4

Diversity of Coronavirus Spikes: Relationship to Pathogen Entry and Dissemination

Edward B. Thorp and Thomas M. Gallagher

Coronaviruses are widespread in the environment, infecting humans, domesticated and wild mammals, and birds. Infections cause a variety of diseases including bronchitis, gastroenteritis, hepatitis, and encephalitis, with symptoms ranging from being nearly undetectable to rapidly fatal. A combination of interacting variables determine the pattern and severity of coronavirus-induced disease, including the infecting virus strain, its transmission strategy, and the age and immune status of the infected host. Coronavirus pathogenesis is best understood by discerning how each of these variables dictates clinical outcomes. This chapter focuses on variabilities amongst the spike (S) proteins of infecting virus strains. Diversity of coronavirus surface proteins likely contributes to epidemic disease, an important and timely topic given the recent emergence of the human SARS coronavirus.

1. Introduction

Coronaviruses circulating in nature exhibit considerable genetic variability, and ongoing virus evolution can generate novel variants capable of epidemic diseases such as the recent SARS-CoV. A central goal in coronavirus research is to pinpoint virus strain variations and relate the differences to epidemiologic and pathogenic potentials. Identifying variations correlating with properties such as viral transmission from animals to humans leads toward mechanistic understanding of epidemics and also points to the relevant targets for antiviral therapeutics. For the coronaviruses, it is clear that pronounced variations are accommodated in the spike (S) genes, which encode the "corona" of virion protrusions (Figure 4.1). Each protrusion is a complex, oligomeric assembly of extremely large. ~1,300 amino acid S protein monomers that are integrated into virion membranes by their C-terminal transmembrane anchors. The spikes are essential for virus binding to cell-surface receptors and for virus–cell membrane fusion. During infection, the spikes also accumulate on cell surfaces,

Edward B. Thorp and Thomas M. Gallagher • Department of Microbiology and Immunology, Loyola University Medical Center, Maywood, IL 60153, Illinois.
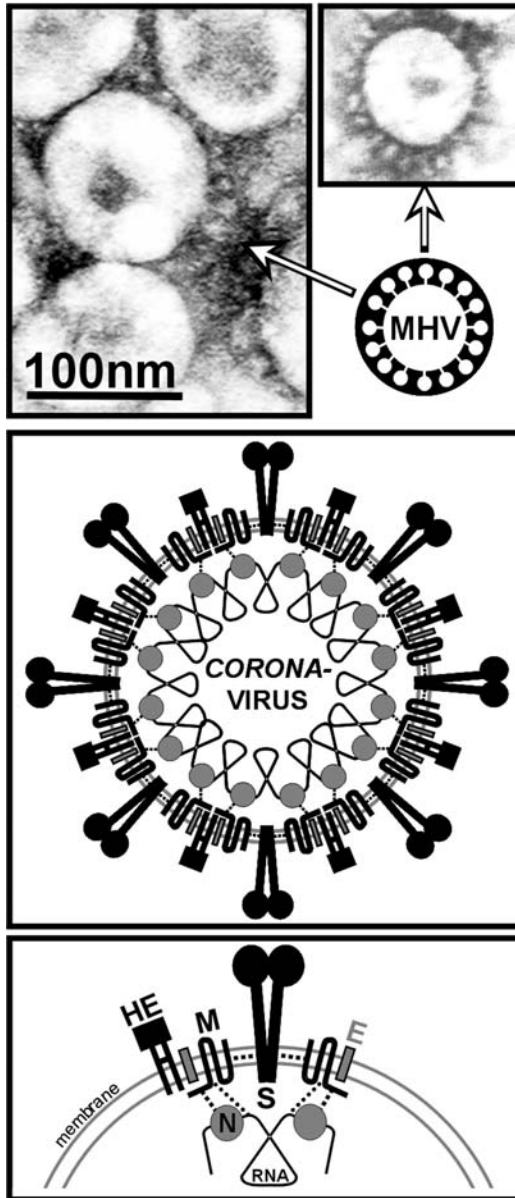
Figure 4.1.   Electron micrographs and model of coronavirus structure. Top: Negatively stained (uranyl acetate) micrographs of murine coronavirus *m*ouse *h*epatitis *v*irus (strain A59). Bottom boxes: Schematic of coronavirus particles. *S*: spike glycoprotein, *HE*: hemagglutinin-esterase glycoprotein, *M*: triple-membrane-spanning membrane glycoprotein, *E*: small envelope glycoprotein, *N*: nucleocapsid phosphoprotein. Dotted lines indicate noncovalent protein–protein or protein–RNA interactions.
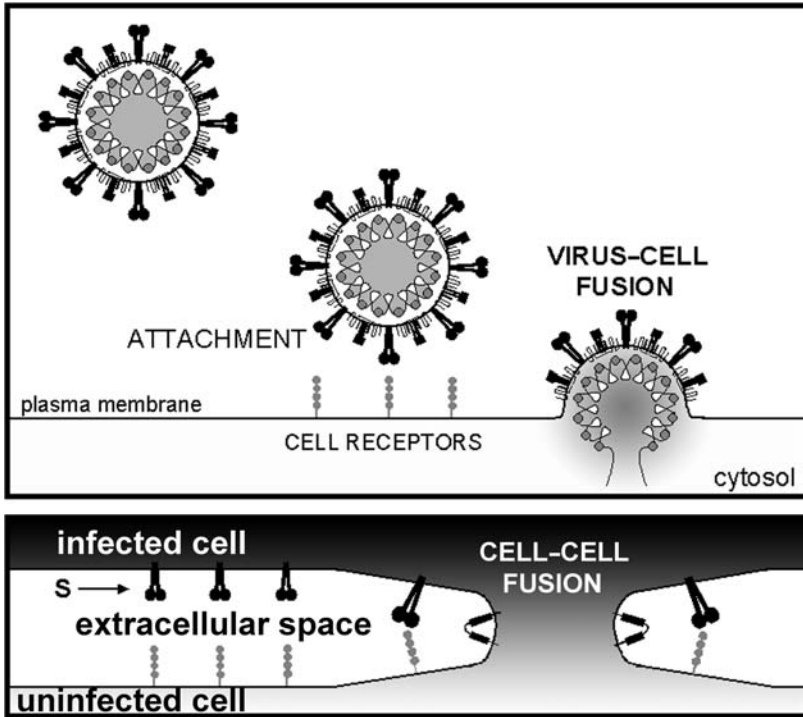
Figure 4.2. Virus entry and dissemination. Top: Depiction of virion attachment to cellular receptors and subsequent fusion of viral envelope with the host cell plasma membrane. Bottom: Depiction of spikes (S) on the infected cell surface recognizing opposing cellular receptors and promoting intercellular fusion and cell–cell spread.

bind to receptors on neighboring uninfected cells, and mediate cell–cell membrane fusion, a process that creates syncytia and causes rapid expansion of infections (Figure 4.2). Thus, the unique characteristics of S proteins from different coronavirus isolates correlate with distinct patterns of virus entry, virus dissemination via syncytia, virus tropism, and pathogenesis.

2. S Functions During Coronavirus Entry

To appreciate the unique characteristics of these S proteins, one must visualize their activities in the context of the infection cycle. We begin with virion binding to susceptible host cells. As mediators of virus attachment to cells, S proteins are set apart by their ability to evolve remarkably varied attachment specificities. Sialic acid, a ubiquitous component of cell-surface carbohydrate complexes, is a documented low-affinity ligand for porcine and bovine coronavirus spikes (Schultze *et al.*, 1991). Aminopeptidase N (APN), a type II-oriented membrane glycoprotein found in abundance on respiratory epithelia, is a receptor for antigenic "group 1" coronaviruses (Delmas *et al.*, 1992; Tresnan *et al.*, 1996). Members of this antigenic cluster include human respiratory viruses such as human CoV 229E, as well as

several devastating animal pathogens such as transmissible gastroenteritis virus of swine and infectious peritonitis virus of cats. *Carcino*Embryonic *A*ntigen-related *C*ell *A*dhesion *M*olecules (CEACAMs), immunoglobulin-like type I-oriented membrane glycoproteins that are prevalent in the liver and gastrointestinal tract, serve as receptors for the prototype member of the antigenic "group 2" coronavirus mouse hepatitis virus (Dveksler *et al.*, 1991; Godfraind *et al.*, 1995). Receptors for group 3 coronaviruses, which include several bird viruses causing severe bronchitis in chickens and turkeys, are currently unknown.

High-resolution structures are predicted for APN (Sjostrom *et al.*, 2000; Firla *et al.*, 2002), and are actually known for CEACAM (Tan *et al.*, 2002). Structural homologies between these two proteins are not readily apparent. Thus, the adaptation of coronaviruses to either receptor likely involves substantial remodeling of binding sites on S proteins. In this regard, it is important to remember that Apn or Ceacam receptor usage correlates with the antigenic and genetic relationships used to divide coronaviruses into groups (Siddell, 1995). Therefore, one can reasonably infer that S variations adapt viruses to particular receptor usage, that receptor usage dictates the ecological niche of infection, and that coronaviruses in distinct niches then evolve somewhat independently to create recognizable antigenic/ phylogenetic groups. Suggestions that the SARS-CoV constitutes the first member of a fourth coronavirus group (Marra *et al.*, 2003; Rota *et al.*, 2003) may imply that this pathogen has adapted some time ago to bind a novel receptor set apart from either Apn or CEACAM.

High-resolution crystallographic structures for coronavirus S proteins are not yet known. Therefore, one can only speculate about the detailed architecture of their receptor-binding sites. The S proteins are moderately amenable to protein dissection techniques in which expressed fragments are assayed for receptor-binding potential, and these studies have roughly localized the sites of receptor interaction on primary sequences ((Suzuki and Taguchi, 1996; Bonavia *et al.*, 2003), see Figure 4.3). Current hypotheses suggest that, as coronaviruses diverge into types with particular receptor specificities, amino acid changes are fixed into S proteins at putative receptor-binding sites (Baric *et al.*, 1999). This may be the case; however, S protein variabilities are relatively complex, and while many strain differences cluster in amino-terminal regions where receptors are thought to bind (Matsuyama and Taguchi, 2002a; see Figure 4.3), several changes are also found outside of this area. This complex variability can be appreciated by recalling the multifunctional properties of the S proteins, which contain receptor-binding sites as well as the machinery necessary to fuse opposing membranes (Figure 4.2). For S proteins, this membrane fusion activity is not constitutive, but is (with few exceptions) manifest only after receptor binding. In part, complex variability in S proteins may reflect the fact that receptor-binding and membrane fusion processes are coupled during virus entry. Put another way, the S-receptor interaction releases energy that is then used to create the conformational changes leading to S-induced membrane fusions. This coupling of receptor binding with membrane fusion activity suggests that subtle strain-specific polymorphisms virtually anywhere in the large S proteins might affect either or both of these essential functions, and by doing so, alter the course of coronavirus entry into cells.

The mechanism by which coronavirus S proteins mediate membrane fusion, recently clarified in studies by Bosch *et al.* (2003), involves a process in which the proteins respond to target cell receptor binding by undergoing conformational change (Gallagher, 1997; Matsuyama and Taguchi, 2002; Zelus *et al.*, 2003) (Figure 4.4A). Next, a currently unidentified hydrophobic portion of the protein termed the fusion peptide (FP) harpoons target cell membranes (Figure 4.4B). This is followed by irreversible conformational changes in which

alpha-helical portions of the protein condense into helical bundles (Figure 4.4C), ultimately bringing opposing membranes into sufficient proximity to coalesce them together (Figure 4.4D). This is a well-documented mechanism by which several viral and cellular proteins catalyze membrane coalescence (Weissenhorn *et al.*, 1999; Russell *et al.*, 2001; Jahn, *et al.*, 2003) and is now classified as a "class-1" type fusion reaction. Algorithms predicting
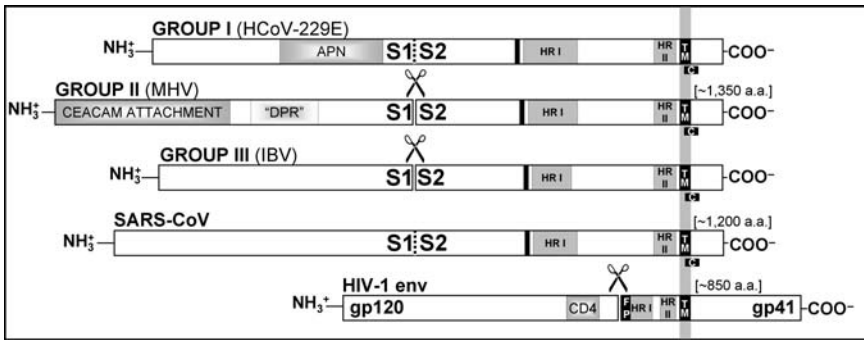
Figure 4.3. Linear depictions of coronavirus spike glycoproteins. Spikes representing groups I–III and the SARS coronavirus are depicted. Scissors indicate proteolytic cleavage between S1 and S2 of spikes. *H*eptad *r*epeat regions (HR) as indicated by *Learn Coil-VMF* and *MultiCoil* are indicated in S2 and are upstream of transmembrane (TM) spans and conserved cysteine-rich stretches (C). Group I spikes recognize aminopeptidase N (CD13) metalloprotease. Group II spikes bind to *c*arcino*e*mbryonic *a*ntigen-related *c*ell *a*dhesion *m*olecule (CEACAM) receptors. Downstream of the CEACAM-binding domain for group II MHV lies a deletion prone region (DPR). Amino acid (a.a.) lengths are drawn approximately to scale and relative to the size of the HIV type I envelope (env) fusion glycoprotein. FP is the hydrophobic fusion peptide for HIV-I. Hydrophobic residues are present N-terminal to HR1 regions in coronaviruses.
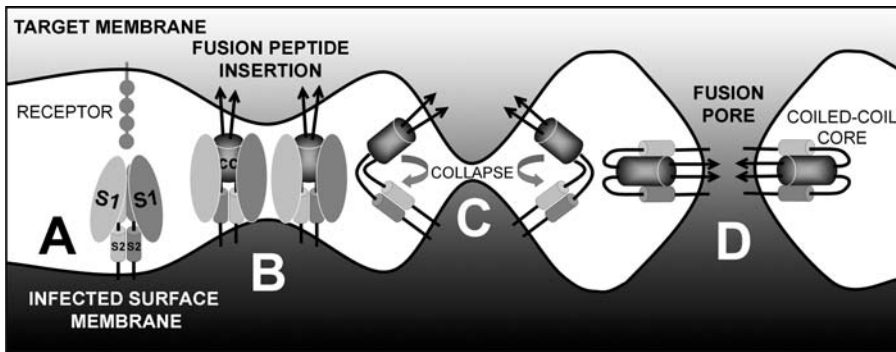
Figure 4.4. Proposed mechanism for S-mediated membrane fusion based on models of class 1-driven viral fusion. (A) Attachment of oligomeric spike S1 domains to receptor. Cylinders represent alpha-helical secondary structure. (B) Arrows depict exposure and insertion of hydrophobic fusion peptides into target membrane, subsequent to displacement of S1. CC indicates formation of alpha-helical *c*oiled *c*oil. (C) Fold-back or collapse of S2 leading to membrane coalescence. S1 domains have been removed for clarity and may in fact be absent as S1 sheds from S2 during fusion activation (see text). (D) Formation of end-stage coiled-coil bundle, fusion pore, and subsequent expansion of pore.

secondary protein structure (Singh *et al.*, 1999) suggest regions of alpha helicity (designated heptad repeat, or HR 1 and 2, see Figure 4.3) in all coronavirus S proteins, including SARS-CoV. Thus, it is generally agreed that the core fusion machinery for all coronaviruses is built in such a way as to catalyze a conserved class-1-type fusion reaction.

It is notable that the fusion module (FP, HR1, HR2, TM span) occupies only about 20% of the inordinately large coronavirus S proteins, and the functional relevance of all but a portion of the remainder is largely unknown. Given that receptor-binding sites are distant from membrane fusion machinery in the primary structures (Figure 4.3), a sensible speculation is that much of the S protein structure is involved in linking receptor binding to the activation of membrane fusion. This is, after all, a crucial coupling that controls the timing and location of virus entry; that is, viral S proteins undergo conformational changes and proceed irreversibly through the class-1-type membrane fusion reaction only when engaged by cellular receptors embedded into the target cell membrane. In considering the activation mechanism, presently available genetic data point toward noncovalent linkages between receptor-binding regions and the fusion machinery (Grosse and Siddell, 1994; Matsuyama and Taguchi, 2002). In the best-studied MHV system, CEACAM binding does cause N-terminal S regions to separate from C-terminal, integral-membrane fragments (Gallagher, 1997), in all likelihood revealing the fusion apparatus (Matsuyama and Taguchi, 2002). This is a process that is augmented by cellular protease(s) that cleave the MHV S proteins at a site between receptor-binding and fusion-inducing domains ((Stauber *et al.*, 1993; Bos *et al.*, 1995; see also Figure 4.3). Proteolytic cleavage likely increases overall S protein conformational flexibility and eases the constraints on exposure of the fusion module, allowing it to advance more readily through the "class-1" pathway (Figure 4.4). These findings are beginning to point toward therapeutic targets interfering with coronavirus entry, and further breakthroughs will likely come from detailed S protein structure determinations.

3. S Functions During Dissemination of Coronavirus Infections

In considering the entire infection cycle, we advance now to describing intracellular events as they pertain to S protein and virion morphogenesis. As stated above, the action of S proteins during entry delivers viral genomes into cells. These genomes are monopartite 27–32 kb single-stranded, positive-sense RNAs (Lai and Stohlman, 1978). The organization of genes on this large RNA has been well characterized, and the mechanisms of gene expression are understood in some detail and are not described here (for reviews, see Lai and Cavanagh, 1997; Sawicki and Sawicki, 1998). As is typical of RNA virus genomes, the vast majority is translated, with the 5′ ~two thirds encoding so-called "nonstructural" proteins that are not found in virions and the remaining ~one third encoding primarily "structural," that is, virion proteins (Figure 4.5). On eclipse, the nonstructural proteins are synthesized without delay, thereby generating RNA-dependent RNA replicase activities that subsequently transcribe antigenomic (negative-sense) RNAs, as well as several subgenomic viral mRNAs. The essential virion proteins S, E (envelope), M (matrix), and N (nucleocapsid) are translated from the newly created set of 3′ proximal subgenomic mRNAs, whose abundance in infected cells is far greater than genomic (virion) RNA, thereby permitting accumulation of virion proteins to the levels required for particle assembly.

A defining characteristic of coronavirion morphogenesis is its intracellular assembly (Figure 4.6), known for some time to take place in the ER-Golgi intermediate compartment
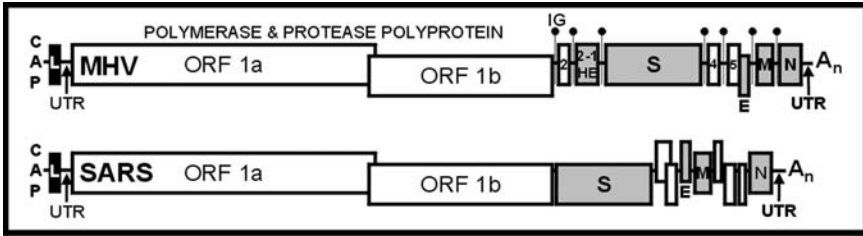
Figure 4.5. Coronavirus genome organization. Depictions of the murine coronavirus MHV (31.2 kb: GenBank accession number NC 001846) and human SARS coronavirus (29.7 kb: # AY278741) positive strand RNA genome. The 5′ end is capped, followed by a leader (L) sequence and *un*translated *r*egion (UTR). The polymerase and protease polyprotein complex is encoded along two open reading frames (1a and 1b) by a ribosomal frame-shifting mechanism and subsequently proteolytically processed into smaller fragments. Vertical lines with globular heads indicate intergenic (IG) sequences. Shaded boxes are structural proteins (sequentially: HE, S, E, M, N) that incorporate into virion particles. Genomes are polyadenylated. Drawn approximately to scale.

(Krijnske-Locker *et al.*, 1994). Infectious virus production requires newly synthesized genome RNA and its associated N proteins, as well as the three integral-membrane proteins S, E, and M. The assembly process, described in greater detail in the legend to Figure 4.6, involves a series of noncovalent interactions; S associating with M (de Haan *et al.*, 1999), M with N (Kuo and Masters, 2002), and N or M with virion RNA (Nelson and Stohlman, 1993; Narayanan *et al.*, 2003). Interestingly, assembly and secretion of intracellular vesicles requires only M and E (Vennema *et al.*, 1996); S and the ribonucleocapsids are dispensable and must be considered to be passive participants in particle morphogenesis. Thus, the S proteins, depending on their affinity for M and their abundance relative to M, may or may not engage in the virion assembly process. In model coronavirus infections, S–M affinities and molar ratios are such that only a portion of S proteins assemble into virions and significant proportions of the population advance as free proteins through the exocytic pathway and on to infected-cell surfaces (Figure 4.6). As there is no evidence that any coronavirus budding takes place at plasma membrane locations, these cell-surface S proteins likely function solely to mediate the cell–cell fusions that facilitate rapid spread of infection. Little is presently known about the relative efficiencies of S–M interactions among the coronaviruses, although the interacting portions of these proteins do indeed differ among virus strains. One might speculate that relatively low S–M affinities reflect adaptations to growth under conditions where syncytial spread of infection provides selective advantages that are greater than those afforded by high S–M affinities favoring efficient infective virion morphogenesis.

4. S Polymorphisms Affect Coronavirus Pathogenesis

It has been known for nearly 20 years that S gene differences correlate with *in vivo* pathogenic potential (Dalziel *et al.*, 1986; Fleming *et al.*, 1986), but only with the recent advent of facile reverse genetics approaches have these S mutations been definitively linked to coronavirus virulence. Definitive links required greater genetic control over the large and heterogeneous coronavirus RNA genome (Figure 4.5), so that one could construct and then characterize panels of recombinant viruses that harbor differences in S genes and nowhere else.
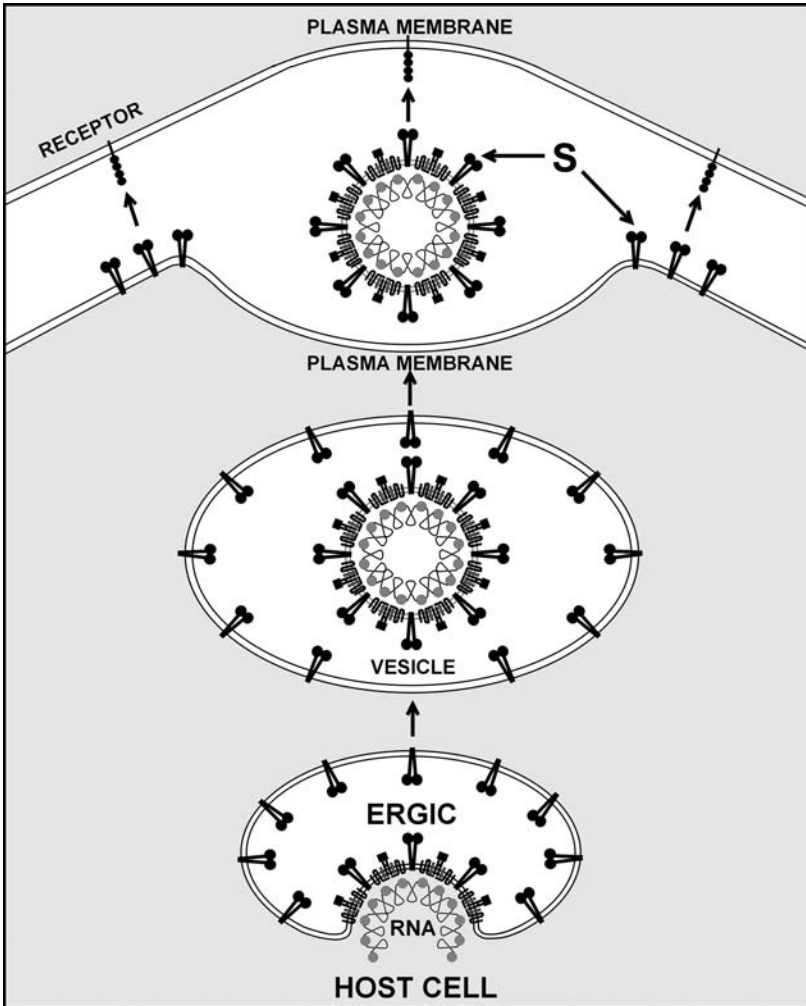
Figure 4.6.   Coronavirus assembly. Depiction of accreting coronavirus structural proteins (HE, S, E, M, and N) and RNA at the *e*ndoplasmic *r*eticulum *g*olgi intermediate *c*ompartment (ERGIC) and the secretion of viral particles along the secretory pathway. Spikes (S) that do not incorporate into particles continue to traffic to the cell surface and promote intercellular spread with neighboring cells.

Studies pioneered in the Masters and Rottier laboratories have yielded creative approaches that are now widely employed to manipulate the 3′ genomic region encompassing "structural" genes (Koetzner *et al.*, 1992; Kuo *et al.*, 2000). The process takes advantage of the fact that coronavirus RNAs tend to recombine, most likely by a copy-choice mechanism (Lai, 1992). Thus, defined site-directed mutant RNAs derived by *in vitro* transcription will recombine with endogenous viral RNAs within infected cells to create site-specific recombinants (Figure 4.7). Isolation of rare recombinants from the far greater parental (nonrecombinant) population

depends on the incorporation of a positive-selection marker on the *in vitro* transcript RNAs, which in current approaches amounts to an S ectodomain with strict specificity for APN or CEACAM receptors. Thus, a mutant RNA encoding a group 2 (CEACAM-specific) S ectodomain will recombine in APN+ cells infected by a group 1 (APN-specific) virus, and desired recombinants can be isolated by plaque assays on CEACAM+ cells (Figure 4.7). A subsequent recombination of *in vitro* transcripts encoding group 1 (APN-specific) S ectodomains with the first-generation CEACAM-specific recombinants can generate additionally mutated second-generation recombinants that can be isolated by plaque assay on APN+ cells. This remains a powerful way to manipulate the 3′ portion of the coronavirus genome despite the recent construction of full-length 27–32 kbp infectious coronavirus cDNAs providing for complete genetic control (Almazan *et al.*, 2000; Casais *et al.*, 2001; Thiel *et al.*, 2001; Yount *et al.*, 2002).

The RNA recombination system has been frequently used to specifically incorporate S gene changes, and the general findings indicate that relatively subtle S alterations strongly influence coronavirus virulence and tropism (Sanchez *et al.*, 1999; Phillips *et al.*, 2001;
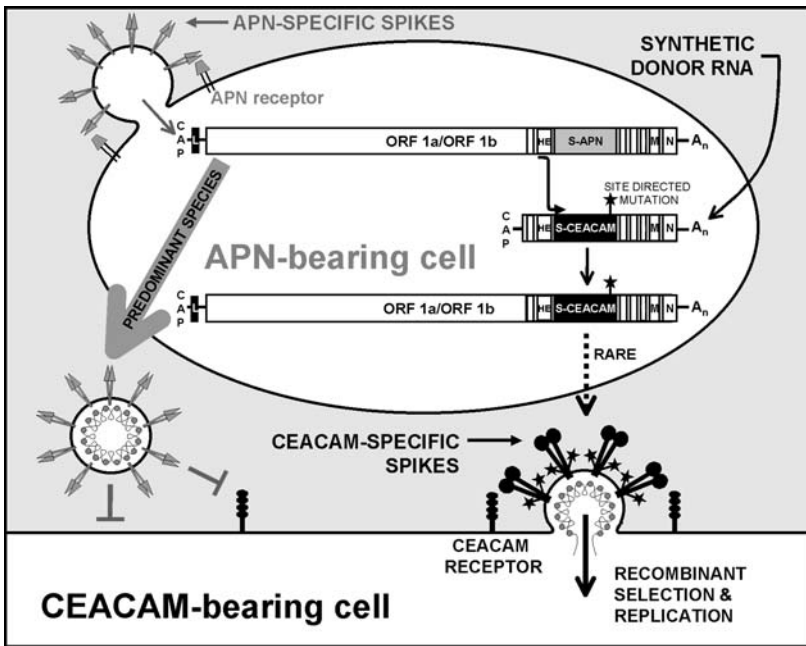

Figure 4.7. Targeted recombination approach to coronavirus reverse genetics. Schematic depicts an input synthetic RNA that harbors a site-directed mutation in the open reading frame (ORF) of a CEACAM-tropic spike (S). The subgenomic transcript recombines during viral replication with the genome of an engineered coronavirus that alternatively harbors APN-tropic spikes. Recombinant full-length genomes encoding CEACAM-tropic spike ORFs are packaged into particles that have incorporated newly synthesized CEACAM-specific S glycoproteins. This small percentage of progeny recombinant virions can enter and replicate in CEACAM-bearing cells. Non-recombinants do not switch tropism and are not selected. (Based on findings by PS Masters & PJ Rottier laboratories.)

Casais *et al.*, 2003; Tsai *et al.*, 2003). Correlating these alterations in virulence with the specific receptor-binding and membrane fusion functions of S proteins has just begun. Recent findings made in the Perlman laboratory have begun to establish these important relationships (Ontiveros *et al.*, 2003). In this laboratory, two variants of the group 2 mouse hepatitis coronavirus were identified with striking differences in neurovirulence, and using the established reverse genetics system, relative virulence was traced to a single S amino acid change, glycine 310 in virulent isolates, serine 310 associated with attenuation. This single difference had global effects on the overall stability of the S proteins, with gly310ser dramatically increasing the stable association of S1 and S2. Concomitant with this stabilization, membrane fusion activities were diminished. In particular, S proteins with the gly310 could mediate cell–cell fusion without the requirement for CEACAM triggering, while those with ser310 could not. These findings indicate that a subtle mutation outside of the core fusion machinery can powerfully influence S-mediated fusion, in this instance affecting its requirements for activation by receptor binding. These findings also point toward S-mediated cell–cell fusion activity as a core agent of coronavirus pathogenicity.

5. Applications to the SARS Coronavirus

As of September 2003, over 30 S_{SARS} sequences were posted in gene banks. To appreciate the sequence variations, one must view the data in the context of known and presumed SARS-CoV epidemiology. This virus is generally considered to be of zoonotic origin. While the natural wild or domestic animal reservoir is probably unknown, isolates strikingly similar to human SARS-CoV has been isolated from exotic animals in Guangdong, China (Guan *et al.*, 2003). These animals included asymptomatic palm civits and raccoon dogs, all housed in a single live animal market. The collection of animal CoV sequences shows some limited diversity (18 nt differences in the 29,709 nt genomes). Speculation is that around November 2002, one or more of these zoonotic "SZ" viruses infected humans and generated SARS fever, dry cough, and pneumonia. Virus from the initially infected human (the true "index" patient) may never be available, but viruses that have been isolated from Guangdong patients have interesting variation relative to "SZ" isolates. In comparing animal SZ viruses with the available collection of human SARS-CoV isolates, 11 clear S polymorphisms were detected (Guan *et al.*, 2003; see Figure 4.8). These appear to be relatively scattered changes throughout the 1,200-residue S ectodomain, and in this regard show some similarity to a collection of
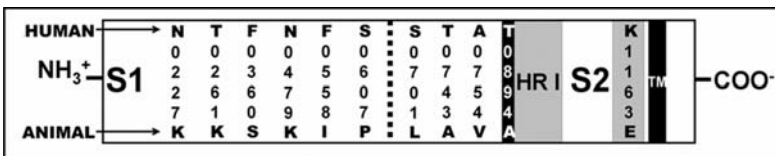
Figure 4.8. Human vs animal spike sequences. Shown are amino acid differences between spikes of human SARS and animal isolates. Numbers indicate location of spike residue. Dashed line demarcates boundary between S1 and S2 regions. Residue 894 resides within a candidate fusion peptide upstream of *h*eptad *r*epeat 1. Residue 1163 is within *h*eptad *r*epeat 2. TM indicates transmembrane region.

16 scattered differences between murine-specific and laboratory-generated zoonotic forms of murine hepatitis coronavirus (Baric *et al.,* 1997, 1999).

Assigning xenotropic potential to a particular combination of these 11 mutations is a challenging but important undertaking. This might be accomplished by employing the approaches used successfully to identify correlates of murine hepatitis virus virulence. S cDNAs encoding SZ or SARS isoforms, as well as SZ/SARS chimeras, can be easily constructed and then used to create recombinant coronaviruses. Tropism of the recombinants for human or animal cells can then be assessed using traditional virological methods. The next challenge will be to correlate S variations to alterations in receptor-binding or membrane fusion potentials. In all likelihood, the successful approaches will again be relatively traditional ones in which soluble S fragments—SZ, SARS, and SZ/SARS chimeras—are developed as mimics of authentic coronaviruses and then used as ligands for binding to human or animal cells, or once identified, SARS cellular receptors and their homologs in animal cells. By titrating soluble S ligands, relative affinities might be obtained. Questions concerning whether SARS polymorphisms specifically affect the membrane fusion reaction can then be addressed by relatively straightforward assays in which S-induced syncytia are measured (Nussbaum *et al.*, 1994). Among the murine coronaviruses, there are S polymorphisms that have no effect on S binding to CEACAM receptors, but yet dramatically impact membrane fusion (Krueger *et al.*, 2001). It will be important to determine whether there are similar variabilities in the SARS S proteins, and whether the membrane fusion process is central to SARS-CoV species transfer and human pathogenicity.

6. Relevance to Antiviral Drug Developments

At present, there are no clinically useful anti-coronavirus drugs, however, the targets for such drugs are clearly in sight. One obvious target is the coronavirus-encoded 3CL protease, as it is essential for the post-translational processing of gene 1 polyproteins into functional subunits ((Ziebuhr *et al.*, 1995; see Figure 4.5). Structure-based, rational anti-3CL protease drug design is at a relatively advanced stage (Anand *et al.*, 2003) and protease inhibitors roughly analogous to those used to combat HIV infection may be forthcoming. A second target, one that is far more relevant to the topic of this chapter, is the S protein. S proteins cause a characteristic syncytial cytopathology in the lung epithelia of SARS patients (Kuiken *et al.*, 2003), and should the S protein dissections described above link syncytial activities with pathogenicity in animal models, investigations would reasonably focus on drugs designed to block S function.

Therapeutics designed to block S-receptor interactions constitute one strategy. Recent structure determinations for a group 2 coronavirus receptor (Tan *et al.*, 2002), and delineation of relevant peptide loops interacting with S proteins (Rao *et al.*, 1997), bring promise to the hypothesis that S-binding receptor fragments might be constructed and used to interfere with virus entry. Such a peptide drug might block infection by inducing the premature triggering of the fusion reaction (Figure 4.4). One must, however, be cautious about advocating this approach because, while many soluble receptors will drive S proteins unproductively into denatured states, some will clearly trigger productive fusion reactions (Matsuyama and Taguchi, 2002). As was found in studies with HIV and soluble CD4 (Moore *et al.*, 1992; Arthos *et al.*, 2002), virus infectivities may be enhanced rather than neutralized.

A second strategy might extend from current hypotheses concerning coronavirus neutralization by antibodies. Several potently neutralizing monoclonal antibodies bind S in regions between CEACAM-binding and fusion-inducing domains (Dalziel *et al.*, 1986). While the mechanisms of neutralization are far from clear, one hypothesis is that the antibodies interfere with conformational transitions linking receptor interaction with fusion activation. High-resolution images of these antibody–S interactions could serve as a guide to construct smaller peptide ligands that neutralize infection by restricting global S conformational change.

Finally, recent convincing evidence that the S proteins of the group 2 mouse hepatitis coronavirus carry out a "class-1" fusion reaction (Bosch *et al.*, 2003) make it probable that several coronaviruses including SARS-CoV will be sensitive to a HR peptide-based fusion inhibition. Peptides derived from the HR regions of structurally similar retroviruses and paramyxoviruses interfere with fusion by associating with complete spikes during the activation reaction, preventing the appropriate collapse into a coiled-coil bundle (Wild *et al.*, 1994; Yao and Compans, 1996; see Figure 4.4). Similarly, a small 38-residue peptide representing mouse hepatitis virus HR2 powerfully inhibited both virus–cell and cell–cell fusion, reducing these activities by several logs when present at 10 μM concentration (Bosch *et al.*, 2003). These HR2 peptides block entry by binding transient intermediate conformations of the fusion protein, depicted in Figure 4.4B, C. It will therefore be important to know whether the genetic variabilities inherent in the coronavirus S proteins alter receptor affinities or fusion kinetics, as these parameters determine the lifespan of the drug-sensitive intermediate structures (Reeves *et al.*, 2002), and by extension they determine whether HR2-based peptides will be effective antiviral agents. By combining comparative studies on S protein receptor binding and membrane fusion with investigations of HR peptide-based antiviral activities, a mechanistic understanding of antiviral action will develop that can lay the groundwork required to develop therapies for human and animal diseases caused by the coronaviruses.

References

Almazan, F., Gonzalez, J.M., Penzes, Z., Izeta, A., Calvo, E., and Plana-Duran, J. *et al.* (2000). Engineering the largest RNA virus genome as an infectious bacterial artificial chromosome. *Proc. Natl. Acad. Sci. USA* **97**(10), 5516–5521.

Anand, K., Ziebuhr, J., Wadhwani, P., Mesters, J.R., and Hilgenfeld, R. (2003). Coronavirus main proteinase (3CLpro) structure: Basis for design of anti-SARS drugs. *Science* **300**(5626), 1763–1767.

Arthos, J., Cicala, C., Steenbeke, T.D., Chun, T.W., Dela Cruz, C., Hanback, D.B. *et al.* (2002). Biochemical and biological characterization of a dodecameric CD4-Ig fusion protein: Implications for therapeutic and vaccine strategies. *J. Biol. Chem.* **277**(13), 11456–11464.

Baric, R.S., Sullivan, E., Hensley, L., Yount, B., and Chen, W. (1999). Persistent infection promotes cross-species transmissibility of mouse hepatitis virus. *J. Virol.* **73**, 638–649.

Baric, R.S., Yount, B., Hensley, L., Peel, S.A., and Chen, W. (1997). Episodic evolution mediates interspecies transfer of a murine coronavirus. *J. Virol.* **71**(3), 1946–1955.

Bonavia, A., Zelus, B.D., Wentworth, D.E., Talbot, P.J., and Holmes, K.V. (2003). Identification of a receptor-binding domain of the spike glycoprotein of human coronavirus HCoV-229E. *J. Virol.* **77**(4), 2530–2538.

Bos, E.C.W., Heijnen, L., Luytjes, W., and Spaan, W.J.M. (1995). Mutational analysis of the murine coronavirus spike protein: Effect on cell-to-cell fusion. *Virology* **214**, 453–463.

Bosch, B.J., van der Zee, R., de Haan, C.A., and Rottier, P.J. (2003). The coronavirus spike protein is a class I virus fusion protein: Structural and functional characterization of the fusion core complex. *J. Virol.* **77**(16), 8801–8811.

Casais, R., Dove, B., Cavanagh, D., and Britton, P. (2003). Recombinant avian infectious bronchitis virus expressing a heterologous spike gene demonstrates that the spike protein is a determinant of cell tropism. *J. Virol.* **77**(16), 9084–9089.

Casais, R., Thiel, V., Siddell, S.G., Cavanagh, D., and Britton, P. (2001). Reverse genetics system for the avian coronavirus infectious bronchitis virus. *J. Virol.* **75**(24), 12359–12369.

Dalziel, R.G., Lampert, P.W., Talbot, P.J., and Buchmeier, M.J. (1986). Site-specific alteration of murine hepatitis virus type 4 peplomer glycoprotein E2 results in reduced neurovirulence. *J. Virol.* **59**, 463–471.

de Haan, C.A.M., Smeets, M., Vernooij, F., Vennema, H., and Rottier, P.J.M. (1999). Mapping of the coronavirus membrane protein domains involved in interaction with the spike protein. *J. Virol.* **73**, 7441–7452.

Delmas, B., Gelfi, J., L'Haridon, R., Vogel, L.K., Sjostrom, H., Noren O. *et al.* (1992). Aminopeptidase N is a major receptor for the entero-pathogenic coronavirus TGEV. *Nature* **357**(6377), 417–420.

Dveksler, G.S., Pensiero, M.N., Cardellichio, C.B., Williams, R.K.,Jiang, G.S., Holmes, K.V., *et al.* (1991). Cloning of the mouse hepatitis virus (MHV) receptor: Expression in human and hamster cell lines confers susceptibility to MHV. *J. Virol.* **65**(12), 6881–6891.

Firla, B., Arndt, M., Frank, K., Thiel, U., Ansorge, S., Tager, M. *et al.* (2002). Extracellular cysteines define ectopeptidase (APN, CD13) expression and function. *Free. Radic. Biol. Med.* **32**(7), 584–595.

Fleming, J.O., Trousdale, M.D., el-Zaatari, F.A., Stohlman, S.A., and Weiner, L.P. (1986). Pathogenicity of antigenic variants of murine coronavirus JHM selected with monoclonal antibodies. *J. Virol.* **58**(3), 869–875.

Gallagher, T.M. (1997). A role for naturally occurring variation of the murine coronavirus spike protein in stabilizing association with the cellular receptor. *J. Virol.* **71**, 3129–3137.

Godfraind, C., Langreth, S.G., Cardellichio, C.B., Knobler, R., Coutelier, J.P., Dubois-Dalcq, M. *et al.* (1995). Tissue and cellular distribution of an adhesion molecule in the carcinoembryonic antigen family that serves as a receptor for mouse hepatitis virus. *Lab. Invest.* **73**(5), 615–627.

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

Guan, Y., Zheng, B.J., He, Y.Q., Liu, X.L., Zhuang, Z.X., Cheung, C.L. *et al.* (2003). Isolation and characterization of viruses related to the SARS coronavirus from animals in Southern China. *Science* **302**, 276–278.

Jahn, R., Lang, T., and Sudhof, T.C. (2003). Membrane fusion. *Cell* **112**(4), 519–533.

Koetzner, C.A., Parker, M.M., Ricard, C.S., Sturman, L.S., and Masters, P.S. (1992). Repair and mutagenesis of the genome of a deletion mutant of the coronavirus mouse hepatitis virus by targeted RNA recombination. *J. Virol.* **66**(4), 1841–1848.

Krijnske-Locker, J., Ericsson, M., Rottier, P.J., and Griffiths, G. (1994). Characterization of the budding compartment of mouse hepatitis virus: Evidence that transport from the RER to the Golgi complex requires only one vesicular transport step. *J. Cell Biol.* **124**, 55–70.

Krueger, D.K., Kelly, S.M., Lewicki, D.N., Ruffolo, R., and Gallagher, T.M. (2001). Variations in disparate regions of the murine coronavirus spike protein impact the initiation of membrane fusion. *J. Virol.* **75**(6), 2792–2802.

Kuiken, T., Fouchier, R.A., Schutten, M., Rimmelzwaan, G.F., van Amerongen, G., van Riel, D. *et al.* (2003). Newly discovered coronavirus as the primary cause of severe acute respiratory syndrome. *Lancet* **362**(9380), 263–270.

Kuo, L., Godeke, G.J., Raamsman, M.J., Masters, P.S., and Rottier, P.J. (2000). Retargeting of coronavirus by substitution of the spike glycoprotein ectodomain: Crossing the host cell species barrier. *J. Virol.* **74**, 1396–1406.

Kuo, L. and Masters, P.S. (2002). Genetic evidence for a structural interaction between the carboxy termini of the membrane and nucleocapsid proteins of mouse hepatitis virus. *J. Virol.* **76**, 4987–4999.

Lai, M.M.C. (1992). Genetic recombination in RNA viruses. *Curr. Top. Microbiol. Immunol.* **176**, 21–32.

Lai, M.M. and Stohlman, S.A. (1978). RNA of mouse hepatitis virus. *J. Virol.* **26**(2), 236–242.

Lai, M.M.C. and Cavanagh, D. (1997). The molecular biology of coronaviruses. *Adv. Virus Res.* **48**, 2–100.

Marra, M.A., Jones, S.J., Astell, C.R., Holt, R.A., Brooks-Wilson, A., Butterfield, Y.S. *et al.* (2003). The genome sequence of the SARS-associated coronavirus. *Science* **300**(5624), 1399–1404.

Matsuyama, S. and Taguchi, F. (2002a). Communication between S1N330 and a region in S2 of murine coronavirus spike protein is important for virus entry into cells expressing CEACAM1b receptor. *Virology* **295**, 160–171.

Matsuyama, S. and Taguchi, F. (2002b). Receptor-induced conformational changes of murine coronavirus spike protein. *J. Virol.* **76**(23): 11819–11826.

Moore, J.P., McKeating, J.A., Huang, Y.X., Ashkenazi, A., and Ho, D.D. (1992). Virions of primary human immunodeficiency virus type 1 isolates resistant to soluble CD4 (sCD4) neutralization differ in sCD4 binding and glycoprotein gp120 retention from sCD4-sensitive isolates. *J. Virol.* **66**(1), 235–243.

Narayanan, K., Chen, C.J., Maeda, J., and Makino, S. (2003). Nucleocapsid-independent specific viral RNA packaging via viral envelope protein and viral RNA signal. *J. Virol.* **77**(5), 2922–2927.

Nelson, G.W. and Stohlman, S.A. (1993). Localization of the RNA-binding domain of mouse hepatitis virus nucleocapsid protein. *J. Gen. Virol.* **74**(Pt 9), 1975–1979.

Nussbaum, O., Broder, C.C., and Berger, E.A. (1994). Fusogenic mechanisms of enveloped-virus glycoproteins analyzed by a novel recombinant vaccinia virus-based assay quantitating cell fusion-dependent reporter gene activation. *J. Virol.* **68**, 5411–5422.

Ontiveros, E., Kim, T.S., Gallagher, T.M., and Perlman, S. (2003). Enhanced virulence mediated by the murine coronavirus, mouse hepatitis virus strain JHM, is associated with a glycine at residue 310 of the spike glycoprotein. *J. Virol.* **77**(19), 10260–10269.

Phillips, J.J., Chua, M., Seo, S.H., and Weiss, S.R. (2001). Multiple regions of the murine coronavirus spike glycoprotein influence neurovirulence. *J. Neurovirol.* **7**, 421–431.

Rao, P.V., Kumari, S., and Gallagher, T.M. (1997). Identification of a contiguous 6-residue determinant in the MHV receptor that controls the level of virion binding to cells. *Virology* **229**, 336–348.

Reeves, J.D., Gallo, S.A., Ahmad, N., Miamidian, J.L., Harvey, P. E., Sharron, M. *et al.* (2002). Sensitivity of HIV-1 to entry inhibitors correlates with envelope/coreceptor affinity, receptor density, and fusion kinetics. *Proc. Natl. Acad. Sci. USA* **99**(25), 16249–16254.

Rota, P.A., Oberste, M.S., Monroe, S.S., Nix, W.A., Campagnoli, R., Icenogle, J.P. *et al.*, (2003). Characterization of a novel coronavirus associated with severe acute respiratory syndrome. *Science* **300**(5624), 1394–1399.

Russell, C.J., Jardetzky, T.S., and Lamb, R.A. (2001). Membrane fusion machines of paramyxoviruses: Capture of intermediates of fusion. *EMBO J.* **20**(15), 4024–4034.

Sanchez, C.M., Izeta, A., Sanchez-Morgado, J.M., Alonso, S., Sola, I., Balasch, M. *et al.* (1999). Targeted recombination demonstrates that the spike gene of transmissible gastroenteritis coronavirus is a determinant of its enteric tropism and virulence. *J. Virol.* **73**, 7607–7618.

Sawicki, S.G., and Sawicki, D.L. (1998). A new model for coronavirus transcription. *Adv. Exp. Med. Biol.* **440**, 215–219.

Schultze, B., Gross, H.J., Brossmer, R., and Herrler, G. (1991). The S protein of bovine coronavirus is a hemagglutinin recognizing 9-O-acetylated sialic acid as a receptor determinant. *J. Virol.* **65**(11), 6232–6237.

Siddell, S.G. (1995). The coronaviridae: An introduction. In S.G. Siddell (ed.), *The Coronaviridae*, Plenum Press, New York and London, pp. 1–10.

Singh, M., Berger, B., and Kim, P.S. (1999). LearnCoil-VMF: Computational evidence for coiled-coil-like motifs in many viral membrane fusion proteins. *J. Mol. Biol.* **290**, 1031–1041.

Sjostrom, H., Noren, O., and Olsen, J. (2000). Structure and function of aminopeptidase N. *Adv. Exp. Med. Biol.* **477**, 25–34.

Stauber, R., Pfleiderara, M., and Siddell, S. (1993). Proteolytic cleavage of the murine coronavirus surface glycoprotein is not required for fusion activity. *J. Gen. Virol.* **74**, 183–191.

Suzuki, H., and Taguchi, F. (1996). Analysis of the receptor-binding site of murine coronavirus spike protein. *J. Virol.* **70**(4), 2632–2636.

Tan, K., Zelus, B.D., Meijers, R., Liu, J.-H., Bergelson, J.M., Duke, N. *et al.* (2002). Crystal structure of murine sCEACAM1a[1,4]: A coronavirus receptor in the CEA family. *EMBO J.* **21**, 2076–2086.

Thiel, V., Herold, J., Schelle, B., and Siddell, S.G. (2001). Infectious RNA transcribed in vitro from a cDNA copy of the human coronavirus genome cloned in vaccinia virus. *J. Gen. Virol.* **82**(Pt 6), 1273–1281.

Tresnan, D.B., Levis, R., and Holmes, K.V. (1996). Feline aminopeptidase N serves as a receptor for feline, canine, porcine, and human coronaviruses in serogroup I. *J. Virol.* **70**(12), 8669–8674.

Tsai, J.C., de Groot, L., Pinon, J.D., Iacono, K.T., Phillips, J.J., Seo, S.H. *et al.* Weiss (2003). Amino acid substitutions within the heptad repeat domain 1 of murine coronavirus spike protein restrict viral antigen spread in the central nervous system. *Virology* **312**(2), 369–380.

Vennema, H., Godeke, G.J., Rossen, J.W., Voorhout, W.F., Horzinek, M.C., Opstelten, D.J. *et al.* (1996). Nucleocapsid-independent assembly of coronavirus-like particles by co-expression of viral envelope protein genes. *EMBO J.* **15**, 2020–2028.

Weissenhorn, W., Dessen, A., Calder, L.J., Harrison, S.C., Skehel, J.J., and Wiley, D.C. (1999). Structural basis for membrane fusion by enveloped viruses. *Mol. Membr. Biol.* **16**(1), 3–9.

Wild, C.T., Shugars, D.C., Greenwell, T.K., McDanal, C.B., and Matthews, T.J. (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**(21), 9770–9774.

Yao, Q., and Compans, R.W. (1996). Peptides corresponding to the heptad repeat sequence of human parainfluenza virus fusion protein are potent inhibitors of virus infection. *Virology* **223**(1), 103–112.

Yount, B., Denison, M.R., Weiss, S.R., and Baric, R.S. (2002). Systematic assembly of a full-length infectious cDNA of mouse hepatitis virus strain A59. *J. Virol.* **76**(21), 11065–11078.

Zelus, B.D., Schickli, J.H., Blau, D.M., Weiss, S.R., and Holmes, K.V. (2003). Conformational changes in the spike glycoprotein of murine coronavirus are induced at 37 degrees C either by soluble murine CEACAM1 receptors or by pH 8. *J. Virol.* **77**(2), 830–840.

Ziebuhr, J., Herold, J., and Siddell, S.G. (1995). Characterization of a human coronavirus (strain 229E) 3C-like proteinase activity. *J. Virol.* **69**(7), 4331–4338.