Polyprotein cleavage mechanism of SARS CoV Mpro and chemical modification of the octapeptide

Qi-Shi Du[a,b], Shu-Qing Wang[a], Yu Zhu[a], Dong-Qing Wei[a,b,c], Hong Guo[d], Suzanne Sirois[b], Kuo-Chen Chou[a,e,f,*]

[a] *Tianjin Normal University and Tianjin Institute of Bioinformatics and Drug Discovery (TIBDD), Tianjin 300074, China*
[b] *Institut Technologique de Montreal, Suite 168, 5253 Boul. Decarie, Montreal, Que., Canada H3W 3C3*
[c] *Center For Research in Molecular Modeling (CERMM), Concordia University, Montreal, Canada*
[d] *University of Tennessee, Department of Biochemistry, Cell and Molecular Biology, Knoxville, TN 37996-0840, USA*
[e] *Institute of Image Processing and Pattern Recognition, Shanghai Jiaotong University, Shanghai 200030, China*
[f] *Gordon Life Science Institute, San Diego, CA 92130, USA*

Abstract

The cleavage mechanism of severe acute respiratory syndrome (SARS) coronavirus main proteinase (Mpro or 3CLpro) for the octapeptide AVLQSGFR is studied using molecular mechanics (MM) and quantum mechanics (QM). The catalytic dyad His-41 and Cys-145 in the active pocket between domain I and II seem to polarize the π-electron density of the peptide bond between Gln and Ser in the octapeptide, leading to an increase of positive charge on C(CO) of Gln and negative charge on N(NH) of Ser. The possibility of enhancing the chemical bond between Gln and Ser based on the "distorted key" theory [Anal. Biochem. 233 (1996) 1] is examined. The scissile peptide bond between Gln and Ser is found to be solidified through "hybrid peptide bond" by changing the carbonyl group CO of Gln to CH_2 or CF_2. This leads to a break of the π-bond system for the peptide bond, making the octapeptide (AVLQSGFR) a "distorted key" and a potential starting system for the design of anti SARS drugs.
© 2004 Elsevier Inc. All rights reserved.

Keywords: SARS; Coronavirus main proteinase; Inhibitor; Distorted key theory; Drug design; Octapeptide; KZ7088

1. Introduction

Spreading of severe acute respiratory syndrome (SARS) [22,26,27,31] in Asia, North America and several countries in Europe prompted an unprecedented global effort to fight the disease. Researchers have identified crucial proteins of SARS-coronavirus and thousands of compounds are being screened in an effort to find new drugs [7,20,32]. Anand et al. [7] described a homology model of SARS coronavirus main proteinase (Mpro or 3CLpro)

based on the experimental Mpro structures of human coronavirus (HCoV) and porcine transmissible gastroenteritis virus (TGEV) complex. Recently, the crystal structures of SARS CoV Mpro at different pH values were reported by Yang et al. [34]. The X-ray structures revealed some differences with the homology model obtained earlier [7]. However, experimental structure of SARS CoV Mpro folds in an arrangement that is similar to the HCoV and TGEV Mpro structures, and both homology model and experimental structure have a His–Cys catalytic dyad between domains I and II [34]. The structures of the three coronavirus main proteinases reveal remarkable degree of conservation on the substrate-binding sites and form the structure basis for rational drug design [7,20].

Abbreviations: SARS, severe acute respiratory syndrome; CoV, coronavirus; Mpro, main proteinase

* Corresponding author.
E-mail address: kchou@san.rr.com (K.-C. Chou).

At the same time, structure-based drug design has been progressing based on the molecular structure of SARS–CoV M^{pro} [20]. AG7088 suggested by Anand et al. [7] could well serve as a starting point to design efficient inhibitors for SARS CoV M^{pro}. AG7088 is a peptide inhibitor designed based on the structure of HCoV M^{pro} and is being clinically tested by Pfizer for the treatment of the human common cold. Chou et al. [20] have found some deficiency of AG7088 for binding to SARS CoV M^{pro} and suggested to use its derivative KZ7088 that could form better interactions with the active pocket of SARS CoV M^{pro}. Chou et al. also proposed an octapeptide inhibitor NH$_2$−AVLQ↓SGFR−COOH (the cleavage site is indicated by ↓) based on molecular modelling. This suggestion was supported by the recent work of Yang et al. [34]. In their structural determination, these authors used a decade peptide NH$_2$−TSAVLQ↓SGFR−COOH, which is quite similar to the octapeptide NH$_2$−AVLQ↓SGFR−COOH. According to a very recent report by Gan et al. [23], the octapeptide originally proposed by Chou et al. [20] has been synthesized and tested as the most active in inhibiting replication of the SARS coronavirus compared with other compounds reported. Moreover, it has been found that the octapeptide had no toxicity in vivo under the physiological concentration [23].

That a peptide is cleavable by a protease means that there is good binding between ligand and receptor on the active region of protease and that the peptide has a scissile bond to be cleaved (see, e.g., [13]) and a comprehensive review [12]. However, one needs to make some further chemical modification for the octapeptide in order to stabilize its inhibition power to SARS CoV M^{pro} and make it become an effective drug. This may be realized based on the "distorted key" theory [12,13]. To reach such a goal, a detailed understanding of the cleavage mechanism of the octapeptide by SARS CoV M^{pro} as well as the 3D structure of the enzyme is essential. Similar strategies have been used to investigate Cdk5–Nc5a*–ATP complex [19,36], apoptosis proteins [10,11,15,17], and beta-secretase zymogen [14]. Many useful insights have been gained through these studies. Accordingly, it is expected that the present study may also provide useful insights for the development of anti SARS drug.

2. Theory and method

The protease-susceptible sites in a given protein or peptide usually extend to an octapeptide region [12,13]. The corresponding amino acid residues are sequentially symbolized by eight subsites R$_4$, R$_3$, R$_2$, R$_1$, R$_{1'}$, R$_{2'}$, R$_{3'}$, R$_{4'}$, and the eight combination positions of protease are noted by S$_4$, S$_3$, S$_2$, S$_1$, S$_1$, S$_{2'}$, S$_{3'}$, S$_{4'}$ (see, e.g., [8,9] as well as Fig. 3 of [12]). Occasionally, the susceptible sites in some proteins may contain one subsite less or more [13], however, eight amino acid residues are the most common cases. Although the protein being cleaved contains much more than eight amino acid residues, usually only the segment of an octapeptide fits in
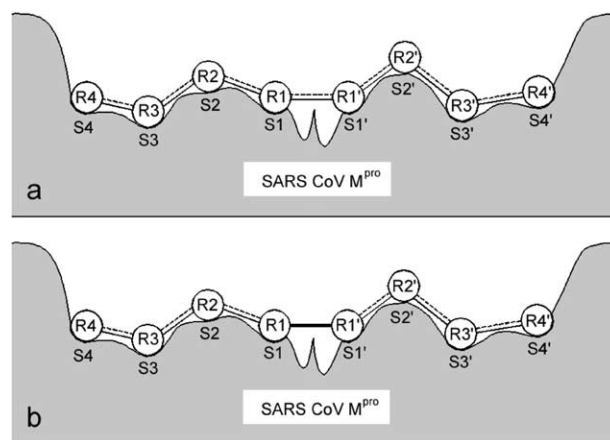

Fig. 1. A schematic drawing to illustrate the "distorted key" theory [12,13]: (a) the cleavage location in the octapeptide by protease is the peptide bond between R$_1$ and R$_{1'}$; (b) after chemical modification, the scissile peptide bond changes to a strong "hybrid peptide bond" and the cleavage is difficult. Adapted from Chou [12] with permission.

the active site region of a protease. Therefore, our research will focus on the cleavability of an octapeptide.

As shown in Fig. 1a, the combination of a thin line and a dashed line is used to represent the conjugate π property in peptide bond. In Fig. 1b, however, the scissile peptide bond is replaced by a strong "hybrid peptide bond" between R$_1$ and R$_{1'}$ through a chemical modification, and the enhanced chemical bond become not cleavable by the protease. According to the "distorted key" theory [12], a cleavable octapeptide can be likened to a key that fits well in binding to the protease active region leading to a cleavage at its scissile bond. The octapeptide after some chemical modification can still bind to the active region but its peptide bond can no longer be cleaved by the protease. Thus, the modified octapeptide can be vividly compared to a "distorted key" that can be inserted into a lock but that can neither open it nor be easily pulled out from the lock [12]. In view of this, the modified octapeptide naturally becomes a stable competitive inhibitor and a potential candidate of drug.

The octapeptide AVLQSGFR is the first designed octapeptide [20] based on the molecular structure of SARS CoV M^{pro} and is proved cleavable experimentally. In this study, we use molecular mechanical and quantum mechanical simulations to investigate the cleavage mechanism, properties of the chemical bonds concerned as well as the catalytic interaction between the octapeptide and SARS CoV M^{pro}. The study is performed in the following four steps: (1) using molecular mechanics to minimize the energy of the SARS CoV M^{pro} complex with the octapeptide from the basis derived from docking studies; (2) computing the atomic charge distribution around the binding pocket of SARS CoV M^{pro} using ab initio quantum mechanics and the minimizing conformational energy; (3) computing the molecular energy, chemical bond properties, and atomic charges of the octapeptide in the background charge distribution [24] of SARS CoV M^{pro} using ab initio quantum mechanics; and (4) in the same back-
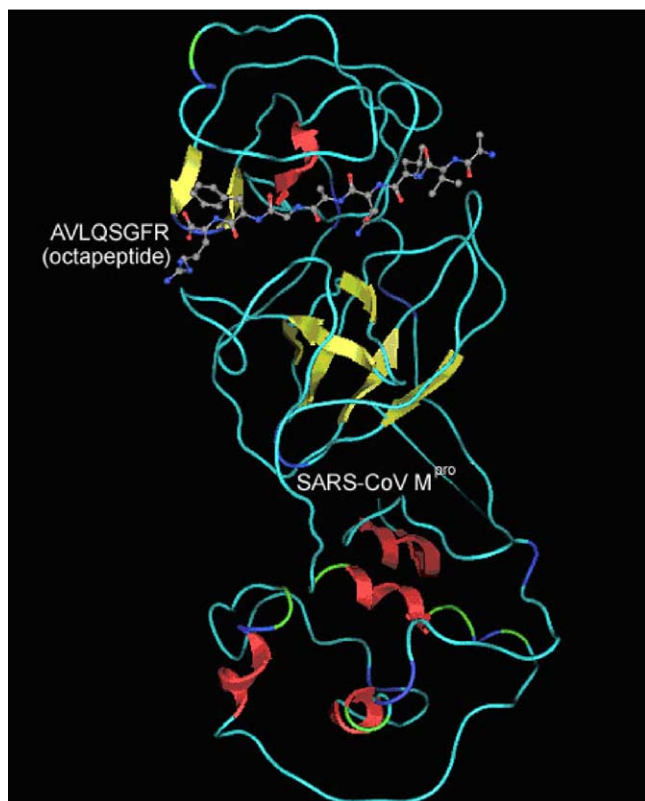
Fig. 2. The energy-refined docked structure of the octapeptide NH_2-AVLQSGFR$-COOH$ with SARS coronavirus main protease (SARS CoV M^{pro}).

ground charge distribution, computing the molecular energy, chemical bond properties, and atomic charges of the modified octapeptide.

3. Calculation results

The docking operation of the octapeptide AVLQSGFR to SARS CoV M^{pro} was performed based on the homology structure [7,20] using MOE (molecular operating environment) program package [28]. Twenty-five docking structures were obtained, and the one with the most optimal docking score was used for further energy minimization. Fig. 2 shows the energy-refined docked structure obtained by the aforementioned step 1. In contrast to the common serine proteases that have a Ser–His–Asp catalytic triad, SARS CoV M^{pro} has a His–Cys catalytic dyad (His-41 and Cys-145), which is similar to TGEV M^{pro} (His-41 and Cys-144) and HCoV M^{pro} (His-41 and Cys-144) [7,29]. According to Chou et al. [20], the catalytic active region is located within the pocket between domain I and II of SARS coronavirus main protease that contains the following 23 amino acid residues: Cys-22, Gly-23, Thr-24, Thr-25, Leu-27, His-41, Val-42, Cys-44, Thr-45, Ala-46, Glu-47, Asp-48, Met-49, Leu-50, Asn-51, Pro-52, Tyr-54, Cys-145, His-164, Met-165, Asp-187, Arg-188, and Gln-189. Fig. 3a shows the location of catalytic dyad

His-41 and Cys-145 in SARS CoV M^{pro}. The active cleft of SARS CoV M^{pro} can well accommodates the octapeptide, and the ligand binds to the receptor through six hydrogen bonds (Fig. 3b), fully in consistent with the results reported by Chou et al. [20].

In order to study the influence of SARS CoV M^{pro} on chemical bonds of the octapeptide, we considered a small region from the catalytic cleft surrounding the octapeptide, as shown in Fig. 3b. The catalytic dyad His-41 and Cys-145 are in the front of peptide bond Gln–Ser on the subsites R_1 and $R_{1'}$. The polar hydrogen $H_{\varepsilon2}$ on $N_{\varepsilon2}$ in midazole group of His-41 points to the peptide bond Gln–Ser and has large influence in the active region.
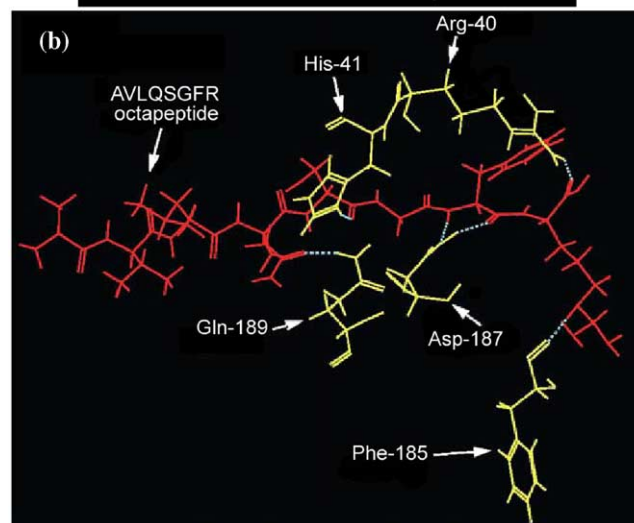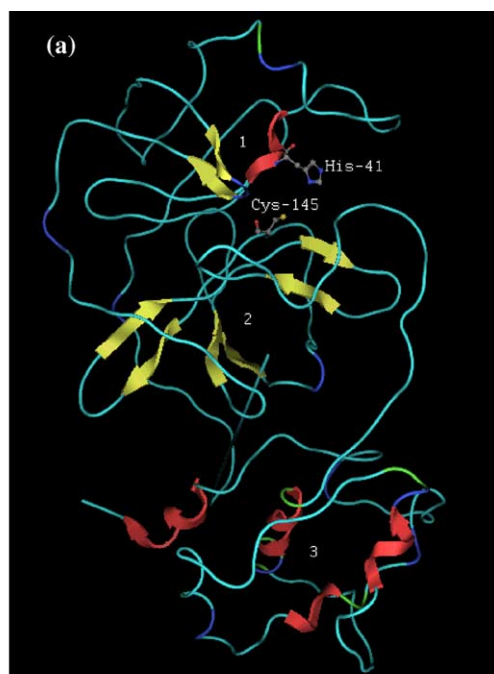

Fig. 3. (a) The catalytic dyad His-41 and Cys-145 are located in the active cleft between domain I and domain II of SARS CoV M^{pro}. (b) The hydrogen bonds between NH_2-AVLQSGFR$-COOH$ and the surrounding amino acid residue of the enzyme.
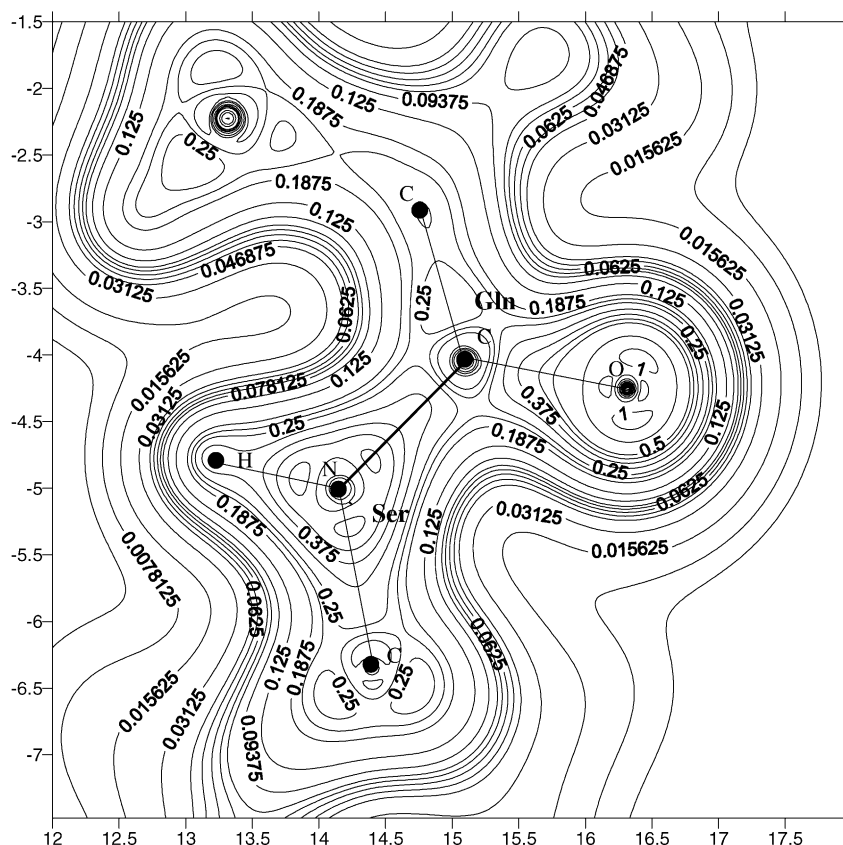
Fig. 4. Electron density counter map of peptide bond Gln–Ser in gaseous phase on the π-plane consisting of carbonyl C and O of Gln and N(NH) of Ser.

Electrostatic interaction plays the dominant role in ligand–receptor combination and must be taken into consideration during the quantum mechanical calculations for the influence of SARS CoV M^{pro} to the chemical bonds of the octapeptide. For this purpose, we divide the amino acid residues in the active cleft between domain I and II of SARS CoV M^{pro} into six segments and compute the atomic charges using ab initio quantum mechanics separately (Fig. 4). The 62 amino acid residues in the six segments are listed in Table 1. After deducting the overlap atoms, there are a total of 953 atomic background charges, including all atoms in catalytic cleft.

In Table 2, we list Mulliken atomic charges q^{Mull} and electrostatic potential equivalent charges q^{ESP}, which are obtained by fitting atomic charges to the electrostatic potential at the van der Waals surface. The atomic charges from semi

empirical method AM1 are quite different from the results by ab initio HF/6-31G* calculations [25], especially q^{ESP} of AM1 are not reasonable. Because q^{ESP} reproduces quantum mechanical electrostatic potential on molecular surface, in this research, we use q^{ESP} from HF/6-31G* calculations to illustrate our points. It can be seen from Table 2 that the two polar hydrogen atoms on imidazole group have large atomic charges q^{ESP}. The atomic charge (0.4201) of polar hydrogen $H_{\varepsilon 2}$ is a little smaller than the atomic charge (0.4288) of proton $H_{\delta 1}^{+}$ on nitrogen $N_{\delta 1}$ and is close to the peptide bond Gln–Ser on subsites R_1 and $R_{1'}$. The polar hydrogen $H_{\varepsilon 2}$ attracts the π-electron density from peptide bond Gln–Ser so as to weaken this chemical bond.

We calculate the cleavage reaction energy from the octapeptide AVLQSGFR to two tetrapeptides AVLQ and SGFR using ab initio HF/6-31G* in the gaseous phase. The molecular energies of the octapeptide and two tetrapeptides are shown in Table 3 and the hydrolyzing reaction energy is 110.8 kJ/mol.

4. Chemical modification of the octapeptide

The peptide bond is considered as a pseudo π-bond, i.e., a partial π-bond consisting of three atoms and four electrons [33]. Table 4 shows the atomic coordinates of the six atoms on

Table 1

Division of amino acid residues in the active cleft of SARS CoV M^{pro}, total 62 amino acid residues and 953 atoms are included

Domain I	Sequential position
D1_1	20–30
D1_2	40–50
D1_3	49–59
Domain II	Sequential position
D2_1	140–150
D2_2	160–169
D2_3	181–190

Table 2
Atomic charges of amino acid His-41in SARS CoV M^{pro}

Atom	AM1 (q^{Mull})	AM1 (q^{ESP})	HF/6-31G* (q^{Mull})	HF/6-31G* (q^{ESP})
N	−0.5403	−5.5624	−0.9145	−0.9637
H(N)	0.3069	0.8209	0.3991	0.3419
C_α	−0.0528	6.1962	−0.0083	0.6296
H_α	0.2347	−1.2792	0.3013	0.0690
C	0.3710	−0.8152	0.8328	0.3524
O	−0.3964	−0.5648	−0.6468	−0.6346
C_β	−0.1888	−4.6016	−0.3596	−0.2660
$H_{\beta3}$	0.1921	1.1542	0.2501	0.0832
$H_{\beta2}$	0.1426	0.6820	0.1807	0.0959
C_γ	−0.0833	1.2703	0.2802	0.2473
$N_{\delta1}$	−0.2937	−0.1343	−0.9053	−0.4730
$H_{\delta1}^+$	0.3777	0.8715	0.4483	0.4288
$C_{\delta2}$	−0.0778	−0.6847	0.1800	0.0619
$H_{\delta2}$	0.3617	0.1458	0.3704	0.1410
$C_{\varepsilon1}$	0.0330	−2.4836	0.5501	0.2165
$H_{\varepsilon1}$	0.3607	1.7427	0.3516	0.2269
$N_{\varepsilon2}$	−0.2724	1.0889	−0.8636	−0.3348
$H_{\varepsilon2}$	0.4199	−0.0400	0.4877	0.4201

Table 3
Chemical reaction energy of the octapeptide cleavage in gaseous phase[a]

Octapeptide P8 (hartree)	−2990.6747
Quatpeptide 1 P4_1 (hartree)	−1461.3733
Quatpeptide 2 P4_2 (hartree)	−1605.2700
Water (hartree)	−76.0107
Energy (ΔE, kJ/mol)	110.8

[a] 1 Hartree = 2625.5 kJ/mol.

the two ends of peptide bond (Gln)CαCO=NHCα(Ser). The X-coordinates of the six atoms are almost the same. The carbonyl group CO of glutamine and the nitrogen atom N(NH) of Serine form a π_3^4 bond. The catalytic dyad His-41 and Cys-145 in the active pocket between domain I and II attracts π-electron density from the peptide bond Gln–Ser, causing the increase of positive charge on C(CO) of glutamine and negative charge on N(NH) of serine, and that the electrophilic proton H$^+$ attacks N(NH) of serine and neucleophilic OH$^-$ attacks C(CO) of glutamine, respectively. The catalytic functional group is the imidazole ring of His-41 and plays the acid–base catalytic role. The pK value of imidazole group of histidine is 6.0, the concentration of [H$^+$] is the same as in water, and hence, His-41 serves as a good proton provider in life condition [33].

Fig. 4 is the electron density counter map of peptide bond Gln–Ser in gaseous phase on the π-plane consisting of C and O in carbonyl group C=O of Gln and N(NH) of Ser. The electron density counters surrounding atom N(NH) and C(CO) form two triangles like sp^2 hybrid orbits, therefore, the 2p$_x$ orbits of the three atoms, which are perpendicular to the plane, form a π_3^4 bond system. In Table 4, we list the atomic charges of the six atoms on both sides of peptide bond Gln–Ser obtained from ab inito HF/6-31G* calculations in gaseous phase with the background charges of SARS CoV M^{pro}. The negative charge of N(NH) in serine increases to −0.8689 in SARS CoV M^{pro} background charges from −0.8344 in gaseous phase. The positive charge of carbonyl carbon C(CO) of glutamine increases to 0.8074 in the background charges from 0.7706 in gaseous phase, and hence, this is favorable to the cleavage reaction.

Fig. 5 is the counter map of electronic density difference of peptide bond Gln–Ser in the octapeptide AVLQSGFR obtained by subtracting the electron density in gaseous phase from the electron density in the background charges [24] of SARS CoV M^{pro}. In Fig. 5, the grey bold line is the 0-value line that means the electronic densities in gaseous phase and in protease background charges are unchanged. The solid thin lines show the regions where the electron densities are greater

Table 4
Atomic charges and coordinates of six atoms on the both sides of peptide bond Gln–Ser in the octapeptide

	Glutamine side			Serine side		
	C_α	C(CO)	O(CO)	C_α	N(NH)	H(NH)
q^{ESP} (gas)[a]	0.0247	0.7706	−0.6743	−0.0128	−0.8343	0.4120
q^{ESP} (M^{pro})[b]	0.0805	0.8074	−0.6545	−0.2073	−0.8689	0.4212
X (Å)	−34.348	−34.698	−34.813	−34.916	−34.878	−34.732
Y (Å)	14.691	15.100	16.316	14.399	14.156	13.185
Z (Å)	−2.681	−4.041	−4.254	−6.348	−4.998	−4.769

[a] Atomic charges in the gaseous phase.
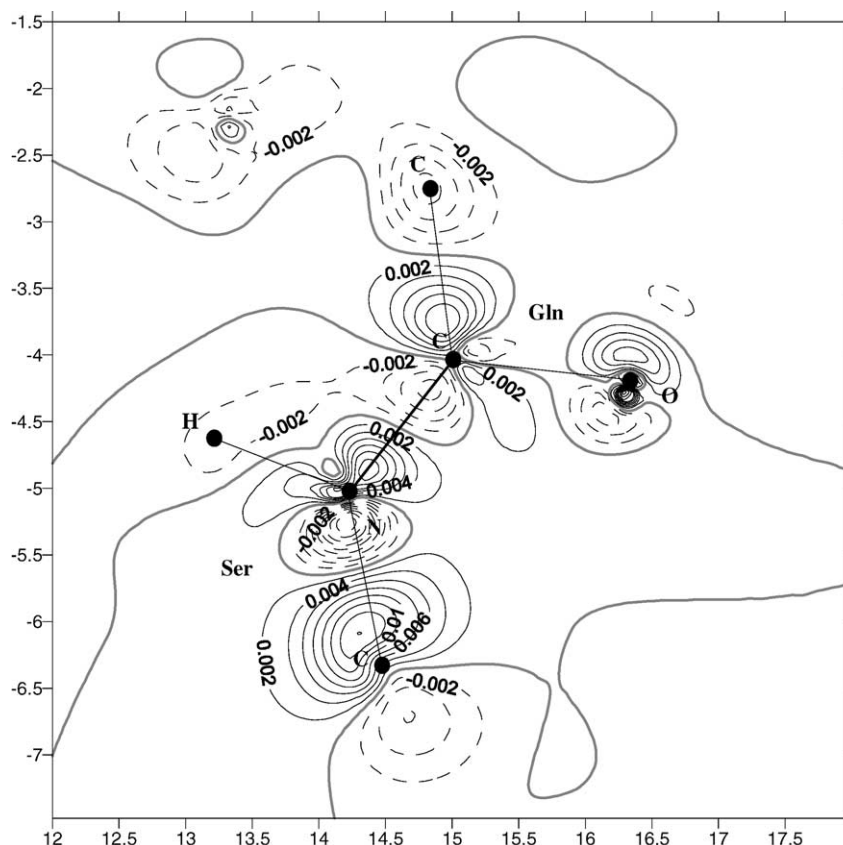[b] Atomic charges in SARS CoV M^{pro} background charges.

Fig. 5. The counter map of electron density difference of peptide bond Gln–Ser in the octapeptide AVLQSGFR obtained by subtracting the electronic density in gaseous phase from the electronic density in background charges of SARS CoV M^{pro}.

Table 5
Atomic charges of the six atoms on two sides of peptide bond Gln–Ser after chemical modification

	Glutamine side			Serine side		
	C_α	C(CH$_2$)	H(CH$_2$)	C_α	N(NH)	H(NH)
q^{ESP} (gas)[a]	0.0578	−0.1040	0.1275	−0.0081	−0.7690	0.3887
q^{ESP}(M^{pro})[b]	0.1127	−0.1211	0.1489	−0.2164	−0.7771	0.4014

[a] Atomic charges in gaseous phase.
[b] Atomic charges in SARS CoV M^{pro} background charges.

in SARS CoV M^{pro} background charges than in gaseous phase and the dashed thin lines show the areas where the electron densities are smaller in protease background charges than in gaseous phase. We find that along the peptide bond between (Gln)C–N(Ser) the electron densities increase on N(Ser) side and decrease on (Gln)C side. This change is favorable for the neucleophilic attack of anion OH$^-$ on (Gln)C and electrophilic attack of cation H$^+$ on N(Ser).

SARS CoV M^{pro} has a very high selectivity [7,34], and in the polyprotein cleavage sites, the subsite R$_1$ is invariably occupied by Gln. A simple routine way to make the octapeptide AVLQSGFR to an effective inhibitor is to change the cleavable scissile peptide bond to the solid single bond by some chemical modification [12,30]. If we replace the carboxyl group CO of glutamine on subsite R$_1$ to CH$_2$ or CF$_2$ group, the π-bond system is broken and the modified octapep-

tide AVLQSGFR [20] may become a competent inhibitor for SARS CoV M^{pro} and an effective drug candidate against SARS. Here we show the possibility of chemical modification to the octapeptide AVLQSGFR through computational modeling.

Table 5 lists atomic charges of the six atoms on the both sides of hybrid peptide bond Gln–Ser after changing carbonyl group CO of Gln to CH$_2$ group. Comparing with Table 4, we find that the carbon atom in CH$_2$ group of hybrid peptide bond turns to be negative −0.1040 from positive charge 0.7706 in CO group (see Table 4), and hence, the neucleophilic attack by OH$^-$ is impossible. On the other hand, the negative charge of N(NH) of the Ser side decreases from −0.8689 to −0.7771 in Table 4, and hence, the electrophilic attack by H$^+$ is more difficult. The third row in Table 5 is the atomic charges in the background

charge distribution of SARS CoV M^{pro}. The atomic charge of C(CH$_2$) on Gln side is down to -0.1211, and hence, the neucleophilic attack by OH$^-$ is, indeed, more difficult.

5. Discussion and conclusion

The octapeptide NH$_2$−Ala–Val–Leu–Gln–Ser–Gly–Phe–Arg−COOH is the first one designed based on the molecular structure of SARS CoV M^{pro} [20] and is proved to be a cleavable octapeptide experimentally. The last eight amino acid residues in the decapeptide used by Yang et al. [34] are exactly the same as the octapeptide AVLQSGFR originally investigated by Chou et al. [20]. A cleavable octapeptide could be changed to an effective inhibitor of SARS CoV M^{pro} or a candidate for anti SARS drug after a proper chemical modification to replace the scissile peptide bond to a hybrid peptide bond. The modified octapetide loses its cleavability but it can still bind to the active site, thus becoming a stable inhibitor [16] or a "distorted key" [12].

The number of possible octapeptides is huge ($20^8 = 2.56 \times 10^{10}$) [12,13] and the octapeptide AVLQSGFR may be not the best cleavable one. Accordingly, in searching for the potential inhibitors, a matter of paramount importance is to discern what kind of peptides can be cleaved by SARS CoV M^{pro} and what kind cannot. Even limited in the range of octapeptides, it is by no means easy to answer the question. It would be exhausting to experimentally test out such a huge number of possible octapeptides. A good cleavable octapeptide could be found by using some existing prediction algorithms, such as discriminant function algorithm [12,18], vectorized sequence-coupled algorithm [13], and some other relevant algorithms [8,9,35]. These prediction methods have been successfully applied to help the inhibitor design for HIV protease [1–6,12,16].

Any statistical prediction methods depend on training sets of known samples [21,37]. For the prediction of cleavable octapeptides for SARS CoV M^{pro}, we need two types of sets: one is positive training set consisting of cleavable octapeptides and other is negative training set consisting only of uncleavable octapeptides. The building of negative training set is easy. For example, if we know a protein consisting of 129 amino acid residues is not cleavable by SARS CoV M^{pro}, we can get $122 = 129 - 7$ non cleavable octapeptides immediately. However, building a positive training set needs some experiments. Computer simulation and molecular modeling are a good ways to accelerate this procedure. The method and conclusion from this research is helpful for this purpose.

Acknowledgements

References

[1] Althaus IW, Chou JJ, Gonzales AJ, Diebel MR, Chou KC, Kezdy FJ, et al. Kinetic studies with the nonnucleoside HIV-1 reverse transcriptase inhibitor U-88204E. Biochemistry 1993;32:6548–54.

[2] Althaus IW, Chou JJ, Gonzales AJ, Diebel MR, Chou KC, Kezdy FJ, et al. Steady-state kinetic studies with the non-nucleoside HIV-1 reverse transcriptase inhibitor U-87201E. J Biol Chem 1993;268:6119–24.

[3] Althaus IW, Chou JJ, Gonzales AJ, Diebel MR, Chou KC, Kezdy FJ, et al. Steady-state kinetic studies with the polysulfonate U-9843, an HIV reverse transcriptase inhibitor. Experientia 1994;50:23–8.

[4] Althaus IW, Chou JJ, Gonzales AJ, Diebel MR, Chou KC, Kezdy FJ, et al. Kinetic studies with the non-nucleoside HIV-1 reverse transcriptase inhibitor U-90152E. Biochem Pharmacol 1994;47:2017–28.

[5] Althaus IW, Chou KC, Franks KM, Diebel MR, Kezdy FJ, Romero DL, et al. The benzylthio-pyrididine U-31355 is a potent inhibitor of HIV-1 reverse transcriptase. Biochem Pharmacol 1996;51:743–50.

[6] Althaus IW, Gonzales AJ, Chou JJ, Diebel MR, Chou KC, Kezdy FJ, et al. The quinoline U-78036 is a potent inhibitor of HIV-1 reverse transcriptase. J Biol Chem 1993;268:14875–80.

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

[8] Chou JJ. A formulation for correlating properties of peptides and its application to predicting human immunodeficiency virus protease-cleavable sites in proteins. Biopolymers 1993;33:1405–14.

[9] Chou JJ. Predicting cleavability of peptide sequences by HIV protease via correlation-angle approach. J Protein Chem 1993;12:291–302.

[10] Chou JJ, Li H, Salvesen GS, Yuan J, Wagner G. Solution structure of BID, an intracellular amplifier of apoptotic signaling. Cell 1999;96:615–24.

[11] Chou JJ, Matsuo H, Duan H, Wagner G. Solution structure of the RAIDD CARD and model for CARD/CARD interaction in caspase-2 and caspase-9 recruitment. Cell 1998;94:171–80.

[12] Chou KC. Review: prediction of HIV protease cleavage sites in proteins. Anal Biochem 1996;233:1–14.

[13] Chou KC. A vectorized sequence-coupling model for predicting HIV protease cleavage sites in proteins. J Biol Chem 1993;268:16938–48.

[14] Chou KC, Howe WJ. Prediction of the tertiary structure of the beta-secretase zymogen. Biochem Biophys Res Commun 2002;292:702–8.

[15] Chou KC, Jones D, Heinrikson RL. Prediction of the tertiary structure and substrate binding site of caspase-8. FEBS Lett 1997;419:49–54.

[16] Chou KC, Kezdy FJ, Reusser F. Review: Steady-state inhibition kinetics of processive nucleic acid polymerases and nucleases. Anal Biochem 1994;221:217–30.

[17] Chou KC, Tomasselli AG, Heinrikson RL. Prediction of the tertiary structure of a caspase-9/inhibitor complex. FEBS Lett 2000;470:249–56.

[18] Chou KC, Tomasselli AL, Reardon IM, Heinrikson RL. Predicting HIV protease cleavage sites in proteins by a discriminant function method. Proteins: Struct Funct Genet 1996;24:51–72.

[19] Chou KC, Watenpaugh KD, Heinrikson RL. A Model of the complex between cyclin-dependent kinase 5(Cdk5) and the activation domain of neuronal Cdk5 activator. Biochem Biophys Res Commun 1999;259:420–8.

[20] Chou KC, Wei DQ, Zhong WZ. Binding mechanism of coronavirus main proteinase with ligands and its implication to drug design against SARS (erratum: ibid.vol.310, 675). Biochem Biophys Res Commun 2003;308:148–51.

[21] Chou KC, Zhang CT. Review: Prediction of protein structural classes. Crit Rev Biochem Mol Biol 1995;30:275–349.

[22] 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;348:1967–76.

[23] Gan YR, Huang YD, Huang H. Synthesis and biological evaluation of a novel SARS CoV Mpro inhibitor. Personal Communication.

[24] Foresman JB, Frisch A. Exploring chemistry with electronic structure methods: a guide to using Gaussian. Gaussian Inc., Pittsburgh, PA; 1993.

[25] Hehre WJ, Radom L, Schleyer PR, Pople JA. Ab initio molecular orbital theory. New York: Wiley; 1986, chapter 2.

[26] 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;348:1953–66.

[27] Miura HS, Nakagaki K, Taguchi F. N-terminal domain of the murine coronavirus receptor CEACAM1 is responsible for fusogenic activation and conformational changes of the spike protein. J Virol 2004;78:216–23.

[28] Molecular operating environment (MOE). Montreal, Quebec, Canada: Chemical Computing Group; 2003.

[29] Qin E, Zhu Q, Yu M, Fan B, Chang G, Si B, Yang B, Peng W, Jiang T, Liu B, et al. SARS coronavirus BJ03 isolate genome sequence. Submitted APR-2003 to the EMBL GenBank DDBJ databases 2003.

[30] Schulz GE, Schirmer RH. Principles of protein structure. New York: Springer–Verlag; 1985, Chapter 2.

[31] Shortridge KF. Severe acute respiratory syndrome and influenza: virus incursions from southern China. Am J Respir Crit Care Med 2003;168:1416–20.

[32] Sirois S, Wei DQ, Du Q, Chou KC. Virtual screening for SARS CoV protease based on KZ7088 pharmacophore points. J Chem Inf Comput Sci 2004;44:1111–22.

[33] Stryer L. Biochemistry. 3rd ed. New York: W.H. Freeman and Company; 1988.

[34] 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;100:13190–5.

[35] Zhang CT, Chou KC. An alternate-subsite-coupled model for predicting HIV protease cleavage sites in proteins. Protein Eng 1993;7:65–73.

[36] Zhang J, Luan CH, Chou KC, Johnson GVW. Identification of the N-terminal functional domains of Cdk5 by molecular truncation and computer modeling. Proteins: Struct Funct Genet 2002;48:447–53.

[37] Zhou GP, Doctor K. Subcellular location prediction of apoptosis proteins. Proteins: Struct Funct Genet 2003;50:44–8.