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

Virology 507 (2017) 101–109

Contents lists available at ScienceDirect

Virology

journal homepage: www.elsevier.com/locate/yviro

Structure of the S1 subunit C-terminal domain from bat-derived coronavirus HKU5 spike protein

Xue Han[a,b], Jianxun Qi[a], Hao Song[a,c], Qihui Wang[d,e], Yanfang Zhang[a,b], Ying Wu[e,f], Guangwen Lu[g], Kwok-Yung Yuen[h,i,j], Yi Shi[a,b,e,k,*], George F. Gao[a,b,c,e,j,k,l,*]

[a] CAS Key Laboratory of Pathogenic Microbiology and Immunology, Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, China
[b] University of Chinese Academy of Sciences, Beijing 100049, China
[c] Research Network of Immunity and Health (RNIH), Beijing Institutes of Life Science, Chinese Academy of Sciences, Beijing 100101, China
[d] CAS Key Laboratory of Microbial Physiological and Metabolic Engineering, Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, China
[e] Shenzhen Key Laboratory of Pathogen and Immunity, Shenzhen Third People's Hospital, Shenzhen 518112, China
[f] School of Basic Medical Sciences, Wuhan University, Wuhan 430071, China
[g] West China Hospital Emergency Department (WCHED), State Key Laboratory of Biotherapy, West China Hospital, Sichuan University, and Collaborative Innovation Center of Biotherapy, Chengdu, Sichuan 610041, China
[h] State Key Laboratory for Emerging Infectious Diseases, The University of Hong Kong, Pokfulam 999077, Hong Kong Special Administration Region
[i] Department of Microbiology, The University of Hong Kong, Pokfulam 999077, Hong Kong Special Administration Region
[j] Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, Zhejiang University, Hangzhou 310003, China
[k] Center for Influenza Research and Early-warning (CASCIRE), Chinese Academy of Sciences, Beijing 100101, China
[l] National Institute for Viral Disease Control and Prevention, Chinese Center for Disease Control and Prevention (China CDC), Beijing 102206, China

ARTICLE INFO

Keywords:
MERS-CoV
BatCoV HKU5
CTD
Crystal structure
Evolution

ABSTRACT

Accumulating evidence indicates that MERS-CoV originated from bat coronaviruses (BatCoVs). Previously, we demonstrated that both MERS-CoV and BatCoV HKU4 use CD26 as a receptor, but how the BatCoVs evolved to bind CD26 is an intriguing question. Here, we solved the crystal structure of the S1 subunit C-terminal domain of HKU5 (HKU5-CTD), another BatCoV that is phylogenetically related to MERS-CoV but cannot bind to CD26. We observed that the conserved core subdomain and those of other betacoronaviruses (betaCoVs) have a similar topology of the external subdomain, indicating the same ancestor of lineage C betaCoVs. However, two deletions in two respective loops located in HKU5-CTD result in conformational variations in CD26-binding interface and are responsible for the non-binding of HKU5-CTD to CD26. Combined with sequence variation in the HKU5-CTD receptor binding interface, we propose the necessity for surveilling the mutation in BatCoV HKU5 spike protein in case of bat-to-human interspecies transmission.

1. Introduction

Coronaviruses (CoVs) are spherical enveloped viruses with single positive-strand RNA genomes of ~30 kb in length, which is the largest among RNA viruses (Saif, 1993). CoVs are divided into four genera: alpha-, beta-, gamma-, and deltaCoVs (de Groot et al., 2013). BetaCoVs are further subdivided into four lineages/subgroups: A, B, C, and D (Chan et al., 2015). To date, both alpha- and betaCoVs are found to infect humans (Lu et al., 2015), causing subclinical or very mild symptoms and accounting for 10–15% of common colds (Heikkinen and Jarvinen, 2003). In addition, CoVs can also be life-threatening and have pandemic potential. The epidemic of severe acute respiratory syndrome coronavirus (SARS-CoV), which belongs to lineage B of the betaCoVs, originated in southern China in 2002 and spread to 28 countries, infecting over 8000 and leading to almost 800 related deaths (WHO, 2004). The outbreak of MERS-CoV, a member of the lineage C betaCoVs (Cotten et al., 2013; Zaki et al., 2012), has caused 1832 laboratory-confirmed cases since 2012, including 651 related deaths as of Nov. 28, 2016 (WHO, 2016). Unlike the SARS-CoV, which suddenly disappeared after a massive global disease control effort, especially in China, the number of MERS-CoV infections is still on the rise.

Mounting evidence indicates that CoVs circulating in bats (BatCoVs) are the gene sources of alphaCoVs and betaCoVs (W. Li et al., 2005; Woo et al., 2012), including SARS-CoV (Ge et al., 2013; Lau et al., 2005; W. Li et al., 2005). The data also underscore that bats are the likely natural reservoir for MERS-CoV (Annan et al., 2013; Ithete et al., 2013; Memish et al., 2013; Wang et al., 2014; Yang et al., 2014). For instance, viral gene fragments identical or quite similar to

* Corresponding authors at: CAS Key Laboratory of Pathogenic Microbiology and Immunology, Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, China.
E-mail addresses: shiyi@im.ac.cn (Y. Shi), gaof@im.ac.cn (G.F. Gao).

http://dx.doi.org/10.1016/j.virol.2017.04.016
Received 12 March 2017; Received in revised form 10 April 2017; Accepted 15 April 2017
Available online 19 April 2017
0042-6822/

those of MERS-CoV have been reported in bats (Annan et al., 2013; Ithete et al., 2013; Memish et al., 2013). Moreover, parallel studies from our group and others show that BatCoV HKU4, grouped in lineage C with MERS-CoV, can also use human CD26 (hCD26; the receptor of MERS-CoV) for viral entry (Wang et al., 2014; Yang et al., 2014). In other words, two members in lineage C use the same human receptor. One has caused a human infection epidemic (MERS-CoV), and the other can utilize the same receptor (BatCoV HKU4) and has potential to infect humans. This highlights the necessity of surveillance for lineage C betaCoVs, including BatCoV HKU5, which was first sequenced in 2006 in Japanese pipistrelles (*Pipistrellus abramus*) (Woo et al., 2006) and is circulating in bats (Lau et al., 2013). Whether the virus has the potential to bypass the bat-human barrier needs to be evaluated.

CoV infections initiate with the virus binding to the host receptor. The envelope-interspersed trimeric spike (S) protein plays a pivotal role in this process. S is further divided into two parts: S1, responsible for receptor binding, and S2, which initiates fusion (Belouzard et al., 2012; Dai et al., 2016; Kielian and Rey, 2006). S1 contains two relatively independent structures named the N-terminal domain (NTD) and C-terminal domain (CTD) based on their position. Most betaCoVs use the CTD as the receptor-binding domain (RBD/CTD) except mouse hepatitis virus (MHV), which uses the NTD to bind the cellular receptor carcinoembryonic-antigen-related cell-adhesion molecule 1 (CEACAM1) (Dai et al., 2016; Du et al., 2009; F. Li et al., 2005; Lu et al., 2013; Peng et al., 2011). Two of the RBD/CTDs in lineage C betaCoVs (MERS-RBD/CTD and HKU4-RBD/CTD) bind to the same human receptor CD26 (hCD26) to initiate infection, and the two domains share high sequences identities (55%) in addition to high structural similarities, with a root mean square deviation (rmsd) of 1.114 (193 Cα) (Lu et al., 2013; Wang et al., 2014). Despite the similar sequence identities between HKU5-CTD and MERS-RBD/CTD (54%) or HKU4-RBD/CTD (57%), no detectable binding was found between HKU5-CTD and hCD26. The structural basis for this variation remains to be elucidated.

In this study, we determined the structure of HKU5-CTD. Similar to other solved structures, HKU5-CTD contains two subdomains: the core subdomain homologous to other CTDs in betaCoVs and the external subdomain, which resembles MERS-RBD/CTD and HKU4-RBD/CTD, indicating conservation of the external domain in lineage C. However, two deletions in HKU5-CTD lead to structural shifts in the hCD26-interaction interface and thereby result in its inability to bind this receptor. Our results suggest that the characteristic insertions in βc4 and βc5 among different lineages in betaCoVs result in different receptor engagement, thereby contributing for potential interspecies transmission.

2. Results

2.1. Overall structure of the HKU5-CTD

We first characterized the S protein of BatCoV HKU5 through bioinformatics analysis. BatCoV HKU5 S is composed of 1352 amino acids and exhibits typical features of CoVs S protein, including the predicted hydrophobic residue-rich HR1 and HR2 motifs and a similar concentration of hydrophobic amino acids to SARS-CoV fusion peptide (FP), internal fusion peptide (IFP), and pre transmembrane domain (PTM) (Gao et al., 2013; Mahajan and Bhattacharjya, 2015; Xu et al., 2004; Zhu et al., 2004). Like MERS-CoV S protein, a furin-like protease recognition motif is predicted at position R745/A746 (S1/S2), which separates the S1 and S2 subunits (Millet and Whittaker, 2014). In addition, a second furin cleavage site can be found at R884/S885, which resembles S2′ in the MERS-CoV S protein (Millet and Whittaker, 2014), indicating that the priming process of BatCoV HKU5 S in human cells probably occur in the same way like MERS-CoV (Fig. 1A). Because most betaCoVs use their CTD to bind their respective receptors, we next focused on the evolutionary relationships of the CTDs. Consistent with the phylogenetic relationships, HKU5-CTD, HKU4-RBD/CTD, and MERS-RBD/CTD are grouped in one branch representing lineage C, while HKU1-CTD and MHV-CTD cluster together in lineage A. HKU9-CTD, a member in lineage D, is phylogenetically more related to SARS-RBD/CTD, which belongs to lineage B (Fig. 1B).

The HKU5-CTD was then purified, crystallized and the structure was successfully determined at a resolution of 2.1 Å, with clear electron densities tracing from Q389 to Q586. The structure, solved through the molecular replacement method, contains a single molecule in the crystallographic asymmetric unit, with an R_{work} of 0.2160 and an R_{free} of 0.2585 (Table 1). Like the other solved CTD structures of betaCoVs (Huang et al., 2016; Kirchdoerfer et al., 2016; F. Li et al., 2005; Lu et al., 2013; Walls et al., 2016; Wang et al., 2013, 2014), HKU5-CTD folds into two discrete subdomains, the core and the external (Fig. 2). The core subdomain contains a five-stranded antiparallel scaffold center (core-center), which is decorated by five helices (α or 3_{10}) and two small strands (βp1 and βp2) on the exterior. Three pairs of disulfide bonds help to stabilize the scaffold, namely C391-C415 and C445-C583, located in the peripheral region of the core subdomain (core-peripheral), and C433-C486 in the core-center, linking βc2 and βc4. Notably, two antiparallel β strands, one of which is located in the C-terminus and the other forming the disulfide bond with the N-terminus, help to make keep two termini in proximity. In addition, traceable electron densities can be observed for two glycosylated modifications at N418 and N495, which form two protrusions at the core-peripheral region (Fig. 2).

The external subdomain of HKU5-CTD extends out of βc4 in the core-center, sequentially folds into two antiparallel β strands (β1′ and β2′), an α helix (H1′), and another two antiparallel β strands (β3′ and β4′), and finally proceed into βc5. Between β1′ and H1′, a pair of disulfide bonds (C511-C532) is formed to stabilize the external structure (Fig. 2).

2.2. Conserved core subdomain and variable external subdomain for betaCoVs S protein

To date, seven structures of CTDs covering all four lineages of betaCoVs have been solved. They are HKU1-CTD and MHV-CTD, belonging to lineage A (Kirchdoerfer et al., 2016; Walls et al., 2016), MERS-RBD/CTD (Lu et al., 2013; Wang et al., 2013), HKU4-RBD/CTD (Wang et al., 2014), and HKU5-CTD grouped in lineage C (reported here), and SARS-RBD/CTD (F. Li et al., 2005) and HKU9-CTD (Huang et al., 2016) representing lineages B and D, respectively. All seven betaCoV CTD structures display a conserved core subdomain, with five antiparallel beta strands and a conserved disulfide bond between βc2 and βc4 (Fig. 3).

Despite the different combinations of α helices and β strands, the orientations of the secondary structures are conserved among CTDs in the core-peripheral region. In addition, two highly conserved disulfide bonds exist. One is formed between the N-terminus and the loop/β strand extended from βc1, and the other links the C-terminus and the loop/β strand proceeding to βc3. Thus, through the two disulfide bonds, both termini are brought into close proximity (Fig. 3). Although in the SARS-RBD/CTD electron density at the C-terminus is not clear enough to determine the structures (Fig. 3C), two conserved cysteines are present, indicating the probability of disulfide bond formation (Fig. 1C).

Opposite to the conserved core subdomain, the external subdomain varies considerably among different lineages. In lineage A, the external subdomain of MHV-CTD, which was obtained by density-guided homology modelling due to its large flexibility and poor quality of the density in this region, consists of four β strands and three small helices (PDB code: 3CJL) (Fig. 3A). HKU1-CTD is comprised of a large, variable loop with three inlaid β strands (Fig. 3B). The absence of clear

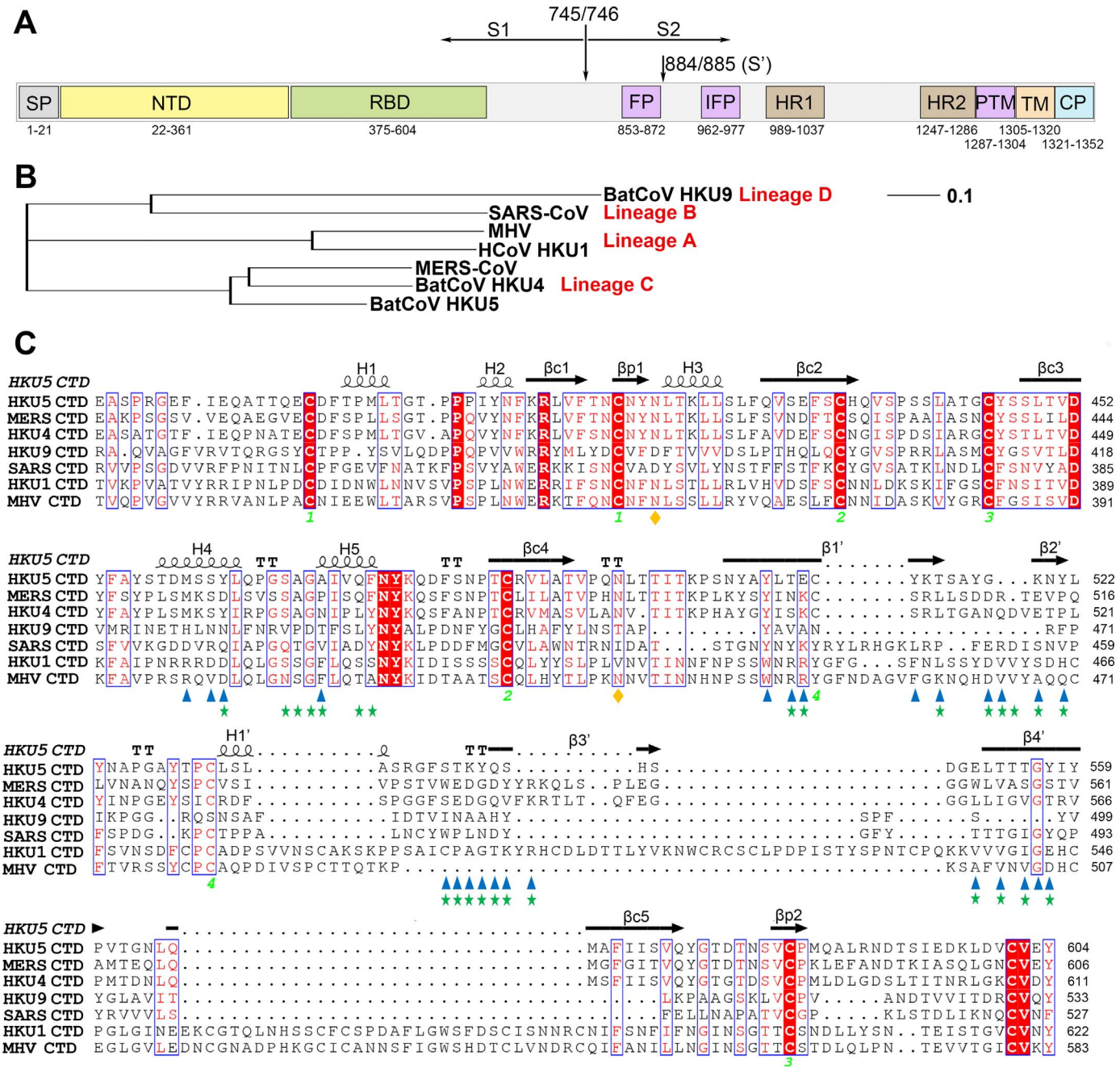

Fig. 1. Sequence features of HKU5-CTD. (A) Schematic representation of BatCoV HKU5 S. The indicated domain elements were defined through pairwise sequence alignments or bioinformatics analyses. The signal peptides (SP), transmembrane domain (TM), and heptad repeats 1 and 2 (HR1 and HR2, respectively) were predicted with the SignalP 4.0 server, TMHMM server, and Learncoil-VMF program, respectively, while the NTD, RBD, and fusion peptides (FP, IFP, and PTM) were deduced by alignment with the N-terminal galectin-like domain of murine hepatitis virus S, MERS-RBD/CTD, and SARS-CoV S, respectively. The S1/S2 and S2′ sites potentially cleaved by furin-like proteases were predicted using the ProP 1.0 server. (B) Phylogenetic tree generated using MEGA with the indicated RBD/CTD sequences. (C) Structure-based sequence alignment. The secondary structure elements are defined based on an ESPript algorithm and are labeled as in Fig. 2. The arrows and spiral line indicate strands and helices, respectively. The conserved cysteine residues that form four disulfide bonds in the structures are marked with Arabic numerals 1–4. The beta strands in the core-center, the elements in the core-peripheral, and the structures in the external domains are marked with c, p, and the character with prime, respectively. The blue triangle and green star represent the key amino acids for binding hCD26 in MERS-RBD/CTD and HKU4-RBD/CTD, respectively, while the yellow rhombus indicates two glycosylation sites in HKU5-CTD.

secondary structure from residues C476-F572 indicates the flexibility of this region (PDB code: 5I08). In lineage C, three CTDs show similar external folds, with rmsd ranging from 0.962 (HKU5-CTD *vs.* MERS-RBD/CTD (PDB code: 4KQZ)) to 1.178 (HKU5-CTD *vs.* HKU4-RBD/CTD (PDB code: 4QZV)). All external subdomains are strand-dominated structures with four anti-parallel β strands and expose a flat strand-face that is stabilized by a conserved disulfide bond (Fig. 3D-F). In lineage B, the SARS-RBD/CTD is dominated by a disulfide bond-stabilized flexible loop that connects two small β strands (PDB code: 2GHV) (Fig. 3C). In BatCoV HKU9, representing lineage D, the external subdomain only contains one large helix in this region (PDB code: 5GYQ) (Fig. 3G).

Although their external subdomain structures differ, all CTDs in betaCoVs extend out from βc4 and proceed back to the core subdomain through βc5 (Fig. 3), indicating that during evolution, different insertions in this region resulted in the different structures of the CTDs. This, then, led to different receptor usage if the CTD is utilized as the RBD.

Table 1
Data collection and refinement statistics.

	HKU5 CTD
Data collection	
Wavelength (Å)	0.97915
Space group	P 21
Cell dimensions	
a, *b*, *c* (Å)	49.612, 212.659, 87.943
α, β, γ (deg)	90.000, 94.756, 90.000
Resolution (Å)	50.00–2.10 (2.18–2.10)[a]
R_{merge}	0.104 (1.038)
$R_{p.i.m.}$	0.052 (0.527)
I / σI	15.355 (1.645)
$CC_{1/2}$	0.998 (0.801)
Completeness (%)	99.9 (99.9)
Redundancy	5.0 (4.8)
Refinement	
Resolution (Å)	37.80–2.10
No. reflections	104555
R_{work} / R_{free}	0.2160/0.2585
No. atoms	
Protein	11032
Ligand/ion	0
Water	527
B-factors	
Protein	52.9
Ligand/ion	–
Water	45.2
R.m.s. deviations	
Bond lengths (Å)	0.006
Bond angles (deg)	1.062
Ramachandran plot	
Favored (%)	95.73
Allowed (%)	3.44
Outliers (%)	0.83

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

2.3. Structural basis for HKU5-CTD not binding to CD26

Both MERS-RBD/CTD and HKU4-RBD/CTD bind to hCD26 to initiate infection. In addition, the structure of the HKU5-CTD displays a similar topology to the two RBD/CTDs in lineage C. Thus, we assayed for binding between HKU5-CTD and hCD26. However, consistent with previous results, no binding was detected, either by fluorescence-activated cell sorting (FACS) or surface plasmon resonance (SPR) (Fig. 4).

According to the two solved complex structures, four concentrations of residues in MERS-RBD/CTD and HKU4-RBD/CTD are involved in binding to hCD26. These residues located in four beta strands and two loops (β1′/β2′ loop and β3′/β4′ loop) in the external subdomain as well as H4 and H5 (for MERS-RBD/CTD) or H5 and H6 (for HKU4-RBD/CTD) positioned in the core subdomain and the loop connecting the two helices (Figs. 1C, 3E and 3F). However, half of these regions (β1′/β2′ loop and β3′) have deletions in HKU5-CTD (Fig. 1C). Due to these deletions, the orientations of two loops (marked 1 and 2, respectively, in Fig. 3D-F) in HKU5-CTD vary compared to MERS-RBD/CTD and HKU4-RBD/CTD, which leads to conformational shifts in HKU5-CTD at the hCD26-binding interface (Fig. 5A and E). The β1′/β2′ loop in both MERS-RBD/CTD and HKU4-RBD/CTD inserts into the groove formed by two helices on the side and β strands on the bottom (Fig. 5B and F). Sixty-five (328 in total) and 49 (214 in total) van der Waals contacts, including 5 (16 in total) and 4 (13 in total) hydrogen bonds, are formed in MERS-RBD/CTD/hCD26 and HKU4-RBD/CTD/hCD26, respectively. In contrast, this loop in HKU5-CTD is tilted away by ~6 and 9 Å (Fig. 5B and F) compared to MERS-RBD/CTD and HKU4-RBD/CTD, respectively, which results in the loss of binding to hCD26 at this region.

Moreover, a six-residue deletion in β3′ causes large discrepancies in the assemblies of β3′, β4′ and their connecting loop, compared with MERS-RBD/CTD and HKU4-RBD/CTD (Fig. 3D-F). In β3′ of both hCD26-binding RBD/CTDs, the side chains of Y540 and R542 face the receptor, conferring a strong hydrophilic interaction between the ligand and the receptor. In contrast, in HKU5-CTD, the orientation of β3′ is opposite. In addition, Y544 in HKU5-CTD likely sterically clashes Q286, which further pushes HKU5 away from hCD26 (Fig. 5C and G). In the other beta strand of β4′, both MERS-RBD/CTD and HKU4-RBD/CTD form a large hydrophobic interaction patch with hCD26. In HKU5-CTD, aside from the shift of the β3′/β4′ loop away from the receptor, the deletion of hydrophobic residues (*e.g.*, W535) compared to MERS-RBD/CTD (Fig. 5D) and the substitution of hydrophilic residues (*e.g.*, T553 and T555) instead of hydrophobic ones (I560 and V562 in HKU4-RBD/CTD) (Fig. 5H) likely inhibit HKU5-CTD binding to hCD26. In total, the conformational variations between HKU5-CTD and hCD26-binding RBD/CTDs explain the lack of hCD26 binding by HKU5-CTD. However, various deletions in HKU5-CTD loop 1 are present in nature (Fig. 6), and might contribute to evolve for receptor binding.

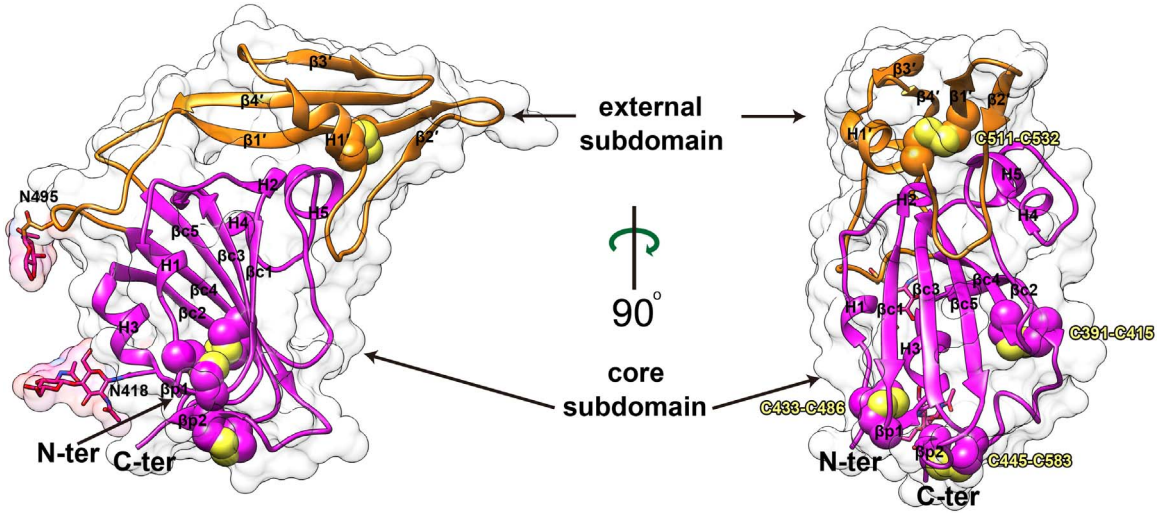

Fig. 2. Crystal structure of the HKU5-CTD. The core and external subdomains are colored orange and magenta, respectively. The core subdomain is further divided into a center region (core-center) and a peripheral region (core-peripheral). The core-center strands and helices are labeled βc1-βc5 and H1-H5, respectively, while the core-peripheral strands are marked βp1 and βp2. The glycan-moieties are displayed in sticks and marked in the left panel. The disulfide bonds are presented in spheres and labeled in the right panel. Both N- and C-terminus are indicated with the arrows. To depict the structure clearly, cartoon structures inside the transparent surface are presented at two angles.

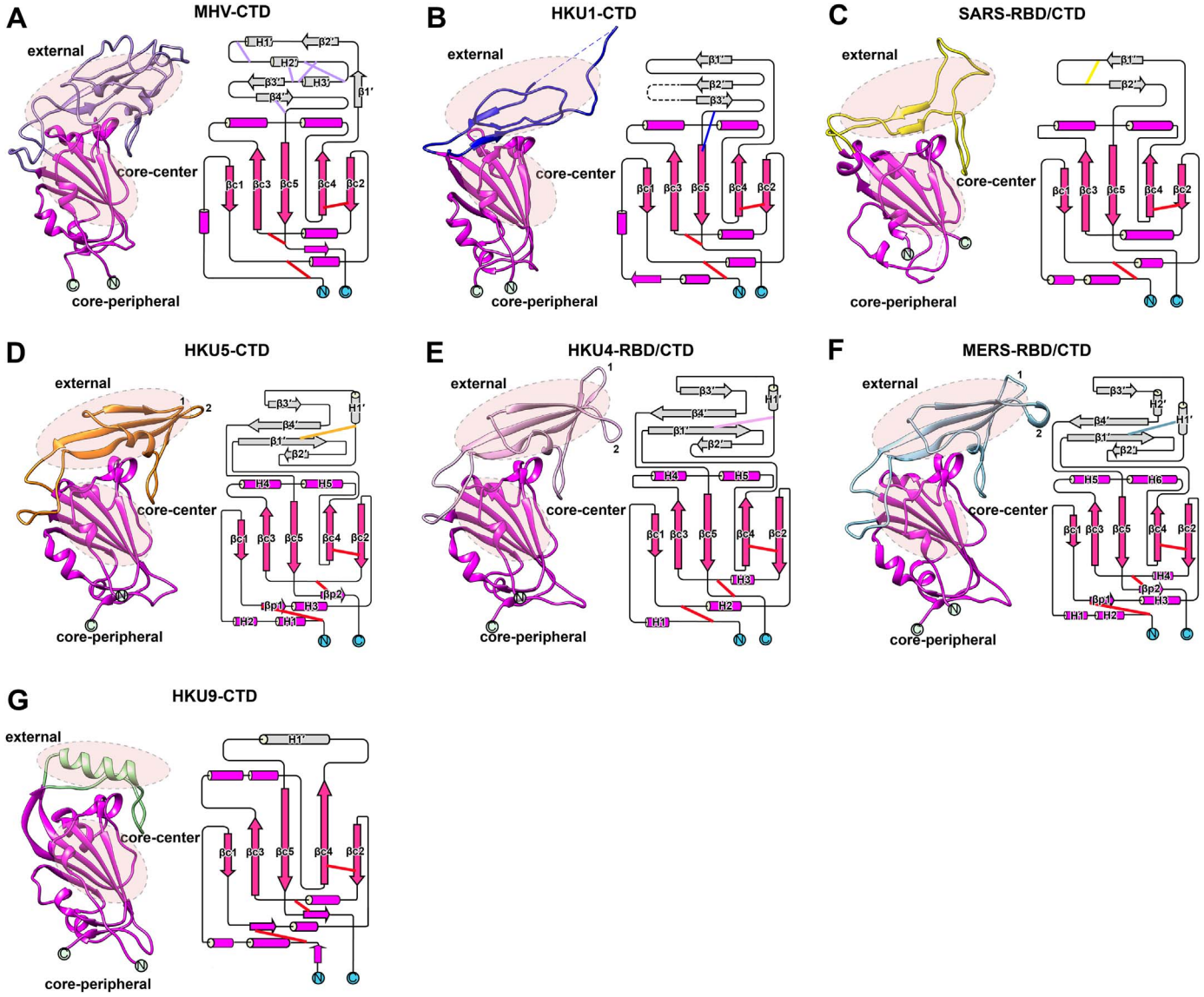

Fig. 3. Topological diagrams of CTDs in betaCoVs. Structural and topological comparison of available betaCoV CTD structures. Seven structures, including those of MHV-CTD (PDB code: 3CJL), HKU1-CTD (PDB code: 5I08), SARS-RBD/CTD (PDB code: 2GHV), HKU5-CTD, HKU4-RBD/CTD (PDB code: 4QZV), MERS-RBD/CTD (PDB code: 4KQZ), and HKU9-CTD (PDB code: 5GYQ) were oriented similarly and are presented as cartoons in parallel. The conserved disulfide bonds are labeled in red lines, while the non-conserved ones are displayed with lines in accordance with the color of indicated external subdomain. Arrows and cylinders represent the strands and helices, respectively.

3. Discussion

In this study, we solved the crystal structure of HKU5-CTD, which represents the seventh structure of a CTD belonging to a betaCoV. Like the other six CTDs (Huang et al., 2016; Kirchdoerfer et al., 2016; F. Li et al., 2005; Lu et al., 2013; Peng et al., 2011; Walls et al., 2016; Wang et al., 2013, 2014), there are two subdomains in HKU5-CTD, the core and the external. Despite the low residue conservation among CTDs (pair-to-pair amino acid identity ranging from 17.2% to 58.7%) and the core subdomains (pair-to-pair amino acid identity ranging from 16.6% to 66.7%) in the four lineages, the topology of the latter ones are highly conserved, with five anti-parallel β strands constituting the core-center and the same orientation of secondary elements in the core-peripheral (Fig. 3). This includes the same region of MHV, which uses the NTD of S1 to bind the receptor.

However, the external subdomains vary considerably among lineages. In lineage A, the MHV-CTD contains several β strands and inlaid helices, while the HKU1-CTD is comprised of loops and three small β strands. However, approximately 100 amino acids (C476-F572) are unclear at this region, likely due to their flexibility. SARS-RBD/CTD, in lineage B is dominated by loops, which are stabilized by a disulfide bond and two anti-parallel β strands. Most CTD structures solved to date are in lineage C, and all three CTDs (MERS-RBD/CTD, HKU4-RBD/CTD, and HKU5-CTD) display conserved structures with β strand-forming platforms decorated with helices. In addition, a disulfide bond is conserved among CTDs in lineage C in the external subdomain. HKU9-CTD, a member of lineage D, is comprised of a helix that is clamped with loops. Although different structures and topologies exist among lineages, all external subdomains extend out from βc4 and proceed back to βc5 (Fig. 3), indicating that different insertions between βc4/βc5 during betaCoV evolution have conferred the betaCoVs with different properties, such as receptor usage, and thereby led to the parallel evolution of lineages.

The ligand–receptor interaction is a key factor determining the tissue tropism and host range of CoVs. For SARS-CoV, MERS-CoV, and BatCoV HKU4, the receptors are clear, and the complex structures demonstrate that the receptor mainly binds to the varied external subdomains of CTDs. Neutralizing antibodies against HCoV HKU1

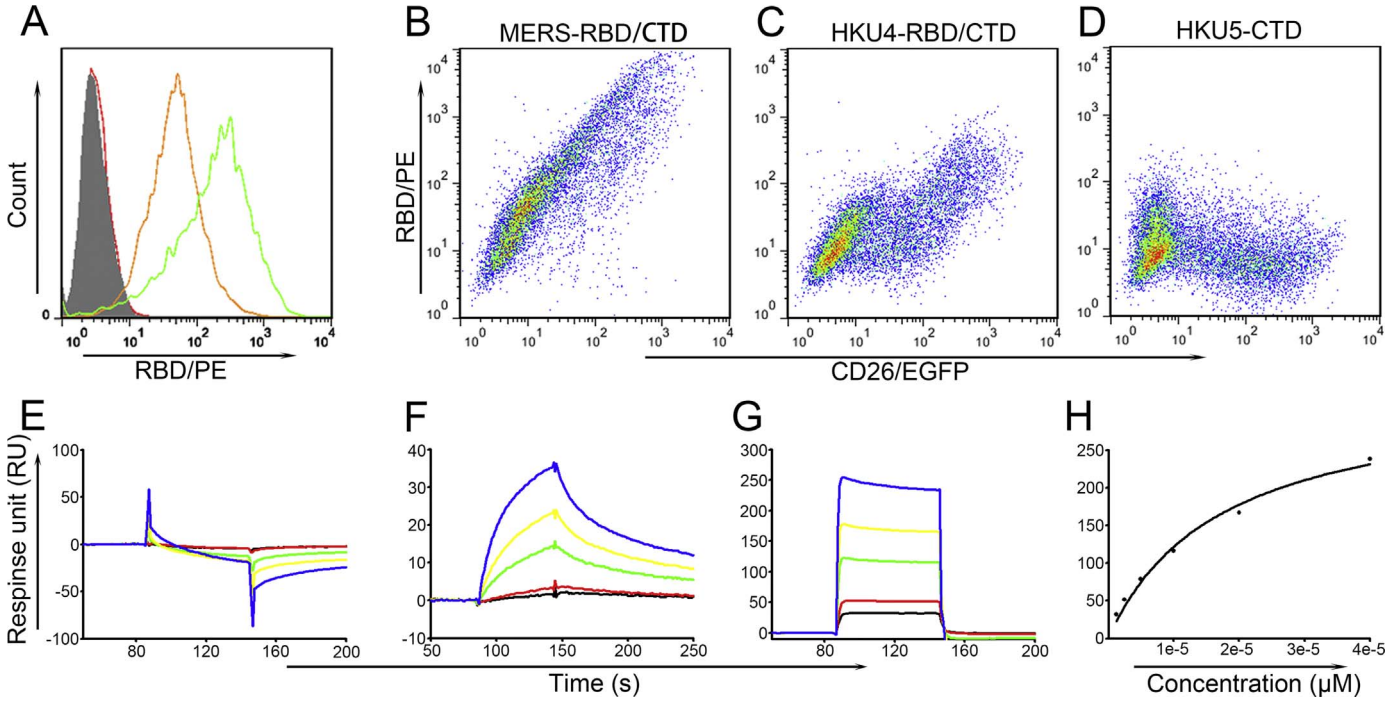

Fig. 4. Characterization of HKU5-CTD by FACS and SPR. (A) Huh7 cells were stained with MERS-RBD/CTD (green), HKU4-RBD/CTD (orange), and HKU5-CTD (red), respectively. (B-D) BHK21 cells transfected with hCD26 (BHK-hCD26) were stained with MERS-RBD/CTD (B), HKU4-RBD/CTD (C), and HKU5-CTD (D), respectively. (E-F) The indicated protein was immobilized on a CM5 chip, and a gradient concentration of hCD26 was flowed through the chip. The RUs were recorded. (E) hCD26 binding to HKU5-CTD. (F) hCD26 binding to MERS-RBD/CTD. (G) hCD26 binding to HKU4-RBD/CTD. (H) The saturation profile for HKU4-RBD/CTD binding to hCD26.

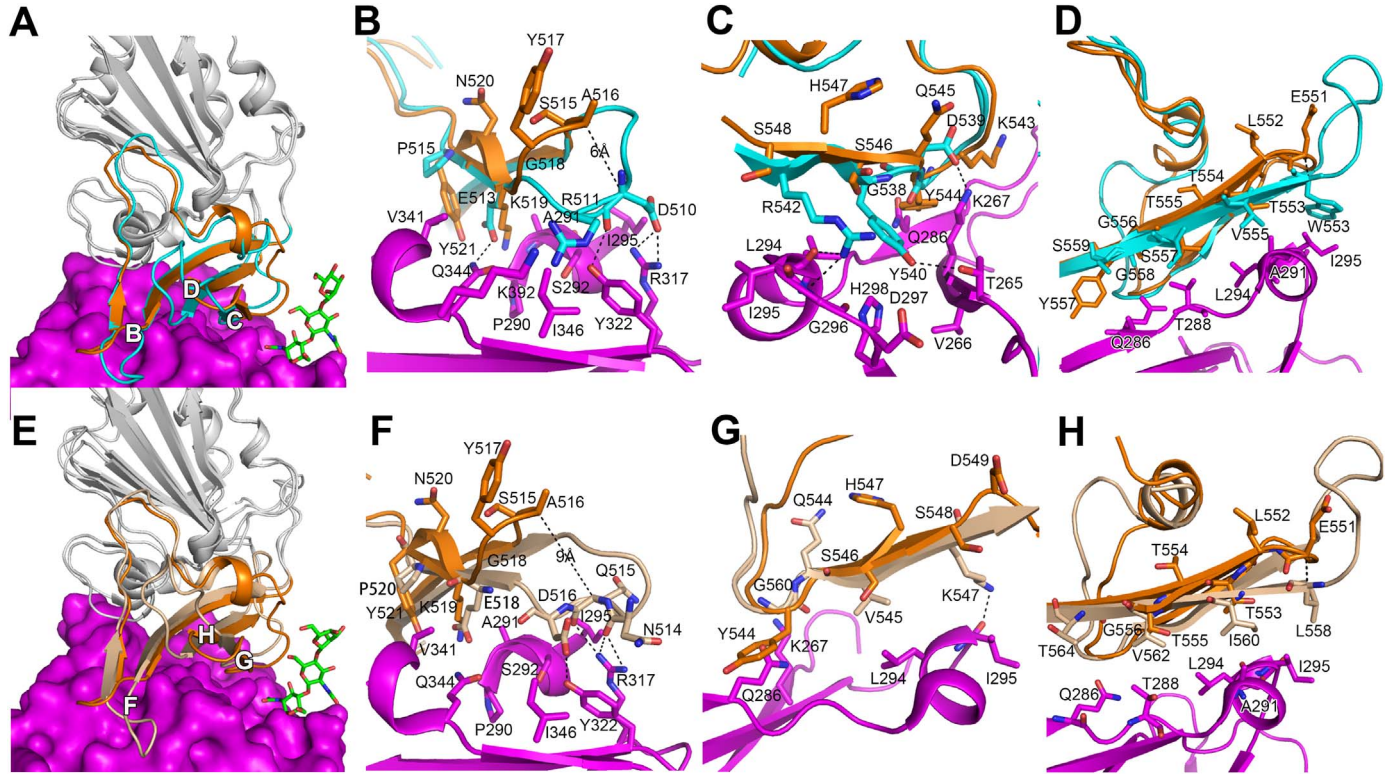

Fig. 5. Structural basis for the lack of binding between HKU5-CTD and hCD26. Superimposition of the structures of HKU5-CTD and hCD26 binding-MERS-RBD/CTD (A-D) or HKU4-RBD/CTD (E-H). The variations in the receptor binding interface of HKU5-CTD compared with MERS-RBD/CTD or HKU4-RBD/CTD are allocated with B-D and F-H and further delineated in B-D and F-H for detailed structural shifts, respectively. The conserved core subdomains in HKU5-CTD, MERS-RBD/CTD, and HKU4-RBD/CTD are colored in grey, while the external subdomains of the three proteins are marked with orange, cyan, and wheat, respectively. The magenta represents hCD26.

bind to the HKU1-CTD, not the HKU1-NTD (Qian et al., 2015), indicating that the CTD in HCoV HKU1 is most likely to be the RBD, though the protein receptor has yet to be identified (Chan et al., 2016; Huang et al., 2015). HKU9-CTD does not bind to ACE2 or hCD26 (Huang et al., 2016). In our study, we found that although HKU5-CTD displays a similar structure and topology to MERS-RBD/CTD and

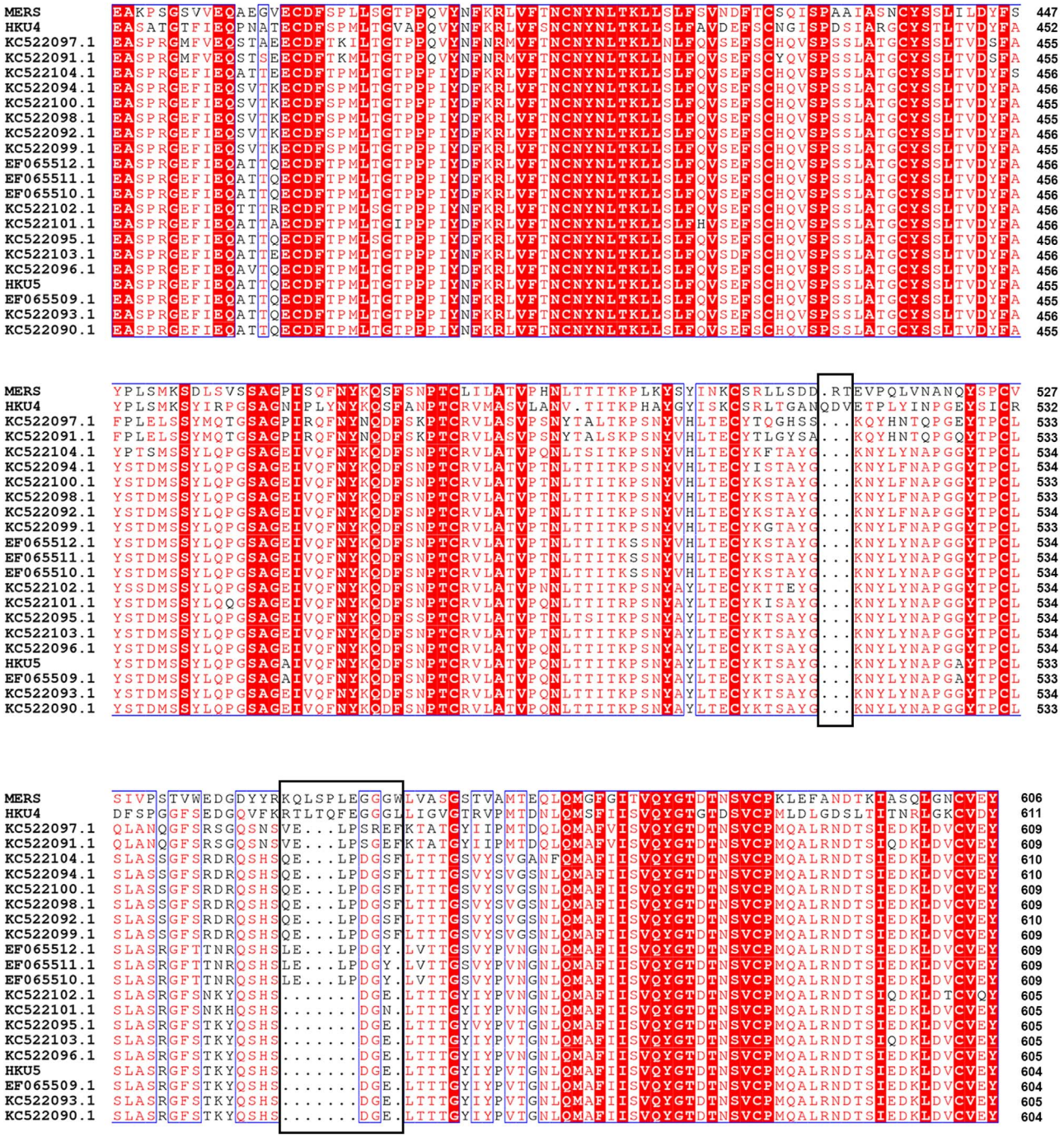

Fig. 6. CTDs of HKU5 show diversities. All referred sequences in HKU5 CTD regions were analyzed by alignment. The two black boxes indicate the sequences of loops 1 and 2 marked in Fig. 3C.

HKU4-RBD/CTD, detailed structural analysis revealed variations at the hCD26-binding interface, which results in the loss of binding to this receptor. Thus, subtle distinctions in external subdomains could determine different receptor usage.

In addition to receptor binding, the priming process, which involves host proteases to liberate S2 and the fusion peptides from the otherwise covalently-linked S1 subunit, is another key factor affecting cell tropism and the entry route of CoVs. A two-step activation mechanism has been proposed for MERS-CoV entry (Millet and Whittaker, 2014). During the secretion of S protein, the proteolysis at S1/S2 occurs in the endoplasmic reticulum (ER)-Golgi compartments where furin is localized, while during virus entry into target cells, S2′ is cleaved. The same proteolysis in S1/S2 and S2′ is also essential for SARS-CoV infection, except that due to the lack of a furin-recognition site at S1/S2, SARS-CoV S remains uncleaved after biosynthesis and a diverse array of proteases are involved in this process (Millet and Whittaker, 2015; Simmons et al., 2011). In contrast, although BatCoV HKU4 can utilize hCD26 as a receptor, the proteolysis is stuck due to the lack of a protease site. Treatment of pseudovirus particles containing BatCoV HKU4 S protein with trypsin or importing the furin-recognition site into S protein enables the particles to infect hCD26-expressing cells (Wang et al., 2014; Yang et al., 2015), indicating the BatCoV HKU4 is less adapted to human cells. However, in the BatCoV HKU5 S protein, both furin-recognition sites are present, as in MERS-CoV. Accordingly, BatCoV HKU5 S is predicted to be cleaved at S1/S2 after biosynthesis and at S2′ during virus entry. Furin is a ubiquitous proteinase and

expressed in nearly all cells lines. The presence of the two furin-recognition sites indicates that the priming process of BatCoV HKU5 is ready to occur.

BatCoV HKU5 has been circulating in bats (Lau et al., 2013). In an epidemiology study over a 7-year period (April 2005 to August 2012), 25% of alimentary specimens of Japanese pipistrelle bats (*Pipistrellus abramus*) collected from 13 locations in Hong Kong were positive for this virus (Lau et al., 2013), indicating that it might target gastro-intestinal tissues. However, BatCoV HKU5 virus has not been isolated and cultured successfully, which is an obstacle to virus transmission research. The problem is largely due to a lack of suitable cell lines for BatCoV HKU5 virus. The emerging but puzzling question is whether this virus could infect humans or not.

An infectious clone of BatCoV HKU5 containing the ectodomain from the SARS-CoV S protein was constructed through reverse genetics and synthetic-genome design, and the recombinant virus replicates efficiently in cell culture and in young and aged mice (Agnihothram et al., 2014). In addition, the key proteins for virus replication, such as the 3C-like protease, polymerase, and exonuclease of BatCoV HKU5 display high amino acid sequence similarity to those in MERS-CoV, indicating that once the genome of BatCoV HKU5 is released into host cells, genome replication, virus particle assembly, and release can readily occur. Therefore, the receptor would be the last barrier for BatCoV HKU5 to infect humans. Our data show that BatCoV HKU5-CTD does not use hCD26 as a receptor, though it folds into a very similar structure as MERS-RBD/CTD and HKU4-RBD/CTD. In other words, the cellular receptor of BatCoV HKU5 is still a mystery that requires further study.

Evolutionally, BatCoV HKU5 S protein is more diverse than BatCoV HKU4 S protein (Lau et al., 2013), and various deletions in loop 1 (Fig. 3D and Fig. 6) have been sequenced. This indicates that BatCoV HKU5 is able to generate variants to occupy new ecological niches and might acquire the ability to bind to hCD26 by accumulating mutations and ultimately cause human respiratory infections like MERS-CoV and SARS-CoV. Accordingly, it is very important to perform long-lasting surveillance of BatCoV HKU5 evolution, especially the variety of S protein in the event that the virus breaks the inter-species and/or inter-tissue transmission barriers.

4. Materials and methods

4.1. Gene construction, protein expression, and purification

The coding region for HKU5-CTD (Q389-Q586) with a 6×His-tag at its C-terminus was cloned into the *Eco*RI and *Xho*I sites of pFastBac. HKU5-CTD, MERS-RBD/CTD, HKU4-RBD/CTD, and hCD26 were expressed and purified as previously reported (Lu et al., 2013; Wang et al., 2014). Briefly, the proteins were expressed in baculovirus-infected Hi5 cells. After 48 h of culturing, the supernatants were collected and purified through a 5 mL HisTrap HP column (GE Healthcare). The bound protein was eluted by a gradient of imidazole. Fractions containing the target protein as determined by SDS-PAGE were pooled and further subjected to gel filtration using a Superdex® 75 column (GE Healthcare) in a buffer composed of 20 mM Tris-HCl (pH 8.0) and 50 mM NaCl.

The Fc fusion protein used for cell staining was purified following a previously published method (Lu et al., 2013; Wang et al., 2014). In brief, the plasmid containing the target gene was transfected into HEK293T cells. After 3 and 7 d of culturing, the supernatants containing secreted protein were pooled and subjected to HiTrap ProteinA chromatography (5 mL, GE Healthcare). Protein was eluted with 0.1 M sodium citrate (pH 4.5) and further purified by gel filtration. The protein was finally buffer-exchanged into PBS, concentrated to ~1 mg/mL, and stored at −80 °C before further usage.

4.2. Crystallization, data collection, and structure determination

The HKU5-CTD protein expressed in Hi5 cells was crystallized by sitting drop vapor diffusion at 18 °C. Diffractable crystals were obtained in a condition consisting of 0.2 M potassium thiocyanate and 20% (w/v) polyethylene glycol 3350 with a protein concentration of 15 mg/mL. Crystals were cryoprotected in reservoir solution containing 20% (v/v) glycerol and flash-frozen in liquid nitrogen. Diffraction data were collected at Shanghai Synchrotron Radiation Facility (SSRF) BL17U and processed with HKL2000 (Otwinowski and Minor, 1997).

The HKU5-CTD structure was solved by the molecular replacement method using Phaser (McCoy et al., 2007) from the CCP4 program suite (Winn et al., 2011) with the structure of HKU4-RBD/CTD (PDB: 5GJ4) as the search model. Initial restrained rigid-body refinement and manual model building were performed using REFMAC5 (Murshudov et al., 2011) and COOT (Debreczeni and Emsley, 2012), respectively. Further refinement was performed using Phenix (Adams et al., 2010). Final statistics for data collection and structure refinement are represented in Table 1. Atomic coordinates and structure factors have been deposited in the Protein Data Bank with accession code 5XGR.

4.3. SPR analysis

The BIAcore experiments were performed at 25 °C using a BIAcore 3000 machine with CM5 chips (GE Healthcare). All proteins used in the experiment were expressed in insect cells. After purification, the protein was exchanged into PBS (pH7.4) containing 0.005% (v/v) Tween 20. The MERS-RBD/CTD, HKU4-RBD/CTD, and HKU5-CTD proteins were immobilized on the chip at ~1000 response units (RUs), respectively. Gradient concentrations of hCD26 (ranging from 0.195 to 200 μM) were then injected at 30 μL/min. After each cycle, the sensor surface was re-generated using 7 μL of 10 mM NaOH. Measurements from the reference flow cell (immobilized with BSA) were subtracted from experimental values. The binding kinetics were analyzed with BIAevaluation Version 4.1 software using the 1:1 Langumuir binding and/or the steady state affinity models.

4.4. Flow cytometry assay

Human hepatocellular carcinoma Huh7 or BHK-21 cells transfected with hCD26 (BHK-hCD26) were used for the binding test. Cells were stained with mouse Fc-fusion proteins, including MERS-RBD/CTD, HKU4-RBD/CTD, and HKU5-CTD, expressed using HEK 293T cells. BHK-21 cells were transfected with hCD26-containing plasmids 24 h before staining. Huh7 and BHK-hCD26 were suspended in PBS and incubated with individual proteins at 4 °C for 30 min, then washed and further incubated at 4 °C for another 30 min with an anti-mIgG/PE antibody. After washing, the cells were analyzed using a BD FACSCalibur. The data were processed with FlowJo software.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (NSFC, Grants 81461168030 and 31502078), the Key Research and Development Program of Ministry of Science and Technology of China (MOST, Grant 2016YFC1200300) and the China National Grand S&T Special Project (Grant 2014ZX10004-001-006). G. L. is supported in part by NSFC (Grant 81522026 and 31400154) and the Sichuan Outstanding Youth Science & Technology Funding (Grant No: 2016JQ0001). Y.S. is supported by the Excellent Young Scientist Program of the Chinese Academy of Sciences (CAS) and the Youth Innovation Promotion Association CAS (2015078). H.S. is supported by the Youth Innovation Promotion Association CAS (2017117). K.Y. Y. is partly supported by the NSFC/RGC Joint Research Scheme (N_HKU728/14) and the Theme-Based Research Scheme (T11/707/15) of the Research Grants Council, Hong Kong Special Administrative Region. G.F.G. is a leading principle investigator of the NSFC Innovative Research Group (Grant 81621091).

References

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., McCoy, A.J., Moriarty, N.W., Oeffner, R., Read, R.J., Richardson, D.C., Richardson, J.S., Terwilliger, T.C., Zwart, P.H., 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solutionActa Crystallogr. Sect. D, Biol. Crystallogr. 66, 213–221.

Agnihothram, S., Yount, B.L., Donaldson, E.F., Huynh, J., Menachery, V.D., Gralinski, L.E., Graham, R.L., Becker, M.M., Tomar, S., Scobey, T.D., Osswald, H.L., Whitmore, A., Gopal, R., Ghosh, A.K., Mesecar, A., Zambon, M., Heise, M., Denison, M.R., Barica, R.S., 2014. A mouse model for betacoronavirus subgroup 2c using a bat coronavirus strain HKU5 variant. mBio 5, e00047-14.

Annan, A., Baldwin, H.J., Corman, V.M., Klose, S.M., Owusu, M., Nkrumah, E.E., Badu, E.K., Anti, P., Agbenyega, O., Meyer, B., Oppong, S., Sarkodie, Y.A., Kalko, E.K., Lina, P.H., Godlevska, E.V., Reusken, C., Seebens, A., Gloza-Rausch, F., Vallo, P., Tschapka, M., Drosten, C., Drexler, J.F., 2013. Human betacoronavirus 2c EMC/2012-related viruses in bats, Ghana and Europe. Emerg. Infect. Dis. 19, 456–459.

Belouzard, S., Millet, J.K., Licitra, B.N., Whittaker, G.R., 2012. Mechanisms of coronavirus cell entry mediated by the viral spike protein. Viruses 4, 1011–1033.

Chan, C.M., Chu, H., Wang, Y.X., Wong, B.H.Y., Zhao, X.Y., Zhou, J., Yang, D., Leung, S.P., Chan, J.F.W., Yeung, M.L., Yan, J.H., Lu, G.W., Gao, G.F., Yuen, K.Y., 2016. Carcinoembryonic antigen-related cell adhesion molecule 5 is an important surface attachment factor that facilitates entry of Middle East respiratory syndrome coronavirus. J. Virol. 90, 9114–9127.

Chan, J.F.W., Lau, S.K.P., To, K.K.W., Cheng, V.C.C., Woo, P.C.Y., Yuen, K.Y., 2015. Middle East respiratory syndrome coronavirus: another zoonotic betacoronavirus causing SARS-like disease. Clin. Microbiol. Rev. 28, 465–522.

Cotten, M., Lam, T.T., Watson, S.J., Palser, A.L., Petrova, V., Grant, P., Pybus, O.G., Rambaut, A., Guan, Y., Pillay, D., Kellam, P., Nastouli, E., 2013. Full-genome deep sequencing and phylogenetic analysis of novel human betacoronavirus. Emerg. Infect. Dis. 19, 736–742B.

Dai, L., Wang, Q., Qi, J., Shi, Y., Yan, J., Gao, G.F., 2016. Molecular basis of antibody-mediated neutralization and protection against flavivirus. IUBMB Life 68, 783–791.

Debreczeni, J.E., Emsley, P., 2012. Handling ligands with Coot. Acta Crystallogr. Sect. D., Biol. Crystallogr. 68, 425–430.

Du, L.Y., He, Y.X., Zhou, Y.S., Liu, S.W., Zheng, B.J., Jiang, S.B., 2009. The spike protein of SARS-CoV - a target for vaccine and therapeutic development. Nat. Rev. Microbiol. 7, 226–236.

Gao, J., Lu, G., Qi, J., Li, Y., Wu, Y., Deng, Y., Geng, H., Li, H., Wang, Q., Xiao, H., Tan, W., Yan, J., Gao, G.F., 2013. Structure of the fusion core and inhibition of fusion by a heptad repeat peptide derived from the S protein of Middle East respiratory syndrome coronavirus. J. Virol. 87, 13134–13140.

Ge, X.-Y., Li, J.-L., Yang, X.-L., Chmura, A.A., Zhu, G., Epstein, J.H., Mazet, J.K., Hu, B., Zhang, W., Peng, C., Zhang, Y.-J., Luo, C.-M., Tan, B., Wang, N., Zhu, Y., Crameri, G., Zhang, S.-Y., Wang, L.-F., Daszak, P., Shi, Z.-L., 2013. Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. Nature 503, 535–538.

de Groot, R.J., Baker, S.C., Baric, R.S., Brown, C.S., Drosten, C., Enjuanes, L., Fouchier, R.A.M., Galiano, M., Gorbalenya, A.E., Memish, Z.A., Perlman, S., Poon, L.L.M., Snijder, E.J., Stephens, G.M., Woo, P.C.Y., Zaki, A.M., Zambon, M., Ziebuhr, J., 2013. Middle East respiratory syndrome coronavirus (MERS-CoV): announcement of the Coronavirus Study Group. J. Virol. 87, 7790–7792.

Heikkinen, T., Jarvinen, A., 2003. The common cold. Lancet 361, 51–59.

Huang, C., Qi, J., Lu, G., Wang, Q., Yuan, Y., Wu, Y., Zhang, Y., Yan, J., Gao, G.F., 2016. Putative receptor binding domain of bat-derived coronavirus HKU9 spike protein: evolution of betacoronavirus receptor binding motifs. Biochemistry 55, 5977–5988.

Huang, X.C., Dong, W.J., Milewska, A., Golda, A., Qi, Y.H., Zhu, Q.K., Marasco, W.A., Baric, R.S., Sims, A.C., Pyrc, K., Li, W.H., Sui, J.H., 2015. Human coronavirus HKU1 spike protein uses O-acetylated sialic acid as an attachment receptor determinant and employs hemagglutinin-esterase protein as a receptor-destroying enzyme. J. Virol. 89, 7202–7213.

Ithete, N.L., Stoffberg, S., Corman, V.M., Cottontail, V.M., Richards, L.R., Schoeman, M.C., Drosten, C., Drexler, J.F., Preiser, W., 2013. Close relative of human Middle East respiratory syndrome coronavirus in bat, South Africa. Emerg. Infect. Dis. 19, 1697–1699.

Kielian, M., Rey, F.A., 2006. Virus membrane-fusion proteins: more than one way to make a hairpin. Nat. Rev. Microbiol. 4, 67–76.

Kirchdoerfer, R.N., Cottrell, C.A., Wang, N., Pallesen, J., Yassine, H.M., Turner, H.L., Corbett, K.S., Graham, B.S., McLellan, J.S., Ward, A.B., 2016. Pre-fusion structure of a human coronavirus spike protein. Nature 531, 118–121.

Lau, S.K., Woo, P.C., Li, K.S., Huang, Y., Tsoi, H.W., Wong, B.H., Wong, S.S., Leung, S.Y., Chan, K.H., Yuen, K.Y., 2005. Severe acute respiratory syndrome coronavirus-like virus in Chinese horseshoe bats. Proc. Natl. Acad. Sci. USA 102, 14040–14045.

Lau, S.K., Li, K.S., Tsang, A.K., Lam, C.S., Ahmed, S., Chen, H., Chan, K.H., Woo, P.C., Yuen, K.Y., 2013. Genetic characterization of Betacoronavirus lineage C viruses in bats reveals marked sequence divergence in the spike protein of pipistrellus bat coronavirus HKU5 in Japanese pipistrelle: implications for the origin of the novel Middle East respiratory syndrome coronavirus. J. Virol. 87, 8638–8650.

Li, F., Li, W., Farzan, M., Harrison, S.C., 2005a. Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. Science 309, 1864–1868.

Li, W., Shi, Z., Yu, M., Ren, W., Smith, C., Epstein, J.H., Wang, H., Crameri, G., Hu, Z., Zhang, H., Zhang, J., McEachern, J., Field, H., Daszak, P., Eaton, B.T., Zhang, S., Wang, L.F., 2005b. Bats are natural reservoirs of SARS-like coronaviruses. Science 310, 676–679.

Lu, G., Hu, Y., Wang, Q., Qi, J., Gao, F., Li, Y., Zhang, Y., Zhang, W., Yuan, Y., Bao, J., Zhang, B., Shi, Y., Yan, J., Gao, G.F., 2013. Molecular basis of binding between novel human coronavirus MERS-CoV and its receptor CD26. Nature 500, 227–231.

Lu, G., Wang, Q., Gao, G.F., 2015. Bat-to-human: spike features determining 'host jump' of coronaviruses SARS-CoV, MERS-CoV, and beyond. Trends Microbiol. 23, 468–478.

Mahajan, M., Bhattacharjya, S., 2015. NMR structures and localization of the potential fusion peptides and the pre-transmembrane region of SARS-CoV: implications in membrane fusion. Biochim. Biophys. Acta 1848, 721–730.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., Read, R.J., 2007. Phaser crystallographic software. J. Appl. Crystallogr 40, 658–674.

Memish, Z.A., Mishra, N., Olival, K.J., Fagbo, S.F., Kapoor, V., Epstein, J.H., Alhakeem, R., Durosinloun, A., Al Asmari, M., Islam, A., Kapoor, A., Briese, T., Daszak, P., Al Rabeeah, A.A., Lipkin, W.I., 2013. Middle East respiratory syndrome coronavirus in bats, Saudi Arabia. Emerg. Infect. Dis. 19, 1819–1823.

Millet, J.K., Whittaker, G.R., 2014. Host cell entry of Middle East respiratory syndrome coronavirus after two-step, furin-mediated activation of the spike protein. Proc. Natl. Acad. Sci. USA 111, 15214–15219.

Millet, J.K., Whittaker, G.R., 2015. Host cell proteases: critical determinants of coronavirus tropism and pathogenesis. Virus Res. 202, 120–134.

Murshudov, G.N., Skubak, P., Lebedev, A.A., Pannu, N.S., Steiner, R.A., Nicholls, R.A., Winn, M.D., Long, F., Vagin, A.A., 2011. REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr. Sect. D., Biol. Crystallogr. 67, 355–367.

Otwinowski, Z., Minor, W., 1997. Processing of X-ray diffraction data collected in oscillation mode. Method Enzymol. 276, 307–326.

Peng, G., Sun, D., Rajashankar, K.R., Qian, Z., Holmes, K.V., Li, F., 2011. Crystal structure of mouse coronavirus receptor-binding domain complexed with its murine receptor. Proc. Natl. Acad. Sci. USA 108, 10696–10701.

Qian, Z., Ou, X., Goes, L.G., Osborne, C., Castano, A., Holmes, K.V., Dominguez, S.R., 2015. Identification of the receptor-binding domain of the spike glycoprotein of human betacoronavirus HKU1. J. Virol. 89, 8816–8827.

Saif, L.J., 1993. Coronavirus immunogens. Vet. Microbiol. 37, 285–297.

Simmons, G., Bertram, S., Glowacka, I., Steffen, I., Chaipan, C., Agudelo, J., Lu, K., Rennekamp, A.J., Hofmann, H., Bates, P., Pohlmann, S., 2011. Different host cell proteases activate the SARS-coronavirus spike-protein for cell-cell and virus-cell fusion. Virology 413, 265–274.

Walls, A.C., Tortorici, M.A., Bosch, B.J., Frenz, B., Rottier, P.J.M., DiMaio, F., Rey, F.A., Veesler, D., 2016. Cryo-electron microscopy structure of a coronavirus spike glycoprotein trimer. Nature 531, 114–117.

Wang, N., Shi, X., Jiang, L., Zhang, S., Wang, D., Tong, P., Guo, D., Fu, L., Cui, Y., Liu, X., Arledge, K.C., Chen, Y.H., Zhang, L., Wang, X., 2013. Structure of MERS-CoV spike receptor-binding domain complexed with human receptor DPP4. Cell Res. 23, 986–993.

Wang, Q., Qi, J., Yuan, Y., Xuan, Y., Han, P., Wan, Y., Ji, W., Li, Y., Wu, Y., Wang, J., Iwamoto, A., Woo, P.C., Yuen, K.Y., Yan, J., Lu, G., Gao, G.F., 2014. Bat origins of MERS-CoV supported by bat coronavirus HKU4 usage of human receptor CD26. Cell Host Microbe 16, 328–337.

WHO, 2004. Cumulative Number of Reported Probable Cases of Severe Acute Respiratory Syndrome (SARS). ⟨http://www.who.int/csr/sars/country/en/⟩.

WHO, 2016. Middle East Respiratory Syndrome Coronavirus (MERS-CoV) – Update. ⟨http://www.who.int/csr/don/7-january-2016-mers-oman/en/⟩.

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G., McCoy, A., McNicholas, S.J., Murshudov, G.N., Pannu, N.S., Potterton, E.A., Powell, H.R., Read, R.J., Vagin, A., Wilson, K.S., 2011. Overview of the CCP4 suite and current developments. Acta Crystallogr. Sect. D., Biol. Crystallogr. 67, 235–242.

Woo, P.C., Lau, S.K., Li, K.S., Poon, R.W., Wong, B.H., Tsoi, H.W., Yip, B.C., Huang, Y., Chan, K.H., Yuen, K.Y., 2006. Molecular diversity of coronaviruses in bats. Virology 351, 180–187.

Woo, P.C., Lau, S.K., Lam, C.S., Lau, C.C., Tsang, A.K., Lau, J.H., Bai, R., Teng, J.L., Tsang, C.C., Wang, M., Zheng, B.J., Chan, K.H., Yuen, K.Y., 2012. Discovery of seven novel mammalian and avian coronaviruses in the genus deltacoronavirus supports bat coronaviruses as the gene source of alphacoronavirus and betacoronavirus and avian coronaviruses as the gene source of gammacoronavirus and deltacoronavirus. J. Virol. 86, 3995–4008.

Xu, Y., Lou, Z., Liu, Y., Pang, H., Tien, P., Gao, G.F., Rao, Z., 2004. Crystal structure of severe acute respiratory syndrome coronavirus spike protein fusion core. J. Biol. Chem. 279, 49414–49419.

Yang, Y., Du, L., Liu, C., Wang, L., Ma, C., Tang, J., Baric, R.S., Jiang, S., Li, F., 2014. Receptor usage and cell entry of bat coronavirus HKU4 provide insight into bat-to-human transmission of MERS coronavirus. Proc. Natl. Acad. Sci. USA 111, 12516–12521.

Yang, Y., Liu, C., Du, L., Jiang, S., Shi, Z., Baric, R.S., Li, F., 2015. Two mutations were critical for bat-to-human transmission of Middle East respiratory syndrome coronavirus. J. Virol. 89, 9119–9123.

Zaki, A.M., van Boheemen, S., Bestebroer, T.M., Osterhaus, A.D., Fouchier, R.A., 2012. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. N. Engl. J. Med. 367, 1814–1820.

Zhu, J., Xiao, G., Xu, Y., Yuan, F., Zheng, C., Liu, Y., Yan, H., Cole, D.K., Bell, J.I., Rao, Z., Tien, P., Gao, G.F., 2004. Following the rule: formation of the 6-helix bundle of the fusion core from severe acute respiratory syndrome coronavirus spike protein and identification of potent peptide inhibitors. Biochem. Biophys. Res. Co. 319, 283–288.