Current Pharmaceutical Design, 2006, *12*, 4555-4564

Quaternary Structure, Substrate Selectivity and Inhibitor Design for SARS 3C-Like Proteinase

Luhua Lai[1,2,*], Xiaofeng Han[1,2], Hao Chen[1,2], Ping Wei[1], Changkang Huang[1], Shiyong Liu[1,2], Keqiang Fan[1], Lu Zhou[1], Zhenming Liu[1], Jianfeng Pei[1] and Ying Liu[1]

[1]*State Key Laboratory for Structural Chemistry of Unstable and Stable Species, College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China and* [2]*Center for Theoretical Biology, Peking University, Beijing 100871, China*

Abstract: The SARS coronavirus 3C-like proteinase is recognized as a potential drug design target for the treatment of severe acute respiratory syndrome. In the past few years, much work has been done to understand the catalytic mechanism of this target protein and to design its selective inhibitors. The protein exists as a dimer/monomer mixture in solution and the dimer was confirmed to be the active species for the enzyme reaction. Quantitative dissociation constants have been reported for the dimer by using analytic ultracentrifuge, gel filtration and enzyme assays. Though the enzyme is a cysteine protease with a chymotrypsin fold, SARS 3C-like proteinase follows the general base catalytic mechanism similar to chymotrypsin. As the enzyme can cut eleven different sites on the viral polyprotein, the substrate specificity has been studied by synthesized peptides corresponding or similar to the cleavage sites on the polyprotein. Predictive model was built for substrate structure and activity relationships and can be applied in inhibitor design. Due to the lack of potential drugs for the treatment of SARS, the discovery of inhibitors against SARS 3C-like proteinase, which can potentially be optimized as drugs appears to be highly desirable. Various groups have been working on inhibitor discovery by virtual screening, compound library screening, modification of existing compounds or natural products. High-throughput *in vitro* assays, auto-cleavage assays and viral replication assays have been developed for inhibition activity tests. Inhibitors with IC_{50} values as low as 60 nM have been reported.

Key Words: SARS 3C-like proteinase, quaternary structure, enzyme catalytic mechanism, substrate selectivity, inhibitor design.

INTRODUCTION

Severe acute respiratory syndrome (SARS) is a respiratory illness that emerged in China and rapidly spread in Asia, North America, and Europe in early 2003 [1, 2]. Many evidences have shown that a novel coronavirus must be the cause of SARS [3]. The genome of SARS coronavirus has been sequenced within a short period of time after the virus was confirmed [4, 5]. Coronaviruses are members of positive-stranded RNA viruses featuring the largest viral RNA genomes up to date. The SARS coronavirus replicase gene encompasses two overlapping translation products, polyprotein 1a (~450kDa) and 1b(~750kD), which are conserved both in length and amino acid sequence in all coronavirus replicase proteins. Polyprotein 1a and 1b are cleaved by the internally encoded 3C-like proteinase, releasing functional proteins necessary for virus replication. The SARS 3C-like proteinase is fully conserved among all the released SARS coronavirus genome sequences and is highly homologous with other coronavirus 3C-like proteinase. Due to its functional importance in the viral life cycle, SARS 3C-like proteinase has been proposed to be a key target for structural based drug design against SARS [6]. Homology modeling for the SARS 3C-like proteinase has been performed by several groups [6-8] and the conformational flexibility of the substrate binding site has been studied [7].

Two crystal structures of coronavirus 3C-like proteinase from transmissible gastroenteritis virus (TGEV) [9] and human coronavirus (hCoV) 229E [6] have been determined. Crystal structure of SARS coronavirus 3CL proteinase was solved in late 2003 by two groups with the coordinates deposited in the Protein Data Bank (http://www.rcsb.org, PDB code: 1Q2W, 1UJ1, 1UK2). In addition, Rao's group also elucidated the complex structure of SARS 3C-like proteinase with a covalent bonded substrate analog, which provided direct insights on the substrate binding site [10]. More recently, Hsu *et al*. published the crystal structure of the C145A mutant and found that the active site in one protomer of the dimer is bound with the C-terminal six residues from another protein molecule [11]. All of the reported structures of coronavirus 3C-like proteinases are topologically similar. The first two domains form a chymotrypsin fold, which is responsible for the catalytic reaction, while the third domain is helical with unclear biological function. Coronavirus 3C-like proteinase shares the chymotrypsin fold with the 3C proteinases from other viruses like rhinovirus, picornavirus and hepatitis A virus [12, 13]. 3C proteinase of rhinovirus has been targeted to develop drugs against common cold [14-18].

*Address correspondence to this author at the State Key Laboratory for Structural Chemistry of Unstable and Stable Species, College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China; Tel: 86-10-62757486; Fax: 86-10-62751725; E-mail: lhlai@pku.edu.cn

1381-6128/06 $50.00+.00 © 2006 Bentham Science Publishers Ltd.

In the present review, we will summarize the major advances on the quaternary structure, catalytic mechanism, substrate specificity and inhibitor design for SARS 3C-like proteinase.

QUATERNARY STRUCTURE OF SARS 3C-LIKE PROTEINASE

As both the crystal structures of the 3C-like proteinase in TGEV and human coronavirus give dimer structure and the residues at dimeric interface are conserved in coronavirus, it has been proposed that the dimer may be the biologically functional form of the protein [6, 9]. Dynamic light scattering experiment shows that both HCoV 229E and TGEV 3C-like proteinases exist as a mixture of monomer (65%) and dimer (35%) at a concentration of 1-2mg/ml [6]. By analytical gel filtration [19], we have reported that SARS 3C-like proteinase showed two peaks corresponding to the monomeric and dimeric form at high concentration and one monomeric peak at low concentration. A peptide cleavage assay was used to test the proteolysis activity of SARS 3C-like proteinase at different enzyme concentrations. The observed k_{cat}/K_m increased linearly as the enzyme concentration was increased. This led to the conclusion that the monomeric form of SARS 3C-like proteinase has no detectable catalytic activity, while the dimer must be the active form.

Several groups have quantitatively measured the dimer dissociation constant of the SARS 3C-like proteinase by analytical ultracentrifuge or enzyme activity dependent assays. The obtained dissociation constants have large variations from different groups and/or using proteins expressed from different constructs. Chou *et al*. reported a dissociation constant of 190 ± 14 nM by using the sedimentation velocity method with the C-terminal His tagged form of SARS 3C-like proteinase [20]. Hsu *et al*. reported a dissociation constant of 0.35nM by using the sedimentation velocity method on the wild-type protein with no tags [11]. Generally, the sedimentation equilibrium test can provide first-principle analysis of the dissociation constants compared to sedimentation velocity method [21, 22]. In addition to sedimentation velocity experiments, we also used sedimentation equilibrium experiments to determine the dissociation constants [23]. Our sedimentation equilibrium experiments give a dissociation constant of 14 μM for the wild-type protein with no tag but one extra residue at the N-terminus. If we assume that the monomers in the assay have no activity, the dissociation constant can also be deduced from the enzyme concentration dependent activity assay. Kuo *et al*. reported a dissociation constant of 15 ± 4 nM for SARS-CoV 3CL proteinase by the enzyme concentration dependent activity assay using a 14mer peptide substrate with FRET groups at both ends [24]. Similarly, using a hexa-peptide chromogenic substrate, we determined the apparent dissociation constant of SARS 3C-like proteinase to be 0.81 μM through the enzyme concentration dependent activity assay [25]. It is well known that substrate can induce and stabilize protein dimerization for dimeric enzymes; the reported difference in apparent dissociation constant could be partially attributed to the effects of different substrates [26, 27].

As the dissociation constants from different experiments have large variance, we also analyzed the dimeric interface and calculated the dissociation constants using a statistical potential [28]. For the crystal structure of SARS 3C-like proteinase at pH 8 [10], the buried surface area of the dimer is 1245 $Å^2$ with the buried hydrophobic area of 732 $Å^2$. Analysis of protein homodimers shows that most homodimers have buried surface between 368 $Å^2$ to 4746 $Å^2$ with a linear increase along with the molecular weight [29]. For a protein with molecular weight around 30K, the buried surface is usually around 1500 $Å^2$; the 1245 $Å^2$ of a dimer interface may imply for relatively weak binding. Based on an atomic based statistical potential function [28] the dissociation constant for the SARS 3CL proteinase dimer was calculated. The potential was shown to give good predictions for dissociation constants of protein-protein complexes. Our calculation gave a dissociation constant of approximately 6 μM, which is in agreement with the measured dissociation constant (14 μM) from our sedimentation equilibrium experiments. Calculations from the PQS server (http://pqs.ebi.ac.uk/) using 1uk2 coordinates [10] gave an approximate dissociation constant of 1.6 μM by estimating the solvation free energy of binding. Both the calculation from potential of mean force and the buried surface analysis imply for a relatively weak dimer.

Although we know from the above studies that only the dimeric form of SARS 3C-like proteinase is active, the reasons why the monomer is inactive and how the dimer regulate the enzyme activity are not clear. We have carried out molecular dynamics simulation on the separate monomers and the dimer. The simulations showed that the active site residues and the substrate binding pocket are not in the right conformation for catalysis in the monomers [25]. The dimer formation is essential for maintaining a right conformation for the enzyme catalysis. Compounds that can interfere with the dimeric interface may act as potent inhibitors. As the N-terminal peptide is buried between the two monomers, we synthesized the N-terminal octa-peptide and tested its activity. The peptide was found to bind with the protein and behave as a weak inhibitor when competing with the large protein monomer molecule [23, 30]. Modifications on the N-terminal peptide or designing other compounds targeting the dimer interface may discover potent and selective inhibitors targeting SARS 3C-like proteinase.

CATALYTIC MECHANISM OF SARS 3C-LIKE PROTEINASE

As 3C proteinases and the catalytic domain in 3C-like proteinases are all cysteine proteinases with a chymotrypsin fold, it is interesting to explore whether they follow a thiolate-imidazolium ion pair catalytic mechanism as found with papain [31-34]. Recently, it was discovered that 3C proteinases from picornain virus, poliovirus and sortase from *S. aureus* do not contain a thiolate-imidazolium ion pair in their active site [35-37]. Compared to the catalytic triad in 3C proteinase, only two catalytic residues were found in 3C-like proteinases of coronaviruses [6, 9, 38]. Whether the 3C-like proteinase of coronavirus follows an ion-pair mechanism or a general base catalytic mechanism still needs consideration. This is important to understand the kinetic mechanism, and to carry out structural and/or mechanism based drug design.

We have used a colorimetric assay with a synthetic substrate for SARS 3C-like proteinase to study the enzyme catalytic mechanism [39]. While His41 and Cys145 were confirmed to be crucial for catalysis by their mutation to Ala, the Cys14/Ser mutation produced an active enzyme. Fig. (**1**) shows the pH-rate profile for SARS 3CL proteinase. The optimum enzymatic reaction pHs are 7.38, 7.60 for the wild-type SARS 3CL proteinase and the C145S mutant. The pKa values for His41 and Cys145 in the wild-type enzyme are calculated to be 6.38±0.02, 8.34±0.02, and the pKa's for His41 and Ser145 in C145S mutant are 6.15±0.05, 9.09± 0.04. The right limb of the pH-rate profile decreases sharply for the wild-type enzyme above 8.6 due to the dissociation of thiol group, which is essential for catalysis. The thiol proton should be transferred to the imidazole ring to act as a general acid, subsequently protonating the leaving group of the tetrahedral intermediate. Due to the higher pKa of the OH group in serine, the pH-rate profile for the C145S variant decreases slowly on the high pH side. This phenomenon is similar to what has been observed for poliovirus proteinase 3C [36, 37], indicating that SARS 3C-like proteinase takes a similar catalytic mechanism.

The pH-dependent catalytic activities, the calculated pKa values for His41, Cys145 in the wild-type enzyme and for His41, Ser145 in the C145S mutant, all agree that the enzyme catalysis follows a general base catalytic mechanism and does not contain a thiolate-imidazolium ion pair in its active site.

Shan *et al*. also studied the enzyme catalytic mechanism by using an auto-cleavage assay [40]. Their study confirmed that His41 and Cys145 are the two active site residues and the Cys145 to Ser145 mutant was active for *cis*-cleavage. From both our and Shan's studies, we can conclude that SARS 3CL proteinase follows general base catalytic mechanism shared by the serine protease family both in *cis*- or in *trans*-cleavage assays.

SUBSTRATE SPECIFICITY OF SARS 3C-LIKE PROTEINASE

Similar to other coronavirus, sequence analysis reveals 11 cleavage sites of the 3C-like proteinase on the SARS polyprotein. In order to study the substrate specificity of SARS 3C-like proteinase, we investigated its activity towards 11 synthesized peptides covering the 11 cleavage sites on the virus polyprotein by using HPLC assay [19]. The relative activity of the 11 peptides are listed in Table **1**. The P1, P2 and P1' positions were found to be the main determinants for substrate specificity of coronavirus 3C-like proteinase [38]. The P1 position has a well conserved Gln residue and P2 a hydrophobic one. Unlike other previously identified coronavirus 3C-like proteinases, which have Leu/Ile at position P2, SARS 3C-like proteinase also tolerates Phe, Val, Met residues at P2 position. Our results showed that the enzyme can cut the 11 peptides covering all the 11 cleavage sites on the viral polyprotein, but with different efficiencies. The two peptides derived from the N-terminal and C-terminal self-cleavage sites are the two most active substrates for SARS 3C-like proteinase cleavage, implying that self-cleavage of SARS 3C-like proteinase occurs first in the polyprotein processing. We studied the relationship between -sheet tendencies of substrates with the cleavage activity by circular dichroism and bioassay testing. Substrates with more -sheet like structure tend to react faster. This has been confirmed by X-ray [10] and NMR studies [41] that substrates bound to the SARS 3C-like proteinase in like structure.

In order to further explore the substrate specificity in detail and provide more information for inhibitor design, we studied the substrate requirements for SARS 3C-like proteinase using 34 truncated and mutated substrate peptides (Table **2**) [42]. This study confirms that the core sequence of the proteinase cleavage site is highly conserved and optimized for enzyme recognition and catalysis. Residues at

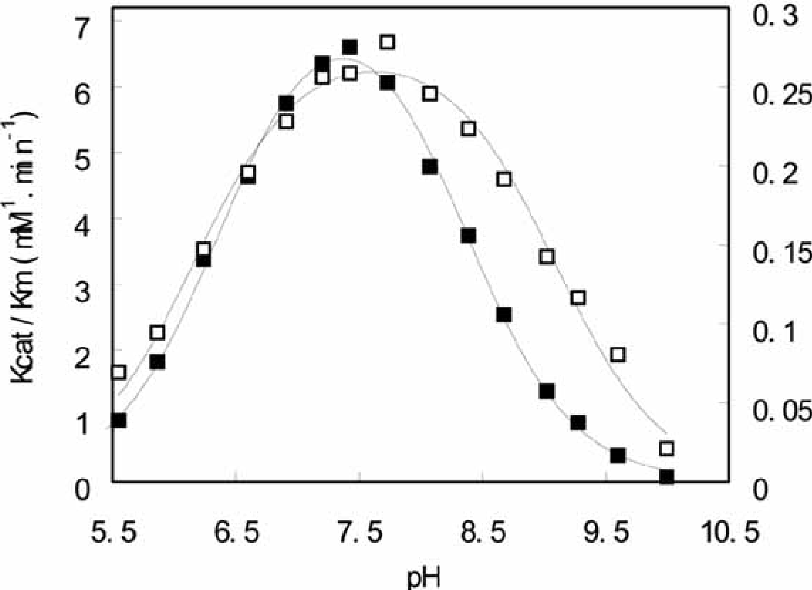

Fig. (1). The pH-rate profile for SARS 3C-like proteinase wild-type (■) and C145S (□) [39]. The left Y axis is for the wild type enzyme and the right Y axis is for the C145S mutant.

position P4, P3 and P3' are critical for substrate recognition and binding, and increment of -sheet conformation tendency at position P4 and P3 is helpful for substrate binding and hydrolysis. In addition, a positive charge at P3 (possibly forming a salt bridge between position P3 and Glu-166 of the enzyme) increases the activity of the substrate.

Table 1. Relative SARS 3C-Like Proteinase Cleavage Efficiencies of Eleven Peptides that Represents All the Eleven Cleavage Sites in SARS Coronavirus Polyprotein Tested by HPLC Assay [19]

Peptide	Cleavage site	Sequence	$(k_{cat}/K_m)_{rel}$
S01	P1/P2	TSAVLQ/SGFRK-NH_2	1.00
S02	P2/P3	SGVTFQ/GKFKK	0.41
S03	P3/P4	KVATVQ/SKMSD	0.03
S04	P4/P5	NRATLQ/AIASE	0.05
S05	P5/P6	SAVKLQ/NNELS	0.02
S06	P6/P7	ATVRLQ/AGNAT	0.22
S07	P7/P8	REPLMQ/SADAS	0.002
S08	P8/P9	PHTVLQ/AVGAC	0.08
S09	P9/P10	NVATLQ/AENVT	0.09
S10	P10/P11	TFTRLQ/SLENV	0.28
S11	P11/P12	*FYPKLQ/ASQAW*	0.27

We built a comparative molecular field analysis (CoMFA) model based on the sequences and activities of the 22 octapeptides (Fig. (**2**)) [42]. The steric interaction is shown to be the major driving force for substrate recognition. Positive charge at P3 and large hydrophobic residue at position P2 are favored. Based on the activity information of all the tested peptide substrates and the CoMFA model, a new octapeptide substrate S24 was designed by including tandem-placed -sheet-preferred residues at position P5 and P4, a positive-charged Lys at P3, and an Ala at P1' (Table **2**). The substrate S24 shows the highest hydrolysis activity among all the 34 designed substrates. The CoMFA model can be used to predict binding strengths of substrates and appropriate inhibitors. Inhibitors can be designed based on this predictive model to have hydrophobic substituents at P2 and P3' and bear positive charge at P3 positions.

INHIBITOR DISCOVERY FOR SARS 3C-LIKE PROTEINASE

Many therapies including different drugs have been tested and used for SARS. For example, ribavirin, a nucleoside analog with broad antiviral activity has been used in combination with interferon and was shown to decrease the death rate [43, 44]. However, there is no evidence that they are valid and reliable measures [45]. Thus, drug design against SARS is urgent and a significant challenge. As SARS 3C-like proteinase is a key step for the replication of

Table 2. Relative SARS 3C-Like Proteinase Cleavage Activities of Mutated Substrate Peptides Tested by HPLC Assay [42]

Substrate	Sequence[a]	$(k_{cat}/K_m)_{rel}$
S12	SAVLQ SGF-$CONH_2$	1.00
S13	AVLQ SG-$CONH_2$	0.152
S14	VLQ SG-$CONH_2$	0.0189
S15	AVLQ S-$CONH_2$	0.129
S16	VLQ S-$CONH_2$	0.0167
S17	AVLQ SGF-$CONH_2$	0.763
S18	LQ SG-$CONH_2$	0.0017
P5L	LAVLQ SGF-$CONH_2$	3.90
P5T	TAVLQ SGF-$CONH_2$	3.66
P5V	VAVLQ SGF-$CONH_2$	3.59
P5A	AAVLQ SGF-$CONH_2$	3.33
P4L	SLVLQ SGF-$CONH_2$	0.15
P4T	STVLQ SGF-$CONH_2$	1.47
P4V	SVVLQ SGF-$CONH_2$	2.44
P3L	SALLQ SGF-$CONH_2$	0.87
P3T	SATLQ SGF-$CONH_2$	1.19
P3A	SAALQ SGF-$CONH_2$	0.19
P3K	SAKLQ SGF-$CONH_2$	2.68
P1'A	SAVLQ AGF-$CONH_2$	2.04
P1'G	SAVLQ GGF-$CONH_2$	0.83
P1'L	SAVLQ LGF-$CONH_2$	ND[b]
P2M	SAVMQ SGF-$CONH_2$	0.208
P2F	SAVFQ SGF-$CONH_2$	0.046
P2I	SAVIQ SGF-$CONH_2$	0.0063
P2V	SAVVQ SGF-$CONH_2$	0.0056
P2A	SAVAQ SGF-$CONH_2$	0.0046
P2R	SAVRQ SGF-$CONH_2$	ND[b]
P1N	SAVLN SGF-$CONH_2$	ND[b]
P1E	SAVLE SGF-$CONH_2$	ND[b]
P1K	SAVLK SGF-$CONH_2$	ND[b]
S21	TVVLQ SGF-$CONH_2$	3.98
S22	TVTLQ SGF-$CONH_2$	2.62
S23	VVTLQ SGF-$CONH_2$	2.43
S24	**TVKLQ AGF-$CONH_2$**	**4.31**

a: Mutated residues are underlined. Cleavage sites are indicated by .

b: Not detectable in HPLC-based peptide cleavage assay.

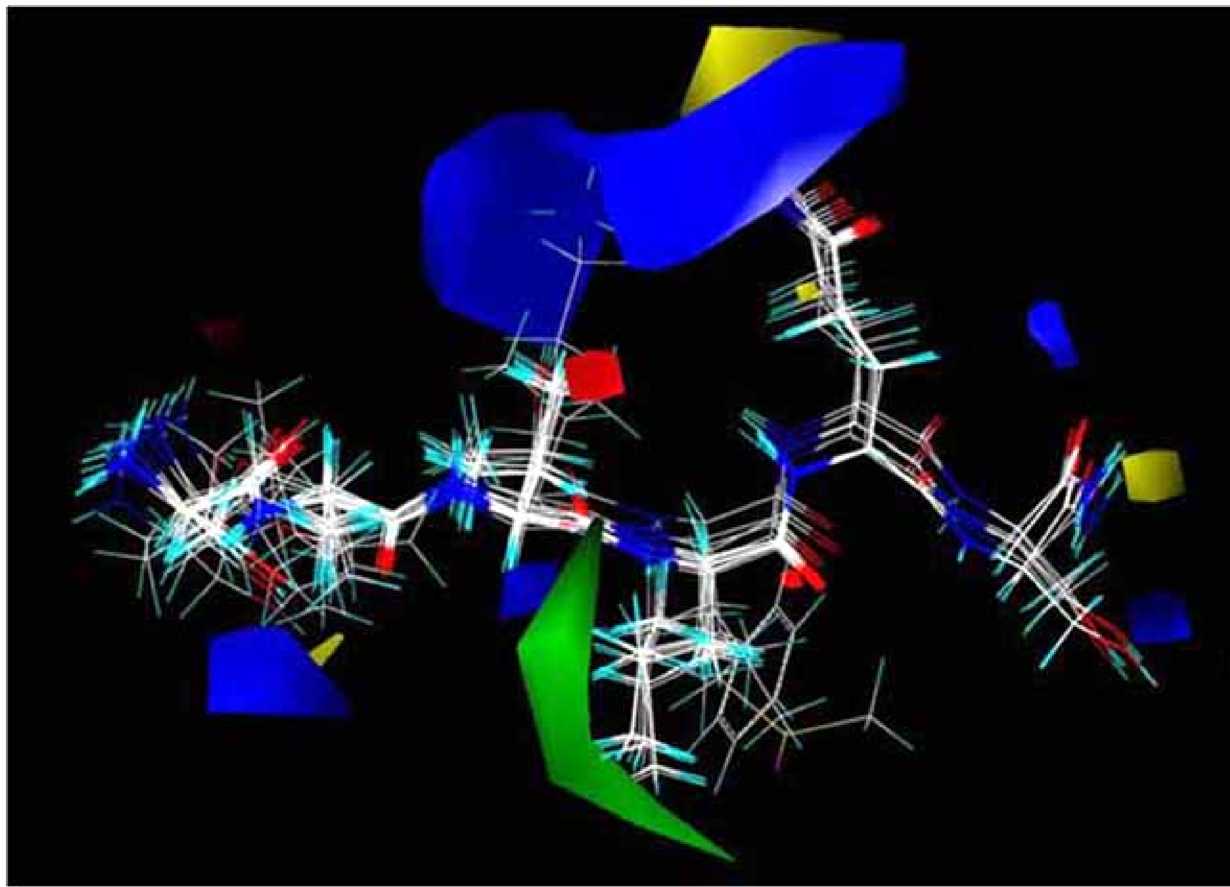

Fig. (2). The superimposed 22 SARS 3C-like proteinase substrate structures and the contour plot of the CoMFA model. This model indicates that increasing positive charge at position P3 is favored (blue), and large hydrophobic residue at position P2 is favored (green) [42].

SARS coronavirus, much efforts have been done on inhibitor discovery for this enzyme [45].

Modeling and Virtual Screening

Shortly after the genome sequencing was completed, models for the SARS 3C-like proteinase were built by comparative modeling based on the structure of the TGEV 3C-like proteinase by various groups [6-8] and the conformational flexibility of the substrate-binding site was studied [7]. Based on the complex crystal structure of their hexapeptidyl chloromethylketone inhibitor with the TGEV 3C-like proteinase, Anand *et al*. discovered that the binding mode of the inhibitor is similar to the binding mode of AG7088 in complex with the 3C proteinase of human rhinovirus [6], although the two enzymes only resemble to each other in the neighborhood of the active site. AG7088 is an enzyme substrate analog that is in phase II/III clinical trials as an inhalation treatment for the common cold caused by human rhinovirus [6, 45]. Anand *et al*. proposed that AG7088 can serve as a good starting point for inhibitor design against SARS 3C-like proteinase [6]. Inspired by this proposal, an octapeptide was docked on the model of SARS 3C-like proteinase to study the possible interactions of the protein and the substrate and for substrate analog design [46]. Optimization of AG7088 was also carried out experimentally [11]. *In vitro* assay showed that AG7088 only weakly binds to SARS 3C-like proteinase; further modifications of this compound improved binding.

Virtual screening on libraries of chemical compounds has predicted possible inhibitors [8, 47]. A few existing drugs and HIV protease inhibitors were also proposed to be possible inhibitors of SARS 3C-like proteinase [45]. Most early papers published in 2003 and 2004 only reported a list of possible inhibitors which need further experimental verifications. So, convenient and high-throughput assays for inhibitor screen are necessary.

Assays for Inhibitor Screening

High performance liquid chromatography (HPLC) assay using natural peptide substrates was applied to study the enzyme activity and substrate specificity [19]. As no modifications on the peptides are necessary, the substrates can be easily synthesized by conventional solid phase peptide synthesis. As the peptides only contain natural sequences, the assay is close to *in vivo* situation. HPLC assay is relatively easy to realize in the laboratory, but can only do the measurements at fixed time points and is difficult to be used in high throughput inhibitor screening.

In order to develop high throughput *in vitro* assays for inhibitor screening, chromogenic and fluorescence resonance energy transfer (FRET) substrates were used, which gener-

ally contain a short natural peptide sequence modified at the C-terminal end or both N- and C-terminal ends. We developed a high throughput colorimetric assay using a modified hexapeptide: Thr-Ser-Ala-Val-Leu-Gln-pNA, with the first six residues corresponding to the N-terminal self-cleavage site of the SARS 3C-like proteinase [39]. Colorimetric measurements of SARS 3C-like proteinase activity were performed in 96-well microtiter plates using a multi-well ultraviolet spectrometer. The substrate cleaved at the Gln-pNA bond to release free pNA, resulting in an increase of absorbance at 390 nm. This method was successfully used in inhibitor screening [46]. Kuo *et al.* [24] developed a continuous FRET assay using a fluorogenic substrate of a 14-mer peptide with the central 12-mer peptide of Thr-Ser-Ala-Val-Leu-Gln-Ser-Gly-Phe-Arg-Lys-Met corresponding to the N-terminal self-cleavage site plus Lys and Glu for attachment of 4-(4-dimethylaminophenyl-azo)benzoic acid (Dabcyl) and 5-[(2-aminoethyl)amino]naphthalene-1-sulfonic acid (Edans) in N- and C-termini. The two fluorophores formed a quenching pair and exhibited FRET within the peptide. When the peptide as is cleaved by the SARS 3C-like proteinase, the FRET signal disappeared. Kuang *et al*. [49] used two internally quenched fluorogenic peptides corresponding to the N-terminal and C-terminal auto-cleavage sites. The two 11mer peptides were conjugated with an *ortho*-aminobenzoic acid (Abz) and a lysine-2,4-dinitrophenylamide (Lys-DNP). Chen *et al.* [50] used a 12mer fluorogenic substrate peptide with the sequence of Edans-Val-Asn-Ser-Thr-Leu-Gln-Ser-Gly-Leu-Arg-Lys(Dabcyl)-Met for the FRET assay. Blanchard *et al.* [51] developed a FRET assay using a modified 9mer peptide: Abz-Ser-Val-Thr-Leu-Gln-Ser-Gly-Tyr(NO2)-Arg and reported that the anthranilate-nitrotyrosine donor-acceptor pair was over an order of magnitude more sensitive in FRET-based assays than the equivalent Edans-Dabcyl peptide. For all the FRET assays, substrates with sequence corresponding to the N-terminal auto-cleavage site of SARS 3C-like proteinase are more active and specific than others. Both colorimetric and FRET assays are convenient and highly efficient, but all with some deficiencies. As colorimetric assay detects the yellow color of the product, compounds with colors, especially yellow ones are hard to be tested. For the FRET assay, compounds with fluorescent properties can easily interfere with the assay, thus giving false positive or negative results.

Auto-cleavage assays to test the in-*cis* activity of SARS 3C-like proteinase were also established by adding extra in-frame amino acid residues at the ends and monitoring the cleavage products on gel [40]. These assays are useful in studying the *cis* enzyme catalytic mechanism, but are difficult to be applied to large scale inhibitor screen. Parera *et al*. developed a genetic assay based on the lambda regulatory circuit to monitor the activity of SARS 3C-like proteinase, which may be used for inhibitor screening [52]. Viral replication assays have also been used in screening possible 3C-like proteinase inhibitors [53, 54].

Active Inhibitors Discovered

Several series of inhibitors from virtual screening, compound library screen or optimization of existing inhibitors of other proteases have been confirmed to be active in the *in vitro* assays or viral replication assay. Table **3** gives a summary for a selective collection of active compounds. Wu *et al.* [53] screened more than 10,000 compounds to inhibit viral replication and found that two anti-HIV compounds can target SARS viral entry and 3C-like proteinase, respectively. One of the compounds, which is a peptidic inhibitor designed as a transition state analog of the HIV protease, inhibits SARS 3C-like proteinase with a K_i of 0.6 μM in the FRET assay. Kao *et al*. screened 50,240 structurally diverse small molecules and identified 104 compounds with anti-SARS coronavirus activity [55]. Among the 104 compounds, 2 of them were inhibitors of SARS 3C-like proteinase. One compound, MP576 displayed potent inhibitory activity with IC_{50} of 2.5 μM in an HPLC assay and EC_{50} of 7 μM in the Vero cell-based SARS-CoV plaque reduction assay. Blanchard *et al.* [51] screened 50,000 drug-like small molecules with a FRET assay of 7mer peptide and identified five compounds that showed potent inhibitory activity with IC_{50} from 0.5 to 7μM. Kuo *et al*. [24] screened a panel of compounds and found 1-hydroxypyridine-2-thione zinc exhibited IC_{50} of 0.8 μM towards SARS 3C-like proteinase in an *in vitro* FRET assay. Hsu *et al.* [56] screened 960 commercially available drugs and biologically active compounds and found mercury-containing compounds, the above zinc compound, and hexachlorophene are potent inhibitors. However, as the SARS 3C-like proteinase is a cysteine enzyme, heavy metals like mercury and zinc react with the active site Cys and be toxic to other cysteine enzymes, too. Chen *et al*. [54] performed virtual screening by molecular docking of more than 8,000 existing drugs onto the SARS 3C-like proteinase crystal and model structures. They found that, cinanserin, a well-characterized serotonin antagonist that has undergone preliminary clinical test in the 1960s, can inhibit SARS 3C-like proteinase with an IC_{50} of 5 μM and shows potent activity in the viral replicon system with IC_{50} values ranging from 19 to 34 μM. Liu *et al.* [48] did virtual screening over 600,000 compounds from Available Chemical Database and National Cancer Institute Diversity Set and found three potent inhibitors from the *in vitro* assay using chromogenic substrates with apparent K_i from 61 to 178 μM. The most active compound they found was calmidazolium, a well-known antagonist of calmodulin.

Based on their previous work on picornaviral 3C proteinase inhibitors, Jain *et al*. [57] synthesized a series of phthalhydrazide-substituted keto-glutamine analogs and discovered potent inhibitors with IC_{50} ranging from 0.60 to 70 μM with a FRET assay of 7mer peptide substrate. Hsu [11] *et al.* synthesized AG7088 related compounds and tested their inhibitory activity. Although AG7088 has no detectable or very weak inhibitory activity against SARS 3C-like proteinase, some of the derivatives bind to the enzyme with IC_{50} of around 11 μM in the *in vitro* FRET assay using 14mer peptide substrate. Shie *et al.* [58] discovered that anilide compounds act as potent inhibitors in the *in vitro* FRET assay and the most active compound gives K_i of 0.03 μM. Bacha *et al.* [59] designed bifunctional boronic acid compounds to bind with the serine cluster (Ser139, Ser144 and Ser147) near the active site cavity and the compounds inhibited SARS 3C-like proteinase as strong as 40 nM. However, they used a commercially available compound: Dabcyl-Leu-Ala-Gln-Ala-Val-Arg-Ser-Ser-Ser-Arg-Edans as the substrate,

Table 3. A Selective Collection of Active Inhibitors Discovered for SARS 3C-Like Proteinase

Compound	Structure	Ki(μM)	IC50(μM)	EC50(μM)	Assay	Ref.
2		0.6			FRET	[53]
MP576			2.5	7	HPLC Viral	[55]
MAC-5576			0.5		FRET	[51]
MAC-8120			4.3		FRET	[51]
MAC-13985			7		FRET	[51]
MAC-22272			2.6		FRET	[51]
MAC-30731			7		FRET	[51]
Hexachlorophene		13.7			FRET	[56]
cinanserin		49.4 4.92			SPR FRET	[54]

(Table 3) contd….

Compound	Structure	Ki(μM)	IC50(μM)	EC50(μM)	Assay	Ref.
C3930	Cl, Cl, N^+, O, Cl, N, Cl, Cl, Cl	61			Colorimetric	[48]
keto-glutamine analogues 8c	O, H N, BnO, O, O, O, AcHN, N H, H N, N H, N, O, O, HN, O, NO_2		0.60		FRET	[57]
AG7088 related compound 4d	O, O, R, N H, H N, N H, COOEt, O, X		11			[11]
2a	O, Cl, H N, N H, O, Me_2N, NO_2	0.03	0.06		FRET	[58]
FL-166	HO, B, OH, H N, O_2N, O, O, NO_2, N H, HO, B, OH	0.04			FRET	[59]
Isatin 4o	O, I, O, N, S		0.95		FRET	[60]

(Table 3) contd….

Compound	Structure	Ki(μM)	IC50(μM)	EC50(μM)	Assay	Ref.
Isatin 5f	H_2NOC … O … O … N		0.37		Colorimetric	[61]

which might have very low activity towards SARS 3C-like proteinase. Our experience with this substrate gave no detectable cleavage activity (unpiblished data). Chen *et al*. [60] started from known rhinovirus 3C protease inhibitors, 2,3-dioxindole compounds, to make a series of derivatives and tested their activity against SARS 3C-like proteinase. Some of the compounds were potent inhibitors with IC_{50} values ranging from 0.95 to 17.5 μM. We have also synthesized a series of 2,3-dioxindole compounds and tested their activity against SARS 3C-like proteinase. Instead of iodine substitute, a carboxamido-group occupies the P1 substrate binding site and increases the IC_{50} to 0.37 μM. The binding mode of these compounds were confirmed to be non-covalent [61]. It should be noted that IC_{50} and K_i values from assays using different substrates cannot be directly compared.

Although several series of compounds have been discovered to inhibit SARS 3C-like proteinase, the binding modes and reaction mechanism need further study.

In summary, after the completion of the genome sequencing for SARS coronavirus in early 2003, much work has been done on the structure, enzyme mechanism, substrate specificity and inhibitor screening for the SARS 3C-like proteinase. Several series of compounds were found to be potent inhibitors in micro-mole range. All these researches have laid the ground for further highly effective and selective inhibitor design against SARS coronavirus, and at the same time, set successful examples for drug design against new infectious diseases.

REFERENCES

References 62-64 are related articles recently published in Current Pharmaceutical Design.

[1] Lee N, Hui D, Wu A, Chan P, Cameron P, Joynt GM, *et al*. A major outbreak of severe acute respiratory syndrome in Hong Kong. N Engl J Med 2003; 348(20): 1986-1994.

[2] Poutanen SM, Low DE, Henry B, Finkelstein S, Rose D, Green K, *et al*. Identification of severe acute respiratory syndrome in Canada. N Engl J Med 2003; 348(20): 1995-2005.

[3] Holmes KV. SARS coronavirus: a new challenge for prevention and therapy. J Clin Invest 2003; 111(11): 1605-1609.

[4] 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): 1399-1404.

[5] Ruan YJ, Wei CL, Ee AL, Vega VB, Thoreau H, Su ST, *et al*. Comparative full-length genome sequence analysis of 14 SARS coronavirus isolates and common mutations associated with putative origins of infection. Lancet 2003; 361(9371): 1779-1785.

[6] Anand K, Ziebuhr J, Wadhwani P, Mesters JR, Hilgenfeld R. Coronavirus main proteinase (3CL(pro)) structure: Basis for design of anti-SARS drugs. Science 2003; 300(5626): 1763-1767.

[7] Liu S, Pei J, Chen H, Zhu X, Liu Z, Ma W, *et al*. Modeling of the SARS coronavirus main proteinase and conformational flexibility of the active site. Beijing Da Xue Xue Bao 2003; 35 Suppl: 62-65.

[8] 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 Pharm Sinica 2003; 24(6): 497-504.

[9] 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): 3213-3224.

[10] 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): 13190-13195.

[11] Hsu MF, Kuo CJ, Fang JM, Shie JJ, Chang KT, Chang HC, *et al*. Understanding the maturation process and inhibitor design of SARS-CoV 3CLpro from the crystal structure of C145A in a product-bound form. J Biol Chem 2005; 280(35): 31257-31266.

[12] Bergmann EM, Mosimann SC, Chernaia MM, Malcolm BA, James MN. The refined crystal structure of the 3C gene product from hepatitis A virus: specific proteinase activity and RNA recognition. J Virol 1997; 71(3): 2436-2448.

[13] 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): 761-771.

[14] Dragovich PS, Prins TJ, Zhou R, Brown EL, Maldonado FC, Fuhrman SA, *et al*. Structure-based design, synthesis, and biological evaluation of irreversible human rhinovirus 3C protease inhibitors. 6. Structure-activity studies of orally bioavailable, 2-pyridone-containing peptidomimetics. J Med Chem 2002; 45(8): 1607-1623.

[15] Dragovich PS, Prins TJ, Zhou R, Johnson TO, Brown EL, Maldonado FC, *et al*. Structure-based design, synthesis, and biological evaluation of irreversible human rhinovirus 3C protease inhibitors. Part 7: structure-activity studies of bicyclic 2-pyridone-containing peptidomimetics. Bioorg Med Chem Lett 2002; 12(5): 733-738.

[16] Dragovich PS, Zhou R, Webber SE, Prins TJ, Kwok AK, Okano K, *et al*. Structure-based design of ketone-containing, tripeptidyl human rhinovirus 3C protease inhibitors. Bioorg Med Chem Lett 2000; 10(1): 45-48.

[17] Johnson TO, Hua Y, Luu HT, Brown EL, Chan F, Chu SS, *et al*. Structure-based design of a parallel synthetic array directed toward the discovery of irreversible inhibitors of human rhinovirus 3C protease. J Med Chem 2002; 45(10): 2016-2023.

[18] 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): 11000-11007.

[19] Fan KQ, Wei P, Feng Q, Chen SD, Huang CK, Ma L, *et al*. Biosynthesis, purification, and substrate specificity of severe acute respiratory syndrome coronavirus 3C-like proteinase. J Biol Chem 2004; 279(3): 1637-1642.

[20] 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): 14958-14970.

[21] Laue TM, Stafford WF III. Modern applications of analytical ultracentrifugation. Annu Rev Biophys Biomol Struct 1999; 28: 75-100.

[22] Lebowitz J, Lewis MS, Schuck P. Modern analytical ultracentrifugation in protein science: a tutorial review. Protein Sci 2002; 11(9): 2067-2079.

[23] Wei P, Fan KQ, Chen H, Ma L, Huang CK, Tan L, *et al*. The N-terminal octapeptide acts as a dimerization inhibitor of SARS coronavirus 3C-like proteinase. Biochem Biophys Res Comm 2006; 339(3): 865-872.

[24] Kuo CJ, Chi YH, Hsu JT, Liang PH. Characterization of SARS main protease and inhibitor assay using a fluorogenic substrate. Biochem Biophys Res Commun 2004; 318(4): 862-867.

[25] Chen H, Wei P, Huang C, Tan L, Liu Y, Lai L. Only one protomer is active in the dimer of SARS 3C-like proteinase. J Biol Chem 2006; 281(20): 13894-13898.

[26] Holzman TF, Kohlbrenner WE, Weigl D, Rittenhouse J, Kempf D, Erickson J. Inhibitor stabilization of human immunodeficiency virus type-2 proteinase dimer formation. J Biol Chem 1991; 266(29): 19217-19220.

[27] Todd MJ, Semo N, Freire E. The Structural Stability of the HIV-1 Protease. J Mol Biol 1998; 283(2): 475-488.

[28] Jiang L, Gao Y, Mao FL, Liu ZJ, Lai LH. Potential of mean force for protein-protein interaction studies. Proteins 2002; 46(2): 190-196.

[29] Jones S, Thornton JM. Principles of protein-protein interactions. Proc Natl Acad Sci USA 1996; 93(1): 13-20.

[30] Ding L, Zhang XX, Wei P, Fan KQ, Lai LH. The interaction between severe acute respiratory syndrome coronavirus 3C-like proteinase and a dimeric inhibitor by capillary electrophoresis. Anal Biochem 2005; 343(1): 159-165.

[31] Asboth B, Stokum E, Khan IU, Polgar L. Mechanism of action of cysteine proteinases: oxyanion binding site is not essential in the hydrolysis of specific substrates. Biochemistry 1985; 24(3): 606-609.

[32] Brocklehurst K, O'Driscoll M, Kowlessur D, Phillips IR, Templeton W, Thomas EW, *et al*. The interplay of electrostatic and binding interactions determining active centre chemistry and catalytic activity in actinidin and papain. Biochem J 1989; 257(1): 309-310.

[33] Dardenne LE, Werneck AS, de Oliveira NM, Bisch PM. Electrostatic properties in the catalytic site of papain: a possible regulatory mechanism for the reactivity of the ion pair. Proteins 2003; 52(2): 236-253.

[34] Johnson FA, Lewis SD, Shafer JA. Perturbations in the free energy and enthalpy of ionization of histidine-159 at the active site of papain as determined by fluorescence spectroscopy. Biochemistry 1981; 20(1): 52-58.

[35] Sarkany Z, Szeltner Z, Polgar L. Thiolate-imidazolium ion pair is not an obligatory catalytic entity of cysteine peptidases: the active site of picornain 3C. Biochemistry 2001; 40(35): 10601-10606.

[36] Sarkany Z, Polgar L. The unusual catalytic triad of poliovirus protease 3C. Biochemistry 2003; 42(2): 516-522.

[37] Connolly KM, Smith BT, Pilpa R, Ilangovan U, Jung ME, Clubb RT. Sortase from Staphylococcus aureus Does Not Contain a Thiolate-Imidazolium Ion Pair in Its Active Site. J Biol Chem 2003; 278(36): 34061-34065.

[38] 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): 581-593.

[39] Huang CK, Wei P, Fan KQ, Liu Y, Lai LH. 3C-like proteinase from SARS coronavirus catalyzes substrate hydrolysis by a general base mechanism. Biochemistry 2004; 43(15): 4568-4574.

[40] Shan YF, Li SF, Xu GJ. A novel auto-cleavage assay for studying mutational effects on the active site of severe acute respiratory syndrome coronavirus 3C-like protease. Biochem Biophys Res Comm 2004; 324(2): 579-583.

[41] Shi JH, Wei Z, Song JX. 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): 24765-24773.

[42] Fan KQ, Ma L, Han XF, Liang HH, Wei P, Liu Y, *et al*. The substrate specificity of SARS coronavirus 3C-like proteinase. Biochem Biophys Res Comm 2005; 329(3): 934-940.

[43] Koren G, King S, Knowles S, Phillips E. Ribavirin in the treatment of SARS: A new trick for an old drug? CMAJ 2003; 168(10): 1289-1292.

[44] Cinatl J, Morgenstern B, Bauer G, Chandra P, Rabenau H, Doerr HW. Treatment of SARS with human interferons. Lancet 2003; 362(9380): 293-294.

[45] Skowronski DA, Astell C, Brunham RC, Low DE, Petric M, Roper RL, *et al*. Severe acute respiratory syndrome (SARS): A year in review. Ann Rev Med 2005; 56: 357-381.

[46] Chou KC, Wei DQ, Zhong WZ. Binding mechanism of coronavirus main proteinase with ligands and its implication to drug design against SARS. Biochem Biophys Res Comm 2003; 308(1): 148-151.

[47] Liu B, Zhou JJ. SARS-CoV protease inhibitors design using virtual screening method from natural products libraries. J Comp Chem 2005; 26(5): 484-490.

[48] Liu ZM, Huang CK, Fan KQ, Wei P, Chen H, Liu SY, *et al*. Virtual screening of novel noncovalent inhibitors for SARS-CoV 3C-like proteinase. J Chem Inf Mod 2005; 45(1): 10-17.

[49] 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 Comm 2005; 331(4): 1554-1559.

[50] Chen S, Chen LL, Luo HB, Sun T, Chen J, Ye F, *et al*. Enzymatic activity characterization of SARS coronavirus 3C-like protease by fluorescence resonance energy transfer technique. Acta Pharm Sinica 2005; 26(1): 99-106.

[51] 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): 1445-1453.

[52] Parera M, Clotet B, Martinez MA. Genetic screen for monitoring severe acute respiratory syndrome coronavirus 3C-like protease. J Virol 2004; 78(24): 14057-14061.

[53] Wu CY, Jan JT, Ma SH, Kuo CJ, Juan HF, Cheng YS, *et al*. Small molecules targeting severe acute respiratory syndrome human coronavirus. Proc Natl Acad Sci USA 2004; 101(27): 10012-10017.

[54] 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): 7095-7103.

[55] Kao RY, Tsui WHW, Lee TSW, 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): 1293-1299.

[56] 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): 116-120.

[57] 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): 6113-6116.

[58] 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): 4469-4473.

[59] Bacha U, Barrila J, Velazquez-Campoy A, Leavitt SA, Freire E. Identification of novel inhibitors of the SARS coronavirus main protease 3CL(pro). Biochemistry 2004; 43(17): 4906-4912.

[60] 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): 3058-3062.

[61] Zhou L, Liu Y, Zhang WL, Wei P, Huang CK, Pei JF, *et al*. Isatin compounds as noncovalent SARS coronavirus 3C-like protease inhibitors. J Med Chem 2006; in press.

[62] Evangelista S. Quaternary ammonium derivatives as spasmolytics for irritable bowel syndrome. Curr Pharm Des 2004; 10(28): 3561-8.

[63] Gomtsyan A, Lee CH. Nonnucleoside inhibitors of adenosine kinase. Curr Pharm Des 2004; 10(10): 1093-103.

[64] Sakoff JA, McCluskey A. Protein phosphatase inhibition: structure based design. Towards new therapeutic agents. Curr Pharm Des 2004; 10(10): 1139-59.

Copyright of Current Pharmaceutical Design is the property of Bentham Science Publishers Ltd. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.