
D
APN-binding affinity (%)
120
100
80
60
40
20
0
***
***
***
Domains
4-5
NGR
NGA
AGR
AGA
E
N. S.
APN WT
APN E384Q
APN-binding affinity (%)
120
100
80
60
40
20
0
*
**
***
**
***
***
**
***
**
***
***
0 mM
1 mM
10 mM
Inhibitor
Methionine
Bestatin
CNGRCG
Methionine
Bestatin
CNGRCG

**Figure 1.6 Interactions between APN and the NGR motif in fibronectin.**

(A) Schematic structure of the free N terminus and five N-terminal domains of fibronectin. The free N terminus is 18 residues long. The NGR motif is located in domain 5. (B) Interactions between pAPN and different fibronectin domains. AlphaScreen assay was performed. The AlphaScreen signal measured between pAPN and fibronectin domains 1–5 with the N terminus was taken as 100%. (C) Interactions between pAPN and GST or fibronectin domains 4 –5 fused to the N terminus of GST. AlphaScreen signal measured between pAPN and GST-tagged fibronectin domains 4 –5 was taken as 100%. (D) Interactions between pAPN and fibronectin NGR domain (domains 4 –5) containing mutations in the NGR motif. AlphaScreen signal measured between pAPN and fibronectin NGR domain (domains 4–5) without any mutation was taken as 100%. (E) Inhibition of the interactions between pAPN (wild type or E384Q mutant) and the fibronectin NGR domain using APN active site inhibitors methionine, bestatin, or CNGRCG peptide. The AlphaScreen signal measured between wild-type pAPN and fibronectin NGR domain in the absence of any inhibitor was used as the standard and taken as 100%. Error bars indicate S.E. (compared with the standard; two-tailed t test; not significant (N.S.), $p > 0.05$; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; $n = 3$).

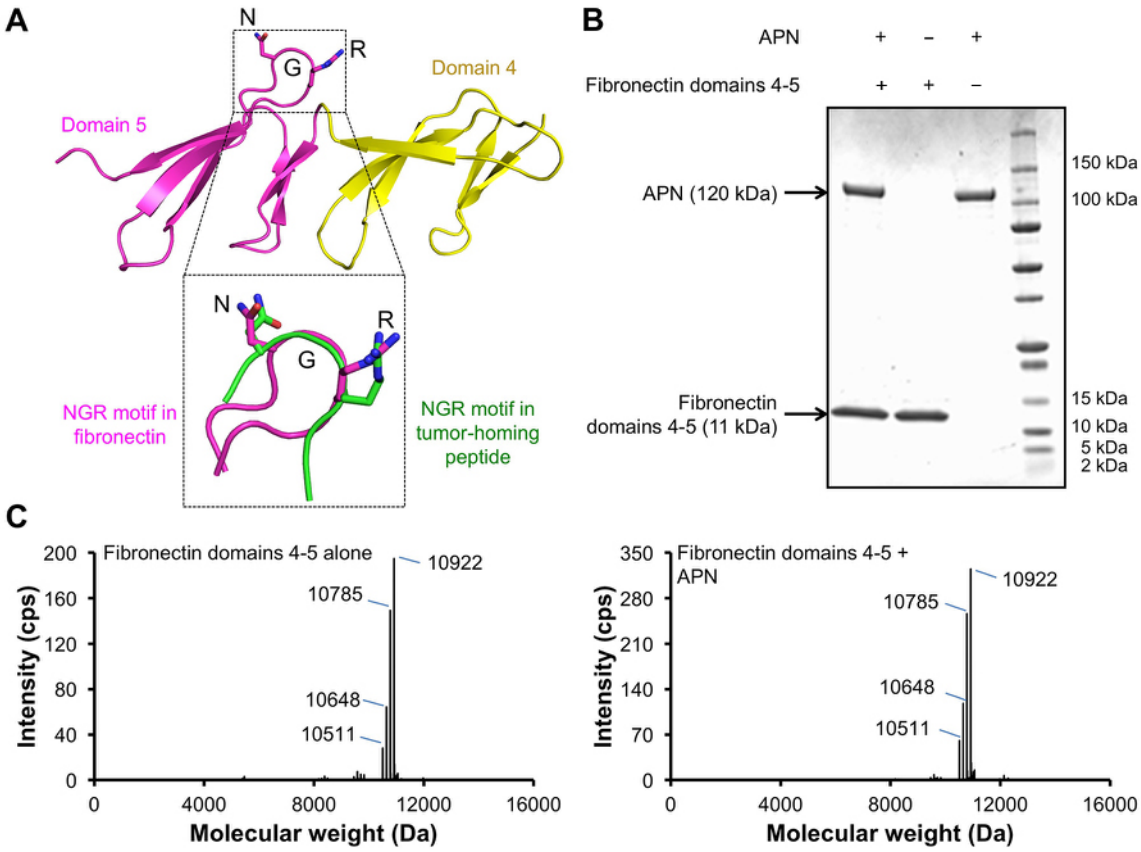



**Figure 1.7 Stable interactions between APN and fibronectin NGR domain.**

(A) Structure of fibronectin NGR domain (domain 5; in magenta), which is stabilized by inter-domain interactions with its neighboring domain (domain 4; in yellow) (PDB code 1FBR)[50]. Also shown is the superposition of the NGR motifs in fibronectin domain 5 (in magenta) and in the tumor-homing peptide CNGRCG (in green). (B) SDS-PAGE showing that pAPN did not cleave the NGR motif in the fibronectin NGR domain after they were incubated together. If the NGR motif in fibronectin domains 4 –5 (molecular mass of 11 kDa) was cleaved by pAPN, two cleavage products (molecular mass of 9 and 2 kDa) would be detected on SDS-PAGE of the reaction mixture. This was not supported by SDS-PAGE. (C) Mass spectrometry also showing that pAPN did not cleave the NGR motif in the fibronectin NGR domain.

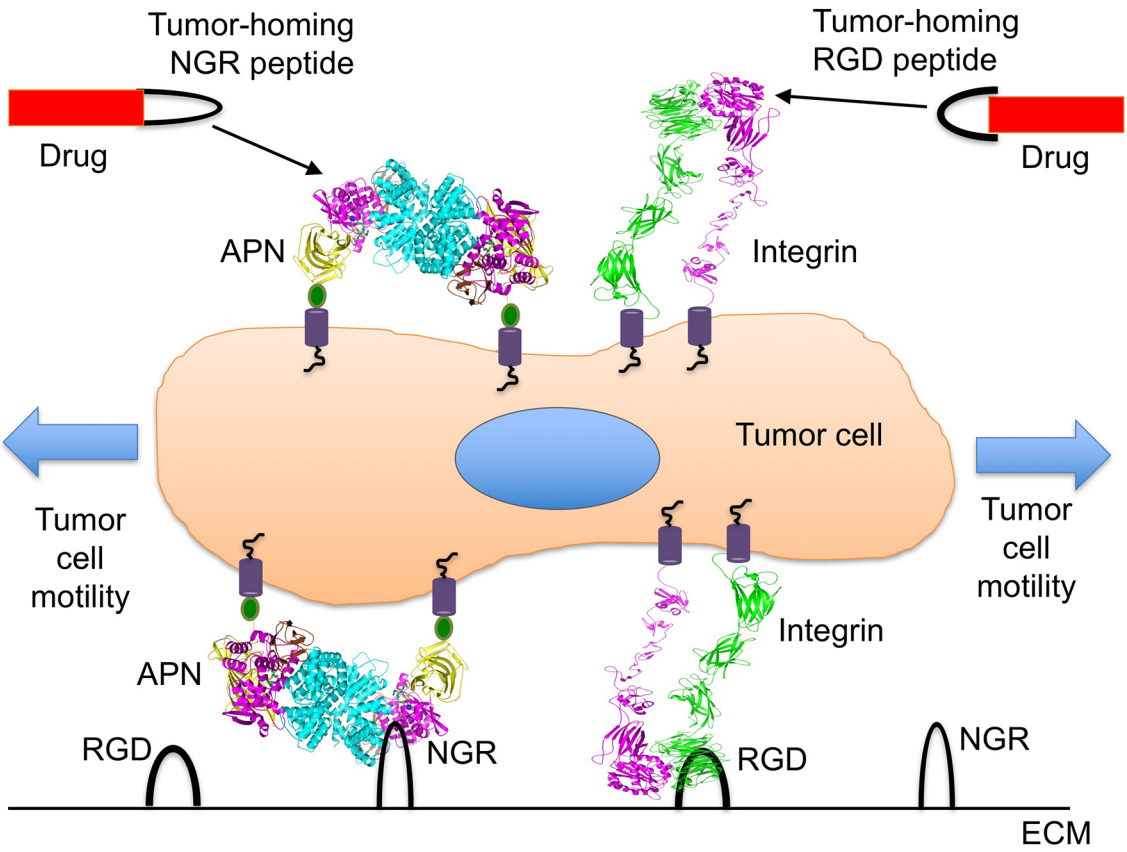



**Figure 1.8 A unified mechanism for APN-based tumor cell motility and tumor-homing therapy.**

ECM is shown as a black line on the bottom. Black loops in two different shapes indicate the signature NGR motif recognized by APN and the signature RGD motif recognized by integrins, respectively. APN and integrins interact with their respective signature motifs in ECM proteins to provide tractions for tumor cell motility. Anti-tumor drugs (in red) are con- nected either to a tumor-homing NGR peptide that targets tumor cell surface APN or to a tumor-homing RGD peptide that targets tumor cell surface integrins.

# 2 Chapter Two

## Receptor usage and cell entry of porcine epidemic diarrhea coronavirus

**Liu C***, Tang J*, Ma Y*, Liang X, Yang Y, Peng G, Qi Q, Jiang S, Li J, Du L, Li F (*Co-first authors)

***J Virol*** 89: 6121-6125

## 2.1 Introduction

Porcine epidemic diarrhea coronavirus (PEDV) causes large-scale outbreaks of diarrhea in pigs and an 80 to 100% fatality rate in suckling piglets[51-53]. Since 2013, PEDV has swept throughout the United States, wiped out more than 10% of America's pig population in less than a year, and significantly damaged the U.S. pork industry[54-56]. No vaccine or antiviral drug is currently available to keep the spread of PEDV in check. PEDV belongs to the α genus of the coronavirus family[57,58], which also includes porcine transmissible gastroenteritis coronavirus (TGEV), bat coronavirus 512/2005 (BtCoV/512/2005), and human NL63 coronavirus (HCoV-NL63). Although both PEDV and TGEV infect pigs, PEDV is genetically more closely related to BtCoV/512/ 2005 than to TGEV, leading to the hypothesis that PEDV originated from bats[59].

Receptor binding and cell entry are essential steps in viral infection cycles, critical determinants of viral host range and tropism, and important targets for antiviral interventions. An envelope-anchored spike protein mediates coronavirus entry into cells. The spike ectodomain consists of a receptor-binding subunit, S1, and a membrane fusion subunit, S2. S1 contains two domains, an N-terminal domain (S1-NTD) and a C-terminal do- main (S1-CTD), both of which can potentially function as receptor-binding domains (RBDs) (Fig. 2.1A)[60,61]. The ability of coronavirus RBDs to recognize receptor orthologs from different species is one of the most important determinants of coronavirus host range and tropism[57,62-64]. HCoV-NL63 S1-CTD recognizes human angiotensin-converting enzyme 2 (ACE2), whereas TGEV S1-CTD recognizes porcine aminopeptidase N (APN), and its S1-NTD recognizes two sugar coreceptors, N-acetylneuraminic acid

(Neu5Ac) and N-glycolylneuraminic acid (Neu5Gc)[3,65-67]. Usage of sugar coreceptors is linked to the enteric tropism of coronaviruses[66,68]. It has been shown that PEDV uses porcine APN as its receptor[69]. However, it is not known whether PEDV recognizes APN from other species or whether it uses sugar coreceptors. Addressing these questions will be critical for understanding the host range, tropism, and evolutionary origin of PEDV, for evaluating its potential risk to other species, particularly humans, and for developing effective vaccines and antiviral drugs to curb the spread of PEDV in pigs and to other species.

## 2.2 Materials and Methods

### Cell lines and plasmids

HEK293T (human embryonic kidney), MRC-5 (human lung), Tb1-Lu (bat lung), ST (porcine testis), MDCK (canine kidney), Vero CCL81 (monkey kidney) and PK-15 (porcine kidney) cells were obtained from American Type Culture Collection. Huh-7 (human hepatoma) cells were kindly provided by Dr. Charles M. Rice at Rockefeller University. These cell lines were cultured in Dulbecco's modified Eagle medium (DMEM) supplemented with 10% fetal bovine serum (FBS), 2mM L-glutamine, 100 units/mL penicillin, and 100 μg/mL streptomycin. (Life Technologies).

The genes of TGEV spike (GenBank accession number CAA29175.1) and PEDV spike (GenBank accession number AGO58924.1) were each cloned into pcDNA3.1(+) vector (Life Technologies) with a C-terminal C9 tag. The DNA fragments encoding S1-NTD-

CTD domains of PEDV spike (residues 19-638) and TGEV spike (residues 17-673) were each cloned into pFUSE-hIgG1-Fc (InvivoGen) expression vector. The DNA fragments encoding extracellular domains of human APN (residues 66-967; GenBank accession number AAA51719.1) and porcine APN (residues 62-963; GenBank accession number CAA82641.1) were each cloned into pFastBact1 vector (Life Technologies) a C-terminal $His_6$ tag.

**Protein preparation and purification**

TGEV and PEDV spike S1-NTD-CTD domains were expressed and purified as previously described [48]. Briefly, the IgG1-Fc-fusion proteins were transiently expressed in HEK293T cells. The protein fragments were harvested from the cell culture supernatants and purified using protein A Sepharose beads (GE Heathcare). The extracellular domains of human or porcine APN containing an N-terminal honeybee melittin signal peptide and a C-terminal $His_6$ tag was expressed in insect cells using the Bac-to-Bac expression system (Life Technologies), secreted into cell culture medium, and subsequently purified on a HiTrap Chelating HP affinity column (GE Healthcare) and a Superdex 200 gel filtration column (GE Healthcare). The proteins were concentrated to 10 mg/mL and stored in buffer containing 20 mM Tris (pH 7.4) and 200 mM NaCl.

**Dot blot binding assay**

Dot blot binding assay was carried out as previously described in Chapter One. Briefly,

10 μg each of the Fc-tagged TGEV or PEDV S1-NTD-CTD domain was dotted onto a nitrocellulose membrane. The membranes were dried completely and blocked with 5% skim milk at 37 °C for 1 h. Next, 25 μg/ml $His_6$-tagged human or porcine APN extracellular domain was added to the membrane, and incubated at 4 °C overnight. The membrane was then washed five times with phosphate-buffered saline with Tween-20 (PBST), incubated with anti-$His_6$ mouse monoclonal HRP conjugate antibody (Santa Cruz biotechnology) at 37 °C for 2 h, and washed five times with PBST. Finally, the bound proteins were detected using ECL plus (GE Healthcare). To detect the binding between TGEV or PEDV S1-NTD-CTD domain with mucin, 10 μg bovine mucin or porcine mucin was preincubated alone or with 200 units of α2-3,6,8,9 neuraminidase A (New England Biolabs) at 37 °C overnight, dotted onto nitrocellular membrane, and incubated with 25 μg/ml Fc-tagged TGEV or PEDV S1-NTD-CTD domain. Bound coronavirus S1-NTD-CTD domains were detected by anti-IgG1-Fc monoclonal HRP conjugate antibody (Santa Cruz biotechnology).

**Glycan Screen Array**

To determine the sugar-binding specificity of PEDV S1-NTD-CTD domain, a glycan screen array was performed at the Consortium for Functional Glycomics. The glycan array (CFG version 5.0) was composed of 609 different natural and synthetic mammalian glycans. In the binding assay, array slides were incubated with PEDV S1-NTD-CTD with a C-terminal IgG1-Fc tag. The slides were then washed, and bound PEDV S1-NTD-CTD was detected with mouse anti-IgG1-Fc HRP-conjugated antibody

**Pseudovirus production**

Retroviruses pseudotyped with PEDV or TGEV spike were generated as previously described [70]. Briefly, HEK293T cells were co-transfected with a plasmid carrying Env-defective, luciferase-expressing HIV-1 genome (pNL4-3.luc.R-E-) and a plasmid encoding PEDV spike or TGEV spike using Lipofectamine 3000 reagent (Life Technologies) according to the manufacturer's instructions. Supernatants containing pseudoviruses were harvested 72 h after transfection, and centrifuged at 1200 × g for 10 min to remove cell debris. Pseudoviruses were concentrated using Amicon Ultra centrifugal filter units with a 100 kDa molecular weight cut-off, aliquoted and frozen at -80 °C for later use.

**Inhibition of PEDV spike-meidated pseudovirus entry by anti-APN antibody or mucin**

MDCK cells expressing human APN, porcine APN, or no APN, Huh-7 cells, or PK-15 cells in 96-well plates were preincubated with 0, 2 μg/ml, or 20 μg/ml anti-APN antibody (Santa Cruz biotechnology) at 37 °C for 1 h, and then transduced by retroviruses pseudotyped with PEDV spike, TGEV spike, or empty vector (mock). For mucin inhibition, PEDV spike- or TGEV spike-pseudotyped retroviruses were preincubated with 500 μg/ml porcine or bovine mucin at 37 °C for 1 h before they were used to transduce Huh-7 cells or PK-15 cells. Cells were incubated with pseudoviruses at 37°C for 5 h, and medium was replaced with fresh DMEM. 60 h later, cells were washed with phosphate-buffered saline (PBS) and lysed. Aliquots of cell lysates were transferred to

Optiplate-96 plate (PerkinElmer), and luciferase substrate (Promega) was addeded. Relative luciferase units (RLU) were measured using EnSpire plate reader (PerkinElmer).

**Isolation and propagation of authentic PEDV VBS2 strain**

In August 2013, intestinal contents were collected from a 7-day-old piglet with severe diarrhea on a farm in Ohio, USA. PEDV RNA was detected in the samples by reverse transcription polymerase chain reaction (RT-PCR). This strain was named PEDV VBS2. Intestinal contents were homogenized and vortexed in cell culture medium (DMEM) on ice. The homogenate was clarified by centrifugation at 5,000 g for 10 min at 4 °C. The supernatant was collected and filtered through a 0.45 μm syringe filter. The supernatant was further filtered through a 0.22-μm-pore-size filter and used as an inoculum for virus isolation. This inoculum was designated as passage 0 (P0). To propagate this virus, T25 flasks of Vero CCL81 cells were inoculated with 0.5 ml of filtered supernatant. Before inoculation, the cells were washed twice using DMEM with 5 μg/ml trypsin 1:250 (Life Technologies). After 2 h infection, the inoculum was removed and washed one time with DMEM. The infected Vero cells were maintained in 3 ml of DMEM supplemented with 0.018% (w/v) Tryptose Phosphate Broth (TPB) (Sigma-Aldrich), 0.02% yeast extract (Sigma-Aldrich), 5 μg/ml Trypsin 1:250, 10 units/ml penicillin-streptomycin, 0.05 mg/ml gentamicin, and 0.05 mg/ml Kanamycin. When an extensive cytopathic effect (CPE) was observed, the cells were subjected to freeze-thaw three times. The cell culture mixtures were centrifuged at 5,000 g for 10 min at 4 °C. The supernatant was collected and designated as P1. The P1 supernatants were used to inoculate on new Vero CCL81 cells

for a second passage (P2). After 3 passages, typical CPEs such as cell enlargement, cell-cell fusion, syncytium formation, and death were observed. The presence of PEDV at each passage was confirmed by RT-PCR.

**Enumeration of infectious PEDV by plaque assay**

PEDV plaque assays were performed in Vero CCL81 cells. Briefly, cells were seeded into six-well plates (BD) at a density of $2 \times 10^6$ cells per well. After 24 h of incubation, cell monolayers were infected with 400 µl of a 10-fold dilution series of PEDV, and the plates were incubated for 1 h at 37 °C, with agitation every 10 min. The cells were overlaid with 2.5 ml of DMEM containing 0.25% agarose, 0.018% TPB, 0.02% yeast extract (Sigma-Aldrich), 5 µg/ml Trypsin 1:250, 10 units/ml penicillin-streptomycin, 0.05 mg/ml gentamicin, and 0.05 mg/ml Kanamycin. After 48 h incubation, the plates were fixed in 10% formaldehyde, and the plaques were visualized by staining with 0.05% (wt/vol) crystal violet.

**Visualization of PEDV nucleocapsid protein expression by immunofluorescence assay**

Confluent cells were infected with PEDV VBS2 at a multiplicity of infection (MOI) of 1.0. Before inoculation, the cells were washed twice using DMEM with 5ug/ml trypsin 1:250 (Life Technologies). After 1 h of absorption, the inoculum was removed, the cells were washed one time with DMEM, and fresh DMEM (supplemented with 5 µg/ml trypsin 1:250) was added, and the infected cells were incubated at 37 °C. When CPE

were observed, the cells were washed twice with PBS and fixed with 4.0% (v/v) paraformaldehyde-0.2% (v/v) glutaraldehyde in 0.1 M potassium phosphate buffer (PPB), pH 7.4, for 15 min, 22 °C, followed by washing 3 times with PBS. After permeabilization with 0.1% Triton X-100 in PBS for 15 min, the cells were washed with PBS, blocked with PBS containing 2% bovine serum albumin for 1 h at room temperature. Then cells were incubated with FITC-labeled mouse anti-PEDV nucleocapsid protein antibody (Medgene) at dilution of 1:200 overnight at 4°C. After 3 washes with PBS, samples were examined under Olympus fluorescent microscope system at The Ohio State University.

## 2.3 Results

To characterize the receptor usage of PEDV, here we identified the two S1 domains of PEDV based on the sequence similarity between PEDV and TGEV S1 subunits (Fig. 2.1B). The S1-NTD and S1-CTD of PEDV cover residues 19 to 252 and residues 509 to 638, respectively. However, expression of the two domains individually gave low yields. Instead, we expressed and purified a longer fragment (residues 19 to 638) using a previously described procedure[48,71]. This fragment contains both of the S1 domains and is termed S1-NTD-CTD (Fig. 2.2A). For comparison studies, we prepared TGEV S1-NTD-CTD (residues 17 to 675) using the same procedure. We also expressed and purified human and porcine APN as previously described[6,72]. These purified recombinant proteins were subsequently used in biochemical studies.

We investigated the receptor binding capabilities of PEDV S1- NTD-CTD. First, using a dot blot hybridization assay as previously described[72], we showed that PEDV S1-NTD-CTD binds both porcine and human APN efficiently (Fig. 2.2B). Thus, both porcine and human APN serve as efficient receptors for PEDV. In contrast, TGEV S1-NTD-CTD binds porcine APN much more tightly than it binds human APN (Fig. 2.2B). Second, using the dot blot hybridization assay as previously described[47,73], we demonstrated that PEDV S1-NTD-CTD binds bovine and porcine mucins, both of which contain a mixture of different types of sugar (Fig. 2.2C).

Treatment of mucins with neuraminidase re- moved part of the coated sugars, reducing the binding by PEDV S1-NTD-CTD. Hence, sugar serves as a coreceptor for PEDV. As a comparison, TGEV S1-NTD-CTD also binds these mucins. Third, using a glycan screen array as previously described[73], we identified Neu5Ac as the type of sugar most favored by PEDV (Fig. 2.2D). Taken together, PEDV uses both porcine and human APNs as its protein receptors and Neu5Ac as a sugar coreceptor, whereas TGEV uses porcine APN and sugar, but not human APN, as its receptors.

To further understand the receptor usage and also to investigate the cell entry of PEDV, we performed a PEDV spike-mediated pseudovirus entry[70]. Retroviruses pseudotyped with PEDV spike (i.e., PEDV pseudoviruses) efficiently entered MDCK (ca- nine kidney) cells exogenously expressing human or porcine APN, and these entries could be blocked by anti-APN antibody (Fig. 2.3A). As a control, PEDV pseudoviruses could not enter MDCK cells not expressing human or porcine APN, consistent with a previous

report that MDCK is nonpermissive to PEDV infection[69]. In contrast, TGEV pseudoviruses efficiently entered MDCK cells exogenously expressing porcine APN but not those expressing human APN. Additionally, PEDV pseudoviruses efficiently entered both PK-15 (pig kidney) and Huh-7 (human lung) cells that endogenously express porcine and human APN, respectively[74,75], and these entries could be blocked by anti-APN antibody and mucins (Fig. 2.3B and C). In contrast, TGEV pseudoviruses efficiently entered PK-15 cells but not Huh-7 cells. These data collectively confirmed that human and porcine APN and sugar serve as receptors for PEDV and play important roles in PEDV spike-mediated cell entry, whereas porcine APN and sugar, but not human APN, are receptors for TGEV.

To further examine PEDV entry into host cells, we carried out live-PEDV infection in the following cell lines: PK-15 (pig kidney), ST (pig testis), Huh-7 (human liver), MRC-5 (human lung), Vero CCL-81 (monkey kidney), and Tb1-Lu (bat lung) cells. To this end, PEDV strain Ohio VBS2 was isolated from a piglet in Ohio, USA, and propagated in Vero CCL-81 cells using a procedure as previously described[76]. Vero CCL-81-adapted PEDV was used to infect each of the above-named cell lines at a multiplicity of infection (MOI) of 1.0. The results showed that PEDV efficiently infects cells from pig, human, monkey, and bat (Fig. 2.4).

It is worth noting that whereas pseudovirus entry is determined by receptor recognition and cell entry, the infection efficiency of live PEDV in cell culture is determined not only

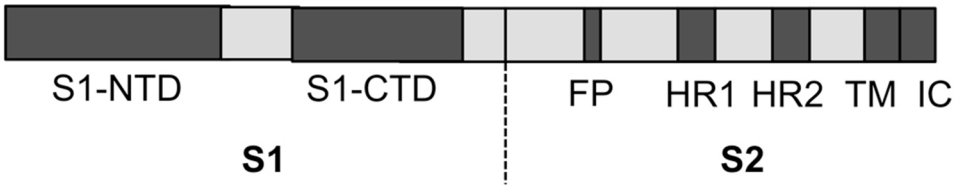
by receptor recognition and cell entry but also by postentry factors, such as viral replication and release[77].

## 2.4 Discussion

PEDV is a highly pathogenic and lethal pig coronavirus. This study investigated how PEDV recognizes host receptors from different species and how it infects cells from different species. First, we verified that PEDV recognizes porcine APN and infects pig cells. Second, for the first time to our knowledge, we showed that PEDV recognizes a sugar coreceptor, Neu5Ac, which explains the enteric tropism of PEDV. Because TGEV also recognizes porcine APN and Neu5Ac, PEDV and TGEV are evolutionarily closely related despite the relatively low sequence similarity in their spikes (Fig. 2.1B). Third, we demonstrated that PEDV infects bat cells, providing evidence that PEDV originated from bats. Finally, un- like TGEV, which does not use human APN as its receptor, PEDV recognizes human APN and infects human cells. Thus, neither receptor recognition nor other host cellular factors (e.g., cellular restrictions of viral replication) pose a hurdle for PEDV to infect humans. It remains to be seen whether systemic factors (e.g., the host immune system) can prevent or clear in a timely manner PEDV infections in humans. Nevertheless, these results suggest that PEDV may be a potential threat to other species, including humans. Overall, our study provides insight into the host range, tropism, and evolution of PEDV.

Our study also has implications for the development of antiviral strategies against PEDV. The S1-NTD-CTD fragment as identified in this study may serve as a subunit vaccine candidate. Monoclonal antibodies against S1-NTD-CTD may serve as immunotherapeutic agents to block PEDV attachment to both the APN receptor and the sugar coreceptor. In addition, sugar or sugar analogues may serve as antiviral drugs to block PEDV attachment to its sugar coreceptor. The development of these anti-viral strategies is urgent because of the damaging impact that PEDV exerts on the U.S. pork industry and the potential threat that PEDV poses to other species.

A




B

| PEDV spike | S1-NTD | S1-CTD | S1 | S2 |
|---|---|---|---|---|
| TGEV | 19% | 23% | 31% | 56% |
| BtCoV/ 512/2005 | 48% | 36% | 48% | 74% |
| HCoV-NL63 | 21% | 29% | 36% | 60% |

**Figure 2.1 PEDV spike protein.**

(A) Domain structure of PEDV spike. It contains a receptor-binding S1 subunit, a membrane fusion S2 subunit, a single-pass trans- membrane anchor (TM), and a short intracellular tail (IC). S1 contains an N-terminal domain (S1-NTD) and a C-terminal domain (S1-CTD). S2 contains the fusion peptide (FP), heptad repeat 1 (HR1), and heptad repeat 2 (HR2), all of which are essential structural elements for the membrane fusion process. (B) Amino acid sequence identities between PEDV spike and the spikes from TGEV, BtCoV/512/2005, and HCoV-NL63 in different regions. GenBank accession num- bers are AGO58924.1 for PEDV spike, CAA29175.1 for TGEV spike, ABG47078.1 for BtCoV/512/2005 spike, and AAS58177.1 for HCoV-NL63 spike.

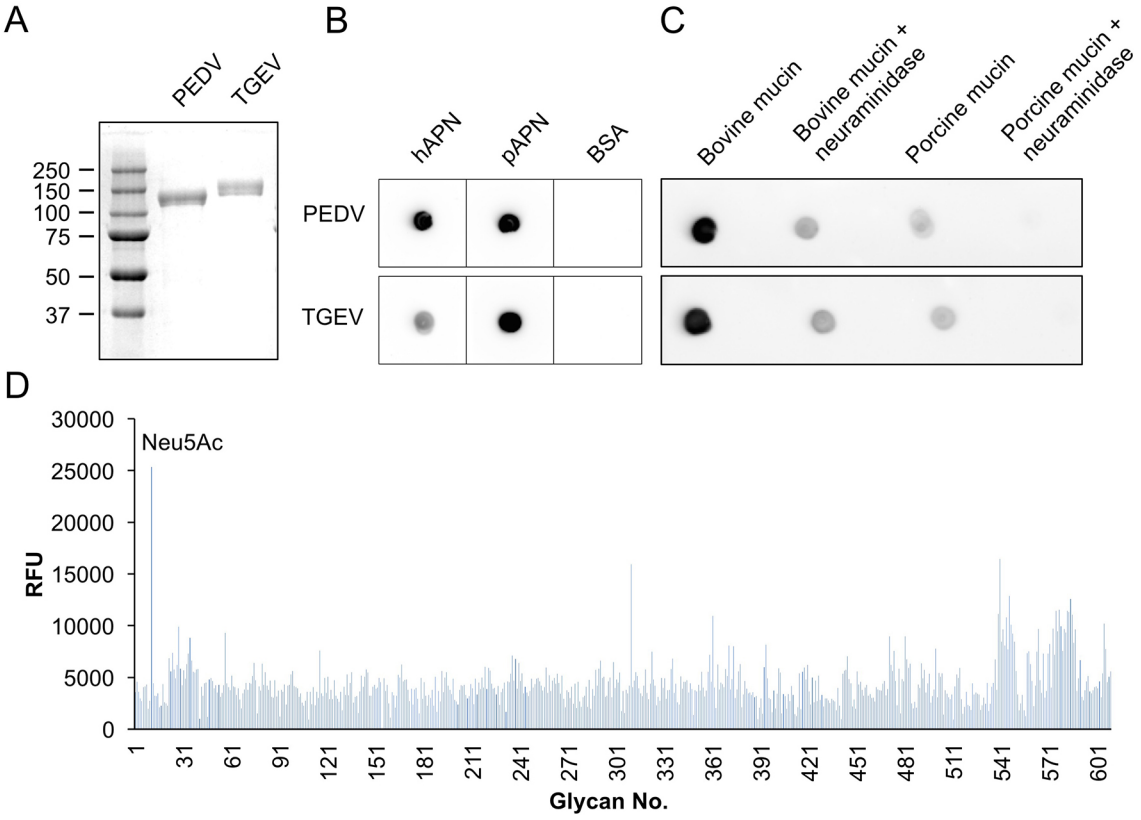

A
PEDV
TGEV
250
150
100
75
50
37
B
hAPN
pAPN
BSA
PEDV
TGEV
C
Bovine mucin
Bovine mucin + neuraminidase
Porcine mucin
Porcine mucin + neuraminidase
D
Neu5Ac
RFU
30000
25000
20000
15000
10000
5000
0
1
31
61
91
121
151
181
211
241
271
301
331
361
391
421
451
481
511
541
571
601
Glycan No.

**Figure 2.2 PEDV spike binds porcine APN, human APN, and sugar receptors.** (A) SDS-PAGE analysis of recombinant PEDV S1-NTD-CTD and TGEV S1-NTD-CTD. The gel was stained using Coomassie blue. Numbers at the left are molecular masses (in kilodaltons). (B) Dot blot hybridization assay showing the interactions between PEDV or TGEV S1-NTD-CTD (with a C-terminal human IgG1 Fc tag) and porcine APN (pAPN) or human APN (hAPN) (with a C-terminal $His_6$ tag). Bovine serum albumin (BSA) was used as a negative control. (C) Dot blot hybridization assay showing the interactions between PEDV or TGEV S1-NTD-CTD and sugar moieties on mucin-spotted nitrocellulose membranes. Mucin was either mock treated or treated with neuraminidase (New England BioLabs). Sugar-binding S1-NTD-CTDs were detected using antibodies against their C-terminal Fc tag. (D) A glycan screen array was performed to identify the type(s) of sugar most favored by PEDV S1-NTD-CTD. Glycan-binding S1-NTD-CTD was detected using antibodies against its C-terminal Fc tag. The readout was described arbitrarily as relative fluorescence units (RFU). Among these glycans, *N*-acetylneuraminic acid (Neu5Ac) shows the highest binding affinity for PEDV S1-NTD-CTD.

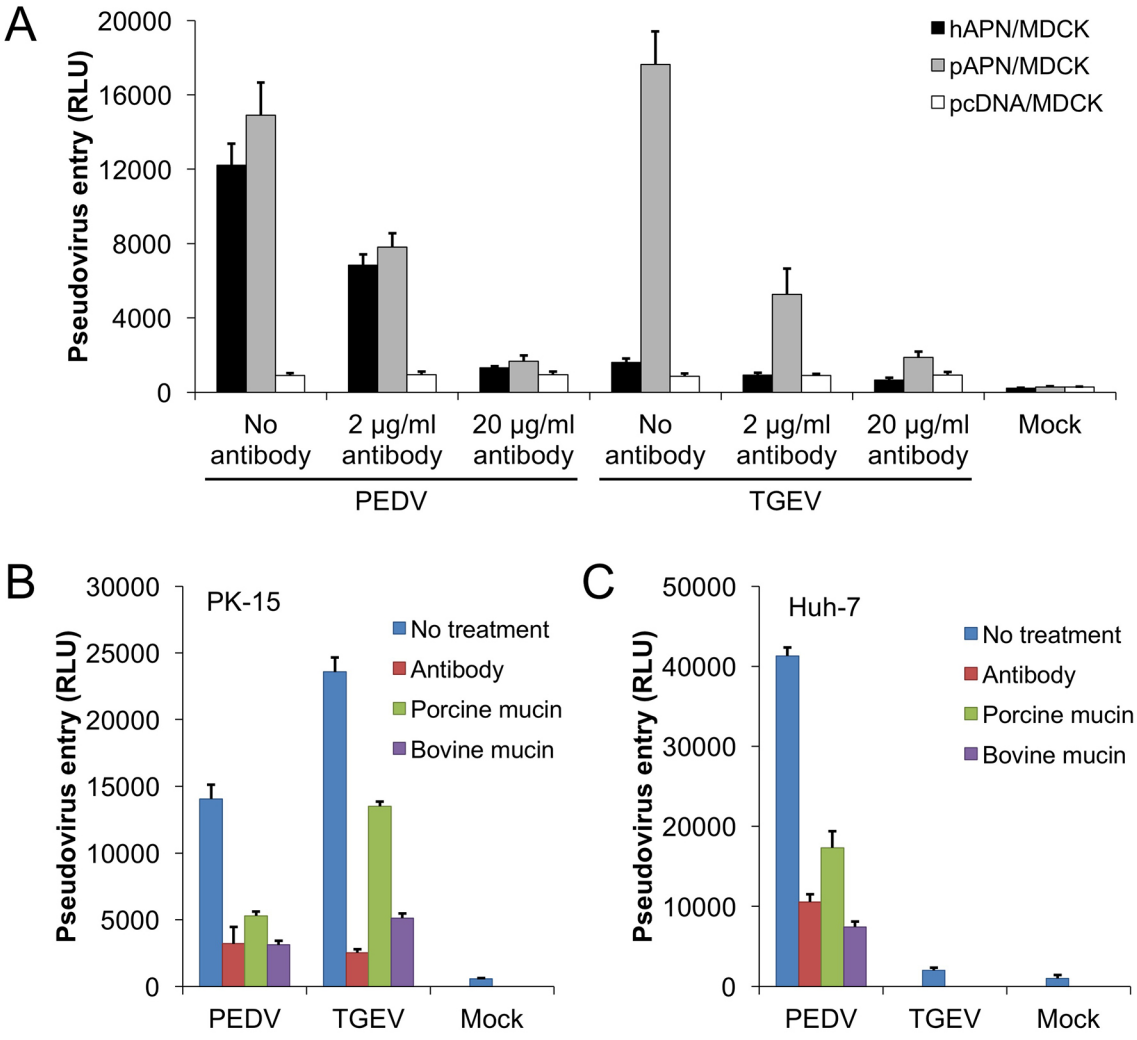

A
Pseudovirus entry (RLU)
20000
16000
12000
8000
4000
0
hAPN/MDCK
pAPN/MDCK
pcDNA/MDCK
No antibody
2 μg/ml antibody
20 μg/ml antibody
No antibody
2 μg/ml antibody
20 μg/ml antibody
Mock
PEDV
TGEV
B
PK-15
Pseudovirus entry (RLU)
30000
25000
20000
15000
10000
5000
0
No treatment
Antibody
Porcine mucin
Bovine mucin
PEDV
TGEV
Mock
C
Huh-7
Pseudovirus entry (RLU)
50000
40000
30000
20000
10000
0
No treatment
Antibody
Porcine mucin
Bovine mucin
PEDV
TGEV
Mock

**Figure 2.3 PEDV spike-mediated pseudovirus entry into host cells.**

PEDV spike- and TGEV spike-pseudotyped retroviruses were used to transduce MDCK cells exogenously expressing human APN (hAPN), porcine APN (pAPN), or an empty vector (A), PK-15 cells (B), and Huh-7 cells (C). For antibody inhibition, cells were preincubated with anti-human APN antibody at indicated concentrations before pseudovirus transduction. For mucin inhibition, PEDV spike- or TGEV spike-pseudotyped retroviruses were preincubated with 500 μg/ml porcine or bovine mucin before they were used to transduce cells. The pseudovirus entry efficiency was characterized as luciferase activity accompanying the entry. Error bars indicate standard errors of the means (SEM) ($n$ = 4).

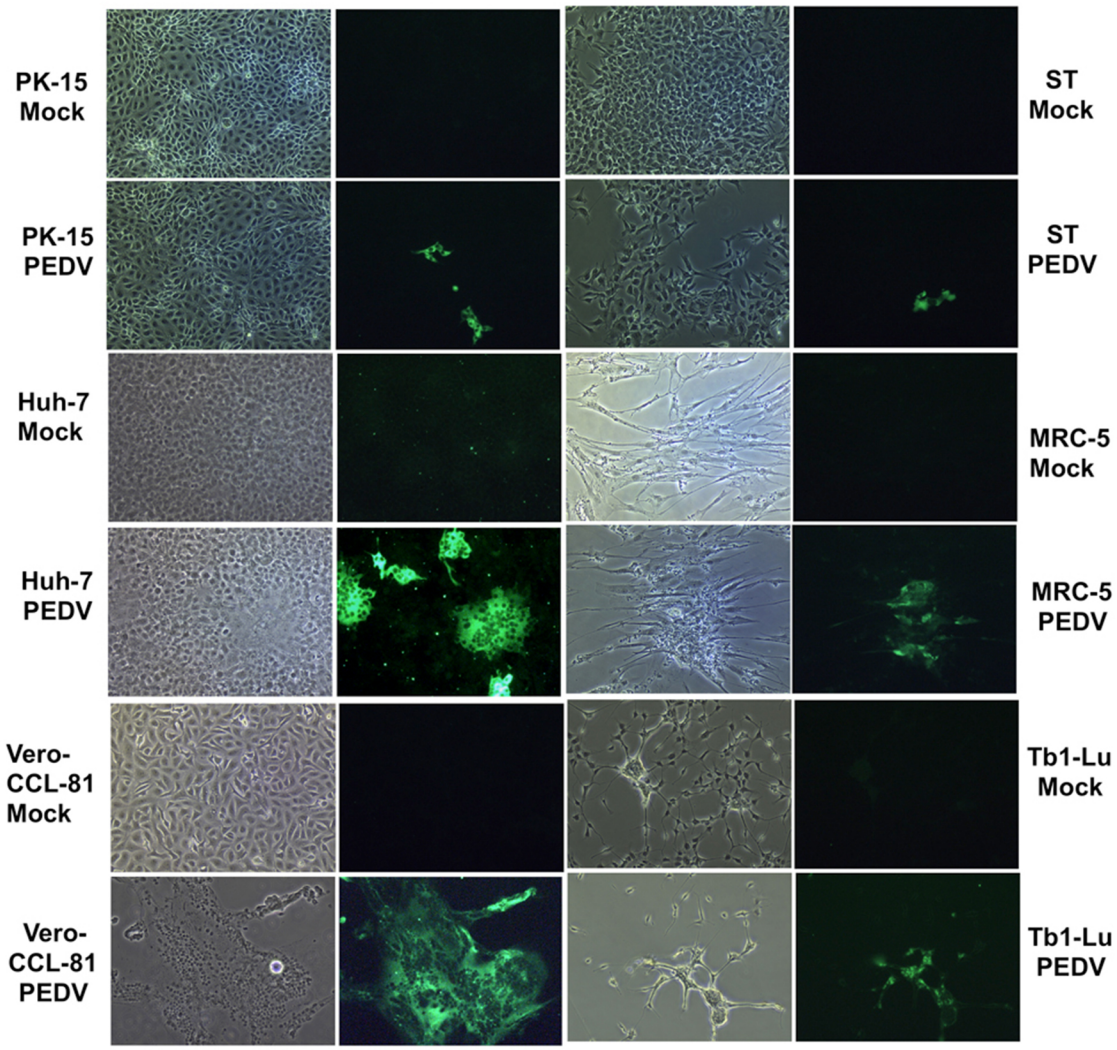

PK-15
Mock
ST
Mock
PK-15
PEDV
ST
PEDV
Huh-7
Mock
MRC-5
Mock
Huh-7
PEDV
MRC-5
PEDV
Vero-
CCL-81
Mock
Tb1-Lu
Mock
Vero-
CCL-81
PEDV
Tb1-Lu
PEDV

**Figure 2.4 PEDV infections in cell culture.**

PEDV strain Ohio VBS2 was used to infect different cell lines at an MOI of 1.0 using a procedure as previously described[76]. Trypsin (5 μg/ml) was included in the cell culture medium to facilitate live-PEDV infections. Twenty-four hours postinoculation, cells were fixed with 4.0% (vol/vol) paraformaldehyde and 0.2% (vol/vol) glutaraldehyde. PEDV was detected with fluorescein isothiocyanate (FITC)-labeled mouse anti-PEDV N protein antibody and observed under a fluorescence microscope.

# 3 Chapter Three

## Entry of porcine epidemic diarrhea coronavirus is activated by lysosomal proteases

**Liu C***, Yuanmei Ma*, Yang Y*, Zheng Y, Zhou Y, Jiang S, Du L, Li J, Li F (*co-first authors)

Manuscript in Preparation

## 3.1 Introduction

Since 2013, porcine epidemic diarrhea coronavirus (PEDV) has swept throughout the United States, causing 80-100% fatality rate in piglets and wiping out more than 10% of America's pig population in less than a year [54-56]. Currently there is no effective strategy available to keep the spread of PEDV in check. PEDV belongs to the α-genus of the coronavirus family. An envelope-anchored spike protein guides coronavirus entry into host cells [61,78,79]. The ectodomain of the spike protein contains a receptor-binding subunit S1 and a membrane-fusion subunit S2. The spike protein exists in two different conformations: the pre-fusion conformation is a clove-shaped trimer with three individual S1 heads and a trimeric S2 stalk; the post-fusion conformation is a trimeric S2 that has been structurally rearranged to fuse the viral and host membranes [80-82]. During viral entry, S1 first binds to a receptor on host cell surface for viral attachment, and then S2 transitions to the post-fusion conformation for membrane fusion. For the spike protein to undergo conformational changes, it needs to be cleaved at or near the S1/S2 boundary by one or more host proteases.

The host proteases mainly come from four different stages of the virus infection cycle: (i) proprotein convertases during virus packaging in virus-producing cells; (ii) extracellular proteases after virus release from virus-producing cells; (iii) cell-surface proteases after virus attachment to virus-targeting cells; (iv) lysosomal proteases after virus endocytosis in virus-targeting cells [81,82]. In addition, proprotein convertases were shown to activate MERS-CoV spike after virus endocytosis in virus-targeting cells [83]. PEDV is unique

among coronaviruses in that its propagation in cell culture requires exogenous trypsin, and thus it has been commonly believed that extracellular proteases in pig intestines are essential for cell entry of PEDV [76,84,85]. However, these cell culture studies used live PEDV particles for cell entry, and did not differentiate PEDV entry from other steps of the PEDV infection cycle such as PEDV replication or release. Therefore, it remains to be a conundrum what proteases activate cell entry of PEDV.

We recently characterized the receptor usage and cell entry of PEDV [86]. We found that PEDV S1 uses human and porcine APN as its receptor and sugar as its co-receptor. In addition, PEDV infects both human and porcine cells. Importantly, we established a PEDV-spike-mediated pseudovirus entry assay. In this assay, a replication-deficient retrovirus becomes pseudotyped with PEDV spike and is used to enter PEDV-susceptible host cells. The pseudovirus assay only concerns virus entry, but not virus replication or release. Hence the pseudovirus entry assay has advantages over the live PEDV infection assay in virus entry studies. Here using this PEDV pseudovirus entry assay and also biochemical assays, we systematically investigated what proteases process and activate PEDV spike, revealing the molecular mechanism of PEDV entry and solving a long-standing puzzle in the field. Identification of the PEDV-spike-processing proteases also provided a potential target for development of antiviral drugs to block PEDV entry.

## 3.2 Materials and Methods

### Cell lines and plasmids

HEK293T (human embryonic kidney), PK-15 (porcine kidney) and Vero CCL-81 (monkey kidney) cells were obtained from American Type Culture Collection. Huh-7 (human hepatoma) cells were kindly provided by Dr. Charles M. Rice at Rockefeller University. These cell lines were cultured in Dulbecco's modified Eagle medium (DMEM) supplemented with 10% fetal bovine serum (FBS), 2mM L-glutamine, 100 units/mL penicillin, and 100 μg/mL streptomycin. (Life Technologies).

The genes of MERS-CoV spike (GenBank accession number AFS88936.1) and PEDV spike (GenBank accession number AGO58924.1) were each cloned into pcDNA3.1(+) vector (Life Technologies) with a C-terminal C9 tag. The genes of cathepsin L (GenBank accession number CAA30981.1) and cathepsin B (GenBank accession number AAA52129.1) were synthesized from GenScript, and were each cloned into pFastBac1 vector (Life Technologies) with a C-terminal $His_6$ tag. The plasmid of human TMPRSS2 was kindly provided by Dr. Tom Gallagher at Loyola University Medical Center.

**Protein expression and purification**

Cathepsin L and cathepsin B were expressed as inactive proform in insect cells and purified as previously described [70]. Briefly, the full length proteases containing a C-terminal $His_6$ tag were expressed in Sf9 insect cells using the Bac-to-Bac expression system (Life Technologies), secreted to cell culture medium, and purified using HiTrap Chelating HP column and Superdex 200 gel filtration column (GE Healthcare), sequentially. Purified Pro-cathepsin L and pro-cathepsin B were auto-activated as

previously described [87]. Briefly, purified pro-cathepsin L or pro-cathepsin B were diluted 10 fold using 100 mM sodium acetate, pH 4.0, and incubated at 37°C for 1 h. The activated mature proteases were further purified on a Superdex 200 gel filtration column.

**Pseudovirus production**

Retroviruses pseudotyped with PEDV spike, MERS-CoV spike, or VSV envelope glycoprotein were generated as previously described [70,86]. Briefly, HEK293T cells were co-transfected with a plasmid carrying Env-defective, luciferase-expressing HIV-1 genome (pNL4-3.luc.R-E-) and a plasmid encoding PEDV spike or MERS-CoV spike using Lipofectamine 3000 reagent (Life Technologies) according to the manufacturer's instructions. Supernatants containing pseudoviruses were harvested 72 h after transfection, and centrifuged at 1200 × g for 10 min to remove cell debris. Pseudoviruses were concentrated using Amicon Ultra centrifugal filter units with a 100 kDa molecular weight cut-off, aliquoted and frozen at -80°C for later use.

**Pseudovirus entry into human and pig cells**

Retroviruses pseudotyped with PEDV spike, MERS-CoV spike, VSV envelope glycoprotein, or empty vector (mock) were used to transduce Huh-7 cells, Huh-7 cells transiently expressing human TMPRSS2, or PK-15 cells in 96-well plates. For trypsin processing of pseudoviruses, the initial DMEM medium was removed after pseudovirus attachment for 2 h, and was subsequently replaced with serum-free DMEM with 0 ug/ml, 10 μg/ml, or 40 μg/mL TPCK-treated trypsin (Sigma-Aldrich). After incubating with

trypsin at 37°C for 10 min, DMEM supplemented with 75 μg/mL soybean trypsin inhibitor (Sigma-Aldrich) was added to neutralize trypsin. Cells were incubated at 37°C for another 12 h, and medium was replaced with fresh DMEM. 48 h later, cells were washed with phosphate-buffered saline (PBS) and lysed. Aliquots of cell lysates were transferred to Optiplate-96 plate (PerkinElmer), and luciferase substrate (Promega) was addeded. Relative luciferase units (RLU) were measured using EnSpire plate reader (PerkinElmer).

**Inhibition of pseudovirus entry into human and pig cells using protease inhibitors**

Inhibition of pseudovirus entry using various protease inhibitors was carried out as previously described [88]. Briefly, target cells were pre-incubated with 20 μg/ml or 100 μg/ml bestatin, 10 μM or 50 μM proprotein convertases inhibitor Dec-RVKR-CMK (Enzo Life Sciences), 20 μM or 100 μM camostat mesylate (Sigma-Aldrich), 20 nM or 100 nM bafilomycin A1 (Sigma-Aldrich), 10 μM or 50 μM E-64d (Sigma-Aldrich), 50 μM cathepsin L inhibitor Z-FY-CHO (Santa Cruz Biotechnology), or 50 μM cathepsin B inhibitor CA-074 Me (Santa Cruz Biotechnology) at 37°C for 1 h. The cells were subsequently transduced by retroviruses pseudotyped with PEDV spike, MERS-CoV spike, or VSV envelope glycoprotein. The cells were incubated at 37°C for 6-8 h, and then medium was replaced with fresh DMEM. 48 h later, the cells were lysed and measured for luciferase activity.

**Western blot analysis of spike cleavage by proprotein convertases**

PEDV and MERS-CoV pseudoviruses were packaged in HEK293T cells, lysed and then subjected to Western blot analysis. The C9-tagged spikes were detected using anti-C9 tag monoclonal antibody (Santa Cruz Biotechnology).

**Western blot analysis of spike cleavage by lysosomal cysteine proteases.**

HEK293T cells were transfected with plasmids encoding PEDV spike or MERS-CoV spike. 48 h after transfection, the cells were harvested, washed with PBS, and lysed by sonication. Cell lysates were then incubated with activated cathepsin L or cathepsin B at gradient concentrations for 30 min at 37°C, and subjected to Western blot analysis. The C9-tagged spikes were detected using anti-C9 tag monoclonal antibody (Santa Cruz Biotechnology).

**Pre-treatment of Vero cells with lysosomal inhibitors and PEDV infection**

Vero CCL81 cells were washed 3 times with DMEM (Life Technologies), and then were pre-incubated with DMSO alone, or in the presence of 5 μg/ml trypsin, 40mM $NH_4Cl$, 50 μM E-64d, 50 μM cathepsin L inhibitor Z-FY-CHO (Santa Cruz Biotechnology), or 50 μM cathepsin B inhibitor CA-074 Me (Santa Cruz Biotechnology) in DMEM at 37°C for 1 h. Cells were infected with PEDV strain Ohio VBS2 at a multiplicity of infection (MOI) of 0.5 using a previously described procedure (cite our previous JV paper, PMID: 25787280). After absorption for 2 h, cells were washed with DMEM 3 times to remove unbound virus particles. Culture medium supplemented with 0.018% (w/v) Tryptose

Phosphate Broth (TPB) (Sigma-Aldrich), 0.02% yeast extract (Sigma-Aldrich), and the respective inhibitors or trypsin at the above concentration were added.

**Detection of PEDV N protein using an immunofluorescence assay**

24 h after infection, cells were washed twice with PBS and fixed with 4.0% (v/v) paraformaldehyde-0.2% (v/v) glutaraldehyde in 0.1 M potassium phosphate buffer (PPB), pH 7.4, for 15 min at 22°C, followed by washing 3 times with PBS. After permeabilization with 0.1% Triton X-100 in PBS for 15 min, the cells were washed with PBS, blocked with PBS containing 2% bovine serum albumin for 1 h at room temperature. Then cells were incubated with fluorescein isothiocyanate (FITC)-labeled mouse anti-PEDV N protein antibody (Medgene labs) in 0.2% BSA in PBS at dilution of 1:100 overnight at 4°C. Cells were examined under Olympus fluorescent microscope system at The Ohio State University.

**Quantification of PEDV viral RNA by RT-qPCR.**

24 h after infection, PEDV infected cells were trypsinized and harvested using 0.25% trypsin (Life Technologies). Harvested cells were centrifuged at 3000 × g for 10 min and viral RNA in the cell pellet was extracted using the RNeasy minikit (Qiagen). Reverse transcription (RT) was conducted using a primer (5' TTTTTTTTTTTTGTGTATCCAT 3') targeting the 3'-end UTR of PEDV and the Superscript III transcriptase kit (Life Technologies). The RT products were then used to perform real-time PCR using primers and probes specifically targeting the N gene of PEDV (forward, 5'

TGTTGATGCCTCAGGCTATGC 3'; reverse, 5' GCAATGCTGCAACATTTGGT 3'; probe, 5'-6FAM- CAGATCGCCAGTTTAG- MGB- 3' [6FAM is 6-carboxyfluorescein] [Applied Biosystems]) in a StepOne realtime PCR system (Applied Biosystems). A standard plasmid for PEDV was constructed by inserting the sequence of the entire PEDV N gene into the pGEM-T Easy vector (Promega). Amplification cycles used were 2 min at 50°C, 10 min at 95°C, and 40 cycles of 15 s at 95°C and 1 min at 60°C. The threshold for detection of fluorescence above the background was set within the exponential phase of the amplification curves. For each assay, 10-fold dilutions of standard plasmid were generated, and negative control samples and double-distilled water ($ddH_2O$) were included in each assay.

### 3.3 Results

To identify the proteases that activate PEDV entry, we examined potential spike-processing proteases from different stages of the virus infection cycle. We first evaluate the possibility that APN cleaves PEDV spike and activates PEDV entry. To this end, we packaged retrovirus particles pseudotyped with PEDV spike (i.e., PEDV pseudoviruses) in HEK293T cells (human embryonic kidney cells), and used it to enter Huh-7 cells endogenously expressing APN. We found that PEDV pseudovirus entry into Huh-7 cells was not affected by the treatment of APN active site inhibitor bestatin, suggesting that APN's catalytic activity is not involved in PEDV entry (Fig. 3.1).

After that, we analyzed whether proprotein convertases cleave PEDV spike during virus packaging. Western blot analysis of PEDV pseudovirus particles revealed that PEDV spike remained intact on the pseudovirus surface (Fig. 3.2A). Hence, proprotein convertases from virus-producing cells did not proteolytically activate PEDV spike. As a positive control, MERS-CoV spike had been cleaved on pseudovirus surfaces (Fig. 3.2A) [70,88]. Moreover, we examined whether proprotein convertases from virus-targeting cells cleave PEDV spike during virus endocytosis. Our result showed that the proportion convertase inhibitor, Dec-RVKR-CMK, did not affect PEDV pseudovirus entry into Huh-7 cells (human liver) or PK-15 cells (pig kidney) (Fig. 3.2B, 3.2C). Thus proprotein convertases from virus-targeting cells did not proteolytically activate PEDV spike either. As a positive control, MERS-CoV pseudoviruses demonstrated decreased entry into Huh-7 cells in the presence of Dec-RVKR-CMK (Fig. 3.2D). Overall, proprotein convertases from neither virus-producing cells nor virus-targeting cells play any significant role in processing PEDV spike and activating PEDV entry.

Next we investigated whether cell-surface proteases activate PEDV entry. Although previous studies demonstrated that Huh-7 cells do not express cell-surface protease TMPRSS2 [70,89], here we showed that PEDV pseudoviruses entered Huh-7 cells efficiently (Fig. 3.3A). Furthermore, exogenously expressing TMPRSS2 in Huh-7 cells did not enhance PEDV pseudovirus entry (Fig. 3.3A). Incubating Huh-7 cells or Huh-7 cells exogenously expressing TMPRSS2 with a TMPRSS2 inhibitor, camostat had no impact on PEDV pseudovirus entry either (Fig. 3.3A). These results all suggest that cell-surface proteases do not activate PEDV entry into host cells. As a positive control, camostat

reduced MERS-CoV pseudoviruses entry into Huh-7 cells exogenously expressing TMPRSS2 (Fig. 3.3B). As a negative control, camostat did not affect the entry of MERS-CoV pseudoviruses into Huh-7 cells not expressing TMPRSS2 (Fig. 3.3B). Therefore, we can rule out the role of cell-surface proteases in processing PEDV spike and activating PEDV entry.

Then we examined whether lysosomal cysteine proteases activate PEDV entry. To this end, we carried out PEDV pseudovirus entry into Huh-7 or PK-15 cells in the presence of lysosomal acidification inhibitor, bafilomycin A1, or lysosomal cysteine protease inhibitor, E-64d. We found that both inhibitors significantly reduced PEDV pseudovirus entry into host cells in a dose-dependent manner (Fig. 3.4A and 3.4B). As a control, only bafilomycin A1, but not E-64d, significantly reduced VSV pseudovirus entry into Huh-7 and PK-15 cells (Fig. 3.4C and 3.4D). The result from the control experiment is consistent with previous reports that VSV entry into host cells depends on endocytosis, but not lysosomal cysteine proteases [90,91]. The control experiment also showed that the inhibitors did not have non-specific cytotoxic effects on target cells. Thus, lysosomal cysteine proteases play a critical role in PEDV entry. We went further to pinpoint the specific lysosomal cysteine proteases that cleave PEDV spike and activate PEDV entry. We focused on cathepsin L and cathepsin B because both of these cathepsins have been previously identified to process the spike proteins from other coronaviruses including SARS and MERS coronaviruses [70,91-93]. To identify the role of cathepsin L and cathepsin B in PEDV entry, we carried out PEDV pseudovirus entry in the presence of inhibitors that are specific for cathepsin L or cathepsin B, respectively. The result indicated that

both inhibitors dramatically reduced PEDV pseudovirus entry into Huh-7 and PK-15 cells (Fig. 3.5A, 3.5B). To provide direct biochemical evidence that cathepsin L and cathepsin B cleave PEDV spike, we treated cell-surface-expressed PEDV spike with recombinant cathepsin L and cathepsin B, respectively, and detected the cleavage state of PEDV spike using Western blot. Our result showed that both cathepsin L and cathepsin B cleave PEDV spike to S2 (Fig. 3.5C). In sum, host lysosomal cysteine proteases, particularly cathepsin L and cathepsin B, process PEDV spike and activate PEDV entry.

Last we tackled the confounding role of extracellular protease trypsin in PEDV entry. Previous studies showed that exogenous trypsin could activate the entry of SARS and MERS coronaviruses into host cells after the viruses had already been attached to host cells [70,93,94]. Hence we added trypsin after PEDV pseudoviruses had been attached to Huh-7 or PK-15 cells. Our result revealed that trypsin slightly reduced PEDV pseudovirus entry into Huh7 or PK-15 cells (Fig. 3.6A and 3.6B). On the other hand, in the presence of cathepsin L or cathepsin B inhibitor, the dramatically reduced entry of PEDV pseudoviruses into host cells could be partially rescued by extracellular trypsin (Fig. 3.5A, 3.5B). Taken together, extracellular trypsin has the potential to process PEDV spike when lysosomal cysteine proteases are inhibited or unavailable; however, when available, lysosomal cysteine proteases play the major role in PEDV entry into host cells.

To further validate the role of lysosomal cysteine proteases and extracellular protease trypsin in activating PEDV entry, we conducted live-PEDV infection in Vero CCL-81 cells in the presence of above lysosome inhibitors or extracellular trypsin. The resulted

showed that all four lysosome inhibitors ($NH_4Cl$, E-64d, cathepsin L inhibitor and cathepsin B inhibitor) greatly blocked PEDV infection (Fig. 3.7A, 3.7B). Interestingly, the presence of trypsin drastically enhanced PEDV infection in Vero CCL-81 cells. The different effects of extracellular trypsin on PEDV pseudovirus entry and live-PEDV infection are likely due to its role in other stages of PEDV infection cycle.

## 3.4 Discussion

Our study has elucidated a long-standing puzzle regarding what proteases activate PEDV entry into host cells. Previous studies identified extracellular protease trypsin as required for PEDV infection in cell culture, which led to the conclusion that intestinal proteases are essential for PEDV entry [76,84,85]. However, these previous studies all used PEDV live virus particles, and thereby were unable to differentiate between PEDV entry and other steps in the PEDV infection cycle such as replication or release. Indeed, an electron microscopic study showed that PEDV release is a limiting step in the PEDV infection cycle and that trypsin is required for PEDV release [77]. To separate PEDV entry from other steps of the PEDV infection cycle, we performed a PEDV pseudovirus entry assay in which PEDV-spike-packaged pseudovirus particles can only enter host cells, but cannot replicate or be released. Thus the PEDV pseudovirus entry assay provides a simplified system for studying PEDV entry [86]. Using this assay, we showed that PEDV entry does not depend on proprotein convertases or cell-surface proteases. Instead, PEDV entry is activated by lysosomal cysteine proteases. Using pseudovirus entry, live PEDV infection and direct biochemical assays, we further identified cathepsin L and cathepsin B

as the specific lysosomal cysteine proteases that can process PEDV spike. Our study also revealed the puzzling role of extracellular trypsin in PEDV entry. When cathepsins are available, trypsin is not as efficient as cathepsins in activating PEDV entry. However, when cathepsins are inhibited or unavailable, trypsin can fill in the role of cathepsins by activating PEDV entry. Intriguingly, in contrast to its minor negative effect on PEDV pseudovirus entry, trypsin dramatically enhanced the infection of authentic PEDV, indicating a role of trypsin in other steps of the PEDV infection cycle such as PEDV release, which remains to be investigated using other research approaches. Nevertheless, our study has laid out a blueprint to systematically examine the roles of proteases in virus entry, and provided insight into the puzzling molecular mechanism for PEDV entry.

PEDV is currently sweeping through America's pig populations with little hindrance, as neither vaccines nor antiviral drugs are available to curb its spread. Our study has identified cysteine protease inhibitors, such as MDL 28170, as a class of antiviral agents that can potentially block PEDV infections. However, our finding also suggests that cysteine protease inhibitors alone may not be sufficient to block PEDV entry because trypsin can serve a backup role in activating PEDV entry. Instead, a combinational use of cysteine protease and trypin inhibitors may be more effective to block PEDV entry and treat infected pigs.

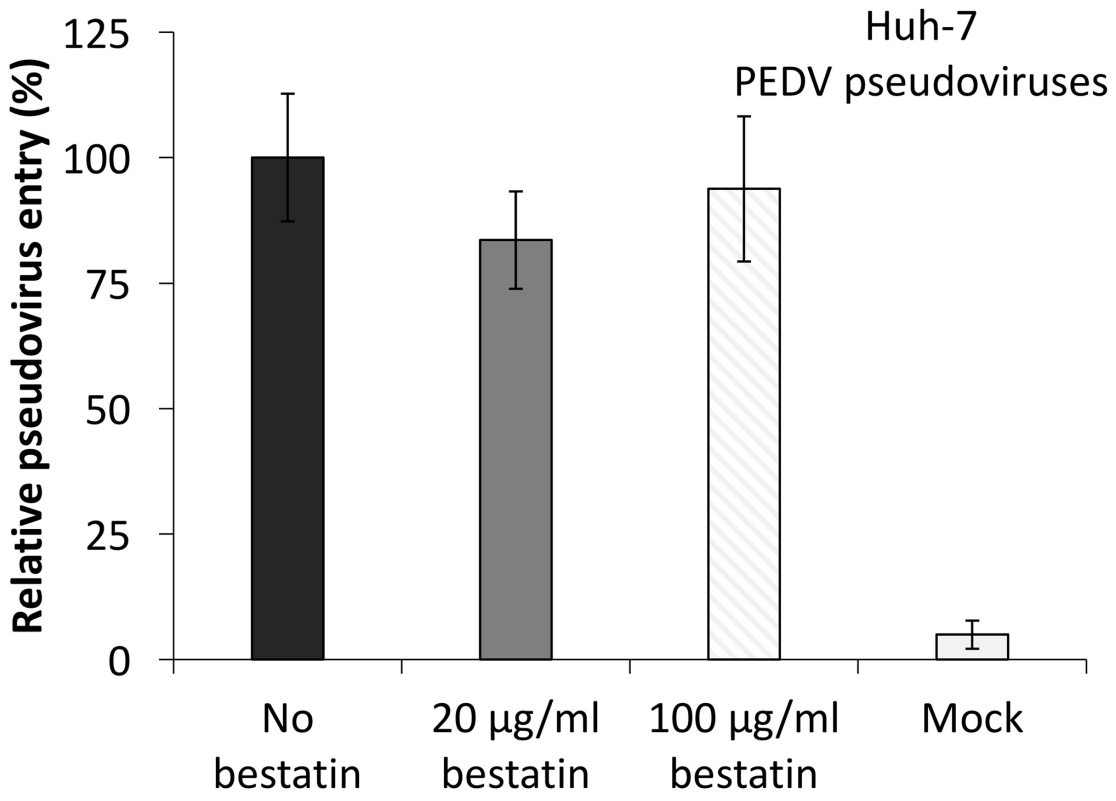



**Figure 3.1 The catalytic activity of APN is not involved in PEDV entry.**

Huh-7 cells were pre-incubated with APN active site inhibitor bestatin at indicated concentrations, and transduced by PEDV pseudoviruses. Empty vector-packaged pseudoviruses (mock) were used as a negative control. The pseudovirus entry efficiency was characterized as luciferase activity accompanying the entry. The pseudovirus entry in target cells without any inhibitor treatment was taken as 100%. Error bars indicate SEM (n = 5).

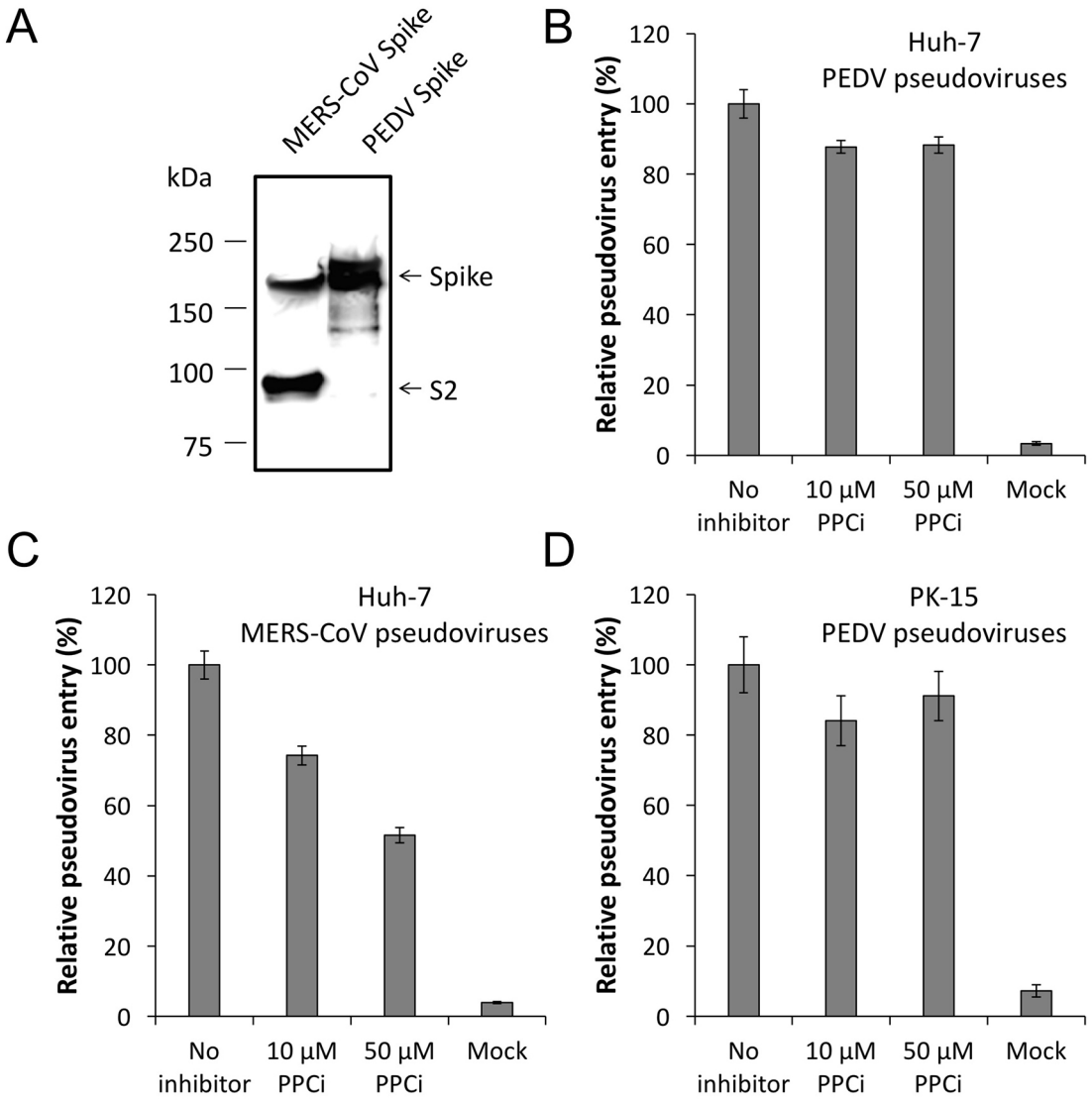

A
MERS-CoV Spike
PEDV Spike
kDa
250
150
100
75
← Spike
← S2
B
Huh-7
PEDV pseudoviruses
Relative pseudovirus entry (%)
120
100
80
60
40
20
0
No inhibitor
10 μM PPCi
50 μM PPCi
Mock
C
Huh-7
MERS-CoV pseudoviruses
Relative pseudovirus entry (%)
No inhibitor
10 μM PPCi
50 μM PPCi
Mock
D
PK-15
PEDV pseudoviruses
Relative pseudovirus entry (%)
No inhibitor
10 μM PPCi
50 μM PPCi
Mock

**Figure 3.2 Proprotein convertases do not activate PEDV entry.**

(A) Western blot analysis of PEDV spike in pseudovirus particles. Retroviruses pseudotyped with PEDV spike (i.e., PEDV pseudovirus) were produced in HEK293T cells and then subjected to Western blot analysis using antibody against its C-terminal C9 tag. Retroviruses pseudotyped with MERS-CoV spike (i.e. MERS-CoV pseudovirus) were used as a control. (B) and (C) Huh-7 cells were pre-incubated with PPCi (proprotein convertases inhibitor, Dec-RVKR-CMK) at indicated concentrations, and then transduced by PEDV pseudoviruses (B) or MERS-CoV pseudoviruses (C). Empty vector-packaged pseudoviruses (mock) were used as a negative control. (D) PK-15 cells were pre-incubated with PPCi at indicated concentrations, and transduced by PEDV pseudoviruses. The pseudovirus entry efficiency was characterized as luciferase activity accompanying the entry. The pseudovirus entry in target cells without any inhibitor treatment was taken as 100%. Error bars indicate SEM (n = 5).

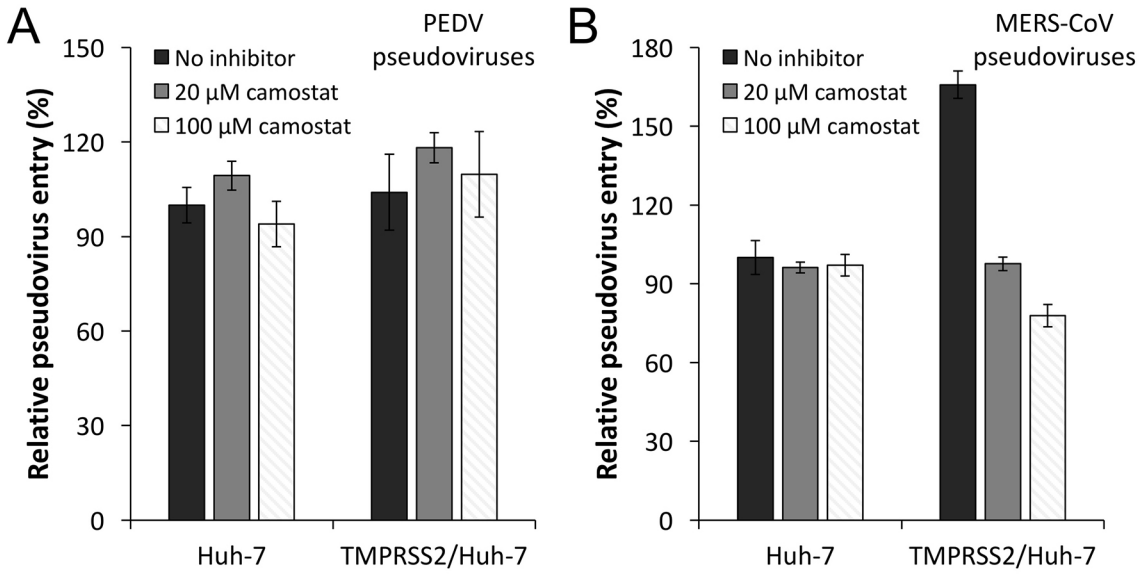



**Figure 3.3 Cell-surface serine proteases do not activate PEDV entry.**

Huh-7 cells transiently transfected with empty pCAGGS vector or TMPRSS2 in pCAGGS vector were pre-incubated with camostat (cell-surface serine proteases inhibitor) at indicated concentrations, and were transduced by PEDV pseudoviruses (A) or MERS-CoV pseudoviruses (B). The pseudovirus entry in empty pCAGGS vector-transfected Huh-7 cells without any inhibitor treatment was taken as 100%. Error bars indicate SEM (n = 5).

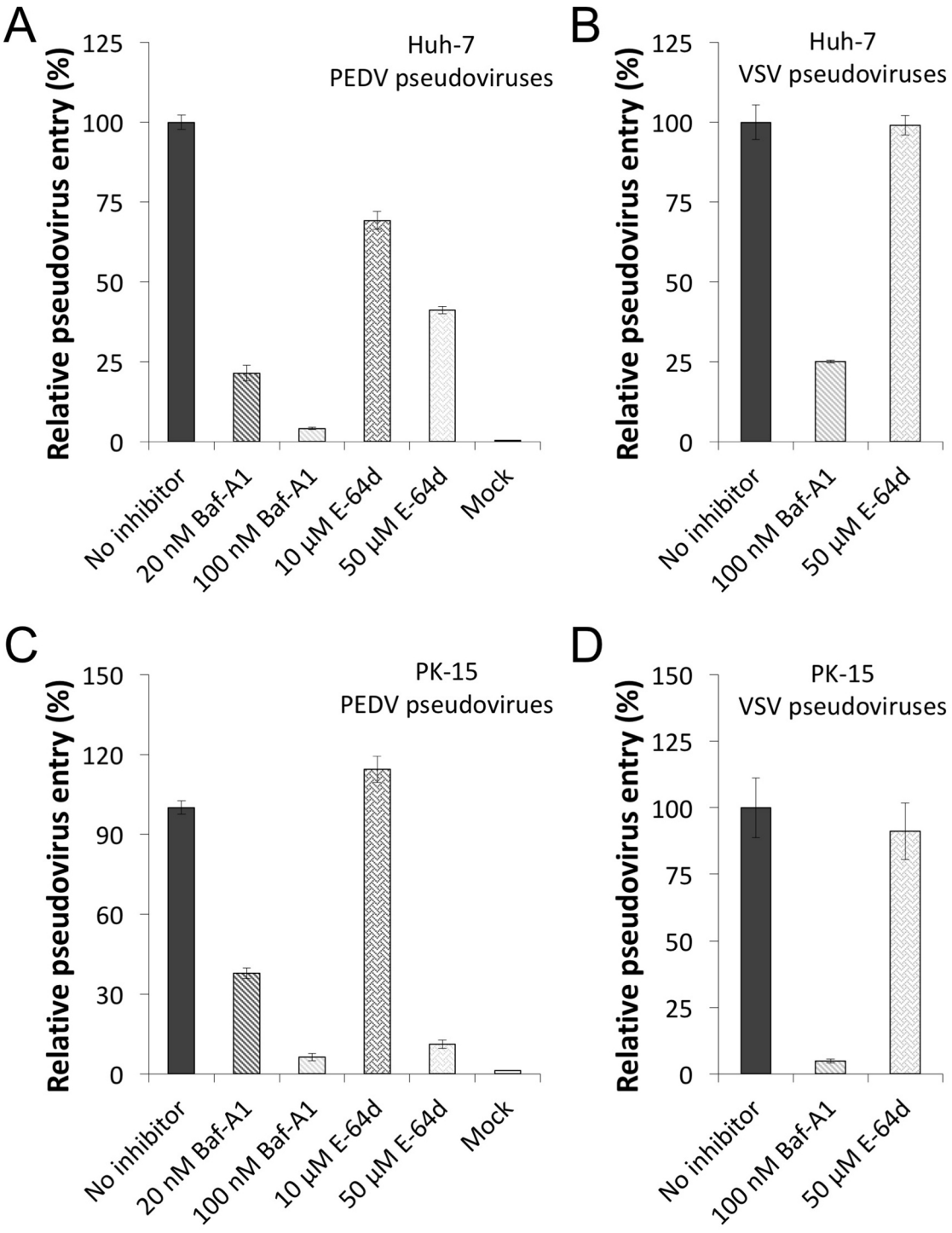

A
Huh-7
PEDV pseudoviruses
Relative pseudovirus entry (%)
125
100
75
50
25
0
No inhibitor
20 nM Baf-A1
100 nM Baf-A1
10 μM E-64d
50 μM E-64d
Mock
B
Huh-7
VSV pseudoviruses
Relative pseudovirus entry (%)
125
100
75
50
25
0
No inhibitor
100 nM Baf-A1
50 μM E-64d
C
PK-15
PEDV pseudovirues
Relative pseudovirus entry (%)
150
120
90
60
30
0
No inhibitor
20 nM Baf-A1
100 nM Baf-A1
10 μM E-64d
50 μM E-64d
Mock
D
PK-15
VSV pseudoviruses
Relative pseudovirus entry (%)
150
125
100
75
50
25
0
No inhibitor
100 nM Baf-A1
50 μM E-64d

**Figure 3.4 Lysosomal cysteine proteases activate PEDV entry.**

(A) Huh-7 cells or (B) PK-15 cells were pre-incubated with Baf-A1 (Bafilomycin A1) (endosomal acidification inhibitor) or E-64d (lysosomal cysteine protease inhibitor) at indicated concentrations, and then transduced by PEDV pseudoviruses. (C) and (D) Retroviruses pseudotyped with VSV envelop glycoprotein (i.e. VSV pseudoviruses) were used as a control. The pseudovirus entry in target cells without any inhibitor treatment was taken as 100%. Error bars indicate SEM (n = 5).

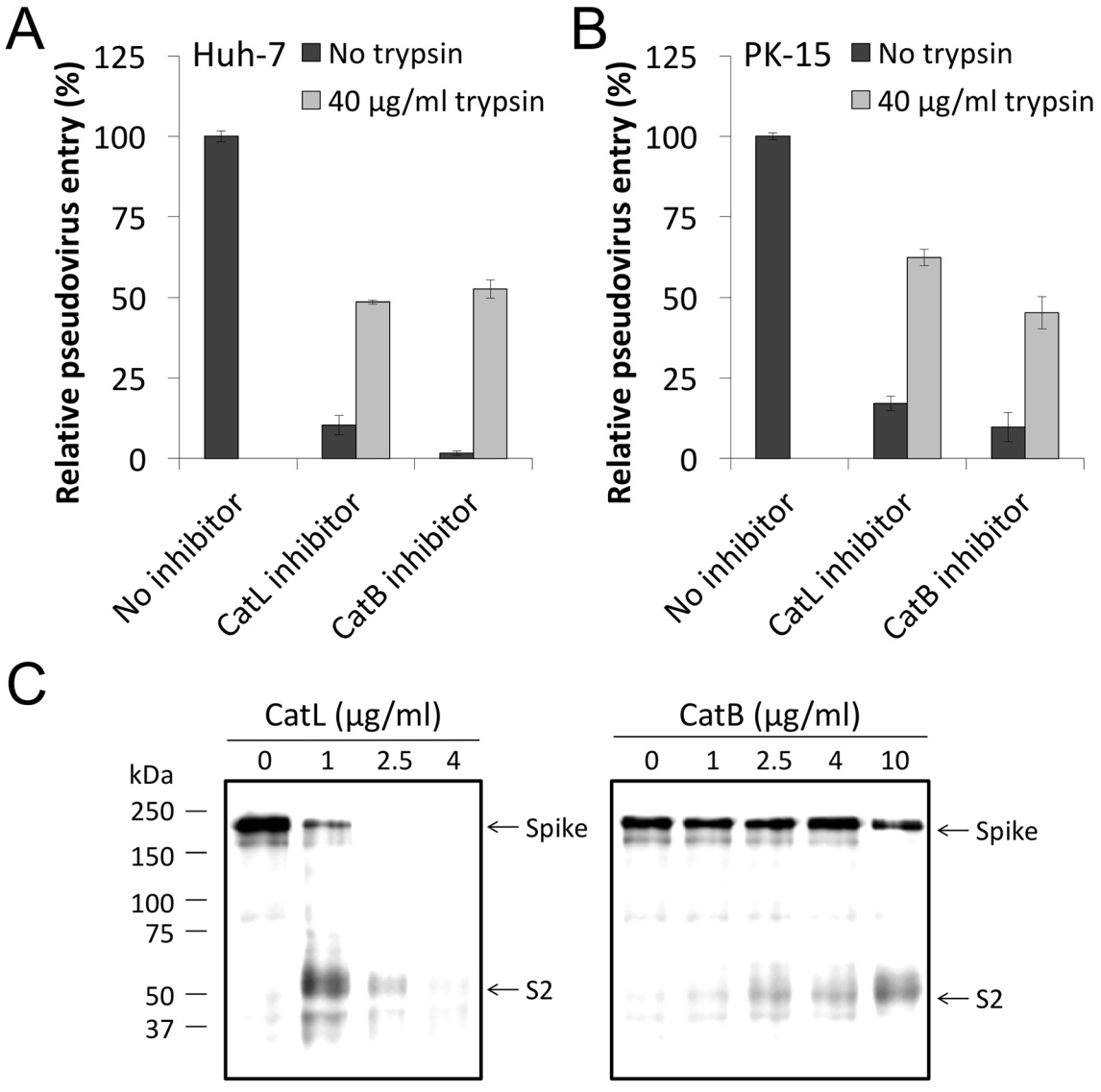

A
Huh-7
No trypsin
40 μg/ml trypsin
Relative pseudovirus entry (%)
125
100
75
50
25
0
No inhibitor
CatL inhibitor
CatB inhibitor
B
PK-15
No trypsin
40 μg/ml trypsin
Relative pseudovirus entry (%)
125
100
75
50
25
0
No inhibitor
CatL inhibitor
CatB inhibitor
C
CatL (μg/ml)
0
1
2.5
4
CatB (μg/ml)
0
1
2.5
4
10
kDa
250
150
100
75
50
37
Spike
S2
Spike
S2

**Figure 3.5 Cathepsin L and cathepsin B activate PEDV entry.**

Before being infected by PEDV pseudoviruses, Huh-7 cells (A) or PK-15 cells (B) were pre-incubated with 50 μM CatL inhibitor (cathepsin L inhibitor Z-FY-CHO) or 50 μM CatB inhibitor (cathepsin B inhibitor CA-074 Me). After pseudovirus attachment to target cells for 3 h, unbound pseudovirus particles were removed and bound pseudovirus particles were either treated or not treated with 40 μg/ml exogenous trypsin. PEDV pseudovirus entry in the absence of inhibitor or exogenous trypsin was taken as 100% in each cell line. Error bars indicated SEM (n = 4). (C) PEDV spike was transiently expressed on the surface of HEK293T cells, and subsequently subjected to cathepsin L or cathepsin B cleavage at gradient concentrations. The spike was detected by an antibody against its C-terminal C9 tag.

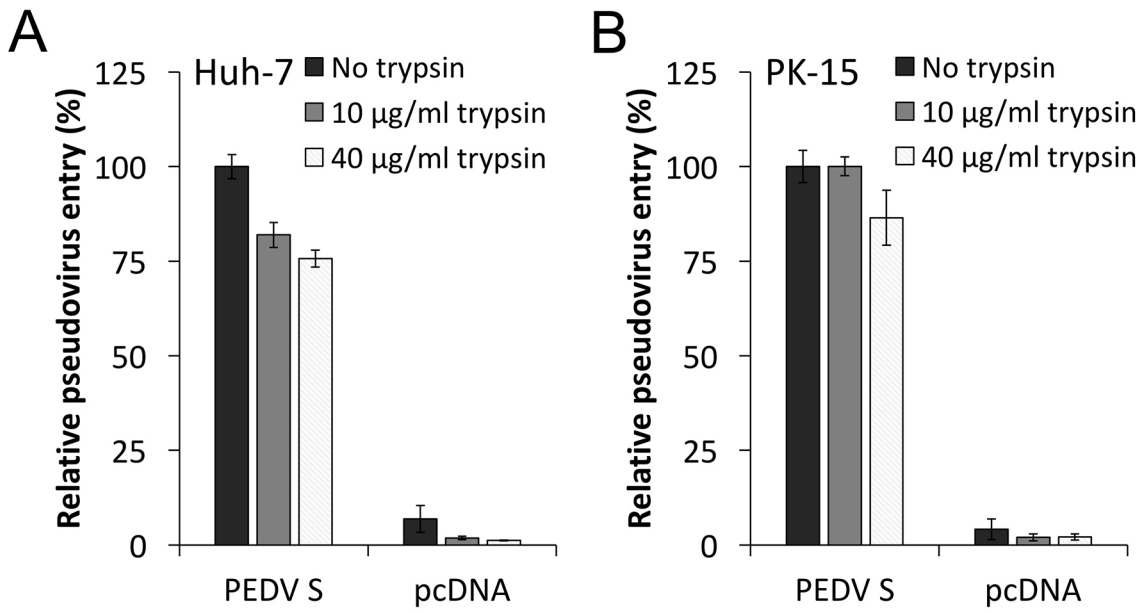



**Figure 3.6 Extracellular protease trypsin serves a backup role in activating PEDV entry.**

After PEDV pseudoviruses were incubated with Huh-7 cells (A) or PK-15 cells (B) for 3 h, unbound PEDV pseudovirus particles were removed and bound PEDV pseudovirus particles were either treated or not treated with exogenous trypsin at indicated concentrations. Empty pcDNA vector-packaged pseudoviruses were used as a negative control. The pseudovirus entry mediated by PEDV spike in the absence of exogenous trypsin was taken as 100%. Error bars indicate SEM (n = 4).

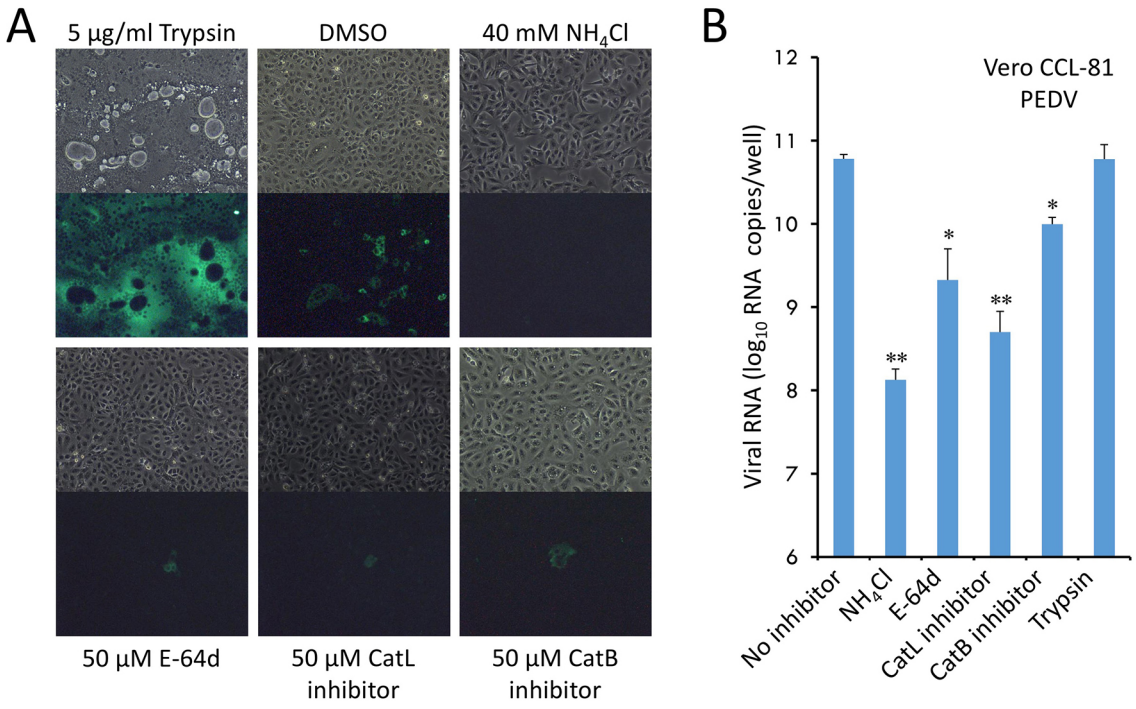



**Figure 3.7 Lysosomal dysfunction serverly impaired PEDV infection in cell culture.** Vero CCL-81 cells were treated DMSO alone, or DMSO in the presence of 5 µg/ml trypsin or various lysosomal inhibitors at indicated concentrations, followed by PEDV infection. (A) 24 h post-infection, cells were fixed and detected with fluorescein isothiocyanate (FITC)-labeled moouse anti-PEDV N protein antibody. Virus positive cells were observed using a fluorescence microscope. (B) 24 h post-infection, total RNA was extracted from cells. Viral RNA genome copies were quantified by real-time RT-PCR. Virus titers are expressed as mean $\log_{10}$ RNA copies/well ± 1 standard deviation. Differences were considered significant when $P \leq 0.05$ (*) and highly significant when $P \leq 0.01$ (**) compared to the DMSO treatment alone.

# 4 Chapter Four

## Conclusions: novel insights into the multi-functional roles of aminopeptidase N in cancer and coronavirus infection

Cancer and virus infection are two major threats to human health. APN plays critical roles in both diseases. Despite substantial advances in cancer prevention and treatment in the past decades, cancer is still the second leading cause of death in United States, partially because current cancer therapies are largely limited by their side effects associated with non-specificity. Thus, continuous improvement in specific cancer treatment is highly demanded. More recently, targeted cancer therapy emerged as a promising new era of cancer therapeutics due to its selectivity for cancer cells. APN is one such target. It serves as a receptor for NGR tumor-homing peptides [2]. By solving the crystal structure of APN in complex with a NGR tumor-homing peptide, my work unraveled the molecular details of this interaction at the cancer cell-drug interface. The NGR motif in tumor-homing peptide is reminiscent of the NGR sequence in some ECM proteins, such as fibronectin. This sequence similarity raises the possibility that APN interacts with the NGR sequences in ECM proteins to fulfill its role in tumor cell motility. Using cellular biology and biochemical approaches, my work demonstrated that APN bound to the NGR motifs in both tumor-homing peptide and ECM proteins in a specific and non-degrading manner [72]. This study not only solved the puzzles regarding APN's function in tumor-homing therapy and tumor cell motility, but also facilitates the development of APN-targeting cancer therapies.

Since the first identification of APN as cellular receptors for TGEV and HCoV-229E two decades ago [3,5], more and more coronaviruses, such as procine respiratory coronavirus (PRCV), feline coronavirus (FCoV) and canine coronavirus (CCoV), were found to

recognize APN during their cell entry process. Recent crystal structure of APN in complex with the PRCV RBD has revealed the molecular basis of the binding between APN and APN-recognizing coronaviruses [95]. In 2013, a new strain of PEDV, a TGEV-related porcine coronavirus, emerged in the United States, and has since spread rapidly through the America's pig populations. PEDV causes 80-100% fatality rate in infected piglets, and has led to substantial economic losses and public health concerns. However, no effective treatment is currently available to keep the spread of PEDV in check partly due to the gap in the knowledge about how PEDV infects cells. To fill this gap, I investigated two essential steps in PEDV infection, receptor binding and cell entry. I was surprised to discover that PEDV not only possessed the capability to infect pig cells, it was also able to infect human cells because both human and porcine APN can serve as a receptor for PEDV. These findings suggest PEDV as a potential threat to human. This study also provides a basis for design and development of effective vaccines and antiviral drugs to curb the spread of PEDV in pigs and to other species.

In addition to APN, host proteases are also required for efficient cell entry of PEDV. It has been commonly believed that PEDV entry depends on extracellular proteases in pig intestines because PEDV replication in cell culture requires trypsin [76,84,85]. However, using live PEDV particles in cell culture cannot differentiate between PEDV entry and other steps of the PEDV infection cycle such as PEDV replication and release. Therefore, I examined PEDV entry using a PEDV pseudovirus entry assay, which only concerns PEDV entry, but not PEDV replication or release. I found that PEDV entry is mainly

activated by lysosomal cysteine proteases, more specifically, cathepsin L and cathepsin B, while trypsin only serves as a backup role in activating PEDV entry when lysosomal cysteine proteases are not accessible or inhibited. This study provides molecular insights into PEDV entry mechanisms and therapeutic implications for the development of anti-PEDV strategies.

The broad tissue distribution of APN determines its ubiquitous role in different biological processes. My thesis studies contribute to a better understanding of these multiple functions of APN, can facilitate the design of drugs targeting specific aspects of APN for treating particular diseases, with much less interference of its other biological functions.

## Bibliography


1 Mina-Osorio, P. The moonlighting enzyme CD13: old and new functions to target. *Trends Mol Med* **14**, 361-371, doi:10.1016/j.molmed.2008.06.003 (2008).

2 Pasqualini, R. *et al.* Aminopeptidase N is a receptor for tumor-homing peptides and a target for inhibiting angiogenesis. *Cancer Res* **60**, 722-727 (2000).

3 Delmas, B. *et al.* Aminopeptidase N is a major receptor for the entero-pathogenic coronavirus TGEV. *Nature* **357**, 417-420, doi:10.1038/357417a0 (1992).

4 Tresnan, D. B., Levis, R. & Holmes, K. V. Feline aminopeptidase N serves as a receptor for feline, canine, porcine, and human coronaviruses in serogroup I. *Journal of virology* **70**, 8669-8674 (1996).

5 Yeager, C. L. *et al.* Human aminopeptidase N is a receptor for human coronavirus 229E. *Nature* **357**, 420-422, doi:10.1038/357420a0 (1992).

6 Chen, L., Lin, Y. L., Peng, G. & Li, F. Structural basis for multifunctional roles of mammalian aminopeptidase N. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 17966-17971, doi:10.1073/pnas.1210123109 (2012).

7 Tusell, S. M., Schittone, S. A. & Holmes, K. V. Mutational analysis of aminopeptidase N, a receptor for several group 1 coronaviruses, identifies key determinants of viral host range. *Journal of virology* **81**, 1261-1273, doi:10.1128/JVI.01510-06 (2007).

8 Wentworth, D. E. & Holmes, K. V. Molecular determinants of species specificity in the coronavirus receptor aminopeptidase N (CD13): influence of N-linked

glycosylation. *Journal of virology* **75**, 9741-9752, doi:10.1128/JVI.75.20.9741-9752.2001 (2001).

9 Wickstrom, M., Larsson, R., Nygren, P. & Gullbo, J. Aminopeptidase N (CD13) as a target for cancer chemotherapy. *Cancer Sci* **102**, 501-508, doi:10.1111/j.1349-7006.2010.01826.x (2011).

10 Bauvois, B. Transmembrane proteases in cell growth and invasion: new contributors to angiogenesis? *Oncogene* **23**, 317-329, doi:10.1038/sj.onc.1207124 (2004).

11 Arap, W., Pasqualini, R. & Ruoslahti, E. Cancer treatment by targeted drug delivery to tumor vasculature in a mouse model. *Science* **279**, 377-380 (1998).

12 Carl-McGrath, S., Lendeckel, U., Ebert, M. & Rocken, C. Ectopeptidases in tumour biology: a review. *Histol Histopathol* **21**, 1339-1353 (2006).

13 Curnis, F. *et al.* Enhancement of tumor necrosis factor alpha antitumor immunotherapeutic properties by targeted delivery to aminopeptidase N (CD13). *Nat Biotechnol* **18**, 1185-1190, doi:10.1038/81183 (2000).

14 Petrovic, N. *et al.* CD13/APN regulates endothelial invasion and filopodia formation. *Blood* **110**, 142-150, doi:10.1182/blood-2006-02-002931 (2007).

15 Fujii, H. *et al.* Human melanoma invasion and metastasis enhancement by high expression of aminopeptidase N/CD13. *Clin Exp Metastasis* **13**, 337-344 (1995).

16 Kehlen, A., Lendeckel, U., Dralle, H., Langner, J. & Hoang-Vu, C. Biological significance of aminopeptidase N/CD13 in thyroid carcinomas. *Cancer Res* **63**, 8500-8506 (2003).

17 Chang, Y. W. *et al.* CD13 (aminopeptidase N) can associate with tumor-associated antigen L6 and enhance the motility of human lung cancer cells. *Int J Cancer* **116**, 243-252, doi:10.1002/ijc.21089 (2005).

18 Hashida, H. *et al.* Aminopeptidase N is involved in cell motility and angiogenesis: its clinical significance in human colon cancer. *Gastroenterology* **122**, 376-386 (2002).

19 Menrad, A., Speicher, D., Wacker, J. & Herlyn, M. Biochemical and functional characterization of aminopeptidase N expressed by human melanoma cells. *Cancer Res* **53**, 1450-1455 (1993).

20 Saiki, I. *et al.* Role of aminopeptidase N (CD13) in tumor-cell invasion and extracellular matrix degradation. *Int J Cancer* **54**, 137-143 (1993).

21 Fukasawa, K. *et al.* Aminopeptidase N (APN/CD13) is selectively expressed in vascular endothelial cells and plays multiple roles in angiogenesis. *Cancer Lett* **243**, 135-143, doi:10.1016/j.canlet.2005.11.051 (2006).

22 Carlsson, L., Ronquist, G., Eliasson, R., Egberg, N. & Larsson, A. Flow cytometric technique for determination of prostasomal quantity, size and expression of CD10, CD13, CD26 and CD59 in human seminal plasma. *Int J Androl* **29**, 331-338, doi:10.1111/j.1365-2605.2005.00601.x (2006).

23 Gonzalez Buitrago, J. M., Navajo, J. A., Garcia Diez, L. C. & Herruzo, A. Seminal plasma leucine aminopeptidase in male fertility. *Andrologia* **17**, 139-142 (1985).

24 Irazusta, J. *et al.* Enkephalin-degrading enzymes in normal and subfertile human semen. *J Androl* **25**, 733-739 (2004).

25 Mina-Osorio, P., Shapiro, L. H. & Ortega, E. CD13 in cell adhesion: aminopeptidase N (CD13) mediates homotypic aggregation of monocytic cells. *J Leukoc Biol* **79**, 719-730, doi:10.1189/jlb.0705425 (2006).

26 Mina-Osorio, P. *et al.* CD13 is a novel mediator of monocytic/endothelial cell adhesion. *J Leukoc Biol* **84**, 448-459, doi:10.1189/jlb.1107802 (2008).

27 Shimizu, T. *et al.* CD13/aminopeptidase N-induced lymphocyte involvement in inflamed joints of patients with rheumatoid arthritis. *Arthritis Rheum* **46**, 2330-2338, doi:10.1002/art.10517 (2002).

28 Tani, K. *et al.* CD13/aminopeptidase N, a novel chemoattractant for T lymphocytes in pulmonary sarcoidosis. *Am J Respir Crit Care Med* **161**, 1636-1642, doi:10.1164/ajrccm.161.5.9902008 (2000).

29 Arnaout, M. A., Goodman, S. L. & Xiong, J. P. Coming to grips with integrin binding to ligands. *Curr Opin Cell Biol* **14**, 641-651 (2002).

30 Desgrosellier, J. S. & Cheresh, D. A. Integrins in cancer: biological implications and therapeutic opportunities. *Nat Rev Cancer* **10**, 9-22, doi:10.1038/nrc2748 (2010).

31 Guo, W. & Giancotti, F. G. Integrin signalling during tumour progression. *Nat Rev Mol Cell Biol* **5**, 816-826, doi:10.1038/nrm1490 (2004).

32 Xiong, J. P. *et al.* Crystal structure of the extracellular segment of integrin alpha Vbeta3. *Science* **294**, 339-345, doi:10.1126/science.1064535 (2001).

33 Xiong, J. P. *et al.* Crystal structure of the extracellular segment of integrin alpha Vbeta3 in complex with an Arg-Gly-Asp ligand. *Science* **296**, 151-155, doi:10.1126/science.1069040 (2002).

34 Sugahara, K. N. *et al.* Coadministration of a tumor-penetrating peptide enhances the efficacy of cancer drugs. *Science* **328**, 1031-1035, doi:10.1126/science.1183057 (2010).

35 Svensen, N., Walton, J. G. & Bradley, M. Peptides for cell-selective drug delivery. *Trends Pharmacol Sci* **33**, 186-192, doi:10.1016/j.tips.2012.02.002 (2012).

36 Eskens, F. A. *et al.* Phase I and pharmacokinetic study of continuous twice weekly intravenous administration of Cilengitide (EMD 121974), a novel inhibitor of the integrins alphavbeta3 and alphavbeta5 in patients with advanced solid tumours. *Eur J Cancer* **39**, 917-926 (2003).

37 Gregorc, V. *et al.* Defining the optimal biological dose of NGR-hTNF, a selective vascular targeting agent, in advanced solid tumours. *Eur J Cancer* **46**, 198-206, doi:10.1016/j.ejca.2009.10.005 (2010).

38 Gregorc, V. *et al.* Phase Ib study of NGR-hTNF, a selective vascular targeting agent, administered at low doses in combination with doxorubicin to patients with advanced solid tumours. *Br J Cancer* **101**, 219-224, doi:10.1038/sj.bjc.6605162 (2009).

39 Gregorc, V. *et al.* Phase II study of asparagine-glycine-arginine-human tumor necrosis factor alpha, a selective vascular targeting agent, in previously treated

patients with malignant pleural mesothelioma. *J Clin Oncol* **28**, 2604-2611, doi:10.1200/JCO.2009.27.3649 (2010).

40 Reardon, D. A. *et al.* Randomized phase II study of cilengitide, an integrin-targeting arginine-glycine-aspartic acid peptide, in recurrent glioblastoma multiforme. *J Clin Oncol* **26**, 5610-5617, doi:10.1200/JCO.2008.16.7510 (2008).

41 Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307-326 (1997).

42 Brunger, A. T. *et al.* Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* **54**, 905-921 (1998).

43 Murshudov, G. N., Vagin, A. A., Lebedev, A., Wilson, K. S. & Dodson, E. J. Efficient anisotropic refinement of macromolecular structures using FFT. *Acta Crystallogr D Biol Crystallogr* **55**, 247-255, doi:10.1107/S090744499801405X (1999).

44 Yang, Y., Liu, C., Lin, Y. L. & Li, F. Structural insights into central hypertension regulation by human aminopeptidase A. *The Journal of biological chemistry* **288**, 25638-25645, doi:10.1074/jbc.M113.494955 (2013).

45 Shor, A. C. *et al.* Dasatinib inhibits migration and invasion in diverse human sarcoma cell lines and induces apoptosis in bone sarcoma cells dependent on SRC kinase for survival. *Cancer Res* **67**, 2800-2808, doi:10.1158/0008-5472.CAN-06-3469 (2007).

46 Chen, Y. *et al.* Combined integrin phosphoproteomic analyses and small interfering RNA--based functional screening identify key regulators for cancer cell adhesion and migration. *Cancer Res* **69**, 3713-3720, doi:10.1158/0008-5472.CAN-08-2515 (2009).

47 Peng, G. *et al.* Crystal structure of mouse coronavirus receptor-binding domain complexed with its murine receptor. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 10696-10701, doi:10.1073/pnas.1104306108 (2011).

48 Du, L. *et al.* A conformation-dependent neutralizing monoclonal antibody specifically targeting receptor-binding domain in Middle East respiratory syndrome coronavirus spike protein. *J Virol* **88**, 7045-7053, doi:10.1128/JVI.00433-14 (2014).

49 Di Matteo, P. *et al.* Immunogenic and structural properties of the Asn-Gly-Arg (NGR) tumor neovasculature-homing motif. *Mol Immunol* **43**, 1509-1518, doi:10.1016/j.molimm.2005.10.009 (2006).

50 Williams, M. J. *et al.* Solution structure of a pair of fibronectin type 1 modules with fibrin binding activity. *J Mol Biol* **235**, 1302-1311, doi:10.1006/jmbi.1994.1083 (1994).

51 Pensaert, M. B. & de Bouck, P. A new coronavirus-like particle associated with diarrhea in swine. *Arch Virol* **58**, 243-247 (1978).

52 Song, D. & Park, B. Porcine epidemic diarrhoea virus: a comprehensive review of molecular epidemiology, diagnosis, and vaccines. *Virus Genes* **44**, 167-175, doi:10.1007/s11262-012-0713-1 (2012).

53 Sun, R. Q. *et al.* Outbreak of porcine epidemic diarrhea in suckling piglets, China. *Emerg Infect Dis* **18**, 161-163, doi:10.3201/eid1801.111259 (2012).

54 Chen, Q. *et al.* Isolation and characterization of porcine epidemic diarrhea viruses associated with the 2013 disease outbreak among swine in the United States. *J Clin Microbiol* **52**, 234-243, doi:10.1128/JCM.02820-13 (2014).

55 Mole, B. Deadly pig virus slips through US borders. *Nature* **499**, 388, doi:10.1038/499388a (2013).

56 Stevenson, G. W. *et al.* Emergence of Porcine epidemic diarrhea virus in the United States: clinical signs, lesions, and viral genomic sequences. *J Vet Diagn Invest* **25**, 649-654, doi:10.1177/1040638713501675 (2013).

57 Li, W. *et al.* Animal origins of the severe acute respiratory syndrome coronavirus: insight from ACE2-S-protein interactions. *Journal of virology* **80**, 4211-4219, doi:10.1128/JVI.80.9.4211-4219.2006 (2006).

58 Perlman, S. & Netland, J. Coronaviruses post-SARS: update on replication and pathogenesis. *Nat Rev Microbiol* **7**, 439-450, doi:10.1038/nrmicro2147 (2009).

59 Huang, Y. W. *et al.* Origin, evolution, and genotyping of emergent porcine epidemic diarrhea virus strains in the United States. *MBio* **4**, e00737-00713, doi:10.1128/mBio.00737-13 (2013).

60 Li, F. Evidence for a common evolutionary origin of coronavirus spike protein receptor-binding subunits. *Journal of virology* **86**, 2856-2858, doi:10.1128/JVI.06882-11 (2012).

61 Li, F. Receptor recognition mechanisms of coronaviruses: a decade of structural studies. *Journal of virology* **89**, 1954-1964, doi:10.1128/JVI.02615-14 (2015).

62 Li, F. Receptor recognition and cross-species infections of SARS coronavirus. *Antiviral Res* **100**, 246-254, doi:10.1016/j.antiviral.2013.08.014 (2013).

63 Li, F., Li, W., Farzan, M. & Harrison, S. C. Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. *Science* **309**, 1864-1868, doi:10.1126/science.1116480 (2005).

64 Li, W. *et al.* Receptor and viral determinants of SARS-coronavirus adaptation to human ACE2. *EMBO J* **24**, 1634-1643, doi:10.1038/sj.emboj.7600640 (2005).

65 Krempl, C., Schultze, B., Laude, H. & Herrler, G. Point mutations in the S protein connect the sialic acid binding activity with the enteropathogenicity of transmissible gastroenteritis coronavirus. *Journal of virology* **71**, 3285-3287 (1997).

66 Schultze, B. *et al.* Transmissible gastroenteritis coronavirus, but not the related porcine respiratory coronavirus, has a sialic acid (N-glycolylneuraminic acid) binding activity. *Journal of virology* **70**, 5634-5637 (1996).

67 Schwegmann-Wessels, C. & Herrler, G. Sialic acids as receptor determinants for coronaviruses. *Glycoconj J* **23**, 51-58, doi:10.1007/s10719-006-5437-9 (2006).

68 Krempl, C., Laude, H. & Herrler, G. Is the sialic acid binding activity of the S protein involved in the enteropathogenicity of transmissible gastroenteritis virus? *Adv Exp Med Biol* **440**, 557-561 (1998).

69 Li, B. X., Ge, J. W. & Li, Y. J. Porcine aminopeptidase N is a functional receptor for the PEDV coronavirus. *Virology* **365**, 166-172, doi:10.1016/j.virol.2007.03.031 (2007).

70 Yang, Y. *et al.* Receptor usage and cell entry of bat coronavirus HKU4 provide insight into bat-to-human transmission of MERS coronavirus. *Proc Natl Acad Sci U S A* **111**, 12516-12521, doi:10.1073/pnas.1405889111 (2014).

71 Ma, C. *et al.* Searching for an ideal vaccine candidate among different MERS coronavirus receptor-binding fragments--the importance of immunofocusing in subunit vaccine design. *Vaccine* **32**, 6170-6176, doi:10.1016/j.vaccine.2014.08.086 (2014).

72 Liu, C., Yang, Y., Chen, L., Lin, Y. L. & Li, F. A unified mechanism for aminopeptidase N-based tumor cell motility and tumor-homing therapy. *The Journal of biological chemistry* **289**, 34520-34529, doi:10.1074/jbc.M114.566802 (2014).

73 Peng, G. *et al.* Crystal structure of bovine coronavirus spike protein lectin domain. *The Journal of biological chemistry* **287**, 41931-41938, doi:10.1074/jbc.M112.418210 (2012).

74 Haraguchi, N. *et al.* CD13 is a therapeutic target in human liver cancer stem cells. *J Clin Invest* **120**, 3326-3339, doi:10.1172/JCI42550 (2010).

75 Weingartl, H. M. & Derbyshire, J. B. Antiviral activity against transmissible gastroenteritis virus, and cytotoxicity, of natural porcine interferons alpha and beta. *Can J Vet Res* **55**, 143-149 (1991).

76 Wicht, O. *et al.* Proteolytic activation of the porcine epidemic diarrhea coronavirus spike fusion protein by trypsin in cell culture. *J Virol* **88**, 7952-7961, doi:10.1128/JVI.00297-14 (2014).

77 Shirato, K., Matsuyama, S., Ujike, M. & Taguchi, F. Role of proteases in the release of porcine epidemic diarrhea virus from infected cells. *Journal of virology* **85**, 7872-7880, doi:10.1128/JVI.00464-11 (2011).

78 Perlman, S. & Netland, J. Coronaviruses post-SARS: update on replication and pathogenesis. *Nature Reviews Microbiology* **7**, 439-450, doi:10.1038/nrmicro2147 (2009).

79 Li, W. H. *et al.* Animal origins of the severe acute respiratory syndrome coronavirus: Insight from ACE2-S-protein interactions. *Journal of Virology* **80**, 4211-4219, doi:10.1128/jvi.80.9.4211-4219.2006 (2006).

80 Li, F. *et al.* Conformational states of the severe acute respiratory syndrome coronavirus spike protein ectodomain. *Journal of Virology* **80**, 6794-6800, doi:10.1128/jvi.02744-05 (2006).

81 Belouzard, S., Millet, J. K., Licitra, B. N. & Whittaker, G. R. Mechanisms of coronavirus cell entry mediated by the viral spike protein. *Viruses* **4**, 1011-1033, doi:10.3390/v4061011 (2012).

82 Heald-Sargent, T. & Gallagher, T. Ready, set, fuse! The coronavirus spike protein and acquisition of fusion competence. *Viruses* **4**, 557-580, doi:10.3390/v4040557 (2012).

83 Millet, J. K. & Whittaker, G. R. Host cell entry of Middle East respiratory syndrome coronavirus after two-step, furin-mediated activation of the spike protein. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 15214-15219, doi:10.1073/pnas.1407087111 (2014).

84 Hofmann, M. & Wyler, R. Propagation of the virus of porcine epidemic diarrhea in cell culture. *J Clin Microbiol* **26**, 2235-2239 (1988).

85 Oka, T. *et al.* Cell culture isolation and sequence analysis of genetically diverse US porcine epidemic diarrhea virus strains including a novel strain with a large deletion in the spike gene. *Vet Microbiol* **173**, 258-269, doi:10.1016/j.vetmic.2014.08.012 (2014).

86 Liu, C. *et al.* Receptor usage and cell entry of porcine epidemic diarrhea coronavirus. *Journal of virology* **89**, 6121-6125, doi:10.1128/JVI.00430-15 (2015).

87 Nomura, T., Fujishima, A. & Fujisawa, Y. Characterization and crystallization of recombinant human cathepsin L. *Biochem Biophys Res Commun* **228**, 792-796, doi:10.1006/bbrc.1996.1733 (1996).

88 Yang, Y. *et al.* Two Mutations Were Critical for Bat-to-Human Transmission of Middle East Respiratory Syndrome Coronavirus. *Journal of virology* **89**, 9119-9123, doi:10.1128/JVI.01279-15 (2015).

89 Bertram, S. *et al.* TMPRSS2 and TMPRSS4 facilitate trypsin-independent spread of influenza virus in Caco-2 cells. *Journal of virology* **84**, 10016-10025, doi:10.1128/JVI.00239-10 (2010).

90 Hofmann, H. *et al.* S protein of severe acute respiratory syndrome-associated coronavirus mediates entry into hepatoma cell lines and is targeted by neutralizing antibodies in infected patients. *Journal of virology* **78**, 6134-6142, doi:10.1128/JVI.78.12.6134-6142.2004 (2004).

91 Qian, Z., Dominguez, S. R. & Holmes, K. V. Role of the spike glycoprotein of human Middle East respiratory syndrome coronavirus (MERS-CoV) in virus entry and syncytia formation. *PLoS One* **8**, e76469, doi:10.1371/journal.pone.0076469 (2013).

92 Simmons, G. *et al.* Inhibitors of cathepsin L prevent severe acute respiratory syndrome coronavirus entry. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 11876-11881, doi:10.1073/pnas.0505577102 (2005).

93 Simmons, G. *et al.* Characterization of severe acute respiratory syndrome-associated coronavirus (SARS-CoV) spike glycoprotein-mediated viral entry. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 4240-4245, doi:10.1073/pnas.0306446101 (2004).

94 Simmons, G., Zmora, P., Gierer, S., Heurich, A. & Pohlmann, S. Proteolytic activation of the SARS-coronavirus spike protein: cutting enzymes at the cutting

edge of antiviral research. *Antiviral Res* **100**, 605-614, doi:10.1016/j.antiviral.2013.09.028 (2013).

95 Reguera, J. *et al.* Structural bases of coronavirus attachment to host aminopeptidase N and its inhibition by neutralizing antibodies. *PLoS Pathog* **8**, e1002859, doi:10.1371/journal.ppat.1002859 (2012).