THE PROTEIN SOCIETY

Influence of hydrophobic and electrostatic residues on SARS-coronavirus S2 protein stability: Insights into mechanisms of general viral fusion and inhibitor design

Halil Aydin, Dina Al-Khooly, and Jeffrey E. Lee*

Department of Laboratory Medicine and Pathobiology, Faculty of Medicine, University of Toronto, Toronto, Ontario M5S 1A8, Canada

Received 7 December 2013; Accepted 10 February 2014
DOI: 10.1002/pro.2442
Published online 12 February 2014 proteinscience.org

Abstract: Severe acute respiratory syndrome (SARS) is an acute respiratory disease caused by the SARS-coronavirus (SARS-CoV). SARS-CoV entry is facilitated by the spike protein (S), which consists of an N-terminal domain (S1) responsible for cellular attachment and a C-terminal domain (S2) that mediates viral and host cell membrane fusion. The SARS-CoV S2 is a potential drug target, as peptidomimetics against S2 act as potent fusion inhibitors. In this study, site-directed mutagenesis and thermal stability experiments on electrostatic, hydrophobic, and polar residues to dissect their roles in stabilizing the S2 postfusion conformation was performed. It was shown that unlike the pH-independent retroviral fusion proteins, SARS-CoV S2 is stable over a wide pH range, supporting its ability to fuse at both the plasma membrane and endosome. A comprehensive SARS-CoV S2 analysis showed that specific hydrophobic positions at the C-terminal end of the HR2, rather than electrostatics are critical for fusion protein stabilization. Disruption of the conserved C-terminal hydrophobic residues destabilized the fusion core and reduced the melting temperature by 30°C. The importance of the C-terminal hydrophobic residues led us to identify a 42-residue substructure on the central core that is structurally conserved in all existing CoV S2 fusion proteins (root mean squared deviation = 0.4 Å). This is the first study to identify such a conserved substructure and likely represents a common foundation to facilitate viral fusion. We have discussed the role of key residues in the design of fusion inhibitors and the potential of the substructure as a general target for the development of novel therapeutics against CoV infections.

Keywords: viral entry; SARS-CoV; viral fusion; coronavirus; MERS-CoV; glycoprotein; S2

Abbreviations: ACE2, angiotensin-converting enzyme 2; APN, aminopeptidase N; ASLV, avian sarcoma leucosis virus; BtCoV, bat coronavirus; CD, circular dichroism; CAECAM1, carcinoembryonic antigen adhesion molecule 1; CoV, coronavirus; DPP4, dipeptidyl peptidase 4; EBOV, ebola virus; Env, retroviral envelope glycoprotein; GP, glycoprotein; GP_1, glycoprotein 1 attachment domain; GP_2, glycoprotein 2 fusion domain; HA, hemagglutinin; HIV-1, human immunodeficiency virus type-1; HKU, Hong Kong university; HR1, Heptad repeat region 1; HR2, heptad repeat region 2; IAV, influenza A virus; LCMV, lymphocytic choriomeningitis virus; MERS-CoV, middle East respiratory syndrome-coronavirus; MHV, mouse hepatitis virus; RBD, receptor-binding domain; RBM, receptor-binding motif; RMSD, root mean square deviation; S, coronavirus spike glycoprotein; S1, coronavirus spike glycoprotein attachment subunit; S2, coronavirus spike glycoprotein fusion subunit; SARS-CoV, severe acute respiratory syndrome-coronavirus.

Grant sponsor: Canadian Institutes of Health Research (CIHR) Open Operating Grant; Grant number: MOP-115066, Grant sponsor: Canada Research Chair in Structural Virology and a CIHR New Investigator Award; Grant number: MSH-113554 to JEL. Grant sponsor: HA was supported by a University of Toronto Graduate Fellowship.

*Correspondence to: J. E. Lee, 1 King's College Circle, Room 6316, Medical Sciences Building, Toronto, ON M5S 1A8, Canada. E-mail: jeff.lee@utoronto.ca

Introduction

Coronaviruses (CoVs) are enveloped, positive-strand RNA viruses responsible for enteric and respiratory diseases in avian and mammalian species.[1] In 2002, the severe acute respiratory syndrome coronavirus (SARS-CoV) emerged in Southeast Asia and rapidly spread worldwide, resulting in more than 8000 cases and almost 800 deaths.[2–4] The unexpected emergence of the highly pathogenic human SARS-CoV revealed the potential for cross-species transmission from circulating strains of CoVs in zoonotic reservoirs.[5,6] Recently, a novel beta-coronavirus, termed Middle East respiratory syndrome (MERS) CoV, was discovered in the Arabian Peninsula.[7] Since then, the virus has now migrated to the United Kingdom, France, Italy, and Africa through infected travelers, and is considered a threat to global health with 42.5% case fatality rate among infected individuals.[8] Genetic sequence analyses show that MERS-CoV belongs to the betacoronavirus genus, along with the bat coronaviruses (BtCoVs) HKU4 and HKU5.[7] Currently, bats host more than 60 CoV species and a number of other SARS-like CoVs were identified from bats in Eurasia, Africa, and North America.[1,9,10] Although much has been discovered in the 10 years since the SARS-CoV discovery, emerging zoonotic CoVs continue to cause deadly outbreaks and threaten human health.

CoV infection is initiated by the spike (S) protein on the viral surface.[11] The SARS-CoV S is synthesized as a 1255-amino acid glycoprotein precursor and is classified as a class I viral fusion protein.[12] Upon proteolytic activation,[12] the S protein is cleaved into a S1 domain (residues 12–680) that is responsible for tropism and cellular attachment, and the S2 domain (residues 681–1255) that facilitates virus and host cell membrane fusion.[12,13] The SARS-CoV S1–S2 heterodimer assembles as a metastable trimer on the viral surface. Similar to other class I viral fusion proteins, such as human immunodeficiency virus type-1 (HIV-1) gp41, Ebola virus glycoprotein (GP_2), and influenza A hemagglutinin (HA_2), conformational changes in three functional elements of S2: the putative fusion peptide, heptad repeat 1 (HR1), and heptad repeat 2 (HR2) are critical for facilitating the fusion process.[12,14] Upon activation, the fusion peptide unfolds and inserts into the target cell membrane forming the pre-hairpin intermediate.[14,15] Subsequently, the HR2 region that anchors the viral membrane folds back near the HR1 trimeric core and triggers the collapse of the pre-hairpin intermediate state.[15] These conformational changes in the HR1 and HR2 regions draw the viral and host cell membranes together and mediate the merger of the two outer leaflets into a hemifusion stalk intermediate.[14,15] A final conformational step results in refolding of both HR1 and HR2 into a low energy postfusion state and allows the fusion pore to form.[14,15] There is a high kinetic barrier for the fusion of the two bilayer membranes; the free energy released during conformational changes of the fusion protein S2 provides the energetics to overcome the kinetic barriers for fusion pore formation.[15]

For class I viral fusion proteins, there are three types of fusion triggers: low pH, receptor binding, and proteolytic cleavage.[16] Some viruses, such as avian sarcoma leucosis virus,[17] SARS-CoV,[18] and perhaps Ebola virus,[19,20] utilize combinations of these triggers. The fusion of SARS-CoV is complex but is thought to require both receptor binding and proteolytic cleavage.[11,12] The proteolytic cleavage event that separates the receptor binding and fusion domains into non-covalently associated fragments depends on the species of CoV. Some CoVs are proteolytically cleaved at the S1–S2 boundary,[21] whereas others remain uncleaved, yet are still infectious.[11,22] For SARS-CoV, a primary proteolytic cleavage at the S1–S2 boundary, followed by a secondary cleavage at a S2′ position is often required to mediate membrane fusion.[22] Protease activation by trypsin-like, thermolysin, elastase, and factor Xa proteases on the plasma membrane[23–25] and cathepsin L proteolytic cleavage in the low pH endosomes were shown to enhance the SARS-CoV infection.[25–27] Regardless of the route of entry, viral fusion proteins require structural rearrangements in the S2 domain to mediate the merger of the virus and host cell lipid bilayers.[24,28]

Recent studies have identified key features that contribute to the function of viral fusion proteins from viruses that enter either at the plasma membrane or low pH endosomes. Although the atomic resolution structures of the SARS-CoV S1 domain and S2 fusion core in the postfusion hairpin conformation have been determined previously, characterization of specific residues involved in the stabilizing the SARS-CoV S2 during membrane fusion remain unclear. Any functional information regarding the molecular identities of these residues is important for the development of novel antiviral therapeutics. In order to understand the structural determinants involved in stabilizing the SARS-CoV S2 fusion protein, we performed site-directed mutagenesis to investigate the roles of electrostatic, polar, and hydrophobic residues on the SARS-CoV S2 extracellular fusion core and the key features necessary for pH-dependent viral fusion. Our results revealed that the SARS-CoV S2 fusion core is stable over a wide pH range and that specific hydrophobic residues at the HR1–HR2 interface play a major role in stabilizing the six-helix bundle. In contrast, three ion-pairs and chloride-binding site residues were shown to play minor roles in stabilizing the postfusion conformation. Specifically, interhelix interactions between the trimeric coiled-coil HR1 inner core and C-terminal portion of the HR2 helices are important determinants of SARS-CoV fusion, whereas those between the tether and inner HR1 core regions are

less important in stabilizing the postfusion state. We also identified a 42-residue conserved substructure within the central heptad repeat region of the SARS-CoV S2 fusion core that we hypothesize will provide a structural foundation for fusion. Our biophysical thermal stability data now explains the inhibition profiles of an array of SARS-CoV S2 peptidomimetics. The results presented here provide insights into the general mechanisms of viral fusion and identify an attractive site for coronavirus fusion inhibitor design.

Results

Generation and characterization of a linked HR1–HR2 trimeric SARS-CoV S2

The SARS-CoV S2 domain contains an extramembrane helical region that transforms into a coiled-coil six-helix bundle structure in the postfusion state. The N-terminus (residues 890–973) forms a long helical strand, often termed the HR1 region, with 22 helical turns. The HR1 region contains the typical heptad repeat motif of coiled-coil structures. Each repeat consists of a seven-residue ***a****bc****d****efg* motif, where hydrophobic residues (leucine, isoleucine, phenylalanine, and valine) are displayed in the ***a*** and ***d*** positions. The C-terminal segment of each protomer extends alongside the inner core in an antiparallel manner. These residues first form the random coil tether (residues 1142–1160) and then second heptad repeat region (HR2) between residues 1161–1179. The HR2 helices make five full turns and pack into the central HR1 trimer to form a highly stable six-helix bundle conformation that coordinates the fusogenic events between the virus and host cell membranes.

Recombinant expression of the full-length SARS-CoV S2 for structural or biophysical studies is challenging. There are no existing structural models of the entire SARS-CoV S2 protein. The SARS-CoV S2 fusion core, consisting of only the HR1 and HR2 helical regions, has been crystallized, and the structure solved. Commonly, the SARS-CoV S2 fusion core is reconstituted through the addition of synthesized peptides corresponding to HR1 and HR2 regions.[29–31] Here, we designed a linked recombinant SARS-CoV S2 fusion protein (SARS-CoV S2 L2H) using N-terminal residues 896–972 and C-terminal residues 1142–1183, with a six-residue glycine–serine linker (GGS–GGS) between the two regions [Fig. 1(A)]. This construct is similar to a previously designed construct.[32] When expressed using an *Escherichia coli* SHuffle T7 expression system, we were able to obtain multi-milligram quantities of protein. The SARS-CoV S2 L2H is soluble and migrates as a stable trimer on size exclusion chromatography. Sedimentation equilibrium analytical ultracentrifugation confirmed the trimeric nature of the SARS-CoV S2 L2H [Fig. 1(B)]. Furthermore, the circular dichroism (CD) spectrum of the protein was characterized by double minima at 208 and 222 nm, the typical CD signature for a predominantly α-helical protein [Fig. 1(C)]. The CD wavelength scans for all SARS-CoV S2 L2H mutants were similar (data not shown), suggesting that the relative α-helical content did not change as a result of the mutations. Moreover, SARS-CoV S2 L2H contains an estimated 50% α-helical content, in line with the secondary structural composition seen in the X-ray crystal structure of SARS-CoV S2.

SARS S2 fusion core is stable over a wide pH range

Some CoV's such as hCoV-229E enter cells via the low pH endosomal environment, whereas others, like mouse hepatitis virus (MHV)−4, directly fuse at the plasma membrane.[33,34] Interestingly, SARS-CoV can enter cells through either pH-dependent or pH-independent entry pathways depending on the presence of proteases.[23,24] In order to investigate the pH dependence of the SARS-CoV S2 fusion core structure, we performed CD thermal denaturation assays in buffers ranging from pH 4.0 to 8.5. The wild-type (WT) SARS-CoV S2 L2H denatured irreversibly with a melting temperature (T_m) of 97.0°C at neutral pH. This is consistent with the previously reported T_m for SARS-CoV S2 of >90°C.[35] The Gibbs free energy of unfolding cannot be calculated from an irreversible denaturation curve; however, the apparent melting temperature may provide a simple measure of protein stability. At all pH levels, the SARS-CoV S2 L2H fusion core was highly stable [Fig. 2(A); Table I]. As pH was increased, the T_m values remained unchanged (~95°C). At lower pH conditions that correspond to early (6.0–6.5) and late (5.0–6.0) endosomal environments, slightly lower melting temperatures from 92.9°C to 94.6°C were observed. Varying pH did not appear to have drastic effects on protein stability [Fig. 2(B)]; the postfusion SARS-CoV S2 L2H is stable from pH 4.0 to 8.5.

Electrostatic interactions play a minor role in the stability of the fusion core

Salt bridges are long-range electrostatic interactions typically formed between an anionic carboxylate ($-COO^-$) functional group of aspartate or glutamate and the cationic ammonium of lysine ($-NH_3^+$), the guanidinium group of arginine, or imidazole ring of histidine. Electrostatic interactions can contribute up to 10 kcal/mol in free energy,[36] and thus are important factors for stabilizing protein structures. The free energy contribution of a salt bridge is pH-dependent, as the ionization of the side chain is affected by the pH of the local environment.[36] The SARS-CoV S2 fusion core structure revealed three sets of electrostatic interactions clustered in two regions of the protein.[30] The HR1–HR2 region contains two salt bridge pairs: between Arg1167 and

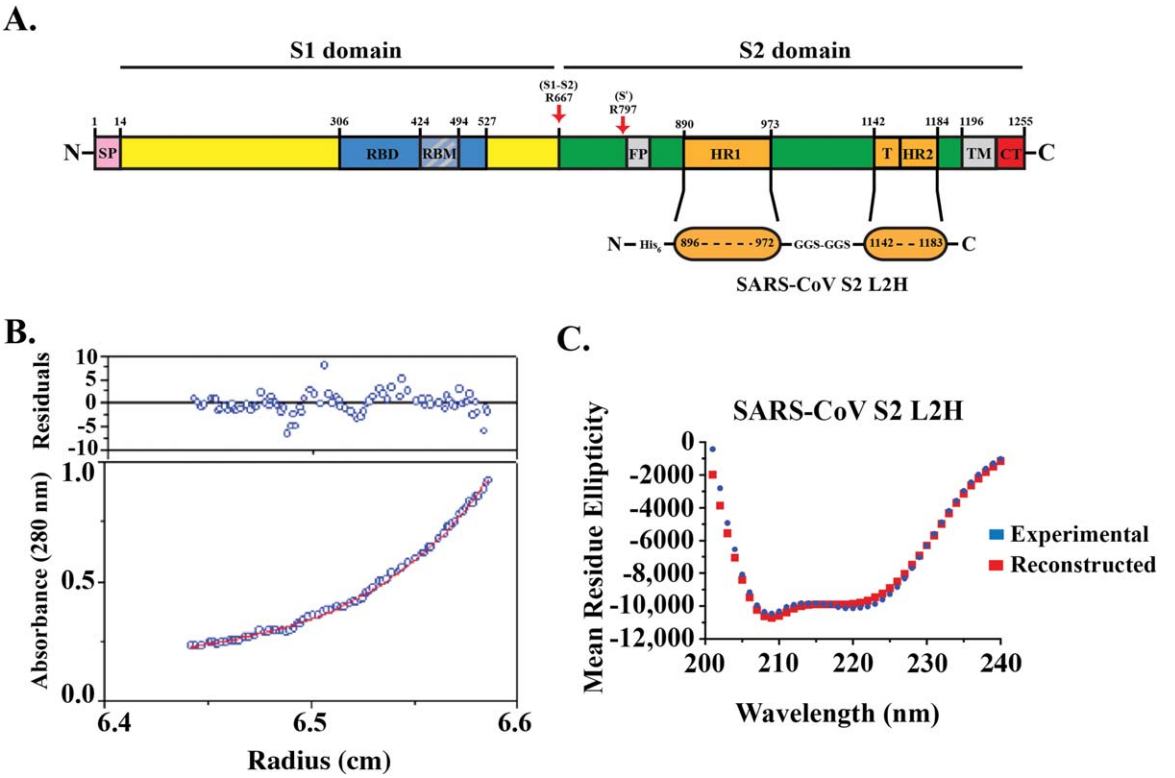

Figure 1. Structural description and biophysical characterization of the SARS-CoV S2 L2H protein. **(A)** Schematic diagram of the SARS-CoV S protein. The S protein exhibits the characteristic domain organization of class I viral proteins. Abbreviations are as follows: S1, CoV attachment subunit; S2, CoV fusion subunit; SP, signal peptide; RBD, receptor binding domain; RBM, receptor binding motif; FP, fusion peptide; HR1, heptad repeat 1 region; HR2, heptad repeat 2 region; T, tether region; TM, transmembrane domain; CT, cytoplasmic tail; L2H, linked two-heptad construct. The positions of the S1 domain (residues 14–667), S2 domain (residues 668–1255), SP (residues 1–14), RBD (residues 306–527), RBM (residues 424–494), HR1 (residues 890–973), tether, and HR2 (residues 1142–1184), TM and CT (residues 1196–1255) are shown above the schematic. *Red arrows* indicate the S1–S2 and S′ proteolytic cleavage sites at residues R667 and R797, respectively. SARS-CoV S2 L2H construct was generated by using HR1 residues 896–972 and tether/HR2 residues 1142–1183 connected by a six amino acid linker at the HR1 C-terminal and HR2 N-terminal ends (colored in *orange*). **(B)** Sedimentation equilibrium data for a 20 μ*M* sample at 4°C and 22,000 rpm in TBS buffer. The curve indicates the distribution of a 48.4-kDa protein. The data fit closely to a trimeric model for SARS-CoV S2 L2H. The deviation in the data from the linear fit for a trimeric model is plotted in the upper panel. **(C)** Experimental CD wavelength scan of SARS-CoV S2 L2H (*blue*) at 25°C reveals minimas at 208 and 222 nm, indicative of strong α-helical secondary structural characteristics. The SARS-CoV S2 L2H is calculated to contain 50% α-helical content. A reconstructed CD wavelength scan (*red*) shows the quality of the fit used in the calculation of secondary structural content.

Glu918 in the middle of the HR2, and between Glu1164 and Lys929 residues at the membrane-distal end of the HR2 [Fig. 3(A)]. A third complex intersubunit electrostatic interaction is formed between Arg965, Asp967, and Glu970 at the base of the HR1 helical core. Alanine mutations to abrogate these salt bridges led to a modest decrease in melting temperatures ($\Delta T_m < 8°C$) [Fig. 3(B); Table II]. Interestingly, charge reversal mutations (Lys929Glu, Arg965Glu, and Arg1167Glu) destabilized the SARS-CoV S2 L2H six-helix bundle to a similar extent ($\Delta T_m < 8°C$). Double and triple charge reversal mutations (Lys929Glu–Arg965Glu, Arg965Glu–Arg1167Glu, and Lys929Glu–Arg965Glu–Arg1167-Glu) resulted in an additive reduction in stability ($\Delta T_m \sim 13°C$); however, the changes were still modest [Fig. 3(C); Table II]. In conclusion, our results

Table I. *Summary of SARS-CoV Fusion Protein Stability Under Various pH Conditions*

Buffered pH	T_m (°C)[a]
10 m*M* NaOAc, pH 4.0	90.6 ± 0.3
10 m*M* NaOAc, pH 4.5	92.7 ± 0.3
10 m*M* NaOAc, pH 5.0	92.9 ± 0.3
10 m*M* NaOAc, pH 5.5	93.4 ± 0.3
10 m*M* NaOAc, pH 6.0	93.5 ± 0.3
10 m*M* NaOAc, pH 6.5	94.6 ± 0.3
10 m*M* Tris-HCl, pH 7.0	95.0 ± 0.5
10 m*M* Tris-HCl, pH 7.5	97.0 ± 0.2
10 m*M* Tris-HCl, pH 8.0	94.9 ± 0.2
10 m*M* Tris-HCl, pH 8.5	95.3 ± 0.2

[a] The midpoint thermal denaturation (T_m) value was estimated from fraction unfolded (F_{unf}) and plotted as a function of temperature. Error values indicate 95% confidence intervals from fitting to a non-linear biphasic sigmoidal curve.

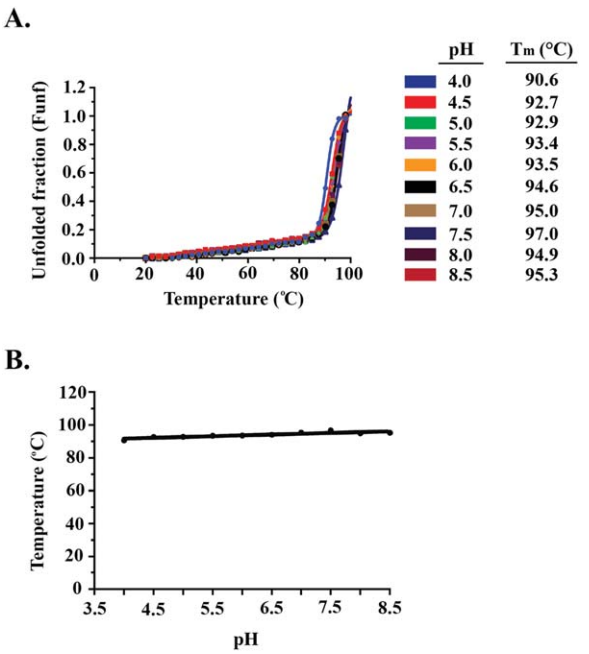

Figure 2. SARS-CoV S2 L2H stability at various pH values. **(A)** Thermal denaturation of SARS-CoV S2 L2H monitored by CD molar ellipticity at 222 nm in sodium acetate buffer (between pH 4.0 and 6.5), and TBS buffer (between pH 7.0 and 8.5). The CD signal was baseline corrected, normalized between 0 (folded) and 1 (unfolded), and fit to a non-linear biphasic sigmoidal curve. The T_m values correspond to the temperature where 50% of the protein has unfolded. **(B)** Plot of SARS-CoV S2 L2H stability as a function of pH. The trimeric SARS-CoV S2 L2H is stable between pH values 4.0 and 8.5.

demonstrated that salt bridges play a small role on the SARS-CoV S2 fusion core stability.

Hydrophobic residues are important for postfusion stability

Given our findings that electrostatic interactions play only a minor role in the stability of the SARS-CoV S2 L2H fusion subunit, we focused on the role of the hydrophobic residues in maintaining structural integrity. The postfusion structure of SARS-CoV S2 reveals a series of hydrophobic residues positioned at the interface of the HR1–tether (Leu1148 and Ile1151) and HR1–HR2 (Ile1161, Leu1168, and Leu1175) [Fig. 4(A)]. We hypothesized that the hydrophobic interactions between the HR1–tether and HR1–HR2 regions play an essential role in stabilizing the outer layer (HR2 and tether) to the inner core (HR1) in the postfusion state. To test our hypothesis, we mutated all five hydrophobic residues to an alanine residue and performed CD thermal denaturation assays. Mutations to hydrophobes at the HR1–tether interface (Leu1148Ala and Ile1151Ala) destabilized SARS-CoV S2 L2H on the order of $\Delta T_m \sim 10°C$ [Fig. 4(B); Table II]. Strikingly, alanine mutations to the hydrophobic residues at the HR1–HR2 interface (Ile1161Ala, Leu1168Ala, and Leu1175Ala) had drastic effects on protein stability, with >20°C decrease each on the apparent T_m, as compared with WT SARS-CoV S2 L2H. Specifically, a single alanine mutation to Leu1168 or Leu1175 residue led to a approximately 30°C decrease on the melting temperature of SARS-CoV S2 L2H [Fig. 4(B); Table II]. Our results indicate that hydrophobic interactions between HR1 and HR2, specifically Leu1168 and Leu1175, are critical for the stability of postfusion SARS-CoV S2. The hydrophobic HR2 residues involved in postfusion stability are well conserved across the coronavirus family (Fig. 5).

Putative chloride binding site reinforces the structural stability of the postfusion core

In many postfusion viral fusion proteins, the heptad repeats in the HR1 helix are broken up by a layer of aspargine or glutamine residues to coordinate a putative chloride ion.[37] This phenomenon is seen in all CX_6CC-containing retrovirus and filovirus fusion proteins, and is suggested to be important as a conformational switch between the prefusion and postfusion states.[37–39] The SARS-CoV S2 extramembrane helical fusion core contains two chloride-binding sites[30] [Fig. 4(C)]. The first chloride-binding site of HR1 is located at the membrane-proximal end of the protein and is coordinated by Gln902, whereas the second site is at the center of the trimeric HR1 coiled coil and is coordinated by Asn937. In order to investigate the significance of these chloride-binding sites, we mutated both Gln902 and Asn937 to an alanine residue and monitored the changes on the stability of the protein. Gln902Ala and Asn937Ala resulted in the decrease of the apparent T_m values by 8°C and 14°C, respectively [Fig. 4(D); Table II]. Thermal unfolding of the Gln902Ala–Asn937Ala double mutant revealed an additive effect in the change of melting temperature ($\Delta T_m = \sim 23°C$). As a control, we mutated an asparagine residue (Asn951) located outside of the chloride-binding sites and assessed its contribution on postfusion stability. As expected, Asn951Ala did not result in a change on the apparent T_m value [Fig. 4(D); Table II]. Taken together, this suggests that conserved polar interactions with the chloride ion, in particular the central chloride-binding site, are important for the postfusion stability of the protein.

Discussion

All class I viral glycoproteins, including those from coronaviruses, utilize a similar mechanism of fusion, in which structural rearrangements of two highly conserved heptad repeats juxtapose the viral and host cell membranes to form the fusion pore. The conformational changes necessary for membrane fusion requires the formation of an energetically

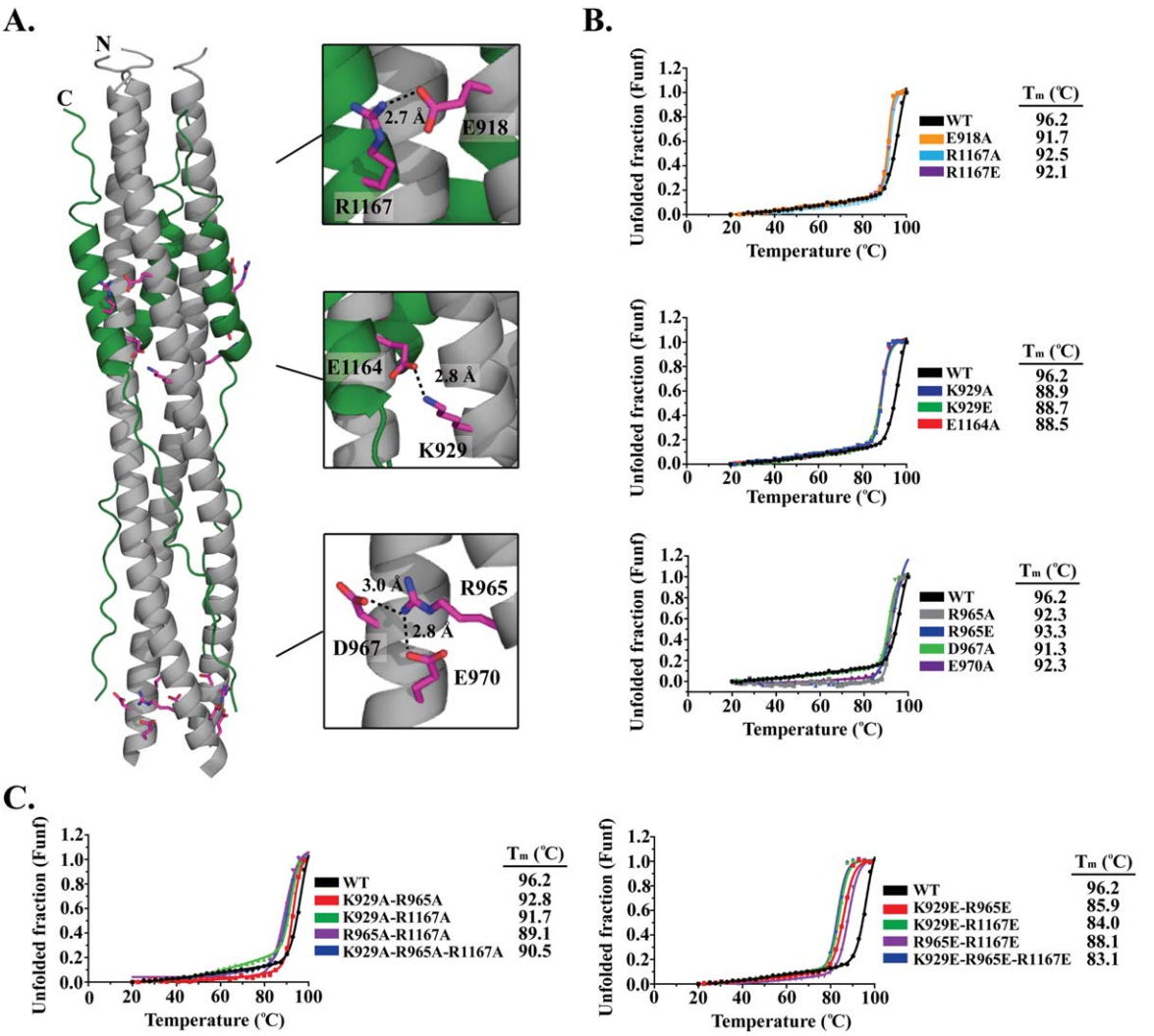

Figure 3. Biophysical characterization of SARS-CoV S2 electrostatic interactions. **(A)** Ribbon diagram of SARS-CoV S2 fusion core (PDB code: 2BEZ) shows electrostatic interactions between the HR1–HR1 and HR1–HR2 regions. The HR1 and tether/HR2 regions are depicted in *gray* and *green*, respectively. The side chains of the ion-pair interactions are colored in *magenta*. The zoomed views of ion-pairs are shown in the inset *boxes* and the distances between the residues are indicated in *Ångstroms* (Å). **(B)** Thermal denaturation profiles of wild-type (WT), single, **(C)** double and triple mutant of electrostatic residues in the SARS-CoV S2 fusion subunit. Thermal stability was recorded at 222 nm. All data were baseline corrected, normalized between 0 (folded) and 1 (unfolded) and plotted as a function of temperature. The T_m values indicate the midpoint melting temperatures for WT and mutant proteins.

stable six-helix bundle structure in the postfusion state. The transition from the higher energy metastable prefusion to the lower energy six-helix bundle postfusion conformation provides the energetics for fusion. Recently, it was suggested that the stability of the postfusion subunit is highest at the pH of the environment of where fusion occurs.[40,41] For example, human T-lymphotropic virus-1 (HTLV-1) fuses at the host plasma membrane at neutral pH. The HTLV-1 gp21 fusion subunit is most stable at pH values above 7.0 (61.0°C at pH 5.0 vs. >99.0°C at pH 7.5).[41] Ebola virus fusion occurs at the endolysosome, and its GP_2 subunit is most stable at pH values below 5.5 (86.8°C at pH 5.3 vs. 49.8°C at pH 6.1).[40] The observation that the stability of the fusion subunit mimics the environment where they fuse also holds true for avian sarcoma leukosis virus (ASLV). ASLV is a retrovirus that undergoes a unique two-step entry mechanism that involves first receptor-binding at the plasma membrane, followed by low pH activation in the endosome.[17] The ASLV TM fusion protein is stable over a broad range of pH values (67.0°C–73.8°C at pH values between 5.0 and 8.5).[41] Although SARS-CoV does not undergo a two-step mechanism of entry, SARS-CoV entry may be promiscuous as it is able to enter target cells through either a low pH endosomal route or direct fusion at the plasma membrane at neutral

Table II. *Summary of Wild-Type and Mutant SARS-CoV Fusion Protein Stabilities*

SARS-CoV S2 mutant	T_m (°C)[a]	Location of mutation
Wild type	96.2 ± 0.2	–
Q902A	89.0 ± 0.2	HR1 top
E918A	91.8 ± 0.4	HR1 top
K929A	88.9 ± 0.3	HR1 top
K929E	88.7 ± 0.2	HR1 top
N937A	82.7 ± 0.3	HR1 central
N951A	94.4 ± 0.2	HR1 central
R965A	92.3 ± 0.4	HR1 bottom
R965E	93.3 ± 0.8	HR1 bottom
D967A	91.3 ± 0.3	HR1 bottom
E970A	92.3 ± 0.1	HR1 bottom
L1148A	88.2 ± 0.2	Tether region
I1151A	86.4 ± 0.2	Tether region
I1161A	79.5 ± 0.3	HR2 bottom
E1164A	88.5 ± 0.3	HR2 bottom
R1167A	92.5 ± 0.4	HR2 middle
R1167E	92.1 ± 0.3	HR2 middle
L1168A	67.0 ± 0.3	HR2 top
L1175A	66.0 ± 0.4	HR2 top
Q902A-N937A	74.3 ± 0.2	HR1 top and central
K929A-R965A	92.8 ± 0.3	HR1 top and bottom
K929E-R965E	85.9 ± 0.2	HR1 top and bottom
K929A-R1167A	91.7 ± 0.3	HR1 top and HR2 middle
K929E-R1167E	84.0 ± 0.2	HR1 top and HR2 middle
R965A-R1167A	~89.1[b]	HR1 bottom and HR2 middle
R965E-R1167E	88.1 ± 0.2	HR1 bottom and HR2 middle
K929A-R965A-R1167A	90.5 ± 0.2	HR1 top, HR1 bottom and HR2 middle
K929E-R965E-R1167E	83.1 ± 0.2	HR1 top, HR1 bottom and HR2 middle

[a] The midpoint thermal denaturation (T_m) value was estimated from fraction unfolded (F_{unf}) and plotted as a function of temperature. Error values indicate 95% confidence intervals from fitting to a non-linear biphasic sigmoidal curve.
[b] The T_m value for this double mutant is an estimate; errors were not calculated

pH.[23,24,26,42,43] Consistent with ASLV TM, the linked SARS-CoV S2 is stable between pH 4.0 and 8.5. This provides further support that the SARS-CoV S2 fusion subunit is able to maintain the core stability regardless of its route of entry and environments encountered.

Structural and biophysical characterization of class I viral fusion subunits have identified general features required for stabilization of the postfusion structure. Mason-Pfizer monkey virus (MPMV), HTLV-1, and xenotropic murine leukemia virus-related virus (XMRV) belong to the β-, δ-, and γ-retrovirus genus, respectively, and fuse at the plasma membrane in a pH-independent entry process.[44,45] Crystal structures of HTLV-1, MPMV, and XMRV transmembrane fusion protein domains identified a series of electrostatic and hydrophobic interactions between the HR1 and HR2 regions.[38,46] Mutations to the negatively and positively charged residues result in a significant decrease in the stability of the postfusion glycoprotein structure and decrease viral infectivity, suggesting a major role for salt bridges in stabilizing the postfusion six-helix bundle.[46,47] Electrostatic interactions may be a common strategy used by viruses to stabilize the fusion protein at the plasma membrane. In ASLV, which has been used extensively as a model virus for both pH-dependent and pH-independent entry, the fusion domain crystal structure also revealed a lining of electrostatic salt bridges. However, these electrostatic interactions do not play a major role in stabilizing the six-helix bundle, as they do in fusion proteins of retroviruses with pH-independent entry. Instead, hydrophobic residues in the ASLV fusion protein play a large role in stabilizing the postfusion state.[41] The use of hydrophobic residues for stabilization is consistent with ASLV's mode of entry requiring low pH,[41] as the strength of hydrophobic interactions are not affected by pH changes in the environment. Other class I viruses such as influenza A virus (IAV) and lymphocytic choriomeningitis virus (LCMV), which solely enter host cells through low pH endosomes, also contain both ionic and hydrophobic interactions within their fusion core in the postfusion state.[48,49] It is not clear whether the ionic residues are important for stability, as no studies have been performed on proteins from these viruses. Like ASLV, these viruses may utilize hydrophobic interactions as their primary mechanism for maintaining postfusion stability within the endosomes.

Hydrophobic residues appear to play a greater role in stabilizing the fusion subunits from viruses that are pH-dependent than do electrostatic

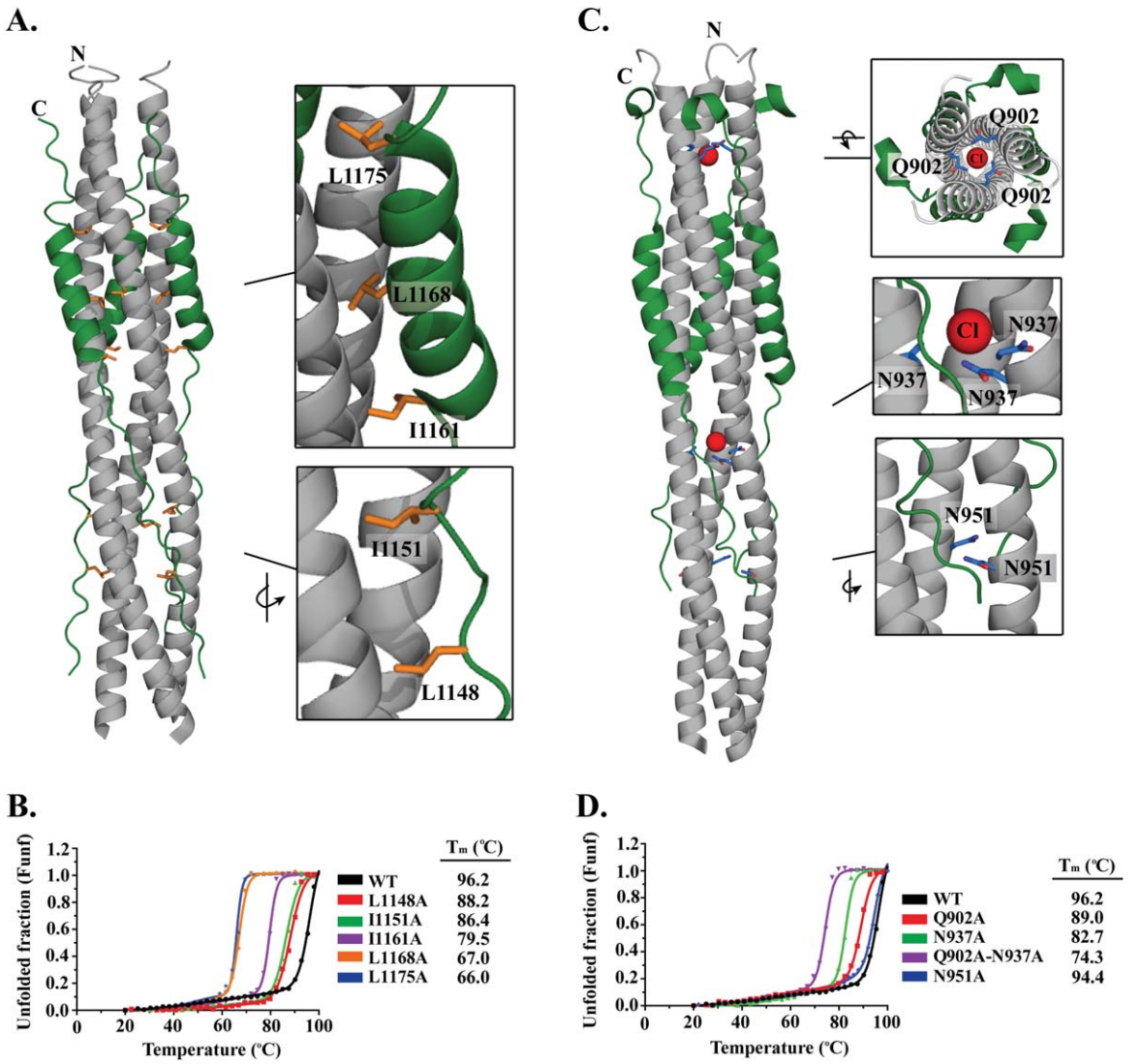

Figure 4. SARS-CoV S2 hydrophobic and polar interactions. **(A)** Ribbon diagram of SARS-CoV S2 fusion core structure (PDB Code: 2BEZ). The HR1 region forms a long helical strand with 22 helical turns (colored in *gray*). The tether and HR2 regions extend alongside the HR1 inner core in an antiparallel manner (colored in *green*). The hydrophobic residues at the HR1–tether and HR1–HR2 interface are depicted in *orange*. The inset boxes show the zoomed view of critical hydrophobic residues positioned at the interfaces. **(B)** Thermal denaturation of wild-type (WT) and mutant hydrophobic residues in the SARS-CoV fusion subunit. **(C)** Ribbon diagram of an extended SARS-CoV S2 fusion core structure (PDB Code: 1WYY) displaying two putative chloride binding sites. Chloride ions observed in the crystal structure of the HR1 inner core are shown in *red*. The polar residues interacting with chloride ion (Q902 and N937) and a single polar residue (N951) at the HR1-tether interface are shown as *blue sticks*. The HR1 and tether/HR2 regions are depicted in *gray* and *green*, respectively. **(D)** Thermal denaturation profiles of wild-type (WT) and chloride binding site mutants. All thermal denaturation profiles are plotted as described in Figure 3.

interactions. On the flip side, electrostatic interactions are important to viruses that fuse at the plasma membrane at neutral pH. SARS-CoV is able to enter through direct fusion at the plasma membrane and the low pH environment of the endosome; therefore, we expected that the fusion subunit would have structural features typical of both types of viral fusion proteins. Our biophysical study on the linked SARS-CoV S2 now provides additional evidence to support this hypothesis. Salt bridge interactions in SARS-CoV S2 play a minor role in stabilizing the six-helix bundle. Moreover, the lack of a role for salt bridges is supported by the poor conservation of some salt bridge residues (i.e., Glu918–Arg1167) in all CoVs.

Coronaviruses are capable of animal-to-human transition, and CoVs that infect pets or animals that frequent urban centers are human health threats due to the potential for mutations that will allow the virus to cross the interspecies barrier. The development of CoV inhibitors will provide a weapon against existing, emerging, or re-emerging CoV outbreaks. The coronavirus S protein plays key roles in facilitating the attachment of the virus to the host

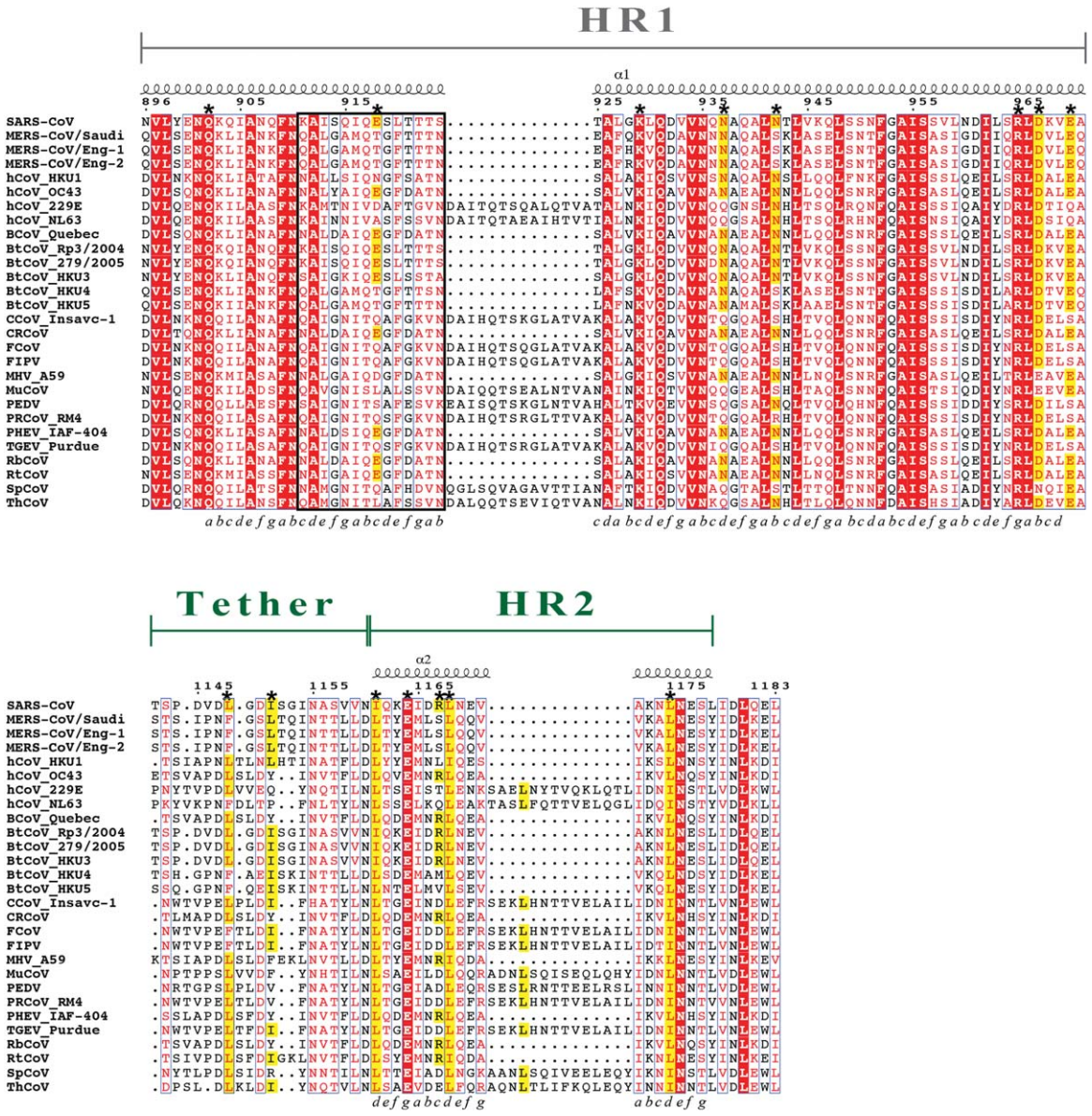

Figure 5. Primary sequence alignment of CoV S2 fusion cores. Multiple sequence alignment of various human and animal CoV fusion proteins. Abbreviations are as follows: SARS-CoV, severe acute respiratory syndrome-coronavirus; MERS-CoV, middle east respiratory syndrome-coronavirus; hCoV, human coronavirus; HKU, Hong Kong University strain; BCoV, bovine coronavirus; BtCoV, bat coronavirus; CCoV, canine coronavirus; FCoV, feline coronavirus; FIPV, feline infectious peritonitis virus; MHV, mouse hepatitis virus; MuCoV, munia coronavirus; PEDV, porcine epidemic bronchitis virus; PRCoV, porcine respiratory coronavirus; PHEV, porcine hemagglutinating encephalomyelitis virus; TGEV, transmissible gastroenteritis virus; RbCoV, rabbit coronavirus; RtCoV, rat coronavirus; SpCoV, sparrow coronavirus; ThCoV, thrush coronavirus. Sequence boundaries for HR1 and tether/HR2 regions are depicted with gray and green lines, respectively. Residue numbers corresponding to the SARS-CoV S2 fusion subunit numbering are indicated above the alignment. Strictly conserved residues are outlined in *red* and residues that are important for the stability of the SARS-CoV S2 fusion core are highlighted in *yellow* and marked with an asterisk (*). Residues (911–924) involved in the formation of the common HR1 substructure are shown in a *black box*. The heptad repeat register (*a, b, c, d, e, f, g*) of the SARS-CoV S2 fusion core is indicated below the alignment.

receptor, and catalyzing the fusion of the virus and host lipid bilayers. While drugs have been developed against the attachment subunit (S1 equivalent) for other viruses, this may not be a good target for coronaviruses, as they utilize a diverse range of cellular receptors for host attachment. For example, SARS-CoV and hCoV-NL63 use angiotensin-converting enzyme 2 (ACE2) as a receptor for infection of target cells,[50,51] whereas MHV and hCoV-229E utilize carcinoembryonic antigen adhesion molecule 1 (CAECAM1) and aminopeptidase N (APN) as receptors, respectively.[52,53] Lastly, Raj *et al*. showed that the recently identified MERS-CoV binds to an exopeptidase, dipeptidyl peptidase 4 (DPP4), as a functional

Table III. *Summary of SARS-CoV S2 HR2 Peptide Mimics*

Peptide	Residue region	Number of residues	IC_{50} (assay)	Reference
N-term HR2 extensions				
sHR2-1	1126–1189	63	43 ± 6.4 μ*M*	Bosch *et al*., 2004
sHR2-2	1130–1189	60	24 ± 2.8 μ*M*	Bosch *et al*., 2004
sHR2–8	1126–1193	68	17 ± 3.0 μ*M*	Bosch *et al*., 2004
sHR2–9	1126–1185	60	34 ± 4.0 μ*M*	Bosch *et al*., 2004
C-term HR2 extensions				
HR2–18	1161–1187	27	3.68 ± 1.5 μ*M*	Yuan *et al*., 2004
CP-1	1153–1189	37	19 μ*M*	Liu *et al*., 2004
HR2–38	1149–1186	38	66.2 n*M*	Zhu *et al*., 2004
HR2–38*	1149–1186	38	0.5–5 n*M*	Zhu *et al*., 2004
HR2–44	1149–1192	44	500 n*M*	Zhu *et al*., 2004
HR2–38	1149–1186	38	1.02 ± 0.02 μ*M*	Ni *et al*., 2005
SR9	1151–1185	35	100 n*M*	Ujike *et al*., 2008
HR2	1151–1185	35	0.34 μ*M*	Chu *et al*., 2008
P1	1153–1189	37	3.04 μ*M*	Liu *et al*., 2009
P4	1153–1182	30	3.17 μ*M*	Liu *et al*., 2009
P6	1153–1175	23	2.28 μ*M*	Liu *et al*., 2009

* = synthetic HR2 peptide.

receptor for entry.[54] Studies conducted by Lu *et al*.,[55] Du *et al*.,[56] and Wang *et al*.[57] revealed that a 286-amino acid fragment within the S1 domain of MERS-CoV interact with the DPP4 receptor. These analyses highlighted notable differences between coronavirus S1 protein-receptor interactions. Furthermore, primary sequence analysis reveals that there is only a <2% sequence identity between all the S1 domains of CoV. Antiviral therapeutics targeting this critical region will likely result in species-specific drugs.

In contrast, the CoV S2 protein is likely an excellent target for the design of more general CoV inhibitors. CoV S2 plays an indispensable role in catalyzing the fusion of the virus and host lipid bilayers and residues involved in fusion are relatively well conserved across all family members (Fig. 5). Targeting the viral fusion subunit is a proven strategy, as demonstrated by the efficacy of the FDA-approved HIV-1 gp41 HR2 mimic enfuvirtide (T-20).[58–60] SARS-CoV S2 HR2 peptides are also effective entry inhibitors based on pseudovirus and cell–cell fusion assays.[32,61–65] The peptides have traditionally been designed blindly by systematic addition of residues to the core HR2 region. Based on our biophysical fusion protein stability data, we are able to rationalize the trends of effectiveness of peptides. Our thermal denaturation data clearly shows that Ile1161, Leu1168, and Leu1175 are critical to the stability of the postfusion six-helix bundle structure, whereas hydrophobic residues that belong to the tether region have modest effects on stability. Hydrophobic residues at the C-terminal end of the HR2 region are more critical to the stability of the postfusion state than those at the tether region. The importance of these residues correlates well with the SARS-CoV HR2 peptide inhibition studies[35,43,61,64–68] (Table III). Peptides that encompass the HR2 hydrophobic residues (Ile1161, Leu1168, and Leu1175) and those corresponding to the HR2 C-terminal ends had better IC_{50} values than other peptides tested. Peptides that contain residues from the N-terminal tether region were less effective. Our data suggest that effective SARS-CoV HR2 peptide inhibitors should encompass the HR2 region and residues C-terminal to HR2.

Recently, crystal structures of MERS-CoV S2 fusion core have been determined at high atomic resolution[69,70] and structural comparison between MERS-CoV and SARS-CoV fusion cores revealed a high degree of overall structural similarity with an root mean squared deviation (RMSD) of approximately 0.9 Å for 204 Cα atoms.[69,70] We hypothesized that interacting residues on the HR1 may form a conserved interface to accommodate the HR2 hydrophobic residues for fusion. Analysis of available coronavirus S2 structures revealed strong structural conservation of the HR1 region that interacts with the key hydrophobic HR2 residues. Superimposition of a 42-residue HR1 region surrounding the HR2 binding site reveals an average RMSD of 0.4 Å between the SARS-CoV, MERS-CoV, MHV, and hCoV-NL63 S2 fusion subunits [Fig. 6(A)]. This is in contrast to an overall superimposition of the entire SARS-CoV, hCoV-NL63, and MERS-CoV, MHV S2 inner HR1 trimeric structures which showed C_α atom RMSDs of ~1.3, ~0.6, and ~0.6 Å, respectively. The surface of the substructure has a long groove (16-Å long × 9-Å wide × 7-Å deep) and a pocket (7-Å long × 9-Å wide × 7-Å deep) at the HR2 interface. The rim of the pocket is surrounded with polar (glutamine and asparagine) and hydrophobic residues (leucine, isoleucine, serine, and alanine) whereas the bottom of the pocket is lined with isoleucine residues [Fig. 6(B)]. HR2 residues that pack into the HR1 pocket are conserved in all fusion core structures

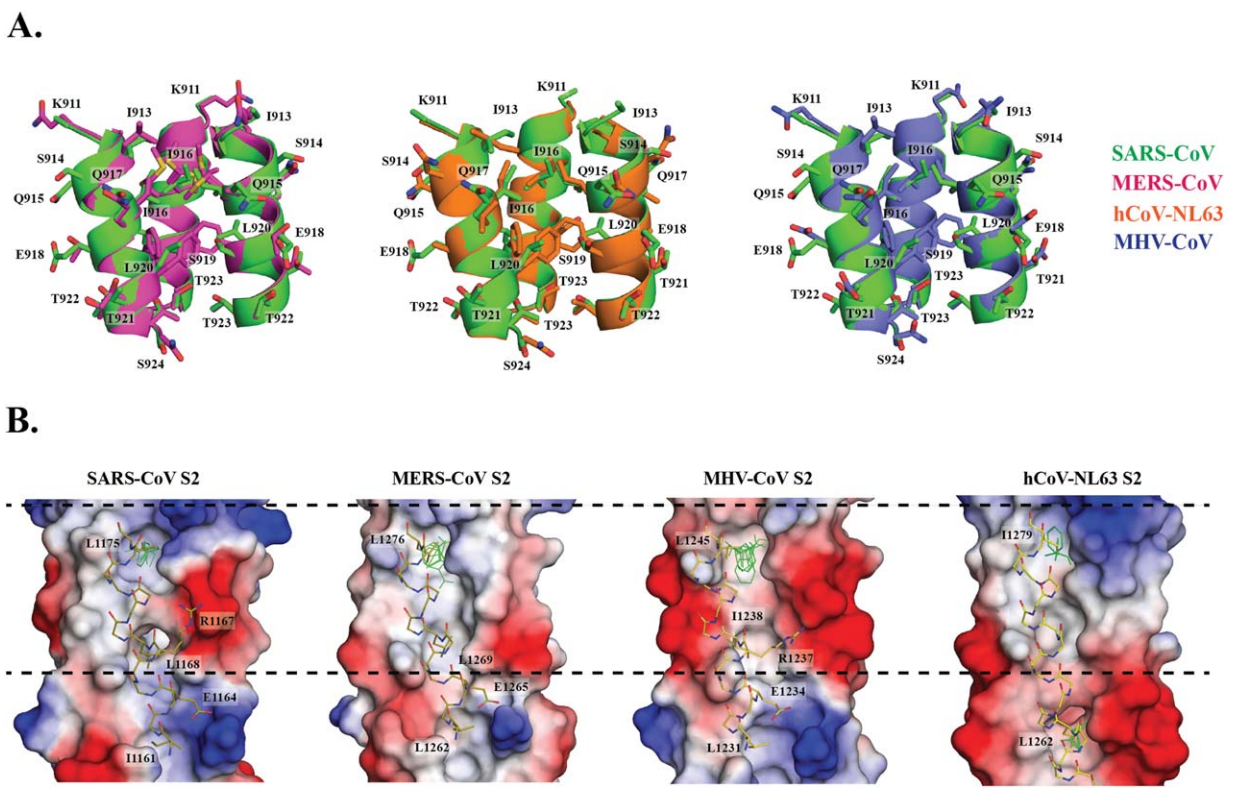

Figure 6. The common HR1 SARS-CoV S2 substructure. **(A)** Ribbon diagram of the structural alignments of the HR1 substructure. The SARS-CoV HR1 substructure (colored in *green*) was superimposed with MHV (colored in *purple*), MERS-CoV (colored in *magenta*), and hCoV-NL63 (colored in *orange*) HR1 substructures (RMSD = ~0.4 Å). Conserved side chains within the HR1 substructure are shown for each virus. SARS-CoV S2 residue numbering was used in both structural alignments. **(B)** Characterization of the groove on the surface of the HR1 inner core. The electrostatic surface potential of the HR1 region at the HR2 interface was depicted for the SARS-CoV S2, MHV S2, MERS-CoV S2, and hCoV-NL63 S2 fusion cores. HR2 helical region residues extending alongside the HR1 inner core are shown as *yellow sticks*. Conserved HR2 residues interacting with the HR1 inner core are labeled accordingly. The structurally conserved HR1 core boundaries are indicated between the *black dashed lines*. Using computational solvent mapping, a hydrophobic pocket was identified, as shown by the green small molecules, on the HR1 substructure.

with the exception of a conservative isoleucine substitution (Ile1279) in the hCoV–NL63 fusion core for the leucine in MHV, MERS-CoV, and SARS-CoV S2 (Fig. 5). The conservation of the 42-residue substructure among coronavirus S2 fusion subunits may represent a common foundation to facilitate viral fusion.

Our discovery of a structurally conserved 42-residue substructure on the HR1 region in SARS-CoV, hCoV-NL63, MERS-CoV, and MHV S2 provides another target for drug development. The conserved HR1 substructure corresponds to the site of interaction for residues Leu1168 and Leu1175 from the HR2 helix [Fig. 6(A)]; these are the key HR2 hydrophobic residues involved in stabilizing the postfusion conformation. More importantly, a large amphipathic cavity with hydrophobic and polar character is present within the HR1 core substructure [Fig. 6(B)] that allows interaction with the critical Leu1175 hydrophobic residue. This HR1 substructure also coincides with the deep hydrophobic grooves previously identified on the HR1 coiled-coil regions of the SARS-CoV and MERS-CoV S2 fusion core crystal structures.[31,70] We rationalize that the conserved core maintains the structural integrity of the viral glycoprotein and acts as a foundation for conformational changes necessary for the fusion of the viral and host cell membranes. Inhibitors designed against this site have the potential to block formation of the fusion core complex of many coronaviruses.

Conclusions

SARS-CoV entry requires binding to specific host cell receptors, and fusion of the viral and host cell membranes in a protease-dependent manner. Biophysical and structural analyses of the S2 protein showed that hydrophobic amino acids occupying the "*d*" position in the C-terminal HR2 region are highly critical for the stability of the six-helix bundle. These residues interact with a conserved substructure within the N-terminal HR1 region and maintain the integrity of the fusion core in the postfusion state. The HR1 substructure is a common feature

among many characterized CoV fusion proteins and provides a tangible target for small molecule and peptide fusion inhibitor design. A general strategy targeting this site could be used to combat disease caused by emerging or re-emerging CoV's.

Materials and Methods

Cloning and site-directed mutagenesis

The gene sequence of SARS-CoV S2 strain BJ302 clone 1 (GenBank accession: AY429072) was codon-optimized and gene synthesized for expression in *E. coli*. The DNA corresponding to a linked two-helix (L2H) SARS-CoV S2 fusion core, containing HR1 residues 896–973 and HR2 residues 1142–1183 connected with a Gly–Gly–Ser–Gly–Gly–Ser linker, was cloned into pET-46 Ek/LIC (EMD Millipore). A QuikChange site-directed mutagenesis-based protocol was used to generate the following single-, double- and triple-site SARS-CoV S2 L2H mutants: Q902A, E918A, K929A, K929E, N937A, N942A, N951A, R965A, R965E, D967A, E970A, L1148A, I1151A, I1161A, E1164A, R1167A, R1167E, L1168A, L1175A, Q902A-N937A, K929A-R965A, K929E-R965E, K929E-R1167E, K929A-R1167A, R965A-R1167A, R965E-R1167E, K929A-R965A-R1167A, and K929E-R965E-R1167E.

Expression and purification of the SARS-CoV S2 linked core

Plasmids containing SARS-CoV S2 L2H WT and mutants were transformed into the *E. coli* SHuffle T7 expression cell line (New England Biolabs). A single colony was inoculated into LB media supplemented with 100 μg/mL ampicillin and grown overnight at 37°C. About 20 mL of the overnight culture was then used to inoculate 1 L of LB media supplemented with 100 μg/mL ampicillin, and grown to OD_{600} of 0.6 at 37°C. Protein expression was induced by the addition of isopropyl β-D−1-thiogalactopyranoside (IPTG) to a final concentration of 0.5 m*M*. The temperature was lowered to 18°C and harvested 18 hours post-induction by centrifugation at 3000*g*. The bacterial cell pellet was resuspended in 25 mL 1× Ni-NTA binding buffer (50 m*M* Tris-HCl pH 7.5, 300 m*M* NaCl, and 20 m*M* imidazole) supplemented with 0.05% (w/v) CHAPS and 1× EDTA-free protease inhibitor cocktail (Bioshop). A hydraulic cell disruption system (Constant Systems) was used to lyse the homogenized cell suspension at 30 kpsi, and the lysate was clarified by centrifugation at 36,500*g* for 45 min at 4°C. The supernatant was then loaded onto a 2 mL Ni-NTA column (Thermo Pierce), and washed with 10 bed volumes of 1× Ni-NTA binding buffer. The SARS-CoV S2 L2H protein was eluted in sequential steps using increasing concentrations of imidazole (125 m*M* imidazole, 250 m*M* imidazole, 375 m*M* imidazole, or 500 m*M* imidazole in 1× Ni-NTA buffer). The eluted SARS-CoV S2 L2H protein was then concentrated and further purified by size exclusion chromatography on a prep grade Superdex-75 10/300 column equilibrated with 10 m*M* K_2HPO_4/KH_2PO_4 pH 7.5, 150 m*M* NaCl, and 0.05% (w/v) CHAPS. Protein concentration was quantified by absorbance at 280 nm, and protein purity was analyzed by SDS-PAGE and electrospray mass spectrometry.

Analytical ultracentrifugation (AUC)

Sedimentation equilibrium experiments were performed in a Beckman Coulter Optima XL-A analytical ultracentrifuge equipped with an An-60 Ti rotor (Beckman-Coulter, Palo Alto, CA) at the Analytical Ultracentrifugation Facility in the Department of Biochemistry at the University of Toronto. The experiments were carried out at 4°C using purified SARS-CoV S2 L2H fusion core in 10 m*M* Tris-HCl, 150 m*M* NaCl pH 7.5. About 20 μ*M* protein sample was centrifuged at three different speeds (18,000, 20,000, and 22,000 rpm) and migration of SARS-CoV S2 L2H was monitored by absorbance at 230 and 280 nm over 60 hours. Data analysis was performed using the Origin MicroCal XL-A/CL-I Data Analysis Software Package Version 4.0.

CD spectroscopy

Purified SARS-CoV S2 L2H proteins were characterized by CD spectroscopy on a Jasco J-810 spectropolarimeter using 1 mm quartz cuvettes (Helma).

Wavelength scans. Wavelength spectra were recorded from 190 to 250 nm for wild-type and mutant SARS-CoV S2 L2H in 10 m*M* K_2HPO_4/KH_2PO_4 pH 7.5, 150 m*M* NaCl, and 0.05% (w/v) CHAPS buffer at 20°C to determine overall protein secondary structural changes due to the mutation. Five spectra were acquired and averaged, with the results reported as molar ellipticity [θ] (units of deg cm^2 $dmol^{-1}$). The α-helical content of wild-type SARS-CoV S2 L2H was calculated from the experimental CD wavelength scans using the SELCON3 algorithm in the program DichroWeb.[71]

Thermal denaturation scans. The relative thermal stabilities of SARS-CoV S2 L2H mutants were performed by heating the sample from 20°C to 99°C at 0.2°C intervals and monitoring the loss of CD signal at 222 nm. Heating alone was insufficient to denature the predominantly helical SARS-CoV S2 L2H fusion core. Similar to previous studies,[72] 4*M* guanidine hydrochloride was added to all samples to facilitate unfolding within a temperature range of 20°C–99°C. Ellipticity readings for thermal denaturation data were baseline corrected, normalized between 0 (folded) and 1 (unfolded) and fit to a nonlinear biphasic sigmoidal curve in Graphpad. Values

of midpoint unfolding transitions (T_m) were calculated from thermal melt curves and they correspond to the temperature where 50% of the protein has unfolded.

pH scans. In order to study the effects of pH on the stability of the SARS-CoV S2 L2H fusion core, the purified proteins were buffer exchanged using an Amicon Ultra-0.5 centrifugal concentrator (10 kDa molecular weight cut off) into the following buffer conditions: pH 4.0–6.5, sodium acetate (NaOAc) buffer (10 m*M* NaOAc, 150 m*M* NaCl, and 0.05% (w/v) CHAPS); pH 7.0–8.5, Tris-HCl buffer (10 m*M* Tris-HCl, 150 m*M* NaCl, and 0.05% (w/v) CHAPS). CD wavelength scans and thermal melts as a function of pH were performed as described above. The pH may exhibit fluctuations within the buffering range during the CD thermal melt experiments due to temperature-dependent phenomenon of the buffer.

Structural comparison and analysis

Coronavirus fusion protein structures used for structural comparisons were obtained from the Protein Data Bank (1WNC, 1WYY, 1ZV7, 1ZV8, 1ZVA, 1ZVB, 2BEQ, and 2BEZ for SARS-CoV S2; 4MOD for MERS-CoV S2; 1WDF and 1WDG for MHV S2; and 2IEQ for hCoV-NL63 S2). Superimpositions of three-dimensional structures and root mean squared deviation (RMSD) calculations were obtained using the pair-wise alignment of the CLICK server.[73] Identification of pockets on the surface of the SARS-CoV conserved substructure was performed by computational solvent mapping using the FTMap server.[74]

Acknowledgments

The authors would like to thank the Molecular Structure & Function program at the Hospital for Sick Children Research Institute (Toronto, ON) for access to the circular dichroism spectrometer, Drs. Walid Houry and Yoshito Kakihara for assistance with the sedimentation equilibrium ultracentrifugation experiments, and Dr. Karen Siu for critical reading of the manuscript.

References

1. Perlman S, Netland J (2009) Coronaviruses post-SARS: update on replication and pathogenesis. Nat Rev Microbiol 7:439–450.
2. Ksiazek TG, Erdman D, Goldsmith CS, Zaki SR, Peret T, Emery S, Tong S, Urbani C, Comer JA, Lim W, Rollin PE, Dowell SF, Ling A, Humphrey CD, Shieh W, Guarner J, Paddock CD, Rota P, Fields B, DeRisi J, Yang J, Cox N, Hughes JM, LeDuc JW, Bellini WJ, Anderson LJ, and the SARS Working Group (2003) A novel coronavirus associated with severe acute respiratory syndrome. N Engl J Med 348:1953–1966.
3. Rota PA, Oberste MS, Monroe SS, Nix WA, Campagnoli R, Icenogle JP, Peñaranda S, Bankamp B, Maher K, Chen M, Tong S, Tamin A, Lowe L, Frace M, DeRisi JL, Chen Q, Wang D, Erdman DD, Peret TCT, Burns C, Ksiazek TG, Rollin PE, Sanchez A, Liffick S, Holloway B, Limor J, McCaustland K, Olsen-Rasmussen M, Fouchier R, Günther S, Osterhaus ADME, Drosten C, Pallansch MA, Anderson LJ, Bellini WJ (2003) Characterization of a novel coronavirus associated with severe acute respiratory syndrome. Science 300:1394–1399.
4. Drosten C, Günther S, Preiser W, van der Werf S, Brodt H, Becker S, Rabenau H, Panning M, Kolesnikova L, Fouchier RAM, Berger A, Burguière A, Cinatl J, Eickmann M, Escriou N, Grywna K, Kramme S, Manuguerra J, Müller S, Rickerts V, Stürmer M, Vieth S, Klenk H, Osterhaus ADME, Schmitz H, Doerr HW (2003) Identification of a novel coronavirus in patients with severe acute respiratory syndrome. N Engl J Med 348:1967–1976.
5. Gerlier D (2011) Emerging zoonotic viruses: new lessons on receptor and entry mechanisms. Curr Opin Virol 1:27–34.
6. Bolles M, Donaldson E, Baric R (2011) SARS-CoV and emergent coronaviruses: viral determinants of interspecies transmission. Curr Opin Virol 1:624–634.
7. Zaki AM, van Boheemen S, Bestebroer TM, Osterhaus ADME, Fouchier RAM (2012) Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. N Engl J Med 367:1814–1820.
8. Centers for Disease Control and Prevention (2013) CDC - Coronavirus - Middle East Respiratory Syndrome - MERS-CoV. http://www.cdc.gov/coronavirus/mers/.
9. Tong S, Conrardy C, Ruone S, Kuzmin IV, Guo X, Tao Y, Niezgoda M, Haynes L, Agwanda B, Breiman RF, Anderson LJ, Rupprecht CE (2009) Detection of novel SARS-like and other coronaviruses in bats from Kenya. Emerg Infect Dis 15:482–485.
10. Ge Y, Li J, Yang X, Chmura AA, Zhu G, Epstein JH, Mazet JK, Hu B, Zhang W, Peng C, Zhang Y, Luo C, Tan B, Wang N, Zhu Y, Crameri G, Zhang S, Wang L, Daszak P, Shi Z (2013) Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. Nature 503:535–538.
11. Heald-Sargent T, Gallagher T (2012) Ready, set, fuse! The coronavirus spike protein and acquisition of fusion competence. Viruses 4:557–580.
12. Du L, He Y, Zhou Y, Liu S, Zheng B-J, Jiang S (2009) The spike protein of SARS-CoV — a target for vaccine and therapeutic development. Nat Rev Microbiol 7: 226–236.
13. Hofmann H, Pöhlmann S (2004) Cellular entry of the SARS coronavirus. Trends Microbiol 12:466–472.
14. Bartlam M, Yang H, Rao Z (2005) Structural insights into SARS coronavirus proteins. Curr Opin Struct Biol 15:664–672.
15. Harrison SC (2008) Viral membrane fusion. Nat Struct Mol Biol 15:690–698.
16. White JM, Delos SE, Brecher M, Schornberg K (2008) Structures and mechanisms of viral membrane fusion proteins. Crit Rev Biochem Mol Biol 43:189–219.
17. Mothes W, Boerger AL, Narayan S, Cunningham JM, Young JAT (2000) Retroviral entry mediated by receptor priming and low pH triggering of an envelope glycoprotein. Cell 103:679–689.
18. Matsuyama S, Taguchi F (2009) Two-step conformational changes in a coronavirus envelope glycoprotein mediated by receptor binding and proteolysis. J Virol 83:11133–11141.
19. Bale S, Liu T, Li S, Wang Y, Abelson D, Fusco M, Woods VL, Ollmann Saphire E (2011) Ebola virus

glycoprotein needs an additional trigger, beyond proteolytic priming for membrane fusion. PLoS Negl Trop Dis 5:e1395.
20. Brecher M, Schornberg KL, Delos SE, Fusco ML, Saphire EO, White JM (2012) Cathepsin cleavage potentiates the Ebola virus glycoprotein to undergo a subsequent fusion-relevant conformational change. J Virol 86:364–372.
21. De Haan CAM, Stadler K, Godeke G-J, Bosch BJ, Rottier PJM (2004) Cleavage inhibition of the murine voronavirus spike protein by a furin-like enzyme affects cell-cell but not virus-cell rusion. J Virol 78: 6048–6054.
22. Belouzard S, Chu VC, Whittaker GR (2009) Activation of the SARS coronavirus spike protein via sequential proteolytic cleavage at two distinct sites. Proc Natl Acad Sci USA 106:5871–5876.
23. Matsuyama S, Ujike M, Morikawa S, Tashiro M, Taguchi F (2005) Protease-mediated enhancement of severe acute respiratory syndrome coronavirus infection. Proc Natl Acad Sci USA 102:12543–12547.
24. Simmons G, Reeves JD, Rennekamp AJ, Amberg SM, Piefer AJ, Bates P (2004) Characterization of severe acute respiratory syndrome-associated coronavirus (SARS-CoV) spike glycoprotein-mediated viral entry. Proc Natl Acad Sci USA 101:4240–4245.
25. Simmons G, Gosalia DN, Rennekamp AJ, Reeves JD, Diamond SL, Bates P (2005) Inhibitors of cathepsin L prevent severe acute respiratory syndrome coronavirus entry. Proc Natl Acad Sci USA 102:11876–11881.
26. Du L, Kao RY, Zhou Y, He Y, Zhao G, Wong C, Jiang S, Yuen K-Y, Jin D-Y, Zheng B-J (2007) Cleavage of spike protein of SARS coronavirus by protease factor Xa is associated with viral infectivity. Biochem Biophys Res Commun 359:174–179.
27. Bosch BJ, Bartelink W, Rottier PJM (2008) Cathepsin L functionally cleaves the severe acute respiratory syndrome coronavirus class I fusion protein upstream of rather than adjacent to the fusion peptide. J Virol 82: 8887–8890.
28. Bosch BJ, Martina BEE, Zee R van der, Lepault J, Haijema BJ, Versluis C, Heck AJR, Groot R de , Osterhaus ADME, Rottier PJM (2004) Severe acute respiratory syndrome coronavirus (SARS-CoV) infection inhibition using spike protein heptad repeat-derived peptides. Proc Natl Acad Sci USA 101:8455–8460.
29. Supekar VM, Bruckmann C, Ingallinella P, Bianchi E, Pessi A, Carfí A (2004) Structure of a proteolytically resistant core from the severe acute respiratory syndrome coronavirus S2 fusion protein. Proc Natl Acad Sci 101:17958–17963.
30. Duquerroy S, Vigouroux A, Rottier PJM, Rey FA, Bosch BJ (2005) Central ions and lateral asparagine/glutamine zippers stabilize the post-fusion hairpin conformation of the SARS coronavirus spike glycoprotein. Virology 335:276–285.
31. Xu Y, Lou Z, Liu Y, Pang H, Tien P, Gao GF, Rao Z (2004) Crystal structure of severe acute respiratory syndrome coronavirus spike protein fusion core. J Biol Chem 279:49414–49419.
32. Xu Y, Zhu J, Liu Y, Lou Z, Yuan F, Liu Y, Cole DK, Ni L, Su N, Qin L, Li X, Bai Z, Bell JI, Pang H, Tien P, Gao GF, Rao Z (2004) Characterization of the heptad repeat regions, HR1 and HR2, and design of a fusion core structure model of the spike protein from severe acute respiratory syndrome (SARS) coronavirus. Biochemistry 43:14064–14071.
33. Nomura R, Kiyota A, Suzaki E, Kataoka K, Ohe Y, Miyamoto K, Senda T, Fujimoto T (2004) Human coronavirus 229E binds to CD13 in rafts and enters the cell through Caveolae. J Virol 78:8701–8708.
34. Nash TC, Buchmeier MJ (1997) Entry of mouse hepatitis virus into cells by endosomal and nonendosomal pathways. Virology 233:1–8.
35. Zhu J, Xiao G, Xu Y, Yuan F, Zheng C, Liu Y, Yan H, Cole DK, Bell JI, Rao Z, Tien P, Gao GF (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 Commun 319: 283–288.
36. Bosshard HR, Marti DN, Jelesarov I (2004) Protein stabilization by salt bridges: concepts, experimental approaches and clarification of some misunderstandings. J Mol Recognit 17:1–16.
37. Fass D, Harrison SC, Kim PS (1996) Retrovirus envelope domain at 1.7 Å resolution. Nat Struct Mol Biol 3: 465–469.
38. Kobe B, Center RJ, Kemp BE, Poumbourios P (1999) Crystal structure of human T cell leukemia virus type 1 gp21 ectodomain crystallized as a maltose-binding protein chimera reveals structural evolution of retroviral transmembrane proteins. Proc Natl Acad Sci USA 96:4319–4324.
39. Weissenhorn W, Carfí A, Lee K-H, Skehel JJ, Wiley DC (1998) Crystal structure of the Ebola virus membrane fusion subunit, GP2, from the envelope glycoprotein ectodomain. Mol Cell 2:605–616.
40. Harrison JS, Higgins CD, Chandran K, Lai JR (2011) Designed protein mimics of the Ebola virus glycoprotein GP2 α-helical bundle: stability and pH effects. Protein Sci 20:1587–1596.
41. Aydin H, Smrke BM, Lee JE (2013) Structural characterization of a fusion glycoprotein from a retrovirus that undergoes a hybrid 2-step entry mechanism. FASEB J 27:5059–5071.
42. Yang Z-Y, Huang Y, Ganesh L, Leung K, Kong W-P, Schwartz O, Subbarao K, Nabel GJ (2004) pH-dependent entry of severe acute respiratory syndrome coronavirus is mediated by the spike glycoprotein and enhanced by dendritic cell transfer through DC-SIGN. J Virol 78:5642–5650.
43. Ujike M, Nishikawa H, Otaka A, Yamamoto N, Yamamoto N, Matsuoka M, Kodama E, Fujii N, Taguchi F (2008) Heptad repeat-derived peptides block protease-mediated direct entry from the cell surface of severe acute respiratory syndrome coronavirus but not entry via the endosomal pathway. J Virol 82:588–592.
44. McClure MO, Sommerfelt MA, Marsh M, Weiss RA (1990) The pH independence of mammalian retrovirus infection. J Gen Virol 71:767–773.
45. Côté M, Zheng Y-M, Liu S-L (2012) Membrane fusion and cell entry of XMRV are pH-independent and modulated by the envelope glycoprotein's cytoplasmic tail. PLoS ONE 7:e33734.
46. Aydin H, Cook JD, Lee JE (2014) Crystal structures of beta- and gamma retrovirus fusion proteins reveal a role for electrostatic stapling in viral entry. J Virol 88: 143–153.
47. Maerz AL, Center RJ, Kemp BE, Kobe B, Poumbourios P (2000) Functional implications of the human T-lymphotropic virus type 1 transmembrane glycoprotein helical hairpin structure. J Virol 74:6614–6621.
48. Igonet S, Vaney M-C, Vonrhein C, Bricogne G, Stura EA, Hengartner H, Eschli B, Rey FA (2011) X-ray structure of the arenavirus glycoprotein GP2 in its postfusion hairpin conformation. Proc Natl Acad Sci USA 108:19967–19972.

49. Bullough PA, Hughson FM, Skehel JJ, Wiley DC (1994) Structure of influenza haemagglutinin at the pH of membrane fusion. Nature 371:37–43.
50. Li W, Moore MJ, Vasilieva N, Sui J, Wong SK, Berne MA, Somasundaran M, Sullivan JL, Luzuriaga K, Greenough TC, Choe H, Farzan M (2003) Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. Nature 426:450–454.
51. Hofmann H, Pyrc K, Hoek L van der, Geier M, Berkhout B, Pöhlmann S (2005) Human coronavirus NL63 employs the severe acute respiratory syndrome coronavirus receptor for cellular entry. Proc Natl Acad Sci USA 102:7988–7993.
52. Williams RK, Jiang GS, Holmes KV (1991) Receptor for mouse hepatitis virus is a member of the carcinoembryonic antigen family of glycoproteins. Proc Natl Acad Sci USA 88:5533–5536.
53. Yeager CL, Ashmun RA, Williams RK, Cardellichio CB, Shapiro LH, Look AT, Holmes KV (1992) Human aminopeptidase N is a receptor for human coronavirus 229E. Nature 357:420–422.
54. Raj VS, Mou H, Smits SL, Dekkers DHW, Müller MA, Dijkman R, Muth D, Demmers JAA, Zaki A, Fouchier RAM, Thiel V, Drosten C, Rottier PJM, Osterhaus ADME, Bosch BJ, Haagmans BL (2013) Dipeptidyl peptidase 4 is a functional receptor for the emerging human coronavirus-EMC. Nature 495:251–254.
55. 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 GF (2013) Molecular basis of binding between novel human coronavirus MERS-CoV and its receptor CD26. Nature 500:227–231.
56. Du L, Zhao G, Kou Z, Ma C, Sun S, Poon VKM, Lu L, Wang L, Debnath AK, Zheng B, Zhou Y, Jiang S (2013) Identification of a receptor-binding domain in the S protein of the novel human coronavirus Middle East respiratory syndrome coronavirus as an essential target for vaccine development. J Virol 87:9939–9942.
57. Wang N, Shi X, Jiang L, Zhang S, Wang D, Tong P, Guo D, Fu L, Cui Y, Liu X, Arledge KC, Chen Y, Zhang L, Wang X (2013) Structure of MERS-CoV spike receptor-binding domain complexed with human receptor DPP4. Cell Res 23:986–993.
58. Jiang S, Lin K, Strick N, Neurath AR (1993) HIV-1 inhibition by a peptide. Nature 365:113–113.
59. Wild CT, Shugars DC, Greenwell TK, McDanal CB, Matthews TJ (1994) Peptides corresponding to a predictive alpha-helical domain of human immunodeficiency virus type 1 gp41 are potent inhibitors of virus infection. Proc Natl Acad Sci USA 91:9770–9774.
60. Kilby JM, Hopkins S, Venetta TM, DiMassimo B, Cloud GA, Lee JY, Alldredge L, Hunter E, Lambert D, Bolognesi D, Matthews T, Johnson MR, Nowak MA, Shaw GM, Saag MS (1998) Potent suppression of HIV-1 replication in humans by T-20, a peptide inhibitor of gp41-mediated virus entry. Nat Med 4:1302–1307.
61. Bosch BJ, Martina BEE, Zee R van der, Lepault J, Haijema BJ, Versluis C, Heck AJR, Groot R de , Osterhaus ADME, Rottier PJM (2004) Severe acute respiratory syndrome coronavirus (SARS-CoV) infection inhibition using spike protein heptad repeat-derived peptides. Proc Natl Acad Sci USA 101:8455–8460.
62. Ingallinella P, Bianchi E, Finotto M, Cantoni G, Eckert DM, Supekar VM, Bruckmann C, Carfi A, Pessi A (2004) Structural characterization of the fusion-active complex of severe acute respiratory syndrome (SARS) coronavirus. Proc Natl Acad Sci USA 101:8709–8714.
63. Yan Z, Tripet B, Hodges RS (2006) Biophysical characterization of HRC peptide analogs interaction with heptad repeat regions of the SARS-coronavirus Spike fusion protein core. J Struct Biol 155:162–175.
64. Chu L-HM, Chan S-H, Tsai S-N, Wang Y, Cheng CH-K, Wong K-B, Waye MM-Y, Ngai S-M (2008) Fusion core structure of the severe acute respiratory syndrome coronavirus (SARS-CoV): in search of potent SARS-CoV entry inhibitors. J Cell Biochem 104:2335–2347.
65. Liu I-J, Kao C-L, Hsieh S-C, Wey M-T, Kan L-S, Wang W-K (2009) Identification of a minimal peptide derived from heptad repeat (HR) 2 of spike protein of SARS-CoV and combination of HR1-derived peptides as fusion inhibitors. Antiviral Res 81:82–87.
66. Yuan K, Yi L, Chen J, Qu X, Qing T, Rao X, Jiang P, Hu J, Xiong Z, Nie Y, Shi X, Wang W, Ling C, Yin X, Fan K, Lai L, Ding M, Deng H (2004) Suppression of SARS-CoV entry by peptides corresponding to heptad regions on spike glycoprotein. Biochem Biophys Res Commun 319:746–752.
67. Liu S, Xiao G, Chen Y, He Y, Niu J, Escalante CR, Xiong H, Farmar J, Debnath AK, Tien P, Jiang S (2004) Interaction between heptad repeat 1 and 2 regions in spike protein of SARS-associated coronavirus: implications for virus fusogenic mechanism and identification of fusion inhibitors. Lancet 363:938–947.
68. Ni L, Zhu J, Zhang J, Yan M, Gao GF, Tien P (2005) Design of recombinant protein-based SARS-CoV entry inhibitors targeting the heptad-repeat regions of the spike protein S2 domain. Biochem Biophys Res Commun 330:39–45.
69. Lu L, Liu Q, Zhu Y, Chan K, Qin L, Li Y, Wang Q, Chan JF, Du L, Yu F, Ma C, Ye S, Yuen K, Zhang R, Jiang S (2014) Structure-based discovery of Middle East respiratory syndrome coronavirus fusion inhibitor. Nat Commun 5, article number 3067.
70. 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 GF (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.
71. Whitmore L, Wallace BA (2004) DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. Nucleic Acids Res 32:W668–W673.
72. Zheng Q, Deng Y, Liu J, van der Hoek L, Berkhout B, Lu M (2006) Core structure of S2 from the human coronavirus NL63 spike glycoprotein. Biochemistry 45: 15205–15215.
73. Nguyen MN, Tan KP, Madhusudhan MS (2011) CLICK—topology-independent comparison of biomolecular 3D structures. Nucleic Acids Res 39:W24–W28.
74. Brenke R, Kozakov D, Chuang G-Y, Beglov D, Hall D, Landon MR, Mattos C, Vajda S (2009) Fragment-based identification of druggable "hot spots" of proteins using Fourier domain correlation techniques. Bioinformatics 25:621–627.