Drug Design Targeting the Main Protease, the Achilles' Heel of Coronaviruses

Haitao Yang, Mark Bartlam and Zihe Rao*

National Laboratory of Biomacromolecules, Institute of Biophysics (IBP), Chinese Academy of Sciences, Beijing 100101, China and Tsinghua-IBP Joint Research Group for Structural Biology, Tsinghua University, Beijing 100084, China

Abstract: Coronaviruses (CoVs), a genus containing about 26 known species to date, cause highly prevalent diseases and are often severe or fatal in humans and animals. In 2003, a previously unknown coronavirus was identified to be the etiological agent of a global outbreak of a form of life-threatening pneumonia called severe acute respiratory syndrome (SARS). No efficacious therapy is currently available, and vaccines and drugs are under development to prevent SARS-CoV infection in many countries. The CoV main protease (M^{pro}), which plays a pivotal role in viral gene expression and replication through a highly complex cascade involving the proteolytic processing of replicase polyproteins, is an attractive target for drug design. This review summarizes the recent advances in biological and structural studies, together with development of inhibitors targeting CoV M^{pro}s. It is expected that inhibitors targeting CoV M^{pro}s could be developed into wide-spectrum antiviral drugs against existing and possible future emerging CoV-associated diseases.

Key Words: Coronavirus, main protease, drug design.

1. INTRODUCTION

Coronaviruses, characterized as enveloped, positive-stranded RNA viruses with the largest known genome, belong to the genus *Coronavirus* of the family *Coronaviridae* [1, 2]. There are approximately 26 species of coronaviruses (CoVs) [3-6] which can be classified into three distinct groups according to their genome sequences and serological reactions [2]. They infect humans and multiple species of animals, causing a variety of highly prevalent and severe diseases. For example, human coronavirus (HCoV) strains 229E (HCoV-229E), NL63 (HCoV-NL63), OC43 (HCoV-OC43), and HKU1 (HCoV-HKU1) are responsible for a significant portion of upper and lower respiratory tract infections in humans, including common colds, bronchiolitis, and pneumonia. They have also been implicated in otitis media, exacerbations of asthma, diarrhea, myocarditis, and neurological disease [1, 7-11]. In 2003, a previously unknown HCoV called severe acute respiratory syndrome coronavirus (SARS-CoV), the emergence of which was most likely as a result of animal-human transmission [12], was identified as the etiological agent of a global outbreak of a life-threatening form of pneumonia called severe acute respiratory syndrome (SARS) [13-16]. SARS ultimately infected more than 8000 people and killed approximately 800 people worldwide, with about 10% mortality rate, before it was effectively brought under control. From the beginning of the SARS epidemic, efforts to identify the natural host for SARS-CoV have never ceased. Recently two independent research groups have discovered a coronavirus

closely related to SARS-CoV in bats, which has prompted a new health alert [5, 6]. Animal coronaviruses, including avian infectious bronchitis virus (IBV), turkey coronavirus (TCV), porcine transmissible gastroenteritis virus (TGEV), porcine hemagglutinating encephalomyelitis virus (HEV), canine coronavirus (CCoV), feline infectious peritonitis virus (FIPV) and bovine coronavirus (BCV), are significant pathogens for chickens, turkeys, pigs, dogs, cats and cattle [2, 17]. Most of the known coronaviruses are highly infectious with high mortality in young animals, resulting in significant economic losses for the animal industry worldwide.

Although vaccines against some animal coronaviruses, such as IBV and CCoV, are routinely used to prevent serious diseases, there are several potential problems [1, 3, 11]. At present, there are no licensed vaccines or specific drugs are available to prevent HCoV infection [1, 11]. Due to lack of efficacious therapies, the mortality rate was high during the SARS outbreak. Consequently, great efforts have been focused on the development of vaccines and drugs against SARS-CoV. Viral-vectored vaccines [18-20] and DNA vaccines [21-23] have been tested in animal models with successful results, and currently an inactivated SARS-CoV vaccine is in clinical trials in China [24]. However, safety remains the major concern. Drug development strategies are focused on two main avenues: inhibitors to block virus entry into the host cells, and compounds to prevent viral replication and transcription. The CoV main protease (M^{pro}), which plays a pivotal role in mediating viral replication and transcription, is a particularly attractive target for anti-SARS drug design [25, 26]. Several reviews published previously have touched on the topic of coronavirus main proteases [25-27]. In this review, recent advances in biological and structural studies, and particularly in development of inhibitors of the CoV M^{pro}s, will be summarized and discussed.

*Address correspondence to this author at the Laboratory of Structural Biology, Tsinghua University, Beijing 100084, China; Tel: +86-10-6277 1493; Fax: +86-10-6277 3145; E-mail: raozh@xtal.tsinghua.edu.cn

2. STRUCTURE AND FUNCTION OF CORONA-VIRUS M^PRO

2.1. Overview

As the largest RNA viruses known to date, the size of coronavirus genomes range from 27 to 31 kb [5, 7, 8, 28-40]. About two-thirds of the genome encodes two overlapping polyproteins, pp1a (450-500 kDa) and pp1ab (750-800 kDa), which undergo extensive proteolytic processing by viral proteases to produce multiple functional subunits. These functional subunits are involved in formation of the replicase complex to carry out viral replication and transcription [25, 26, 41]. The viral proteases mentioned above are classified into *accessory* proteases and the *main* protease (M^pro, also called 3C-like protease). Accessory proteases are papain-like cysteine proteases which cleave the N-proximal polyproteins regions at two or three limited sites [25, 26].

The main protease is a chymotrypsin-like cysteine protease (~33 kDa) [42, 43], which not only processes at its own flanking sites within the polyproteins, but also directs the processing of all downstream domains of the replicase polyproteins *via* at least 11 conserved cleavage sites [25, 26, 44] (see Fig. **1**). It is termed the *main* protease because of its dominant role in processing replicase polyproteins and gene expression. The alternative name of *3C-like* protease was designated after the picornavirus *3C* proteases because of the similar substrate specificities and the identification of cysteine as a catalytic residue in the context of a predicted two- -barrel structure [45, 46].

The existence of the coronavirus M^pro was originally predicted by sequence analysis of IBV replicase polyprotein in 1989 [46]. Although comparative sequence analysis was subsequently extended to include the replicase genes of MHV, HCoV-229E and TGEV [36, 38, 39], the first ex-
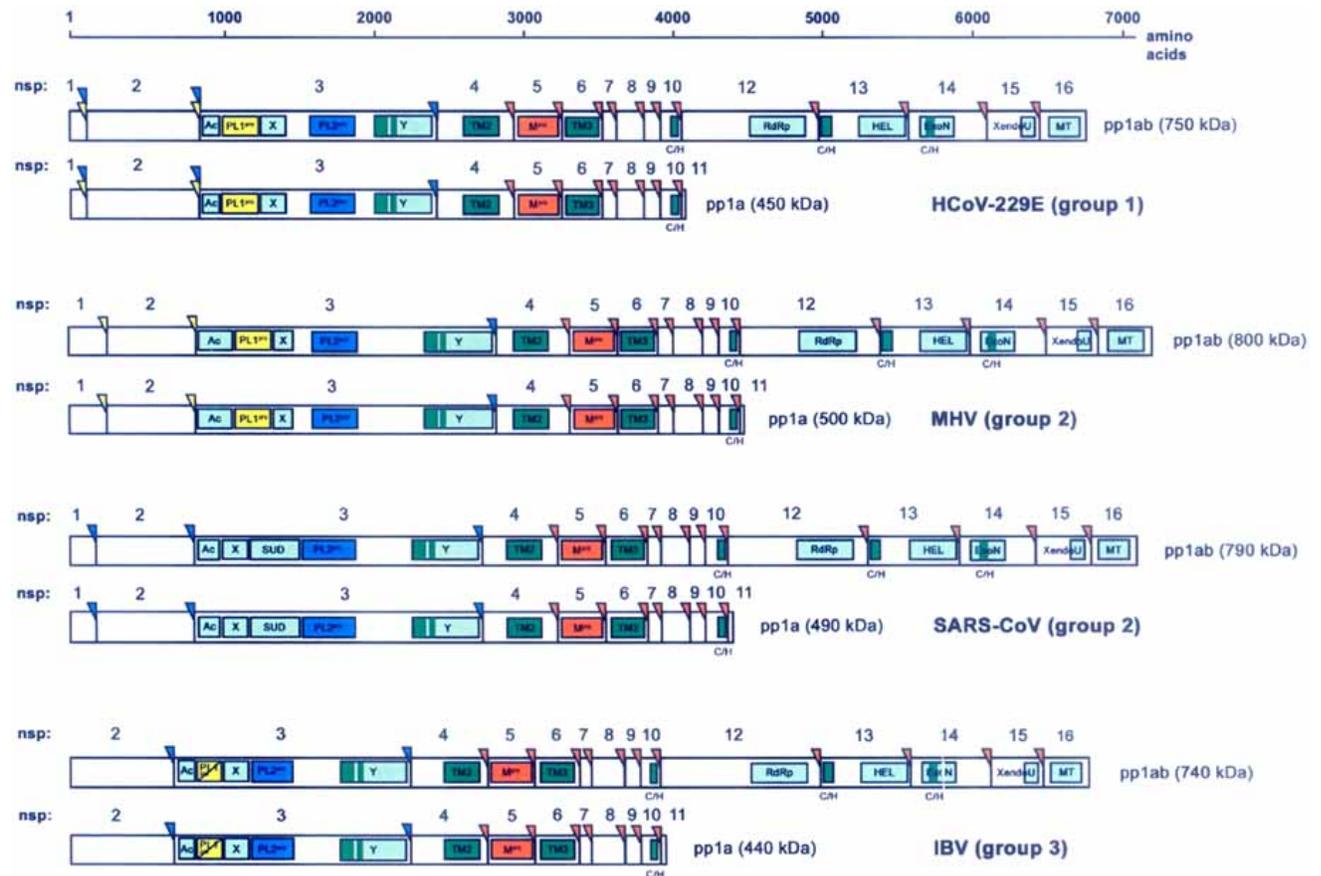

Fig. (1). The domain organization and proteolytic processing of coronavirus replicase polyproteins from human coronavirus 229E (HCoV-229E), murine hepatitis virus (MHV), SARS coronavirus (SARS-CoV) and avian infectious bronchitis virus (IBV). Shown are the replicase polyproteins pp1a and pp1ab; the processing products of pp1a are termed non-structural proteins (nsp) 1-11, and those of pp1ab are termed nsp1-nsp10 and nsp12-nsp16. Cleavage sites for the main protease (M^pro) are marked in *orange* arrows; cleavage sites for the papain-like cysteine protease 1 (PL1^pro) are marked in *yellow* arrows; and cleavage sites for the papain-like cysteine protease 2 (PL2^pro) are marked in *blue* arrows. Key to abbreviations: *Ac*, acidic domain; *PL1^pro*, papain-like cysteine protease 1; *X*, X domain with adenosine diphosphate-ribose 1"-phosphatase activity; *SUD*, SARS-CoV unique domain; *PL2^pro*, papain-like cysteine protease 2; *Y*, Y domain containing a transmembrane domain and a putative Cys/His-rich metal binding domain; *TM1, TM2, TM3*, transmembrane domains 1, 2, 3; *M^pro*, main protease (or 3C-like protease); *RdRp*, RNA-dependent RNA polymerase; *HEL*, helicase; *ExoN*, 3'-to-5' exonuclease; *XendoU*, poly(U)-specific endoribonuclease; *MT*, *S*-adenosylmethionine-dependent ribose 2'-*O*-methyltransferase; *C/H*, Cys/His-rich domains predicted to bind metal ions. IBV pp1a and pp1ab do not possess a counterpart to nsp1 of other coronaviruses. The PL1^pro of IBV is crossed out to indicate that it is proteolytically inactive. Figure adapted from [25].

perimental evidence of protease activity was reported for IBV in 1994 [47]. Afterwards, extensive activity studies of M^{pro} in different expression systems were extended to MHV and HCoV [48, 49]. The information accumulated from different coronaviruses can be used to map the M^{pro} processing of replicase polyproteins from all three coronavirus groups [25] (see Fig. **1**).

2.2. Overall Structure

The crystal structure of TGEV M^{pro} at 1.96 Å, which was the first structure of any M^{pro} to be solved, was reported in 2002 [43]. During the SARS outbreak and shortly after the SARS epidemic in 2003, the structures of HCoV 229E M^{pro} and SARS-CoV M^{pro} were respectively solved at 2.54 Å and 1.9 Å [42, 50]. Later, several structures for SARS-CoV M^{pro} were published one after the other with the aim of drug development [51-53]. TGEV and HCoV 229E both belong to group 1 [2]. SARS-CoV is considered to be an early split-off from group 2 and has been classified as a group 2b CoV [44, 54]. HCoV 229E M^{pro} and TGEV M^{pro} share 61 % primary sequence identity, whereas SARS-CoV M^{pro} has only about 40 % sequence identity with these two M^{pro}s (see Fig. **2**). Although the three M^{pro}s are from two different groups, they share a similar overall structure. It is interesting to find that

although SARS-CoV M^{pro} shares a slightly higher sequence identity with TGEV M^{pro} (44 %) than with the HCoV 229E enzyme (40 %), the structure is significantly more similar to that of the latter (r.m.s. deviations 2.3 Å and 1.6 Å, respectively), which explains why HCoV 229E M^{pro} is a better search model in determining the structure of SARS-CoV M^{pro} by molecular replacement.

All three main proteases are comprised of three domains (see Fig. **3A** and **B**). Domains I (TGEV M^{pro}: residues 8-100; HCoV M^{pro}: 8-99; SARS-CoV M^{pro}: 8-101) and II (TGEV M^{pro}: residues 101-183; HCoV M^{pro}: 100-183; SARS-CoV M^{pro}: 102-184) both have an anti-parallel β-barrel fold, which is similar to the serine proteases of the chymotrypsin family. These two domains have very limited similarity with picornavirus 3C proteases [55-58], which also have a chymotrypsin-related structure, suggesting that the name *main* protease is more appropriate than the somewhat misleading *3C-like* protease. The superposition of domains I and II of the TGEV M^{pro} onto the equivalent domains of the HAV 3C protease yields an r.m.s.d. of 2 Å for 114 equivalent (out of 184 compared) Cα pairs [43]. The substrate-binding site is located in a cleft formed between domains I and II and the catalytic site lies at the center of the cleft. In contrast to picornavirus 3C proteases, there is an
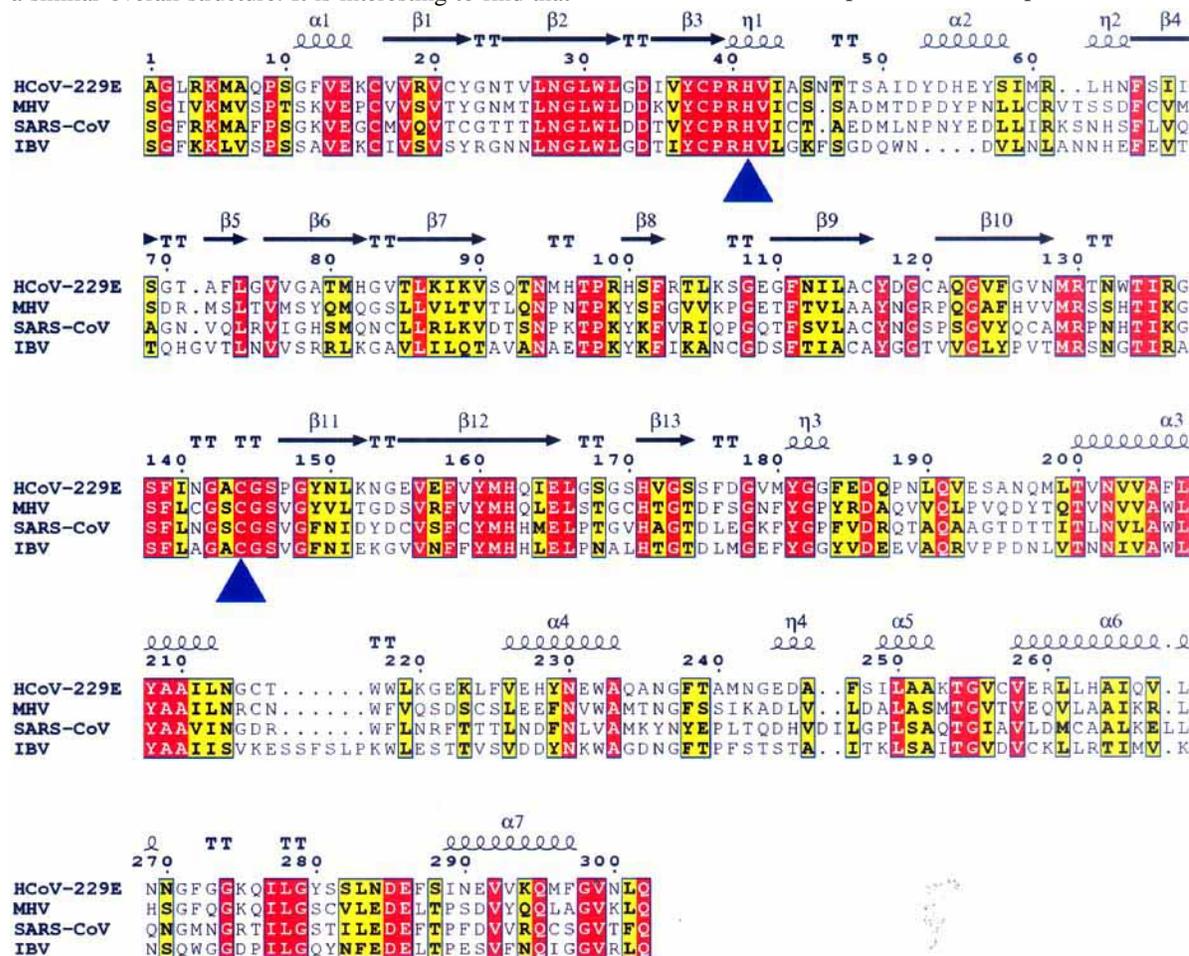

Fig. (2). Multiple sequence alignment of coronavirus M^{pro}s from human coronavirus 229E (HCoV-229E), murine hepatitis virus (MHV), SARS coronavirus (SARS-CoV) and avian infectious bronchitis virus (IBV). The catalytic dyad (41-His and 144/145-Cys) is shown by blue triangles.
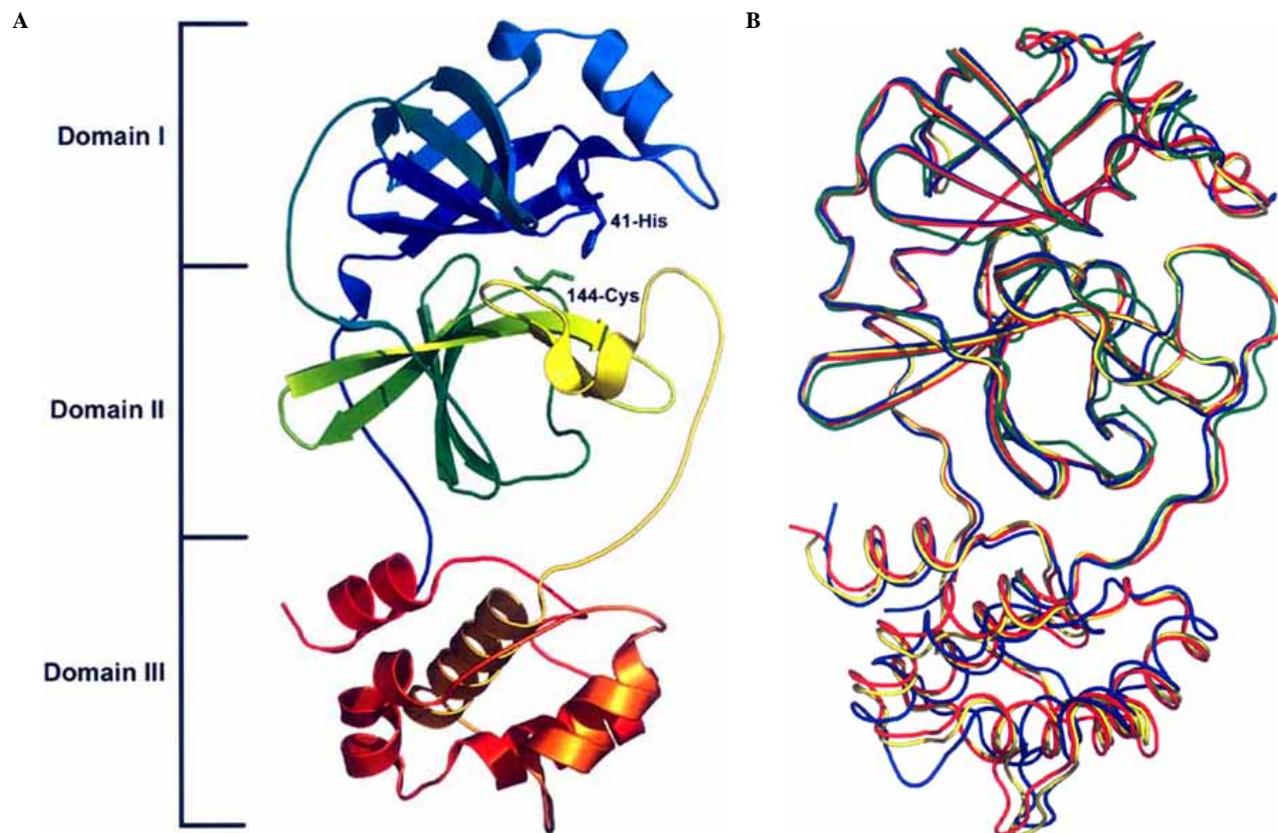
Fig. (3). A). The structure of the porcine transmissible gastroenteritis virus (TGEV) M^{pro} monomer. The M^{pro} monomer structure is drawn in ribbon representation and colored from blue at the N-terminal to red at the C-terminal. The catalytic dyad (41-His and 144-Cys) are shown as sticks and domains I, II and III are marked. **B).** Superposition of coronavirus M^{pro}s from human coronavirus 229E (HCoV-229E; blue), murine hepatitis virus (MHV; yellow), SARS coronavirus (SARS-CoV; red) and avian infectious bronchitis virus (IBV; green). It should be noted that the refinement of IBV M^{pro} is still in progress and the MHV M^{pro} structure is a homology model. Please see [3] for further details.

additional domain III in CoV M^{pro}s with a unique topology [43]. Domain III (TGEV M^{pro}: residues 200-302; HCoV M^{pro}: 200-300; SARS-CoV M^{pro}: 201-303) contains five largely anti-parallel -helices arranged into a globular cluster, and is connected with domain II *via* a long loop region consisting of 16 amino acids. The N-terminal within domain I folds onto domain III, bringing it into close proximity with the C-terminal of domain III.

CoV M^{pro} can form a homodimer both in the crystal and in solution (see Fig. **4**). In the crystal, one protomer is oriented perpendicular to the other and the solvent accessible surface area (per protomer) buried upon dimerization of each protomer ranges from 1300~1600 Å^2. In each protomer, domains II, III and the "N-finger" are involved in dimer interface formation. The "N-finger", defined as the N-terminal residues 1-7, inserts between domains II and III of the partner subunit [50]. Dynamic light scattering results show that both HCoV 229E and TGEV M^{pro}s exist as a mixture of monomers (~65%) and dimers (~35%) in diluted solutions (1-2 mg/ml). Analytical gel filtration, ultracentrifugation and isothermal titration calorimeter experiments both indicate that SARS-CoV M^{pro} is also a mixture of monomers and dimers in solution [59-61], although the values of the dissociation constant (K_d) determined by the dif-

ferent methods are incongruent, ranging from μM to nM [51, 59, 62, 63]. While the N-terminal and domain III are observed to participate in dimerization interactions in the crystal structure, the role of the N-terminal in maintaining the quaternary structure has been reported with inconsistent results. In their SARS-CoV M^{pro} study, Chou *et al.* showed that an ion pair formed by 4-Arg at the N-terminal of one protomer and 290-Glu at the C-terminal of domain III in its partner is crucial in dimer interactions. Mutating either of these two residues will increase the dissociation constant [62]. Another experiment demonstrated that deletion of 1-3 amino acids from the N-terminal will not have a substantial effect on dimerization, whereas the truncation of amino acids 1-4 shows a major form of a monomer [63]. However, Shi *et al.* reported that after the dissection of SARS-CoV M^{pro} into two parts, the chymotrypsin fold containing domains I and II is a monomer while the additional domain III exists as a stable dimer [64]. Chen *et al.* reported that K_d of a N-terminal (1-7 amino acids) truncated SARS-CoV M^{pro} changes slightly compared with the full-length protease by isothermal titration calorimeter analysis, suggesting that the N-terminal should not be indispensable for dimerization [61]. Nonetheless, it is generally accepted that domain III is essential to maintain quaternary structure and the "N-finger"
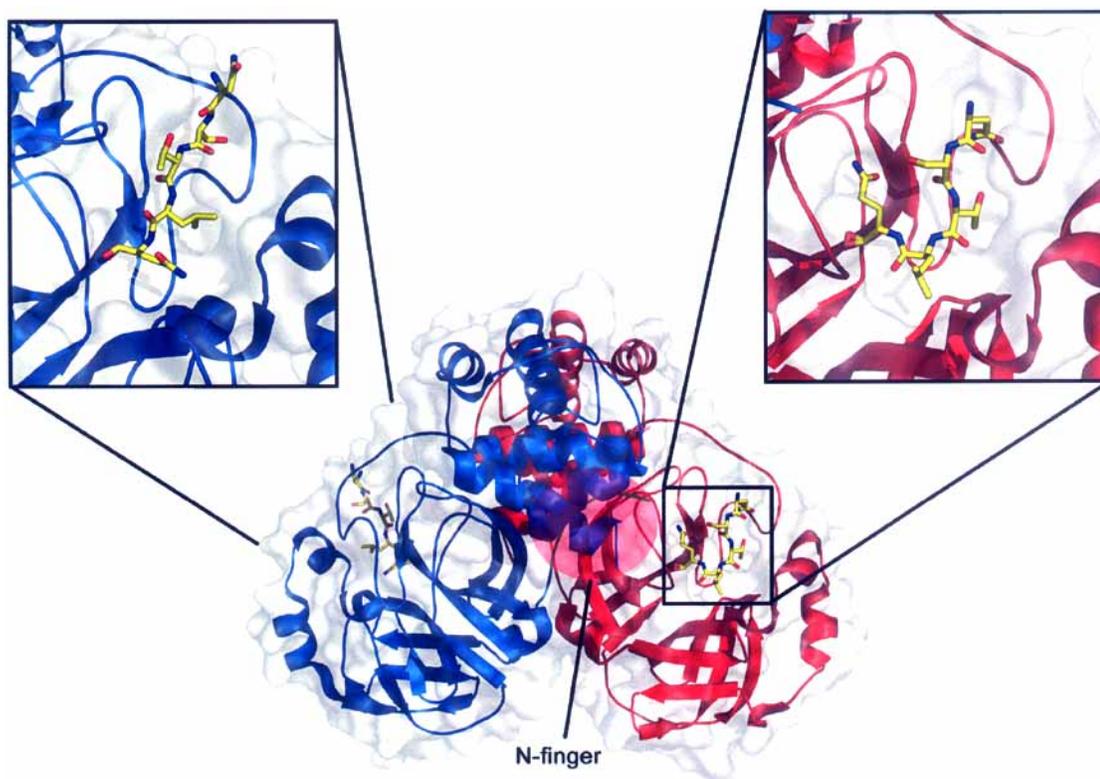
Fig. (4). The SARS-CoV M^pro in complex with a CMK inhibitor, as reported by Yang *et al*. [50]. The M^pro is shown in ribbon representation with protomer A in red and protomer B in blue, and is covered by a transparent molecular surface. The CMK inhibitors bound to protomers A and B are shown in yellow stick representation.

is important for enzymatic activity, the latter of which will be discussed below.

Another question is whether the role of domain III is only limited to stabilization of the dimer, thus maintaining the activity of the M^pro, as has been verified by loss of activity following deletion of part or whole of domain III in TGEV and SARS-CoV M^pros [43, 60, 63, 65]. In characterization experiments of temperature-sensitive MHV mutants, substitution of 291-Phe in domain III of M^pro with Leu causes an RNA-minus phenotype at the restrictive temperature [66]. In the MHV M^pro homology model constructed from SARS-CoV M^pro [3], 291-Phe is located far from the active site and does not participate in dimerization, and so there is little probability of this mutation independently affecting the protease activity. As 291-Phe is situated near the protease surface, it has the opportunity to interact with other domains in polyproteins. One possible role for domain III might be to regulate the proleolytic processing pathways for the polyproteins through interaction of specific regions of downstream domains of the replicase polyproteins, although we can not exclude other functions of domain III in RNA synthesis.

2.3. Catalytic Site and Substrate Binding Pocket

It was originally proposed from IBV sequence analysis that the catalytic center of CoV M^pro resembles that of other viral 3C and 3C-like proteases and might possibly include a His-Asp(Glu)-Cys catalytic triad [46]. Mutagenesis studies have confirmed that His and Cys residues are essential com-

ponents of the catalytic centre in HCoV 229E, IBV, MHV and SARS-CoV M^pros [48, 49, 51, 67-71]. There has been a long-stranding debate on the existence of a third member in the catalytic center of CoV M^pros. Great efforts made in sequence analysis and mutagenesis studies could not completely solve this question [26]. Only when the crystal structure of TGEV M^pro was determined in 2002 was this problem finally clarified [43]. Compared with serine proteases and other cysteine proteases, which adopt a catalytic triad, CoV M^pro only takes His-Cys as a catalytic dyad and completely lacks a third catalytic residue [3, 42, 43, 50]. In contrast to picornavirus 3C proteases, TGEV, HCoV 229E and SARS-CoV M^pros all have a particularly ordered water molecule in the corresponding position to the third catalytic member. This water forms at least three hydrogen bond interactions with surrounding residues (including the catalytic dyad member His) in all M^pro structures determined to date, suggesting its possible role is to stabilize the protonated histidine in the intermediate state during proteolytic cleavage, reminiscent of the function of Asp (Glu) in classic serine protease catalysis. The replacement of cysteine in the charge-relay system with a serine in IBV and SARS-CoV will produce a protease with residual activity, supporting a classical general base mechanism related to serine proteases rather than a thiolate-imidazolium mechanism for papain-like proteases [69, 71].

Coronaviruses only encode one M^pro, which is highly selective for substrates. A typical substrate sequence for M^pro is Leu-Gln (Ser, Ala, Gly) in polyproteins, which is

conserved among different coronavirus M[pro]s [26]. Hegyi *et al.* discovered that a synthetic peptide representing the N-terminal HCoV 229E M[pro] cleavage site was shown to be effectively hydrolysed by non-cognate M[pro]s [72]. In studies on SARS-CoV M[pro], it was reported that the N-terminal SARS-CoV M[pro] processing sequence could be efficiently cleaved by M[pro]s of all other groups of CoV, including the recently identified HCoV-NL63 and HCoV-HKU1 [3, 42, 73]. This is explained by a comparison of representatives from all three groups of CoVs: the structures of TGEV (group 1), HCoV-229E (group 1), SARS-CoV (group 2b) M[pro]s, a homology model of MHV (group 2a) M[pro] and the structure of IBV M[pro] under refinement [3, 27, 42, 43, 50].

Superposition of the structures and model reveal that backbones of the CoV M[pro]s substrate binding pocket superimpose particularly well (see Fig. **3B**), except for a small segment located on the outer wall of S2 [3], and so attention is focused on the variation of side chains forming important subsites S1, S2, S4 and S1'. In coronavirus M[pro]s, the S1 subsite endows the peptidases with unique specificity for substrate recognition and P1 is invariably occupied by Gln in polyprotein processing. Anand and colleagues reported the structure of TGEV M[pro] in complex with a substrate analog, hexapeptidyl chloromethyl ketone inhibitor, Cbz-Val-Asn-Ser-Thr-Leu-Gln-CMK, derived from the N-terminal processing cleavage site [42]. This structure revealed that the side chains of 165-Glu, 162-His, 171-His, and 139-Phe in TGEV M[pro] (also conserved in other M[pro]s) are incorporated with other backbone elements to constitute the S1 site, which has an absolute requirement for Gln at the P1 position *via* two hydrogen bonds. One hydrogen bond is formed by the side chain of Gln and the N atom of His-162, which is situated at the bottom of S1 subsite. The required neutral state of His-162 at physiological pH appears to be maintained by two important interactions: (i) stacking with the phenyl ring of Phe-139 and (ii) acceptance of a hydrogen bond from the hydroxyl group of the buried Tyr-160, which is conserved in other known SARS-CoV and HCoV 229E M[pro] structures [42, 50-52]. The role of His-162 and Tyr-160 was experimentally verified by mutagenesis analysis of corresponding residues in HCoV 229E and FIPV M[pro]s [68, 74]. Inserting in between domain III of the parent protomer and domain II of the neighboring protomer in the dimer, the N-finger plays an important role for the protease activity. In particular, the first amino acid of the N-finger in the parent protomer commonly forms ionic or hydrogen bond interactions with the essential residues constituting S1 site of the partner protomer, helping to maintain the active conformation of this subsite [43, 50, 52]. Deletion of residues 1-5 in the related TGEV M[pro] renders the enzyme almost completely inactive [43]. Site-direct mutation of the crucial residue or deletion of the whole N-finger will result in a great loss of activity in SARS-CoV M[pro] [61-63, 65].

The side chains of 164-Leu, 51-Ile, 41-His, and 53-Tyr, as well as the alkyl portion of side chains of 186-Asp and 47-Thr, are involved in forming a deep hydrophobic S2 subsite that can accommodate the relatively large side chain of Leu in TGEV M[pro]. This same feature can also be observed in the HCoV-229E M[pro]. Several conservative substitutions occur in other CoV M[pro]s (164-Leu   165-Met in SARS-CoV and MHV M[pro]s; 53-Tyr   50-Trp in IBV

M[pro]). Another minor difference is observed in SARS-CoV and MHV M[pro]s, where the outer wall segment is composed of a short 3_{10}-helix from residues 45–50, in contrast to the less regular structure in HCoV and TGEV M[pro]s. With respect to the structure of IBV M[pro] undergoing refinement, no clear electron density was observed in the corresponding stretch of residues 44–47. Apart from a few exceptions, CoV M[pro] has a Leu residue in the P2 position [26]. However, SARS-CoV M[pro] has Phe at the P2 site in its C-terminal cleavage sequence, representing a structural difference in the S2 site compared with TGEV and HCoV M[pro]s [51, 52]. The side chain of the residue at the P3 position in all M[pro]s is solvent-exposed, so this site was expected to tolerate a wide range of functionality. The side chains of 164-Leu, 166-Leu, 184-Tyr, and 191-Gln that form the S4 hydrophobic subsite of TGEV are conserved in other CoV M[pro]s, with the exception of the following conservative substitutions: 184-Tyr 184-Phe in HCoV M[pro]; 164-Leu   165-Met, 184-Tyr 185-Phe in SARS-CoV. The congested S4 site implies that only small amino acid residues such as Ser, Thr, Val or Pro can be accommodated at the P4 position [27]. Although the S1' subsite of TGEV M[pro] formed by 27-Leu, 41-His and 47-Thr is not deep, similar with other CoV M[pro]s, it has sufficient space for common small P1' residues such as Ser, Ala, or Gly to extend into this subsite. This could help in the design of substrate-analogue inhibitors targeting CoV M[pro]s with higher binding affinity [3]. Yang and colleagues have solved the structure of SARS-CoV M[pro] mutant in complex with an 11-mer peptidyl substrate, which could provide whole insight into the interactions between substrate and protease (Rao, personal communication). In the first crystal structure of SARS-M[pro] determined at pH 6.0, the S1 specificity pocket of one protomer is in the active conformation while that of the neighboring protomer is partially collapsed, resulting in only about 50 % protease activity. Activity curves show that SARS-CoV protease activity increases from pH at 6.0 to approximately physiological pH [50, 53, 59], and so a pH-triggered switch for the catalytic activity of the peptidase was proposed [50]. Tan and colleagues have reported similar results by structural analysis and molecular dynamics simulations [53]. The crystal structures solved by Lee and co-workers from crystals with different space groups at pH 6.5 showed that the substrate-binding regions of both protomers are in the catalytically competent conformation. Thus the authors proposed the possibility of an alternative or additional mechanism [52].

Although CoV M[pro] has similar substrate specificity with picornavirus 3C proteases as mentioned above, no reports have shown that any CoV M[pro] could efficiently process the substrate of picornavirus 3C proteases, or vice versa, which can be inferred from the differences substrate binding pockets between CoV M[pro]s and picornavirus 3C proteases.

It is generally agreed that M[pro] commonly processes the downstream cleavage sites in polyproteins *in trans*, which was characterized for M[pro]s of various species [25]; however, it remains largely unknown whether the M[pro] releases itself from polyproteins *in trans* or *in cis*. In TGEV M[pro] structures, it seems reasonable to suggest that the cleavage even occurs on an *inter*molecular basis, which has also been suggested for MHV M[pro] based on biochemical evidence [75]. However, the previously reported structures of M[pro] are

all in the mature form without any clues for the flanking hydrophobic regions (TM domains). A number of MHV and IBV precursors containing M^{pro} were found to require microsomal membranes for efficient autocleavage to release M^{pro} [69, 76], suggesting that the autocleavage event might take place on the membrane and that hydrophobic regions help the cleavage sites to approach the active site by interaction with the membranes. It is difficult to tell whether cleavage *in cis* will occur or not without further consideration of the interaction between TM domains and membranes. In the N-terminal autocleavage event, for instance, no matter whether the autocleavage is *in cis* or *in trans* within a dimer, domain I will undergo a large conformational change between *pre* and *post* cleavage. Hence, the structure of M^{pro} with flanking hydrophobic regions will help to address this question. Recently, Hsu *et al.* reported that they found in a dimeric structure of C145A mutant of SARS-M^{pro}, the active site of one protomer contains the C-terminal of one protomer in another asymmetric unit. Therefore, they proposed an autocleavage mechanism to explain this [51].

3. INHIBITOR DESIGN

3.1. Enzyme Activity Assay

In the development of CoV M^{pro} inhibitors, efficacious methods are required to measure the enzyme activity and the effect of compounds screened. Initially the proteolytic activity of M^{pro} was characterized mainly through processing the expressed polypeptide *in vitro* and analyzing the products by electrophoresis [48, 49]. Later the HPLC method was utilized to analyze the cleavage of M^{pro} on synthesized peptides [42, 72, 77] and determine the kinetic parameters for the proteases [59, 68]. Using an HPLC-based peptide cleavage assay, Fan and colleagues synthesized 34 peptide substrates and evaluated their specificity for SARS-CoV M^{pro} by measuring hydrolytic activity [78]. However, these two methods are not sensitive and convenient enough for large-scale inhibitor screening. Afterwards a continuous colorimetirc assay was reported as a substitute for SARS-CoV M^{pro} activity assay based on cleavage at the Gln-pNA bond for a substrate Thr-Ser-Ala-Val-Leu-Gln-pNA [79]. At present, a fluorescence resonance energy transfer (FRET) technique based on a fluorescence-labeled substrate has been developed and is routinely utilized to monitor the M^{pro} activity *in vitro* for anti-CoV inhibitor screening and drug design [51, 52, 60-63, 73, 80-88]. Parea and co-workers. invented a genetic screen assay to monitor the SARS-CoV M^{pro} activity, in which the protease activity is associated with the proliferation of lambda phage in infected *Escherichia coli* cells [89].

3.2. Design of Inhibitors Targeting CoV M^{pro}

To date, the inhibitors reported for CoV M^{pro} have mostly been compounds targeting the active site of CoV M^{pro}. Although inhibitors preventing M^{pro} dimerization were suggested [27, 64], there have been no related reports. In the early stages of CoV M^{pro} studies, a series of typical serine, cysteine, aspartic acid and metalloprotease inhibitors were screened for CoV M^{pro}s. Ziebuhr and colleagues reported that 3,4-dichloroisocoumarin, phenylmethylsulfonyl fluoride
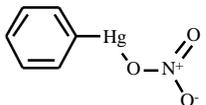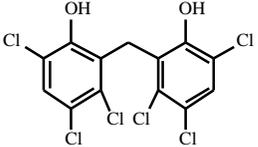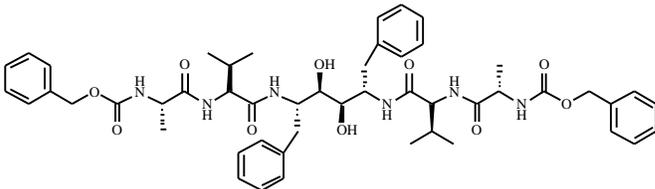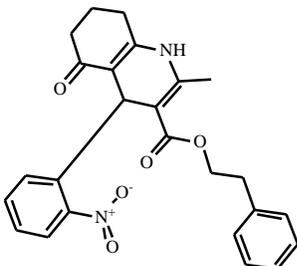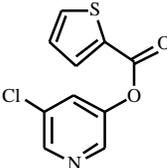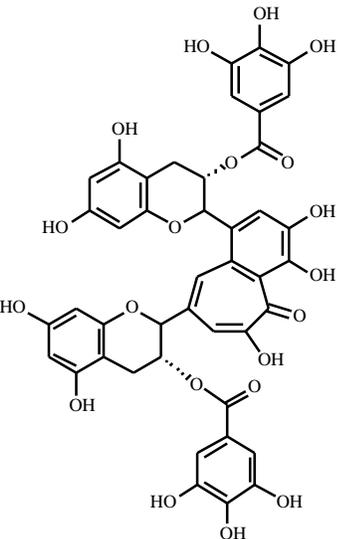
(PMSF), Pefaloc SC, tosyl lysyl chloromethyl ketone (TLCK), antipain and $ZnCl_2$ all function as HCoV 229E M^{pro} inhibitors [68]. A cysteine protease inhibitor E64d was shown to prevent MHV RNA synthesis and virus replication, possibly by inhibition of MHV M^{pro} processing activity [75]. However, these inhibitors were assayed primarily for characterization of the CoV M^{pro} active site and its role in viral RNA synthesis and replication, but not for pharmaceutical use due to their non-specificity.

During the global SARS epidemic, the M^{pro} was considered to be an attractive target [42], which led to an upsurge in the development of anti-SARS-CoV M^{pro} inhibitors. Since the structure of SARS-CoV M^{pro} was not available at the height of the outbreak, a series of homology models were constructed from the TGEV M^{pro} and HCoV 229E M^{pro} for virtual screening of compounds suitable for binding into the substrate binding pocket [42, 90-93]. However, no further reports have shown that the compounds suggested above are effective in inhibiting SARS-M^{pro} activity or preventing viral replication. In the wake of the SARS outbreak, the inhibitors were developed predominantly by three methods: (1) large-scale screening of structurally diverse small molecules from compound libraries or natural sources such as plant extracts; (2) virtual screening of compounds from chemical databases based on the SARS-CoV M^{pro} structure; and (3) *ab initio* design of small molecules directly from the structure, which will be respectively discussed below. A series of representative inhibitors identified against SARS-CoV M^{pro} is listed in Table **1**.
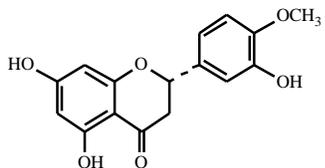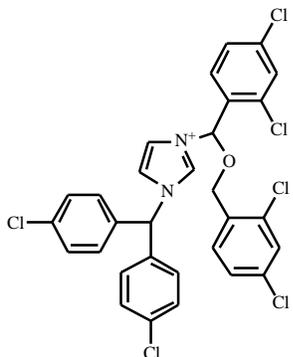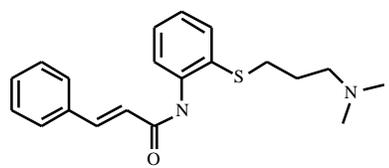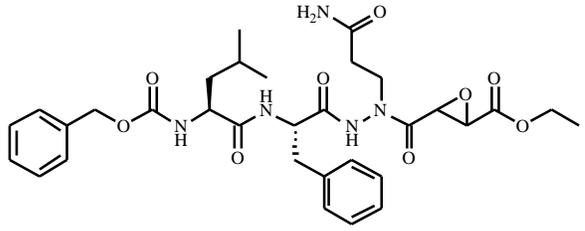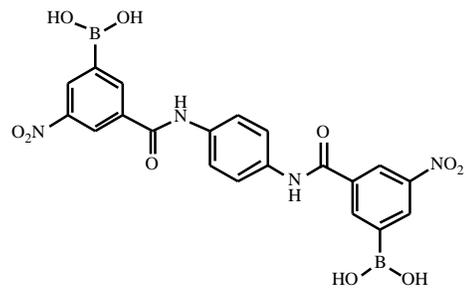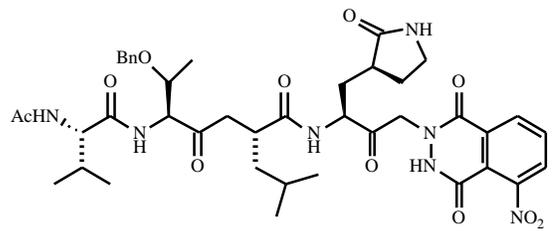
3.2.1. High-Throughput Screening

The rapid, sensitive fluorescence resonance energy transfer (FRET) technique provides the possibility to simultaneously screen a large number of compounds from available drugs, chemical libraries or natural plant extracts. Except for a few cases, most research groups have chosen to adopt this method. Hsu and co-workers reported that metal ions such as Hg^{2+}, Zn^{2+} and Cu^{2+} are SARS-CoV M^{pro} inhibitors, possibly because they have the ability to coordinate to the catalytic cysteine in the protease active site [94]. From a compound library, the authors also found that phenylmercuric nitrate (see Table **1**), thimerosal and hexachlorophene are effective inhibitors against SARS-CoV M^{pro}. However, safety is the main concern of the metal ions and particularly the toxicity of compounds containing mercury. Although preliminary data showed that hexachlorophene can decrease the amount of viral spike protein in the Vero E6 cells infected by SARS-CoV below the cytotoxicity concentration, the effect of this compound needs to be fully characterized in the prevention of viral replication and cytotoxicity. Liu and colleagues reported similar results for hexachlorophene and its derivatives in enzymatic inhibition assays [88]. Another study by Wu and co-workers reported that a transition-state analog inhibitor of HIV protease is active against the SARS-CoV M^{pro} (K_i=0.6 μM) [82]. In a large scale s creening, Kao and colleagues screened 50,240 small molecules, in which they identified 104 compounds with anti-SARS-CoV activity. From these active compounds, the authors discovered one compound named MP576 that displays inhibition activity against SARS-CoV M^{pro} with an IC_{50} of 2.5 μM and an EC_{50} of 7 μM in cell-based assays [81]. Similarly, Blanchard

Table 1.   Inhibition of SARS CoV M^{pro} Activity *In Vitro* and SARS-CoV in Cell Culture by Published Compounds

Compound	Structure	Enzymatic assay	Antiviral activity in cell-based assay (μM)	TC$_{50}$ (μM)
Phenylmercuric nitrate [94]		0.3 μM (K_i)	N/A[a]	N/A
Hexachlorophene [88, 94]		4-13.7 μM (K_i)	~10 (EC$_{90}$)	100
Compound 2 [82]		0.6 μM (K_i)	N/A	N/A
MP576 [81]		2.5 μM (IC$_{50}$)	7 (EC$_{50}$)	>50
MAC-5776 [83]		0.5 μM (IC$_{50}$)	N/A	N/A
Theaflavin-3,3'-digallate (TF3) [95]		9.5 μM (IC$_{50}$)	N/A	N/A

Compound	Structure	Enzymatic assay	Antiviral activity in cell-based assay (µM)	TC$_{50}$ (µM)
Hesperetin [96]		60 µM (IC$_{50}$)	N/A	2718
Calmidazolium [79]		61 µM (K_i)	N/A	N/A
Cinanserin [73]		4.92 µM (IC$_{50}$)	31 (IC$_{50}$)	>134
Aza-peptide epoxides [52]		18 µM (K_i) 35x10^{-3}S^{-1} (k_3)	N/A	N/A
FL-166 [60]		0.04 µM (K_i)	N/A	N/A
8c [80]		0.6 µM (IC$_{50}$)	N/A	N/A

(Table 1) contd….
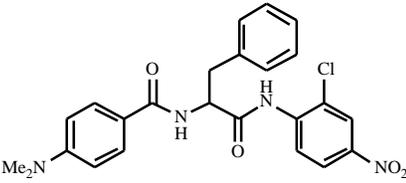
Compound	Structure	Enzymatic assay	Antiviral activity in cell-based assay (μM)	TC$_{50}$ (μM)
2a [87]		0.03 μM (K_i)	N/A	N/A
4o [84]		0.95 μM (IC$_{50}$)	N/A	N/A
N3 [3]		9 μM (K_i) 3.1x10^{-3}S^{-1} (k_3)	6 (IC$_{50}$)^b	>500

[a] N/A : data not available.
[b] Unpublished data.

and colleagues screened 50,000 small molecules and identified 5 compounds exhibiting inhibitory activity (IC$_{50}$= 0.5-7 μM) towards SARS-CoV M^{pro} [83]. In addition to small molecules from chemical libraries, some research groups have turned to natural plant extracts such as Chinese herbs. Two research groups individually claimed that tannic acid, 3-isotheaflavin-3-gallate and theaflavin-3,3'-digallate from black tea and one plant-derived phenolic compound, hesperetin, are SARS-CoV M^{pro} inhibitors [95, 96].

3.2.2. Virtual Screening

In a virtual screening of compounds from several chemical databases, Liu and colleagues identified calmidazolium, an antagonist of calmodulin, as a SARS-CoV M^{pro} inhibitor with a K_i of 61 μM from enzyme inhibition assays [79]. Chen and co-workers identified cinanserin, a serotonin antagonist, as an inhibitor of both SARS-CoV and HCoV-229E M^{pro}s from a database containing more than 8,000 compounds by a docking approach [73]. Cinanserin has an IC$_{50}$ of about 5 μM against both of these two proteases in enzyme inhibition assays, and an EC$_{50}$ ranging from 19 to 34 μM in cell-based assays. The authors also reported that the enzyme was not completely inhibited at the maximum drug concentration, while the amount of virus was significantly reduced in cell culture. This incongruity was explained as additional drug effects by the authors, but could arise if cinanserin acts as an irreversible inhibitor with a

small inactivation rate constant, and so could not completely inactivate the protease in a limited time during the enzyme activity assay.

3.2.3. Ab Initio Inhibitor Design

Many compounds designed directly from the M^{pro} structure are mechanism-based irreversible inhibitors, which could increase their inhibition effects. The inhibitor initially forms a reversible complex with the protease, which then undergoes a chemical step (nucleophilic attack by Cys) leading to the formation of a stable covalent bond. The evaluation of this series of inhibitors requires both the equilibrium-binding constant K_i (designated as k_2/k_1) and the inactivation rate constant for covalent bond formation k_3 (see Equation (1)) [3, 97].

$$E + I \quad \underset{k_2}{\overset{k_1}{\rightleftharpoons}} \quad EI \quad \overset{k_3}{\longrightarrow} \quad E - I \qquad (1)$$
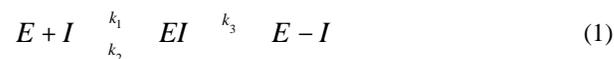
The first reported SARS-CoV M^{pro} irreversible inhibitor was a substrate-analog chloromethyl ketone (CMK) inhibitor, Cbz-VNSTLQ-CMK [50]. The sequence of this substrate-analog was derived from residues P6–P1 of the N-terminal autoprocessing site of TGEV M^{pro}. However, the two protomers of SARS-CoV M^{pro} each exhibited an unexpected binding mode (see Fig. **4**), possibly resulting from the comparatively weak binding of peptidyl elements derived from the substrate of TGEV M^{pro} and from the highly

reactive electrophile CMK. This would suggest that nucleophilic attack might have occurred before a stable noncovalently bound enzyme-inhibitor complex was formed [3]. It should be noted that the CMK compound can react with a variety of proteases, including cellular proteases, without specificity due to their high electrophilicity. Thus, the CMK compound is mainly utilized to explore the active sites of proteases but is not suitable for direct use in drug development. Recently, Lee and colleagues designed an aza-peptide epoxide inhibitor [52], Cbz-Leu-Phe-AGln-EP-COOEt and solved the co-crystal structure in which the Cys145 S' atom of the protease forms a covalent bond with the epoxide C3 atom of the inhibitor. As expected, this substrate analog binds to the substrate binding pocket of the protease in a normal mode, as seen in the TGEV M^pro complex [42]. Bacha and co-workers identified a cluster of serine residues near the active site cavity which were susceptible to targeting by compounds containing boronic acid. In a series of designed bifunctional aryl boronic acid compounds, a compound named FL-166 is effective at inhibiting the protease with a K_i of 40 nM [60]. Jain and colleagues reported a series of keto-glutamine analogues with a phthalhydraido group at the   -position, whose derivatives are known HAV 3C inhibitors, could act as reversible inhibitors against SARS M^pro with IC$_{50}$ ranging from 0.6 to 70 µM [80]. Shie and colleagues reported that an anilide derived from 2-chloro-4-nitroaniline, L-phenylalanine and 4-(dimethylamino) benzoic acid can reversibly inhibit SARS-CoV M^pro with a K_i of 30 nM [87]. Chen and co-workers reported that a series of isatin (2,3-dioxindole) derivatives, which are known covalent inhibitors against rhinovirus 3C protease, could inhibit SARS-CoV M^pro activity [98]. Unfortunately, previous studies have also shown that isatin compounds displayed poor antiviral activity [97], which might limit their use in drug development against SARS-CoV. In a recent study, Yang and colleagues proposed a strategy for preventing infection by existing and possible future emerging coronaviruses. After analyzing the critical subsites of M^pro substrate binding pocket from representatives of all three groups of CoVs M^pro s, the authors concluded that all of those subsites are highly conserved. A compound (designated as N3) incorporating a trans- , -unsaturated ester with the peptidyl portion targeting this conservative region could rapidly inactivate multiple M^pro s covering all three groups of CoVs *in vitro*, and shows potent antiviral activity with extremely low cytotoxicity. A uniform inhibition mechanism was elucidated from the structures of M^pro-inhibitor complexes from SARS-CoV and TGEV. In the SARS-CoV M^pro complex with N3 (see Fig. **5**), for instance, this compound binds to the shallow cleft formed by a portion of the strand eII and a segment of the loop linking domains II and III. The C atom of the Michael acceptor forms a covalent bond with the S atom of 145-Cys. The P1, P2, P4 and P1' groups of the inhibitor insert favorably into thier corresponding subsites. It is of interest that this compound could also rapidly inactivate the M^pro s of HCoV-NL63 and HCoV-HKU1, two recently identified HCoVs associated with bronchiolitis and pneumonia. The authors expect that further modification of these inhibitors could rapidly lead to the discovery of a single agent with clinical potential against existing and possible future emerging CoV-related diseases [3].

4. OTHER THERAPEUTIC TARGETS IN THE SARS-COV

Much of the focus in structure-based drug design targeting the SARS coronavirus has so far been made on the M^pro. However, a slew of recent SARS-CoV protein structures have opened new avenues towards the design of SARS antiviral compounds, as well as for vaccine design. Of particular importance is the spike (S) glycoprotein, one of four structural proteins that are required for viral assembly. The S protein is mainly responsible for binding to the host cell and for subsequent cell entry by virus-cell membrane fusion. Great attention is also being made to identify targets involved in the replication and transcription of the virus, which was mediated by machinery composed of so-called non-structural proteins.

4.1. SARS Spike Protein Fusion Core

The SARS-CoV S protein can be subdivided into an N-terminal half (S1) and C-terminal half (S2), but without proteolytic cleavage [99]. S1 is responsible for binding to cellular receptors, while. S2 contains an internal fusion peptide and has two hydrophobic (heptad) repeat regions designated HR1 and HR2 [100].

The structure of the spike (S) protein fusion core consisting of HR1 and HR2 regions was first determined in 2004 by two groups in the post-fusion (or fusion-active) state (Fig. **6**) [99, 101]. Briefly, the structure of the S fusion core structure is a six-helix bundle in which three HR1 helices form a central coiled-coil surrounded by three HR2 helices in an oblique, anti-parallel manner. HR2 peptides pack into the hydrophobic grooves of the HR1 trimer in a mixed extended and helical conformation, representing a stable post-fusion structure, similar to that observed for HIV-1 gp41 [102].

The classical mechanism of enveloped virus and host-cell membrane fusion mediated by class I fusion proteins was first established by Don Wiley and colleagues from comprehensive studies of the influenza hemagglutinin (HA) [102, 103]. In subsequent years, extensive structural studies on the orthomyxovirus, retrovirus, paramyxovirus, and filovirus families have yielded a common fusion mechanism [102]. To confirm the value of fusion proteins as antiviral targets, one HIV-1 membrane fusion-inhibitory peptide T-20 (Trimeris, Research Triangle Park, NC, USA), targeting the prehairpin intermediate, was recently approved by the US Food and Drug Administration as a new anti-HIV drug [104]. In the case of SARS-CoV, several peptides derived from the HR1 and HR2 regions of SARS-CoV spike proteins have been synthesized to block viral entry targeting the putative pre-hairpin intermediate [104-106]. Importantly, two groups have found that only peptides derived from HR2, and not from HR1, inhibited SARS-CoV infection [104, 105]. Moreover, the efficacy of HR2 peptides derived from SARS-CoV spike protein is lower than those of corresponding HR2 peptides of murine coronavirus mouse hepatitis virus in inhibiting MHV infection [105]. One possibility might be the lower affinity of these peptides for the corresponding HR1 trimer [105], since a larger surface area is buried in the HR1-HR2 interface of MHV S2 than in SARS-CoV S2 [99]. In any case, elucidation of the HR1-HR2
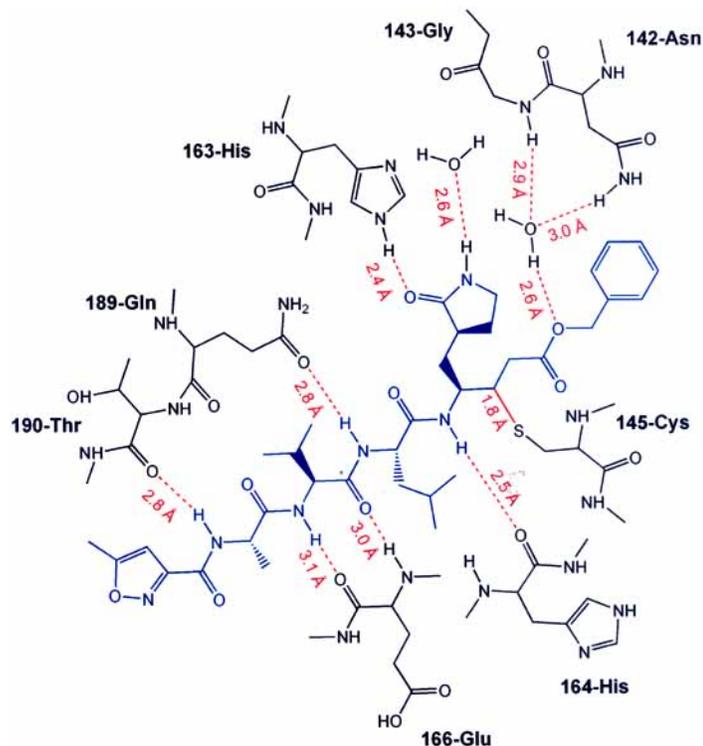
A
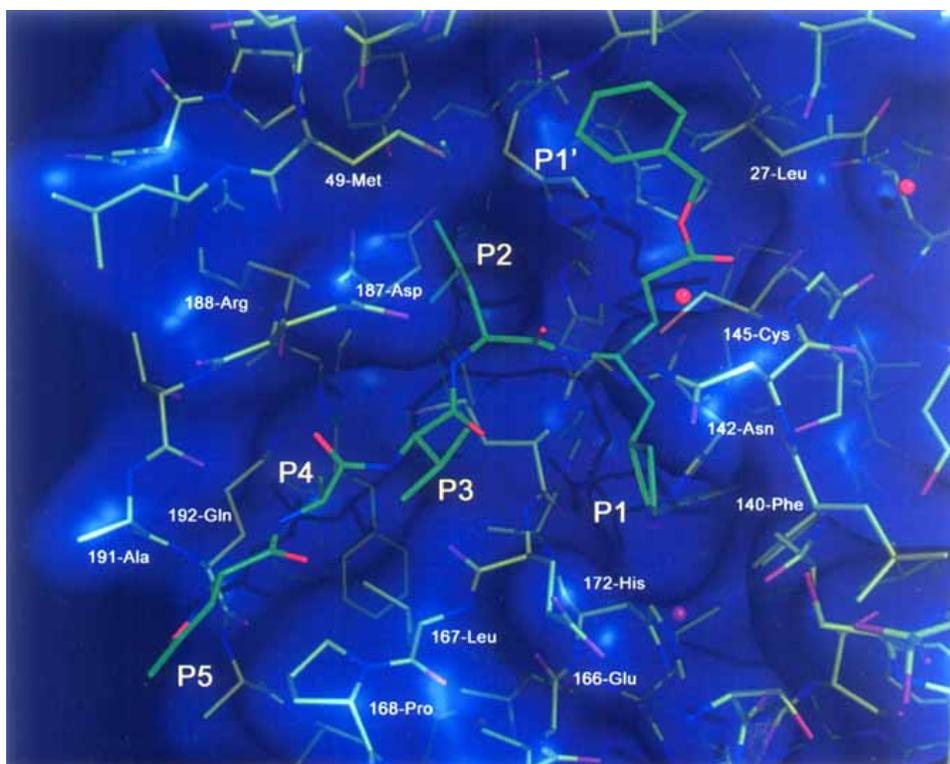
B

Fig. (5). A). Surface representation of SARS-CoV M^pro (blue, with residues shown in yellow) complexed with the N3 inhibitor (green) reported by Yang *et al*. [3]. Water molecules are shown as red spheres. The P1, P2, P3, P4, P5 and P1' groups and residues forming the substrate binding pocket are labeled. **B).** Detailed view of the interactions between SARS-CoV M^pro and the N3 inhibitor reported by Yang *et al*. [3]. The N3 inhibitor is shown in blue. Hydrogen bonds are shown as dashed lines and interaction distances are given. The covalent bond is labeled in red.

fusion core structure should prove beneficial for the discovery of viral entry inhibitors against SARS.

4.2. SARS-CoV Spike Protein Receptor Binding Domain in Complex with the Cellular Receptor ACE2

As previously described, the S1 region within the spike protein features the receptor binding site while the S2 region mediates the fusion activity. One potential SARS-CoV receptor has been identified as angiotensin-converting enzyme 2 (ACE2) [107], and a fragment of the S1 region has been shown to be required for tight binding to the peptidase domain of ACE2 [108-110]. It is this fragment, termed the receptor binding domain (RBD), which mediates the virus-receptor interaction and thus determines the viral host range and tropism. Changes in just a few residues in the RBD can result in efficient cross-species transmission [111, 112]. Furthermore, the RBD also includes important viral-neutralizing epitopes [113-116], suggesting that variants of the RBD could be used for the development of protein-based vaccines. The recent structure of the RBD in complex with the cellular receptor ACE2, determined by Stephen Harrison and colleagues [117], should therefore provide an important means for the design of effective coronavirus vaccines. From the complex structure, the authors observed tight complementarity between the shallow concave surface of the S protein RBD and one lobe of the human ACE2 peptidase domain (Fig. **6**) and identified residues critical residues for species specificity.

4.3. Non-Structural or Replicase Proteins

The SARS-CoV replicase gene encodes 16 non-structural proteins (nsps) with multiple enzymatic functions, which are known or are predicted to include types of enzymes that are common components of the replication machinery of plus-strand RNA viruses (see [25] for a review). A number of recent structures of replicase proteins, both alone and in protein-protein complexes, provide insights into the sophisticated function and assembly of the replication/transcription machinery. Targeting and inhibiting the function of this machinery should be another important focus of structure-based drug design. However, little remains known about the coronavirus replication/transcription machinery, particularly the interactions between the various non-structural protein components, and further work is needed to identify viable targets.

Nsp9, the crystal structures of which were determined simultaneously in 2004 by two groups [118, 119], have ascertained its previously unknown function as a dimeric single-stranded RNA binding protein (Fig. **6**). In addition to its structure, nsp9 has been shown to interact with nsp8, while dual-labeling studies of SARS-CoV replicase proteins have demonstrated co-localization of nsp8 with nsp2 and nsp3. Zihe Rao and colleagues have determined the first crystal structure of a complex between two non-structural proteins, nsp7 and nsp8 (Fig. **6**) [120]. Eight copies of nsp7 and eight copies of nsp8 together form an intricate scaffold that resembles a hollow cylinder. The inner dimensions and electrostatic properties of the cylinder suggest that it should encircle nucleic acid, and an interaction was demonstrated with dsRNA by EMSA and mutagenesis. The architecture

and electrostatic properties are reminiscent of PCNA or the -subunit ring, the processivity factors of DNA polymerase, suggesting that the nsp7-nsp8 complex should be a processivity factor for the RNA-dependent RNA polymerase (nsp12). Finally, the crystal structure of the ADRP domain of nsp3 has also been determined [121].

4.4. Other Targets for Anti-SARS Drug and Vaccine Design

Other targets for anti-viral therapy and vaccine design have been identified in the SARS-CoV by structural biology. For instance, antigenic peptides of the coronavirus nucleocapsid (N) protein can be recognized on the surface of infected cells by T cells [122, 123]. The structure of the MHC-I molecule HLA-A*1101 in complex with such a peptide derived from the SARS-CoV N protein, a nonamer with SARS specific sequence, could be used as a template for peptide-based vaccine design [124]. DNA vaccines targeting the SARS-CoV nucleocapsid have been generated by two different groups [21, 23].

The crystal structure of the rigorously conserved stem-loop II motif (s2m) RNA element from SARS-CoV was determined earlier this year [125]. The unusual structural features form likely surfaces for interaction with conserved host cell components or other reactive sites required for virus function. This, together with the high sequence conservation of s2m in an otherwise rapidly mutable RNA genome, implies its pathogenic importance and signals that it could be another attractive target for the design of antiviral therapeutics.

5. CONCLUSIONS

Prior to the global SARS outbreak in 2003, CoVs had not attracted enough attention from researchers because this genus of viruses causes severe diseases dominantly in animals and only comparatively mild diseases in humans. The SARS epidemic led to extensive studies of this etiological agent, the pathogenesis of SARS, and vaccine and drug development. The CoV M^pro, spike protein, helicase, RNA-dependent RNA polymerase (RdRp) and possibly other viral (or cellular) proteins involved in pathogenesis are considered to be targets for drug design. Among them, the M^pro has been characterized the most in terms of structure and biochemistry. Despite this, there are several open questions: (1) How does the M^pro release itself from the replicase polyproteins, *in cis* or *in trans*? (2) Do alternative proteolytic processing pathways exist for the replicase polyproteins in CoVs, as in equine arteritis virus? If so, how does the M^pro mediate the pathways? (3) Does the M^pro involve in pathogenesis in CoV-associated diseases? To answer these questions requires further structural and biochemical studies.

Evidence suggests that CoVs may have completed at least two animal-to-human interspecies transmissions to date [5, 6, 12, 31], resulting in the emergence of HCoV 229E and SARS-CoV. Although SARS has been contained, the discovery of the natural host for SARS-like CoV has aroused new fears for public health. CoVs, especially those that can infect hosts such as domestic animals and pets in frequent contact with humans, remain a potential threat to human health assuming they cross the interspecies barrier again.
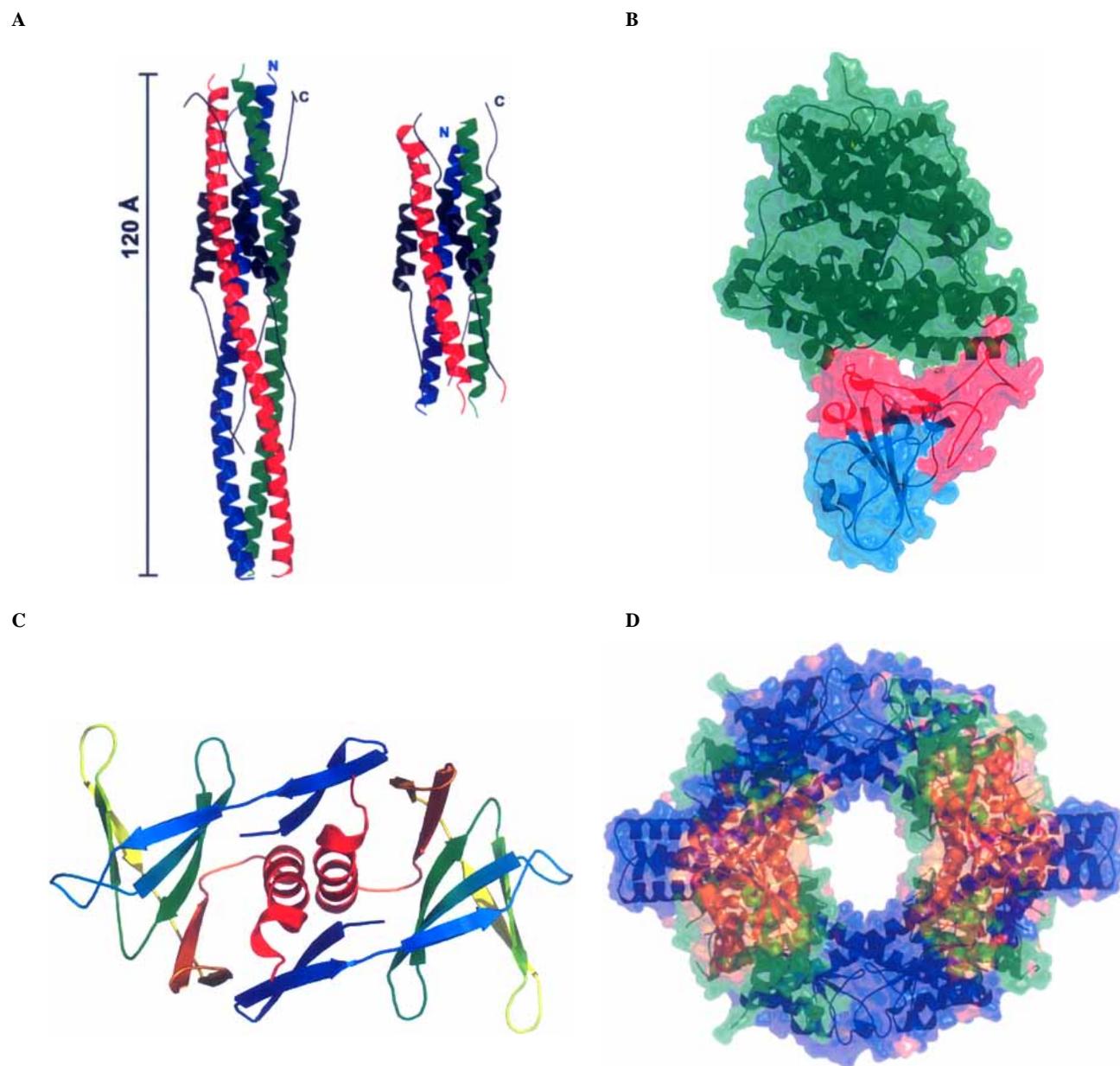
A

B

C

D

Fig. (6). A). The S protein fusion core "six-helix bundle" structures from Supekar *et al*. (left, [99]) and Xu *et al*. (right, [101]). The central HR1 peptides are shown in ribbon representation and colored red, blue and green. The HR2 peptides are shown in black. The N- and C- termini are labeled. **B).** The SARS-CoV S protein receptor binding domain (RBD) in complex with the cellular receptor ACE2 [117]. The complex structure of ACE2 (in green) and the S protein RBD (core structure in cyan and the receptor binding motif (RBM) in red) is shown in ribbon representation and covered by a transparent surface. **C).** The dimer structure of nsp9, a single stranded RNA binding protein [118, 119]. Each monomer is shown in ribbon representation and colored from blue at the N-terminal to red at the C-terminal. **D).** The structure of the nsp7-nsp8 hexadecameric complex. The complex structure is shown in ribbon representation and covered by a transparent surface. Nsp7 is shown in green and the two conformations of nsp8 are shown in blue (I) and orange (II).

Several key factors controlling the host spectrum and viral pathogenicity are highly variable among CoVs, including the requirement of different host receptors for cellular entry, poorly conserved structural proteins (antigens), and diverse accessory genes in their 3'-terminal genome regions that probably contribute to the pathogenicity of CoVs in specific hosts [2, 8, 28, 29, 107, 126-128]. This structural and func-

tional diversity presents a significant obstacle for the design of a versatile compound against all CoVs. For instance, a fusion peptide inhibitor derived from the MHV spike protein can not prevent SARS-CoV replication in cell culture [105]. Identification of the CoV M^{pro} as a conserved target among all CoVs will provide an opportunity for the development of broad-spectrum inhibitors against all CoV-related diseases.

The development of inhibitors targeting SARS-CoV M^{pro} has produced a number of positive results. Some compounds exhibit inhibitory effects on both the protease and antiviral activity. A series of inhibitors whose backbone incorporates a trans- , -unsaturated ester with a peptidyl portion seem particularly promising, and one of their analogues, ruprintrivir, has entered clinical trials against rhinovirus infection [97]. Further modification of these compounds should be expected to rapidly lead to the discovery of a single agent with clinical potential against existing and possible future emerging CoV-related diseases.

ACKNOWLEDGEMENTS

REFERENCES

References 129-131 are related articles recently published in Current Pharmaceutical Design.

[1] Siddell SG, Ziebuhr J, Snijder EJ. Coronaviruses, toroviruses, and arteriviruses. In: Mahy BWJ, ter Meulen V, editors. Topley & Wilson's Microbiology and Microbia Infections, 10th edition. London: Hodder Arnold 2005; p. 823-56.

[2] Lai MMC, Holmes KV. Coronaviridae: The viruses and their replication. In: Knipe DM, Howley PM, editors. Fields virology. 4th ed. Philadelphia: Lippincott Williams & Wilkins, 2001; p. 1163-79.

[3] Yang H, Xie W, Xue X, Yang K, Ma J, Liang W, et al. Design of wide-spectrum inhibitors targeting coronavirus main proteases. PLoS Biol 2005; 3(10): e324.

[4] Spaan WJM, Cavanagh D. Coronaviridae. Virus taxonomy, VIIIth Report of the ICTV. London: Elsevier-Academic Press, 2004; p. 945-62.

[5] Lau SKP, Woo PCY, Li KSM, Huang Y, Tsoi H-W, Wong BHL, et al. Severe acute respiratory syndrome coronavirus-like virus in Chinese horseshoe bats. Proc Natl Acad Sci USA 2005; 102(39): 14040-5.

[6] Li W, Shi Z, Yu M, Ren W, Smith C, Epstein JH, et al. Bats are natural reservoirs of SARS-like coronaviruses. Science 2005; 309: 1864.

[7] van der Hoek L, Pyrc K, Jebbink MF, Vermeulen-Oost W, Berkhout RJ, Wolthers KC, et al. Identification of a new human coronavirus. Nat Med 2004; 10(4): 368-73.

[8] Woo PC, Lau SK, Chu CM, Chan KH, Tsoi HW, Huang Y, et al. Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. J Virol 2005; 79(2): 884-95.

[9] Myint SH. Human coronavirus infections. In: Siddell SG, editor. The Coronaviridae. New York: Plenum Press 1995; p. 389-401.

[10] Gagneur A, Sizun J, Vallet S, Legr MC, Picard B, Talbot PJ. Coronavirus-related nosocomial viral respiratory infections in a neonatal and paediatric intensive care unit: a prospective study. J Hosp Infect 2002; 51(1): 59-64.

[11] Holmes KV. Coronaviruses. In: Knipe DM, Howley PM, editors. Fields' Virology. 4th ed. Philadelphia: Lippincott Williams & Wilkins 2001; p. 1187-203.

[12] Guan Y, Zheng BJ, He YQ, Liu XL, Zhuang ZX, Cheung CL, et al. Isolation and characterization of viruses related to the SARS coronavirus from animals in southern China. Science 2003; 302(5643): 276-8.

[13] Drosten C, Gunther S, Preiser W, van der Werf S, Brodt HR, Becker S, et al. Identification of a novel coronavirus in patients with severe acute respiratory syndrome. N Engl J Med 2003; 348(20): 1967-76.

[14] Ksiazek TG, Erdman D, Goldsmith CS, Zaki SR, Peret T, Emery S, et al. A novel coronavirus associated with severe acute respiratory syndrome. N Engl J Med 2003; 348(20): 1953-66.

[15] Kuiken T, Fouchier RA, Schutten M, Rimmelzwaan GF, van Amerongen G, van Riel D, et al. Newly discovered coronavirus as the primary cause of severe acute respiratory syndrome. Lancet 2003 J; 362(9380): 263-70.

[16] Peiris JS, Lai ST, Poon LL, Guan Y, Yam LY, Lim W, et al. Coronavirus as a possible cause of severe acute respiratory syndrome. Lancet 2003; 361(9366): 1319-25.

[17] Pereira HG. In: Porterfield JS Ed, ANDREWES' Viruses of Vertebrates, London, Baillière Tindall 1989; 42-57.

[18] Bisht H, Roberts A, Vogel L, Bukreyev A, Collins PL, Murphy BR, et al. Severe acute respiratory syndrome coronavirus spike protein expressed by attenuated vaccinia virus protectively immunizes mice. Proc Natl Acad Sci USA 2004; 101(17): 6641-6.

[19] Bukreyev A, Lamirande EW, Buchholz UJ, Vogel LN, Elkins WR, St Claire M, et al. Mucosal immunisation of African green monkeys (Cercopithecus aethiops) with an attenuated parainfluenza virus expressing the SARS coronavirus spike protein for the prevention of SARS. Lancet 2004; 363(9427): 2122-7.

[20] Gao W, Tamin A, Soloff A, D'Aiuto L, Nwanegbo E, Robbins PD, et al. Effects of a SARS-associated coronavirus vaccine in monkeys. Lancet 2003; 362(9399): 1895-6.

[21] Kim TW, Lee JH, Hung CF, Peng S, Roden R, Wang MC, et al. Generation and characterization of DNA vaccines targeting the nucleocapsid protein of severe acute respiratory syndrome coronavirus. J Virol 2004; 78(9): 4638-45.

[22] Yang ZY, Kong WP, Huang Y, Roberts A, Murphy BR, Subbarao K, et al. A DNA vaccine induces SARS coronavirus neutralization and protective immunity in mice. Nature 2004; 428(6982): 561-4.

[23] Zhu MS, Pan Y, Chen HQ, Shen Y, Wang XC, Sun YJ, et al. Induction of SARS-nucleoprotein-specific immune response by use of DNA vaccine. Immunol Lett 2004; 92(3): 237-43.

[24] Jiang S, He Y, Liu S. SARS vaccine development. Emerg Infect Dis 2005; 11(7): 1016-20.

[25] Ziebuhr J. The coronavirus replicase. Curr Top Microbiol Immunol 2005; 287: 57-94.

[26] Ziebuhr J, Snijder EJ, Gorbalenya AE. Virus-encoded proteinases and proteolytic processing in the Nidovirales. J Gen Virol 2000; 81(Pt 4): 853-79.

[27] Anand K, Yang H, Bartlam M, Rao Z, Hilgenfeld R. Coronavirus main proteinase: target for antiviral drug therapy. In: Schmidt A, Wolff MH, Weber O, editors. Coronaviruses with Special Emphasis on First Insights Concerning SARS. Basel: Birkhäuser Verlag 2005; p. 173-99.

[28] Rota PA, Oberste MS, Monroe SS, Nix WA, Campagnoli R, Icenogle JP, et al. Characterization of a novel coronavirus associated with severe acute respiratory syndrome. Science 2003; 300(5624): 1394-9.

[29] Marra MA, Jones SJ, Astell CR, Holt RA, Brooks-Wilson A, Butterfield YS, et al. The Genome sequence of the SARS-associated coronavirus. Science 2003; 300(5624): 1399-404.

[30] Poon LL, Chu DK, Chan KH, Wong OK, Ellis TM, Leung YH, et al. Identification of a novel coronavirus in bats. J Virol 2005; 79(4): 2001-9.

[31] Vijgen L, Keyaerts E, Moes E, Thoelen I, Wollants E, Lemey P, et al. Complete genomic sequence of human coronavirus OC43: molecular clock analysis suggests a relatively recent zoonotic coronavirus transmission event. J Virol 2005; 79(3): 1595-604.

[32] St-Jean JR, Jacomy H, Desforges M, Vabret A, Freymuth F, Talbot PJ. Human respiratory coronavirus OC43: genetic stability and neuroinvasion. J Virol 2004; 78(16): 8824-34.

[33] Penzes Z, Gonzalez JM, Calvo E, Izeta A, Smerdou C, Mendez A, et al. Complete genome sequence of transmissible gastroenteritis coronavirus PUR46-MAD clone and evolution of the purdue virus cluster. Virus Genes 2001; 23(1): 105-18.

[34] Kocherhans R, Bridgen A, Ackermann M, Tobler K. Completion of the porcine epidemic diarrhoea coronavirus (PEDV) genome sequence. Virus Genes 2001; 23(2): 137-44.

[35] Chouljenko VN, Lin XQ, Storz J, Kousoulas KG, Gorbalenya AE. Comparison of genomic and predicted amino acid sequences of respiratory and enteric bovine coronaviruses isolated from the

same animal with fatal shipping pneumonia. J Gen Virol 2001; 82(Pt 12): 2927-33.

[36] Eleouet JF, Rasschaert D, Lambert P, Levy L, Vende P, Laude H. Complete sequence (20 kilobases) of the polyprotein-encoding gene 1 of transmissible gastroenteritis virus. Virology 1995; 206(2): 817-22.

[37] Bonilla PJ, Gorbalenya AE, Weiss SR. Mouse hepatitis virus strain A59 RNA polymerase gene ORF 1a: heterogeneity among MHV strains. Virology 1994; 198(2): 736-40.

[38] Herold J, Raabe T, Schelle-Prinz B, Siddell SG. Nucleotide sequence of the human coronavirus 229E RNA polymerase locus. Virology 1993; 195(2): 680-91.

[39] Lee HJ, Shieh CK, Gorbalenya AE, Koonin EV, La Monica N, Tuler J, *et al*. The complete sequence (22 kilobases) of murine coronavirus gene 1 encoding the putative proteases and RNA polymerase. Virology 1991; 180(2): 567-82.

[40] Bredenbeek PJ, Pachuk CJ, Noten AF, Charite J, Luytjes W, Weiss SR, *et al*. The primary structure and expression of the second open reading frame of the polymerase gene of the coronavirus MHV-A59; a highly conserved polymerase is expressed by an efficient ribosomal frameshifting mechanism. Nucleic Acids Res 1990; 18(7): 1825-32.

[41] Thiel V, Herold J, Schelle B, Siddell SG. Viral replicase gene products suffice for coronavirus discontinuous transcription. J Virol 2001; 75(14): 6676-81.

[42] Anand K, Ziebuhr J, Wadhwani P, Mesters JR, Hilgenfeld R. Coronavirus main proteinase (3CLpro) structure: basis for design of anti-SARS drugs. Science 2003; 300(5626): 1763-7.

[43] Anand K, Palm GJ, Mesters JR, Siddell SG, Ziebuhr J, Hilgenfeld R. Structure of coronavirus main proteinase reveals combination of a chymotrypsin fold with an extra alpha-helical domain. EMBO J 2002; 21(13): 3213-24.

[44] Snijder EJ, Bredenbeek PJ, Dobbe JC, Thiel V, Ziebuhr J, Poon LL, *et al*. Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. J Mol Biol 2003; 331(5): 991-1004.

[45] Gorbalenya AE, Donchenko AP, Blinov VM, Koonin EV. Cysteine proteases of positive strand RNA viruses and chymotrypsin-like serine proteases. A distinct protein superfamily with a common structural fold. FEBS Lett 1989; 243(2): 103-14.

[46] Gorbalenya AE, Koonin EV, Donchenko AP, Blinov VM. Coronavirus genome: prediction of putative functional domains in the non-structural polyprotein by comparative amino acid sequence analysis. Nucleic Acids Res 1989; 17(12): 4847-61.

[47] Liu DX, Brierley I, Tibbles KW, Brown TD. A 100-kilodalton polypeptide encoded by open reading frame (ORF) 1b of the coronavirus infectious bronchitis virus is processed by ORF 1a products. J Virol 1994; 68(9): 5772-80.

[48] Lu Y, Lu X, Denison MR. Identification and characterization of a serine-like proteinase of the murine coronavirus MHV-A59. J Virol 1995; 69(6): 3554-9.

[49] Ziebuhr J, Herold J, Siddell SG. Characterization of a human coronavirus (strain 229E) 3C-like proteinase activity. J Virol 1995; 69(7): 4331-8.

[50] Yang H, Yang M, Ding Y, Liu Y, Lou Z, Zhou Z, *et al*. The crystal structures of severe acute respiratory syndrome virus main protease and its complex with an inhibitor. Proc Natl Acad Sci USA 2003; 100(23): 13190-5.

[51] Hsu M-F, Kuo C-J, Chang K-T, Chang H-C, Chou C-C, Ko T-P, *et al*. Mechanism of the maturation process SARS-CoV 3CL protease. J Biol Chem 2005; 280: 31257-66.

[52] Lee T-W, Cherney MM, Huitema C, Liu J, James KE, Powers JC, *et al*. Crystal structures of the main peptidase from the SARS coronavirus inhibited by a substrate-like aza-peptide epoxide. J Mol Biol 2005; 353(5): 1137-51.

[53] Tan J, Verschueren KHG, Anand K, Shen J, Yang M, Xu Y, *et al*. pH-dependent conformational flexibility of the SARS-CoV main proteinase (Mpro) dimer: Molecular dynamics simulations and multiple X-ray structure analyses. J Mol Biol 2005; 354(1): 25-40.

[54] Gorbalenya AE, Snijder EJ, Spaan WJ. Severe acute respiratory syndrome coronavirus phylogeny: toward consensus. J Virol 2004; 78(15): 7863-6.

[55] Allaire M, Chernaia MM, Malcolm BA, James MN. Picornaviral 3C cysteine proteinases have a fold similar to chymotrypsin-like serine proteinases. Nature 1994; 369(6475): 72-6.

[56] Matthews DA, Smith WW, Ferre RA, Condon B, Budahazi G, Sisson W, *et al*. Structure of human rhinovirus 3C protease reveals a trypsin-like polypeptide fold, RNA-binding site, and means for cleaving precursor polyprotein. Cell 1994; 77(5): 761-71.

[57] Mosimann SC, Cherney MM, Sia S, Plotch S, James MN. Refined X-ray crystallographic structure of the poliovirus 3C gene product. J Mol Biol 1997; 273(5): 1032-47.

[58] Birtley JR, Knox SR, Jaulent AM, Brick P, Leatherbarrow RJ, Curry S. Crystal Structure of Foot-and-Mouth Disease Virus 3C Protease: New insights into catalytic mechanism and cleavage specificity. J Biol Chem 2005; 280(12): 11520-7.

[59] Fan K, Wei P, Feng Q, Chen S, Huang C, Ma L, *et al*. Biosynthesis, purification, and substrate specificity of severe acute respiratory syndrome coronavirus 3C-like proteinase. J Biol Chem 2004; 279(3): 1637-42.

[60] Bacha U, Barrila J, Velazquez-Campoy A, Leavitt SA, Freire E. Identification of novel inhibitors of the SARS coronavirus main protease 3CLpro. Biochemistry 2004; 43(17): 4906-12.

[61] Chen S, Chen L, Tan J, Chen J, Du L, Sun T, *et al*. Severe acute respiratory syndrome coronavirus 3C-like proteinase N terminus is indispensable for proteolytic activity but not for enzyme dimerization. Biochemical and thermodynamic investigation in conjunction with molecular dynamics simulations. J Biol Chem 2005; 280(1): 164-73.

[62] Chou CY, Chang HC, Hsu WC, Lin TZ, Lin CH, Chang GG. Quaternary structure of the severe acute respiratory syndrome (SARS) coronavirus main protease. Biochemistry 2004; 43(47): 14958-70.

[63] Hsu WC, Chang HC, Chou CY, Tsai PJ, Lin PI, Chang GG. Critical assessment of important regions in the subunit association and catalytic action of the severe acute respiratory syndrome coronavirus main protease. J Biol Chem 2005; 280(24): 22741-8.

[64] Shi J, Wei Z, Song J. Dissection study on the severe acute respiratory syndrome 3C-like protease reveals the critical role of the extra domain in dimerization of the enzyme: defining the extra domain as a new target for design of highly specific protease inhibitors. J Biol Chem 2004; 279(23): 24765-73.

[65] Kuang WF, Chow LP, Wu MH, Hwang LH. Mutational and inhibitive analysis of SARS coronavirus 3C-like protease by fluorescence resonance energy transfer-based assays. Biochem Biophys Res Commun 2005; 331(4): 1554-9.

[66] Siddell S, Sawicki D, Meyer Y, Thiel V, Sawicki S. Identification of the mutations responsible for the phenotype of three MHV RNA-negative ts mutants. Adv Exp Med Biol 2001; 494: 453-8.

[67] Liu DX, Brown TD. Characterisation and mutational analysis of an ORF 1a-encoding proteinase domain responsible for proteolytic processing of the infectious bronchitis virus 1a/1b polyprotein. Virology 1995; 209(2): 420-7.

[68] Ziebuhr J, Heusipp G, Siddell SG. Biosynthesis, purification, and characterization of the human coronavirus 229E 3C-like proteinase. J Virol 1997; 71(5): 3992-7.

[69] Tibbles KW, Brierley I, Cavanagh D, Brown TD. Characterization *in vitro* of an autocatalytic processing activity associated with the predicted 3C-like proteinase domain of the coronavirus avian infectious bronchitis virus. J Virol 1996; 70(3): 1923-30.

[70] Seybert A, Ziebuhr J, Siddell SG. Expression and characterization of a recombinant murine coronavirus 3C-like proteinase. J Gen Virol 1997; 78(Pt 1): 71-5.

[71] Huang C, Wei P, Fan K, Liu Y, Lai L. 3C-like proteinase from SARS coronavirus catalyzes substrate hydrolysis by a general base mechanism. Biochemistry 2004; 43(15): 4568-74.

[72] Hegyi A, Ziebuhr J. Conservation of substrate specificities among coronavirus main proteases. J Gen Virol 2002; 83(Pt 3): 595-9.

[73] Chen L, Gui C, Luo X, Yang Q, Gunther S, Scandella E, *et al*. Cinanserin is an inhibitor of the 3C-like proteinase of severe acute respiratory syndrome coronavirus and strongly reduces virus replication *in vitro*. J Virol 2005; 79(11): 7095-103.

[74] Hegyi A, Friebe A, Gorbalenya AE, Ziebuhr J. Mutational analysis of the active centre of coronavirus 3C-like proteases. J Gen Virol 2002; 83(Pt 3): 581-93.

[75] Lu X, Lu Y, Denison MR. Intracellular and *in vitro*-translated 27-kDa proteins contain the 3C-like proteinase activity of the coronavirus MHV-A59. Virology 1996; 222(2): 375-82.

[76] Pinon JD, Mayreddy RR, Turner JD, Khan FS, Bonilla PJ, Weiss SR. Efficient autoproteolytic processing of the MHV-A59 3C-like

proteinase from the flanking hydrophobic domains requires membranes. Virology 1997; 230(2): 309-22.

[77] Ziebuhr J, Siddell SG. Processing of the human coronavirus 229E replicase polyproteins by the virus-encoded 3C-like proteinase: identification of proteolytic products and cleavage sites common to pp1a and pp1ab. J Virol 1999; 73(1): 177-85.

[78] Fan K, Ma L, Han X, Liang H, Wei P, Liu Y, *et al*. The substrate specificity of SARS coronavirus 3C-like proteinase. Biochem Biophys Res Commun 2005; 329(3): 934-40.

[79] Liu Z, Huang C, Fan K, Wei P, Chen H, Liu S, *et al*. Virtual screening of novel noncovalent inhibitors for SARS-CoV 3C-like proteinase. J Chem Inf Model 2005; 45(1): 10-7.

[80] Jain RP, Pettersson HI, Zhang J, Aull KD, Fortin PD, Huitema C, *et al*. Synthesis and evaluation of keto-glutamine analogues as potent inhibitors of severe acute respiratory syndrome 3CLpro. J Med Chem 2004; 47(25): 6113-6.

[81] Kao RY, Tsui WH, Lee TS, Tanner JA, Watt RM, Huang JD, *et al*. Identification of novel small-molecule inhibitors of severe acute respiratory syndrome-associated coronavirus by chemical genetics. Chem Biol 2004; 11(9): 1293-9.

[82] Wu C-Y, Jan J-T, Ma S-H, Kuo C-J, Juan H-F, Cheng Y-SE, *et al*. Small molecules targeting severe acute respiratory syndrome human coronavirus. Proc Natl Acad Sci USA 2004; 101(27): 10012-7.

[83] Blanchard JE, Elowe NH, Huitema C, Fortin PD, Cechetto JD, Eltis LD, *et al*. High-throughput screening identifies inhibitors of the SARS coronavirus main proteinase. Chem Biol 2004; 11(10): 1445-53.

[84] Chen L-R, Wang Y-C, Lin YW, Chou S-Y, Chen S-F, Liu LT, *et al*. Synthesis and evaluation of isatin derivatives as effective SARS coronavirus 3CL protease inhibitors. Bioorg Med Chem Lett 2005; 15(12): 3058-62.

[85] Martina E, Stiefl N, Degel B, Schulz F, Breuning A, Schiller M, *et al*. Screening of electrophilic compounds yields an aziridinyl peptide as new active-site directed SARS-CoV main protease inhibitor. Bioorg Med Chem Lett 2005; 15(24): 5365-9.

[86] Shie J-J, Fang J-M, Kuo T-H, Kuo C-J, Liang P-H, Huang H-J, *et al*. Inhibition of the severe acute respiratory syndrome 3CL protease by peptidomimetic   , -unsaturated esters. Bioorg Med Chem 2005; 13(17): 5240-52.

[87] Shie JJ, Fang JM, Kuo CJ, Kuo TH, Liang PH, Huang HJ, *et al*. Discovery of potent anilide inhibitors against the severe acute respiratory syndrome 3CL protease. J Med Chem 2005; 48(13): 4469-73.

[88] Liu YC, Huang V, Chao TC, Hsiao CD, Lin A, Chang MF, *et al*. Screening of drugs by FRET analysis identifies inhibitors of SARS-CoV 3CL protease. Biochem Biophys Res Commun 2005; 333(1): 194-9.

[89] Parera M, Clotet B, Martinez MA. Genetic screen for monitoring severe acute respiratory syndrome coronavirus 3C-like protease. J Virol 2004; 78(24): 14057-61.

[90] Jenwitheesuk E, Samudrala R. Identifying inhibitors of the SARS coronavirus proteinase. Bioorg Med Chem Lett 2003; 13(22): 3989-92.

[91] Toney JH, Navas-Martin S, Weiss SR, Koeller A. Sabadinine: a potential non-peptide anti-severe-acute-respiratory-syndrome agent identified using structure-aided design. J Med Chem 2004; 47(5): 1079-80.

[92] Chou K-C, Wei D-Q, Zhong W-Z. Binding mechanism of coronavirus main proteinase with ligands and its implication to drug design against SARS. Biochem Biophys Res Commun 2003; 308(1): 148-51.

[93] Xiong B, Gui CS, Xu XY, Luo C, Chen J, Luo HB, *et al*. A 3D model of SARS-CoV 3CL proteinase and its inhibitors design by virtual screening. Acta Pharmacol Sin 2003; 24(6): 497-504.

[94] Hsu JT, Kuo CJ, Hsieh HP, Wang YC, Huang KK, Lin CP, *et al*. Evaluation of metal-conjugated compounds as inhibitors of 3CL protease of SARS-CoV. FEBS Lett 2004; 574(1-3): 116-20.

[95] Chen CN, Lin CP, Huang KK, Chen WC, Hsieh HP, Liang PH, *et al*. Inhibition of SARS-CoV 3C-like Protease Activity by Theaflavin-3,3'-digallate (TF3). Evid Based Complement Alternat Med 2005; 2(2): 209-15.

[96] Lin CW, Tsai FJ, Tsai CH, Lai CC, Wan L, Ho TY, *et al*. Anti-SARS coronavirus 3C-like protease effects of Isatis indigotica root and plant-derived phenolic compounds. Antiviral Res 2005; 68(1): 36-42.

[97] Matthews DA, Dragovich PS, Webber SE, Fuhrman SA, Patick AK, Zalman LS, *et al*. Structure-assisted design of mechanism-based irreversible inhibitors of human rhinovirus 3C protease with potent antiviral activity against multiple rhinovirus serotypes. Proc Natl Acad Sci USA 1999; 96(20): 11000-7.

[98] Chen LR, Wang YC, Lin YW, Chou SY, Chen SF, Liu LT, *et al*. Synthesis and evaluation of isatin derivatives as effective SARS coronavirus 3CL protease inhibitors. Bioorg Med Chem Lett 2005; 15(12): 3058-62.

[99] Supekar VM, Bruckmann C, Ingallinella P, Bianchi E, Pessi A, Carfi A. Structure of a proteolytically resistant core from the severe acute respiratory syndrome coronavirus S2 fusion protein. Proc Natl Acad Sci USA 2004; 101(52): 17958-63.

[100] de Groot RJ, Luytjes W, Horzinek MC, van der Zeijst BA, Spaan WJ, Lenstra JA. Evidence for a coiled-coil structure in the spike proteins of coronaviruses. J Mol Biol 1987; 196(4): 963-6.

[101] Xu Y, Lou Z, Liu Y, Pang H, Tien P, Gao GF, *et al*. Crystal structure of severe acute respiratory syndrome coronavirus spike protein fusion core. J Biol Chem 2004; 279(47): 49414-9.

[102] Eckert DM, Kim PS. Mechanisms of viral membrane fusion and its inhibition. Annu Rev Biochem 2001; 70: 777-810.

[103] Skehel JJ, Wiley DC. Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. Annu Rev Biochem 2000; 69: 531-69.

[104] Liu S, Xiao G, Chen Y, He Y, Niu J, Escalante CR, *et al*. 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 2004; 363(9413): 938-47.

[105] Bosch BJ, Martina BE, Van Der Zee R, Lepault J, Haijema BJ, Versluis C, *et al*. Severe acute respiratory syndrome coronavirus (SARS-CoV) infection inhibition using spike protein heptad repeat-derived peptides. Proc Natl Acad Sci USA 2004; 101(22): 8455-60.

[106] Yuan K, Yi L, Chen J, Qu X, Qing T, Rao X, *et al*. Suppression of SARS-CoV entry by peptides corresponding to heptad regions on spike glycoprotein. Biochem Biophys Res Commun 2004; 319(3): 746-52.

[107] Li W, Moore MJ, Vasilieva N, Sui J, Wong SK, Berne MA, *et al*. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. Nature 2003; 426(6965): 450-4.

[108] Babcock GJ, Esshaki DJ, Thomas WD Jr, Ambrosino DM. Amino acids 270 to 510 of the severe acute respiratory syndrome coronavirus spike protein are required for interaction with receptor. J Virol 2004; 78(9): 4552-60.

[109] Wong SK, Li W, Moore MJ, Choe H, Farzan M. A 193-amino acid fragment of the SARS coronavirus S protein efficiently binds angiotensin-converting enzyme 2. J Biol Chem 2004; 279(5): 3197-201.

[110] Xiao X, Chakraborti S, Dimitrov AS, Gramatikoff K, Dimitrov DS. The SARS-CoV S glycoprotein: expression and functional characterization. Biochem Biophys Res Commun 2003; 312(4): 1159-64.

[111] Li W, Zhang C, Sui J, Kuhn JH, Moore MJ, Luo S, *et al*. Receptor and viral determinants of SARS-coronavirus adaptation to human ACE2. EMBO J 2005; 24(8): 1634-43.

[112] Song HD, Tu CC, Zhang GW, Wang SY, Zheng K, Lei LC, *et al*. Cross-host evolution of severe acute respiratory syndrome coronavirus in palm civet and human. Proc Natl Acad Sci USA 2005; 102(7): 2430-5.

[113] He Y, Lu P, Siddiqui Y, Zhou Y, Jiang S. Receptor-binding domain of severe acute respiratory syndrome coronavirus spike protein contains multiple conformation-dependent epitopes that induce highly potent neutralizing antibodies. J Immunol 2005; 174: 4908.

[114] Pang H, Liu Y, Han X, Xu Y, Jiang F, Wu D, *et al*. Protective humoral responses to severe acute respiratory syndrome-associated coronavirus: implications for the design of an effective protein-based vaccine. J Gen Virol 2004; 85: 3109-13.

[115] Sui J, Li W, Murakami A, Tamin A, Matthews LJ, Wong SK, *et al*. Potent neutralization of severe acute respiratory syndrome (SARS) coronavirus by a human mAb to S1 protein that blocks receptor association. Proc Natl Acad Sci USA 2004; 101: 2536-41.

[116] van den Brink EN, Ter Meulen J, Cox F, Jongeneelen MA, Thijsse A, Throsby M, *et al*. Molecular and biological characterization of human monoclonal antibodies binding to the spike and nucleocap-

sid proteins of severe acute respiratory syndrome coronavirus. J Virol 2005; 79: 1635-44.

[117] Li F, Li W, Farzan M, Harrison SC. Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. Science 2005; 309(5742): 1864-8.

[118] Egloff MP, Ferron F, Campanacci V, Longhi S, Rancurel C, Dutartre H, *et al*. The severe acute respiratory syndrome-coronavirus replicative protein nsp9 is a single-stranded RNA-binding subunit unique in the RNA virus world. Proc Natl Acad Sci USA 2004; 101(11): 3792-6.

[119] Sutton G, Fry E, Carter L, Sainsbury S, Walter T, Nettleship J, *et al*. The nsp9 replicase protein of SARS-coronavirus, structure and functional insights. Structure (Camb) 2004; 12(2): 341-53.

[120] Zhai Y, Sun F, Li X, Pang H, Xu X, Bartlam M, *et al*. Insights into coronavirus transcription and replication from the structure of the SARS-CoV nsp7-nsp8 hexadecamer. Nat Struct Mol Biol 2005; 12(11): 980-6.

[121] Saikatendu KS, Joseph JS, Subramanian V, Clayton T, Griffith M, Moy K, *et al*. Structural basis of severe acute respiratory syndrome coronavirus ADP-ribose-1"-phosphate dephosphorylation by a conserved domain of nsp3. Structure (Camb) 2005; 13(11): 1665-75.

[122] Bergmann C, McMillan M, Stohlman S. Characterization of the Ld-restricted cytotoxic T-lymphocyte epitope in the mouse hepatitis virus nucleocapsid protein. J Virol 1993; 67(12): 7041-9.

[123] Boots AM, Van Lierop MJ, Kusters JG, Van Kooten PJ, Van der Zeijst BA, Hensen EJ. MHC class II-restricted T-cell hybridomas recognizing the nucleocapsid protein of avian coronavirus IBV. Immunology 1991; 72(1): 10-4.

[124] Blicher T, Kastrup JS, Buus S, Gajhede M. High-resolution structure of HLA-A*1101 in complex with SARS nucleocapsid peptide. Acta Crystallogr D Biol Crystallogr 2005; 61(Pt 8): 1031-40.

[125] Robertson MP, Igel H, Baertsch R, Haussler D, Ares M Jr, Scott WG. The structure of a rigorously conserved RNA element within the SARS virus genome. PLoS Biol 2005; 3(1): e5.

[126] de Haan CA, Masters PS, Shen X, Weiss S, Rottier PJ. The group-specific murine coronavirus genes are not essential, but their deletion, by reverse genetics, is attenuating in the natural host. Virology 2002; 296(1): 177-89.

[127] Hofmann H, Pyrc K, van der Hoek L, Geier M, Berkhout B, Pohlmann S. Human coronavirus NL63 employs the severe acute respiratory syndrome coronavirus receptor for cellular entry. Proc Natl Acad Sci USA 2005.

[128] Jeffers SA, Tusell SM, Gillim-Ross L, Hemmila EM, Achenbach JE, Babcock GJ, *et al*. CD209L (L-SIGN) is a receptor for severe acute respiratory syndrome coronavirus. Proc Natl Acad Sci USA 2004; 101(44): 15748-53.

[129] Debnath AK. Application of 3D-QSAR techniques in anti-HIV-1 drug design--an overview. Curr Pharm Des 2005; 11(24): 3091-110.

[130] Imamichi T. Action of anti-HIV drugs and resistance: reverse transcriptase inhibitors and protease inhibitors. Curr Pharm Des 2004; 10(32): 4039-53.

[131] Yusa K, Harada S. Acquisition of multi-PI (protease inhibitor) resistance in HIV-1 *in vivo* and *in vitro*. Curr Pharm Des 2004; 10(32): 4055-64.