3C-like Proteinase from SARS Coronavirus Catalyzes Substrate Hydrolysis by a General Base Mechanism†

Changkang Huang,‡ Ping Wei,‡ Keqiang Fan,‡ Ying Liu,‡ and Luhua Lai*,‡,§

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

Received November 11, 2003; Revised Manuscript Received January 26, 2004

ABSTRACT: SARS 3C-like proteinase has been proposed to be a key enzyme for drug design against SARS. Lack of a suitable assay has been a major hindrance for enzyme kinetic studies and a large-scale inhibitor screen for SARS 3CL proteinase. Since SARS 3CL proteinase belongs to the cysteine protease family (family C3 in clan CB) with a chymotrypsin fold, it is important to understand the catalytic mechanism of SARS 3CL proteinase to determine whether the proteolysis proceeds through a general base catalysis mechanism like chymotrypsin or an ion pair mechanism like papain. We have established a continuous colorimetric assay for SARS 3CL proteinase and applied it to study the enzyme catalytic mechanism. The proposed catalytic residues His41 and Cys145 were confirmed to be critical for catalysis by mutating to Ala, while the Cys145 to Ser mutation resulted in an active enzyme with a 40-fold lower activity. From the pH dependency of catalytic activity, the pK_a's for His41 and Cys145 in the wild-type enzyme were estimated to be 6.38 and 8.34, while the pK_a's for His41 and Ser145 in the C145S mutant were estimated to be 6.15 and 9.09, respectively. The C145S mutant has a normal isotope effect in D_2O for general base catalysis, that is, reacts slower in D_2O, while the wild-type enzyme shows an inverse isotope effect which may come from the lower activation enthalpy. The pK_a values measured for the active site residues and the activity of the C145S mutant are consistent with a general base catalysis mechanism and cannot be explained by a thiolate−imidazolium ion pair model.

A novel form of coronavirus has been attributed as the major cause of severe acute respiratory syndrome (SARS) that broke out in early 2003 (*1*, *2*). The genome of SARS coronavirus has been sequenced within a short period of time after isolation of the virus (*3*, *4*). Coronaviruses are members of positive-stranded RNA viruses featuring the largest viral RNA genomes known to date. Translation of the SARS coronavirus replicase gene produces two proteins with overlapping sequences, polyproteins 1a (∼450 kDa) and 1ab (∼750 kDa), which are conserved both in length and in amino acid sequence to other coronavirus replicase proteins. Polyproteins 1a and 1ab are cleaved by the internally encoded 3C-like proteinase (3CL) to release functional proteins necessary for virus replication. SARS 3CL[1] proteinase is fully conserved among all released SARS coronavirus genome sequences and is highly homologous with other coronavirus 3CL proteinases. Due to its functional importance in the viral life cycle, SARS 3CL proteinase has been proposed to be a key target for structure-based drug design against SARS (*5*). Homology modeling for the SARS 3CL proteinase has been performed by various groups (*5*−*7*), and the conformational flexibility of the substrate binding site has been studied (*7*). Virtual screening of small molecule libraries has predicted possible inhibitors (*6*). An 8-mer peptide has been docked on the model of SARS 3CL proteinase to study the interaction of SARS 3CL with its substrates (*8*).

Two crystal structures of coronavirus 3CL proteinase from transmissible gastroenteritis virus (TGEV) (*9*) and human coronavirus (hCoV) 229E have been reported (*5*). Recently, the crystal structure of SARS coronavirus 3CL proteinase has also been solved by two groups, with the coordinates deposited in the Protein Data Bank (http://www.rcsb.org). In addition, Z. H. Rao's group also elucidated the complex structure of SARS 3CL proteinase with a covalent bonded substrate analogue, which provided insights on the substrate binding site (*10*). All of the reported structures of coronavirus 3CL proteinases are similar. The first two domains form a chymotrypsin fold which is responsible for the catalytic reaction, and the third domain is α helical with unclear biological function. Coronavirus 3CL proteinase shares the chymotrypsin fold with the 3C proteinases from other viruses such as rhinovirus, picornavirus, and hepatitis A virus (*11*−*13*). The 3C proteinase of rhinovirus has been targeted to develop drugs against the common cold (*14*−*20*).

As 3C proteinases and the catalytic domain in 3CL proteinases are all cysteine proteinases with a chymotrypsin fold, it is interesting to know whether they follow a thiolate−imidazolium ion pair catalytic mechanism as found with

† This work was supported by the 863 High-Technology Project, the National Basic Research Project (2003CB715900), and NSFC (10345004, 90103029, 20173001, 20228306).

* Corresponding author. Tel: +86 10 62757486. Fax: +86 10 62751725. E-mail: lhlai@pku.edu.cn.

‡ State Key Laboratory for Structural Chemistry of Unstable and Stable Species, College of Chemistry and Molecular Engineering, Peking University.

§ Center for Theoretical Biology, Peking University.

[1] Abbreviations: SARS 3CL, SARS coronavirus 3C-like proteinase; CD, circular dichroism; BME, 2′-mercaptoethanol; DTT, 1,4-dithiothreitol.

10.1021/bi036022q CCC: $27.50
Published on Web 03/12/2004

Table 1: Primers Used to Generate SARS 3CL Proteinase Mutants

mutants	sequence
H41A	sense primer: 5′-CAGTATACTGTCCAAGAGCAGTCATTTGCACAGCAGAAGAC-3′
	antisense primer: 5′-GTCTTCTGCTGTGCAAATGACTGCTCTTGGACAGTATACTG-3′
C145A	sense primer: 5′-CTTTCCTTAATGGATCAGCAGGTAGTGTTGGTTTTAAC-3′
	antisense primer: 5′-GTTAAAACCAACACTACCTGCTGATCCATTAAGGAAAG-3′
C145S	sense primer: 5′-CTTTCCTTAATGGATCATCTGGTAGTGTTGGTTTTAAC-3′
	antisense primer: 5′-GTTAAAACCAACACTACCAGATGATCCATTAAGGAAAG-3′

papain (*21*−*24*). It was recently shown that 3C proteinases from picornain virus, poliovirus, and sortase from *Staphylococcus aureus* do not contain a thiolate−imidazolium ion pair in their active site (*25*−*27*). Compared to the catalytic triad in 3C proteinase, only two catalytic residues were found in 3CL proteinases of coronaviruses (*5*, *9*, *28*). Whether 3CL proteinase of coronavirus follows an ion pair mechanism or a general base catalytic mechanism is important to understand the kinetic mechanism and to carry out structural and/or mechanism-based drug design. On the basis of our previous study on the cloning, expression, and substrate specificity of SARS 3CL proteinase (*29*), we have established a colorimetric assay using a synthetic substrate for SARS 3CL proteinase and characterized its enzyme catalysis mechanism. While His41 and Cys145 were confirmed to be crucial for catalysis by mutating to Ala, the substitution of Cys145 with Ser resulted in an active enzyme. From pH-dependent catalytic activities, the calculated pK_a values for His41 and Cys145 in the wild-type enzyme and for His41 and Ser145 in the C145S mutant indicated that the enzyme catalysis follows a general base catalysis mechanism and does not contain a thiolate−imidazolium ion pair in its active site.

EXPERIMENTAL PROCEDURES

Site-Directed Mutagenesis of SARS-CoV 3CL Proteinase. The construction of the plasmid pET 3CLP-21x and the expression and purification of the C-terminal His-tagged SARS-CoV 3CL proteinase have been reported previously (*29*). The H41A, C145A, C145S mutants in SARS-CoV 3CL proteinase were prepared with the QuikChange site-directed mutagenesis kit (Stratagene) using pET 3CLP-21x as a template. The nucleotide sequences of primers used for site-directed mutagenesis are listed in Table 1. The mutations were verified by sequencing. The 3CL proteinase mutants were expressed in *Escherichia coli*, purified, and analyzed using a similar procedure for the wild-type protein (*29*). Briefly, appropriate plasmids were transformed into *E. coli* BL21(DE3) cells. Cultures were grown at 37 °C in 1 L of LB containing ampicillin (100 μg/mL) until the OD at 600 nm reached 0.8 and induced with 0.5 mM IPTG at 30 °C for 3 h. Harvested cell pellets were suspended in buffer A (40 mM Tris-HCl, pH 8.0, 100 mM NaCl, 10 mM imidazole, 7.5 mM BME) at $^1/_{50}$th of the original culture volume. After sonication, the cell lysate was centrifuged at 24000*g* for 20 min. After filtration, the supernatant was applied to a Ni-NTA column (Qiagen) equilibrated with 50 mL of buffer A. After being washed with 100 mL of buffer A, 3CL proteinase was eluted with a gradient of 1−100% buffer B (40 mM Tris-HCl, pH 8.0, 100 mM NaCl, 250 mM imidazole, 7.5 mM BME). The eluted enzyme was concentrated and loaded on a gel filtration column, Sephacryl S-200 HR (Pharmacia), equilibrated by 180 mL of buffer C (40 mM Tris-HCl, pH 8.0, 100 mM NaCl, 7.5 mM BME) and eluted with another 180 mL of buffer C. The purity of each mutant was verified by SDS−PAGE. Secondary and tertiary structures of the mutants were determined to be essentially the same as the wild-type enzyme (data not shown) as monitored by far- and near-UV CD.

Construction of Plasmid pET 3CLP-21h, Protein Expression, and Purification. The non-His-tagged SARS 3C-like proteinase was amplified by PCR from the plasmid pET 3CLP-21x using primers 3CLP-*Nhe* (5′-CACTGCTAGCGGTTTTAGGAAAATGGCATTCCC-3′) and 3CLP-*Hin*d (5′-CCTCAAGCTTATTGGAAGGTAACACCAGAGC-3′). The PCR product was inserted into the *Nhe*I and *Hin*dIII sites of pET21a DNA. The resulting plasmid pET 3CLP-21h was transformed into *E. coli* BL21(DE3) cells for expression. The recombinant 3CLP-21h was purified by fractional precipitation with ammonium sulfate and further purified by using an anion-exchange column, HiTrap Q HP (Amersham Bioscience), and a gel filtration column, Sephacryl S-200 HR (Amersham Bioscience).

Substrate Specificity of C145S. Three substrate peptides derived from the natural cleavage sites of the SARS-CoV polyprotein were used in this study and were synthesized as previously described (*29*). Their sequences are listed in Table 3. The SARS 3CL proteinase C145S mutant was incubated with different substrates in 40 mM Tris-HCl buffer, pH 7.3, with a final enzyme concentration of 76.9 μM. The concentrations of substrate peptides were 0.4 mM for S01 and S06 and 0.2 mM for S09. The reaction mixture was analyzed by RP-HPLC using a 15 min, 0−50% linear gradient of acetonitrile with 0.1% trifluoroacetic acid after 25 h. k_{cat}/K_m was calculated using the equation:

$$\ln \mathrm{PA} = \ln \mathrm{PA}_0 - k_{cat}C_E t/K_m \quad (1)$$

where PA was the peak area of the substrate peptide in the reacting mixture after 25 h, PA_0 was the origin peak area of the substrate peptide, and C_E was the total concentration of 3C-like proteinase.

Continuous Colorimetric Enzyme Assay. A colorimetric substrate, Thr-Ser-Ala-Val-Leu-Gln-pNA, of HPLC purity was purchased from GL Biochemistry Ltd. The first six residues of the substrate correspond to the N-terminal self-cleavage site of the SARS 3CL proteinase. Colorimetric measurements of SARS 3CL activity were performed in 96-well microtiter plates using a multiwell ultraviolet spectrometer (Spectra Max 190; Molecular Device). This substrate is cleaved at the Gln-pNA bond to release free pNA, resulting in an increase of absorbacne at 390 nm. To verify the cleavage site on the substrate, the two products were purified by semipreparative RP-HPLC using a 15 min, 0−50% linear gradient of acetonitrile with 0.1% TFA and lyophilized. The

molecular weight of the hexapeptide product was identified by MALDI-TOF-MS (BIFLEX III time-of-flight mass spectrometer; Bruker). The MW of cleaved pNA was confirmed by EIMS (Micromass GCT mass spectrometer, Manchester, England). The amount of pNA released from proteolysis can be calculated with a standard curve generated using analytical grade pNA. A sample of preheated substrate stock solution (2 mM water solution) was added to the preheated reaction mixture, containing 155 μL of standard buffer (75 mM Tris-HCl, 25 mM NaOAc, 25 mM Bis-Tris, 25 mM glycine, 5 mM EDTA, and 1 mM DTT, pH 7.4) and 20 μL of enzyme stock solution. The kinetic curve was recorded for 20 min. The enzyme concentration was calculated from absorbance at 280 nm with 1.08 mg/OD.

Kinetic Measurements. The pH dependence of rate constants was measured at 37 °C in the standard buffer, adjusted to the required pH with HCl or NaOH solution. The second-order rate constant (k_{cat}/K_m) was calculated by dividing the pseudo-first-order constant by the enzyme concentration. The data were fitted by Origin 6.0 (OriginLab) using the equation:

$$k = k_0[1/(1 + 10^{pK_{a1}-pH} + 10^{pH-pK_{a2}})] \quad (2)$$

in which k_0 is the pH-independent rate constant and pK_{a1} and pK_{a2} are the values of catalytically competent functional groups, i.e., side chain of His41 and Cys145 or Ser145.

The isotope effect on catalysis was tested in deuterium oxide. The deuterium oxide content of each reaction was more than 98%. The p^{2}H of heavy water buffer solution (75 mM Tris-HCl, 25 mM NaOAc, 25 mM Bis-Tris, 25 mM glycine, 5 mM EDTA, and 1 mM DTT, p^{2}H 8.1) can be calculated from p^{2}H = pH(meter reading) + 0.4 (*27*).

The temperature dependence of the enzyme reaction was studied. A sample of the preheated substrate was added to a reaction mixture to get a final concentration of 2.8 μM SARS 3CL or 21.4 μM C145S and 125−250 μM substrate. To get k_{cat} and K_m at different temperatures, data at each temperature were fitted using eq 3 by Origin 6.1 (OriginLab):

$$E/v = (K_m/k_{cat})/S + 1/k_{cat} \quad (3)$$

where E and S are the concentrations of enzyme and substrate and v is the initial rate of the reaction, which is gotten from the average velocity before 10% of the substrate was hydrolyzed.

Activation parameters can be calculated using the equation:

$$\ln(k_{cat}/K_mT) = \ln(R/N_Ah) + \Delta S^*/R - \Delta H^*/RT \quad (4)$$

where R is the gas constant (8.314 J·mol^{-1}·K^{-1}), T is the absolute temperature, N_A is Avogadro's number, and h is the Planck constant. The enthalpy of activation ΔH^* = −(slope) × 8.314 J·mol^{-1}, and the entropy of activation ΔS^* = (intercept −23.76) × 8.314 J·mol^{-1}·K^{-1}. The activation free energy can be calculated from the equation:

$$\Delta G^* = \Delta H^* - T\Delta S^* \quad (5)$$

K_m is the dissociation constant of the substrate when the dissociation velocity of the ES complex is much faster than product formation. Since SARS 3CL proteinase is a catalytically inefficient enzyme compared to other proteases, we assumed that K_m for proteolysis represents the affinity of the initial Michaelis−Menten complex. Thus, thermodynamic parameters can be calculated from a linear plot of ln($1/K_m$) versus $1/T$:

$$\ln(1/K_m) = -\Delta G(\text{bind})/RT = \Delta S(\text{bind})/R - \Delta H(\text{bind})/RT \quad (6)$$

The enthalpy of binding ΔH(bind) = −(slope) × 8.314 J.mol^{-1}, and the entropy of binding ΔS(bind) = intercept × 8.314 J·mol^{-1}·K^{-1}.

RESULTS AND DISCUSSION

Development of a Continuous Assay for SARS 3CL Proteinase. Most reported enzyme assays for coronavirus 3CL proteinase use natural peptide substrates for cleavage and HPLC for analysis (*5*, *28*). Rather than using peptides modified with chromophores or fluorophores, HPLC assay allows the use of natural peptide substrate. But the analysis is relatively time-consuming and difficult for kinetic studies and for a high-throughput screen for inhibitors. Clearly, a high-throughput enzyme assay for SARS 3CL proteinase will speed up discovery of novel inhibitors. Fluorescent and colorimetric assays have been successfully developed for 3C proteinases, including rhinovirus, poliovirus, etc. Enzyme kinetic studies and inhibitor screening were carried out successfully using these assay formats (*30*−*32*). We have developed a colorimetric assay for SARS 3CL proteinase which uses a colorimetric peptide with the six amino acid residues mapped from the N-terminal self-cleavage site of SARS 3CL proteinase attached by a pNA group. This modified peptide was found to be successfully cleaved by SARS 3CL proteinase at the designated site (Gln-pNA) as verified by HPLC and mass spectroscopy with experimental $(M + Na)^+$ of 640 and (M) of 138, respectively, which were identical to the theoretical $(M + Na)^+$ of the hexapeptide and (M) of pNA. The measured k_{cat} and K_m (Table 4) for this pNA peptide are comparable to an 11-mer natural peptide substrate mapped from the same site of the polyprotein (*29*). With a multiplate reader, this assay can be applied to enzyme kinetics studies and to high-throughput screening of inhibitors.

pH Dependence of the Specific Activity of SARS 3CL Wild Type and the C145S Variant. The recombinant protein SARS 3CL (21X) contains a C-terminal His tag. This was thought to have no effect on the enzyme activity as both of the C-terminals in the crystal structures of TGEV 3CL and SARS 3CL proteinase are far from the active site and disordered. To confirm that the C-terminal His tag of SARS 3CL (21X) have no effect on the activity of the enzyme, we compared the specific activites with that of non-His-tagged protein [SARS 3CL (21H)] over a large pH range. Table 2 shows that SARS 3CL proteinase with a C-terminal His tag and that without the His tag have similar activity and optimum pH both in water and in heavy water, confirming that the C-terminal His tag does not have notable effects on the enzyme activity. As the enzyme with His tag is easier to express and purify, the C-terminal His-tagged enzyme is used in all of the other experiments in this paper. Figure 1 shows the pH−rate profile of SARS 3CL proteinase. The optimum pHs are 7.38 and 7.60 for the wild type and the C145S variant, respectively. The two edges of the bell-shaped curve are associated with pK_a's of the catalytic histidine and

Table 2: Kinetic Parameters and pK_a Values for SARS 3CL Wild Type and the C145S Mutant from the Continuous Colorimetric Assay[a]

variant	solvent	k_0 ($mM^{-1}\,min^{-1}$)	pK_{a1}	pK_{a2}
SARS 3CL	H_2O	7.8 ± 0.1	6.38 ± 0.02	8.34 ± 0.02
(21X)	D_2O	8.8 ± 0.2	6.24 ± 0.05	8.95 ± 0.06
SARS 3CL	H_2O	7.3 ± 0.2	6.25 ± 0.04	8.29 ± 0.04
(21H)	D_2O	8.6 ± 0.3	6.05 ± 0.06	8.73 ± 0.06
SARS 3CL	H_2O	0.28 ± 0.01	6.15 ± 0.05	9.09 ± 0.04
(21X) C145S	D_2O	0.20 ± 0.01	6.24 ± 0.07	9.58 ± 0.08

[a] All data were collected at 37 °C. The concentration of SARS 3CL (21X) and SARS 3CL (21H) measured in this experiment is 2.8 μM and of SARS 3CL (21X) C145S is 17.1 μM.

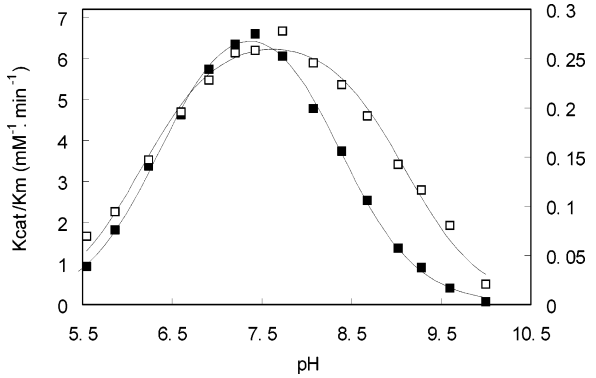

FIGURE 1: The pH−rate profile for SARS 3CL wild type (■) and C145S (□).

cysteine or serine. By fitting the curves with eq 2, the pK_a's for His41 and Cys145 in the wild-type enzyme were calculated to be 6.38 ± 0.02 and 8.34 ± 0.02, and the pK_a's for His41 and Ser145 in the C145S mutant are 6.15 ± 0.05 and 9.09 ± 0.04 (Table 2). Although there is not much difference in pK_a for His41 in the wild type and the mutant, there is a shift of about 0.7 pH unit of the pK_a for C145 when mutating to S145. The right limb of the pH−rate profile decreases sharply for the wild-type enzyme above 8.6 due to the dissociation of the thiol group which is essential for catalysis. The thiol proton should be transferred to the imidazole ring to work as a general acid; subsequently, it will protonate the leaving group of the tetrahedral intermediate. Due to the higher pK_a 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 (*27*), indicating that SARS 3CL proteinase takes a similar catalytic mechanism.

The SARS 3CL C145S Mutant Is Active with Reduced Specific Activity. In the ion pair mechanism of cysteine proteinases, a positive histidine and a negative cysteine are mandatory for the active site. The active site His and Cys are conserved among 3CL proteinases of coronavirus. His41 and Cys145 have been predicted to be the active site residues of SARS 3CL proteinase (*3*). We have mutated His41 and Cys145 to Ala, respectively, and both of the mutants were found to be inactive, which confirmed that these two residues are involved in the enzyme catalysis. We also mutated Cys145 to Ser, but the resulting enzyme remains active with a specific activity 40-fold lower than that of the wild-type enzyme at the concentration of 2.8 μM at 37 °C (data not shown). As we have discussed before that the activity of SARS 3CL proteinase changes with the enzyme concentration due to dimerization (*29*), the relative activity of the mutant and the wild-type enzyme are only compared at a given concentration. The products of the mutant C145S are the same as the wild-type enzyme as verified by HPLC. This implies that SARS 3CL proteinase undergoes a different mechanism other than the ion pair mechanism. Mutational studies for the active site Cys to Ser have been carried out on the 3CL proteinases from other coronaviruses and have shown that the mutants are not active (*28*, *33*). This discrepancy may come from the less sensitive HPLC assay used in the previous study. With a continuous colorimetric assay, we have shown here that the C145S mutant is active over a large range of pH (Figure 1) though with a 40-fold lower activity compared to the wild-type enzyme.

Table 3: Effect of the C145S Mutant on Substrate Specificity of SARS 3CL Determined by the HPLC Assay

substrate	$(k_{cat}/K_m)_{rel}$ SARS 3CL (*29*)	$(k_{cat}/K_m)_{rel}$ SARS 3CL C145S
S01 (P1/P2: TSAVLQSGFRK-NH_2)	1.00	1.00
S06 (P6/P7: ATVRLQAGNAT)	0.22	0.14
S09 (P9/P10: NVATLQAENVT)	0.09	<0.02

We also studied the substrate specificity of the C145S mutant. The k_{cat}/K_m values of substrates S01 and S06 obtained by HPLC assay at room temperature are 5.25×10^{-3} and 7.48×10^{-4} mM^{-1} min^{-1}, respectively. k_{cat}/K_m of substrate S09 is too low to determine. Substrate specificities of the wild-type enzyme and the C145S mutant are summarized in Table 3. The order of substrate utilization by both wild-type and C145S enzymes is S01 > S06 > S09, suggesting that substitution of the active site thiol group by a hydroxyl group does not noticeably modify substrate specificity.

Temperature Dependence and Solvent Isotope Effects. The activity of 3CL was measured at different temperatures. All data were fitted using eq 2 to get K_m and k_{cat} at that temperature. The activity (k_{cat}/K_m) of 3CL increases with temperature in the whole range we studied (Table 4). It is necessary to point out that the enzyme equilibrates between the monomeric and the dimeric forms and the dimer is the main active form (*29*); thus all of the catalytic parameters described here were apparent ones at given enzyme concentrations.

Kinetic deuterium effects are studied to understand the general base-catalyzed processes, which should be slower in deuterium oxide than in water as described for serine proteinase. This is true for the C145S mutant with $(k_{cat}/K_m)_{H_2O}/(k_{cat}/K_m)_{D_2O}$ around 1.3 (Figure 2). It is interesting that the wild-type enzyme has a large inverse isotope effect with $(k_{cat}/K_m)_{H_2O}/(k_{cat}/K_m)_{D_2O}$ around 0.4 (Figure 2), which is similar to picornain 3C proteinase (*27*). The three-dimensional structure of SARS 3CL proteinase also has flexible loops near the active site (*7*) which may change conformation upon substrate binding or get stabilized in heavy water. The activity of SARS 3CL improved significantly (about 2 times) when H_2O is replaced by D_2O (Table 4) in the whole temperature range we studied. The enthalpy of activation in H_2O is higher than that in D_2O, though the entropies of activation are nearly the same (Table 5).

As the K_m value is a measure of the affinity of the substrate for the enzyme when the dissociation velocity of the ES

Table 4: Temperature-Dependent Kinetic Parameters for SARS 3CL Wild Type and C145S[a]

enzyme	temp (°C)	solvent: H_2O K_m (mM)	H_2O k_{cat} (min^{-1})	H_2O k_{cat}/K_m ($mM^{-1}\,min^{-1}$)	D_2O K_m (mM)	D_2O k_{cat} (min^{-1})	D_2O k_{cat}/K_m ($mM^{-1}\,min^{-1}$)
SARS 3CL (21X)	21	0.46 ± 0.05	1.3 ± 0.1	2.86 ± 0.09	0.42 ± 0.11	3.1 ± 0.6	7.30 ± 0.56
	25	0.69 ± 0.14	2.7 ± 0.4	3.95 ± 0.15	0.43 ± 0.09	3.9 ± 0.6	9.06 ± 0.52
	29	0.85 ± 0.41	4.4 ± 1.8	5.16 ± 0.40	0.44 ± 0.09	4.6 ± 0.7	10.47 ± 0.59
	33	0.86 ± 0.2.3	5.5 ± 1.2	6.36 ± 0.26	0.45 ± 0.07	5.9 ± 0.7	13.18 ± 0.56
	37	0.98 ± 0.36	6.4 ± 2.0	6.48 ± 0.33	0.46 ± 0.04	6.3 ± 0.4	13.72 ± 0.29
	41	1.47 ± 0.75	10.6 ± 5.0	7.20 ± 0.36	0.47 ± 0.10	7.3 ± 1.2	15.72 ± 0.01
SARS 3CL (21X) C145S	21	0.68 ± 0.27	0.11 ± 0.04	0.16 ± 0.01	0.56 ± 0.07	0.08 ± 0.01	0.15 ± 0.01
	25	0.75 ± 0.35	0.16 ± 0.06	0.21 ± 0.02	0.60 ± 0.15	0.09 ± 0.02	0.15 ± 0.01
	29	0.89 ± 0.27	0.23 ± 0.06	0.26 ± 0.01	0.61 ± 0.08	0.12 ± 0.01	0.20 ± 0.01
	33	1.03 ± 0.53	0.31 ± 0.14	0.30 ± 0.02	0.65 ± 0.15	0.15 ± 0.03	0.23 ± 0.01
	37	1.14 ± 0.76	0.40 ± 0.23	0.35 ± 0.03	0.64 ± 0.16	0.16 ± 0.03	0.26 ± 0.01
	41	1.28 ± 0.69	0.49 ± 0.24	0.38 ± 0.02	0.67 ± 0.23	0.20 ± 0.06	0.30 ± 0.02

[a] The concentration of SARS 3CL (21X) and of SARS 3CL (21X) C145S was 2.8 and 21.4 μM, respectively.

Table 5: Thermodynamic Parameters for SARS 3CL and C145S[a]

enzyme	solvent	ΔH(bind) (kJ/mol)	ΔS(bind) [J/(mol·K)]	ΔG(bind) (kJ/mol)	ΔH^* (kJ/mol)	ΔS^* [J/(mol·K)]	ΔG^* (kJ/mol)
SARS 3CL (21X)	H_2O	−39.6 ± 4.9	−72 ± 16	−17.3	32.0 ± 5.2	−102 ± 17	63.6
	D_2O	−3.5 ± 0.3	53 ± 1	−19.9	26.7 ± 2.4	−114 ± 8	62.0
SARS 3CL (21X) C145S	H_2O	−24.9 ± 0.8	−24 ± 3	−17.5	29.5 ± 2.5	−134 ± 8	71.0
	D_2O	−6.0 ± 1.0	42 ± 3	−19.0	26.0 ± 2.8	−147 ± 9	71.6

[a] ΔG and ΔG^* are calculated at 37 °C. The concentration of SARS 3CL (21X) measured in this experiment is 2.8 μM and of SARS 3CL (21X) C145S is 21.4 μM.

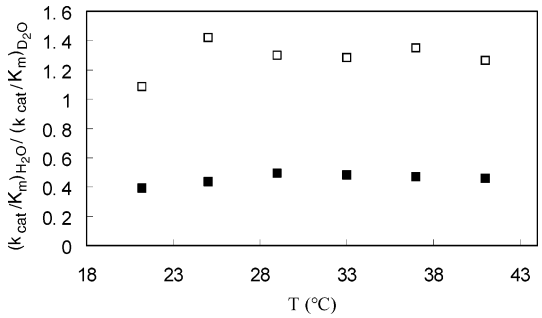

FIGURE 2: $(k_{cat}/K_m)_{H_2O}/(k_{cat}/K_m)_{D_2O}$ measured at different temperatures. The concentration for WT (■) is 2.8 μM, and the concentration for C145S (□) is 21.4 μM.

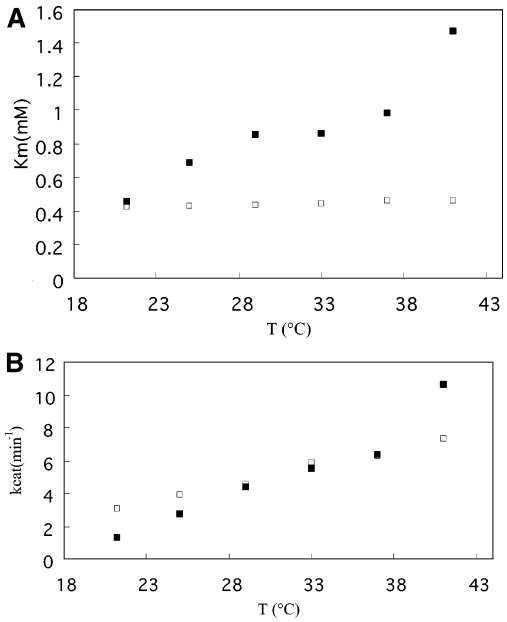

FIGURE 3: Temperature dependence of the SARS 3CL proteinase catalysis parameters. (A) Temperature dependence of K_m in H_2O (■) and D_2O (□). (B) Temperature dependence of k_{cat} in H_2O (■) and D_2O (□).

complex is much faster than product formation (the k_{cat} is quite small in this system), it is clear that the substrate binding to SARS 3CL is better in D_2O buffer. K_m in H_2O increases much faster with temperature than in D_2O (Figure 3A). k_{cat} in H_2O buffer is smaller at low temperatures and increases more quickly than that in D_2O buffer (Figure 3B). The two curves have a cross at about 34 °C. Our study has shown that both the wild-type enzyme and the C145S mutant undergo a general base catalysis mechanism, although the serine proteinase (SARS 3CL C145S) and the cysteine proteinase (SARS 3CL wild type) may differ in the rate-limiting steps.

We have established a continuous colorimetric assay for SARS 3CL proteinase and applied it to study the enzyme catalytic mechanism. Cys145 and His41 were confirmed to be the key catalytic residues. The His tag at the C-terminal does not have a notable effect on the enzyme activity. When Cys145 was mutated to Ser, the resulting enzyme remains active with about 40-fold lower activity. Measurements of pK_a values for Cys145 and Ser145 agree with a general base/acid mechanism for both the wild-type and the mutant enzyme. Deuterium oxide has a normal effect on the C145S, similar to other serine proteinases. The inverse effect of the wild-type enzyme may come from the lower activation enthalpy.

ACKNOWLEDGMENT

The authors thank Prof. Jianguo Chen for providing the clones of SARS-CoV 3CL proteinase and Dr. Zhihong Lai for critical reading of the manuscript and helpful discussions.

REFERENCES

1. Drosten, C., Gunther, S., Preiser, W., van der Werf, S., Brodt, H. R., Becker, S., Rabenau, H., Panning, M., Kolesnikova, L., Fouchier, R. A., Berger, A., Burguiere, A. M., Cinatl, J., Eickmann, M., Escriou, N., Grywna, K., Kramme, S., Manuguerra, J. C., Muller, S., Rickerts, V., Sturmer, M., Vieth, S., Klenk, H. D., Osterhaus, A. D., Schmitz, H., and Doerr, H. W. (2003) Identification of a novel coronavirus in patients with severe acute respiratory syndrome, *N. Engl. J. Med. 348*, 1967−1976.
2. Ksiazek, T. G., Erdman, D., Goldsmith, C. S., Zaki, S. R., Peret, T., Emery, S., Tong, S., Urbani, C., Comer, J. A., Lim, W., Rollin, P. E., Dowell, S. F., Ling, A. E., Humphrey, C. D., Shieh, W. J., Guarner, J., Paddock, C. D., Rota, P., Fields, B., DeRisi, J., Yang, J. Y., Cox, N., Hughes, J. M., LeDuc, J. W., Bellini, W. J., and Anderson, L. J. (2003) A novel coronavirus associated with severe acute respiratory syndrome, *N. Engl. J. Med. 348*, 1953−1966.
3. Rota, P. A., Oberste, M. S., Monroe, S. S., Nix, W. A., Campagnoli, R., Icenogle, J. P., Penaranda, S., Bankamp, B., Maher, K., Chen, M. H., Tong, S., Tamin, A., Lowe, L., Frace, M., DeRisi, J. L., Chen, Q., Wang, D., Erdman, D. D., Peret, T. C., Burns, C., Ksiazek, T. G., Rollin, P. E., Sanchez, A., Liffick, S., Holloway, B., Limor, J., McCaustland, K., Olsen-Rasmussen, M., Fouchier, R., Gunther, S., Osterhaus, A. D., Drosten, C., Pallansch, M. A., Anderson, L. J., and Bellini, W. J. (2003) Characterization of a novel coronavirus associated with severe acute respiratory syndrome, *Science 300*, 1394−1399.
4. Marra, M. A., Jones, S. J., Astell, C. R., Holt, R. A., Brooks-Wilson, A., Butterfield, Y. S., Khattra, J., Asano, J. K., Barber, S. A., Chan, S. Y., Cloutier, A., Coughlin, S. M., Freeman, D., Girn, N., Griffith, O. L., Leach, S. R., Mayo, M., McDonald, H., Montgomery, S. B., Pandoh, P. K., Petrescu, A. S., Robertson, A. G., Schein, J. E., Siddiqui, A., Smailus, D. E., Stott, J. M., Yang, G. S., Plummer, F., Andonov, A., Artsob, H., Bastien, N., Bernard, K., Booth, T. F., Bowness, D., Czub, M., Drebot, M., Fernando, L., Flick, R., Garbutt, M., Gray, M., Grolla, A., Jones, S., Feldmann, H., Meyers, A., Kabani, A., Li, Y., Normand, S., Stroher, U., Tipples, G. A., Tyler, S., Vogrig, R., Ward, D., Watson, B., Brunham, R. C., Krajden, M., Petric, M., Skowronski, D. M., Upton, C., and Roper, R. L. (2003) The Genome sequence of the SARS-associated coronavirus, *Science 300*, 1399−1404.
5. Anand, K., Ziebuhr, J., Wadhwani, P., Mesters, J. R., and Hilgenfeld, R. (2003) Coronavirus main proteinase (3CLpro) structure: basis for design of anti-SARS drugs, *Science 300*, 1763−1767.
6. Xiong, B., Gui, C. S., Xu, X. Y., Luo, C., Chen, J., Luo, H. B., Chen, L. L., Li, G. W., Sun, T., Yu, C. Y., Yue, L. D., Duan, W. H., Shen, J. K., Qin, L., Shi, T. L., Li, Y. X., Chen, K. X., Luo, X. M., Shen, X., Shen, J. H., and Jiang, H. L. (2003) A 3D model of SARS-CoV 3CL proteinase and its inhibitors design by virtual screening, *Acta Pharmacol. Sin. 24*, 497−504.
7. Liu, S., Pei, J., Chen, H., Zhu, X., Liu, Z., Ma, W., He, F., and Lai, L. (2003) Modeling of the SARS coronavirus main proteinase and conformational flexibility of the active site, *Beijing Daxue Xuebao 35* (*Suppl*.), 62−65.
8. Chou, K. C., Wei, D. Q., and Zhong, W. Z. (2003) Binding mechanism of coronavirus main proteinase with ligands and its implication to drug design against SARS, *Biochem. Biophys. Res. Commun. 308*, 148−151.
9. Anand, K., Palm, G. J., Mesters, J. R., Siddell, S. G., Ziebuhr, J., and Hilgenfeld, R. (2002) Structure of coronavirus main proteinase reveals combination of a chymotrypsin fold with an extra alpha-helical domain, *EMBO J. 21*, 3213−3224.
10. Yang, H., Yang, M., Ding, Y., Liu, Y., Lou, Z., Zhou, Z., Sun, L., Mo, L., Ye, S., Pang, H., Gao, G. F., Anand, K., Bartlam, M., Hilgenfeld, R., and Rao, Z. (2003) The crystal structures of severe acute respiratory syndrome virus main protease and its complex with an inhibitor, *Proc. Natl. Acad. Sci. U.S.A. 100*, 13190−13195.
11. Bergmann, E. M., Mosimann, S. C., Chernaia, M. M., Malcolm, B. A., and James, M. N. (1997) The refined crystal structure of the 3C gene product from hepatitis A virus: specific proteinase activity and RNA recognition, *J. Virol. 71*, 2436−2448.
12. Matthews, D. A., Smith, W. W., Ferre, R. A., Condon, B., Budahazi, G., Sisson, W., Villafranca, J. E., Janson, C. A., McElroy, H. E., and Gribskov, C. L., (1994) Structure of human rhinovirus 3C protease reveals a trypsin-like polypeptide fold, RNA-binding site, and means for cleaving precursor polyprotein, *Cell 77*, 761−771.
13. Seipelt, J., Guarne, A., Bergmann, E., James, M., Sommergruber, W., Fita, I., and Skern, T. (1999) The structures of picornaviral proteinases, *Virus Res. 62*, 159−168.
14. Dragovich, P. S., Prins, T. J., Zhou, R., Brown, E. L., Maldonado, F. C., Fuhrman, S. A., Zalman, L. S., Tuntland, T., Lee, C. A., Patick, A. K., Matthews, D. A., Hendrickson, T. F., Kosa, M. B., Liu, B., Batugo, M. R., Gleeson, J. P., Sakata, S. K., Chen, L., Guzman, M. C., Meador, J. W., III, Ferre, R. A., and Worland, S. T. (2002) 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. 45*, 1607−1623.
15. Dragovich, P. S., Prins, T. J., Zhou, R., Johnson, T. O., Brown, E. L., Maldonado, F. C., Fuhrman, S. A., Zalman, L. S., Patick, A. K., Matthews, D. A., Hou, X., Meador, J. W., Ferre, R. A., and Worland, S. T. (2002) 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. 12*, 733−738.
16. Dragovich, P. S., Zhou, R., Webber, S. E., Prins, T. J., Kwok, A. K., Okano, K., Fuhrman, S. A., Zalman, L. S., Maldonado, F. C., Brown, E. L., Meador, J. W., III, Patick, A. K., Ford, C. E., Brothers, M. A., Binford, S. L., Matthews, D. A., Ferre, R. A., and Worland, S. T. (2000) Structure-based design of ketone-containing, tripeptidyl human rhinovirus 3C protease inhibitors, *Bioorg. Med. Chem. Lett. 10*, 45−48.
17. Dragovich, P. S., Webber, S. E., Prins, T. J., Zhou, R., Marakovits, J. T., Tikhe, J. G., Fuhrman, S. A., Patick, A. K., Matthews, D. A., Ford, C. E., Brown, E. L., Binford, S. L., Meador, J. W., III, Ferre, R. A., and Worland, S. T. (1999) Structure-based design of irreversible, tripeptidyl human rhinovirus 3C protease inhibitors containing *N*-methyl amino acids, *Bioorg. Med. Chem. Lett. 9*, 2189−2194.
18. Dragovich, P. S., Prins, T. J., Zhou, R., Webber, S. E., Marakovits, J. T., Fuhrman, S. A., Patick, A. K., Matthews, D. A., Lee, C. A., Ford, C. E., Burke, B. J., Rejto, P. A., Hendrickson, T. F., Tuntland, T., Brown, E. L., Meador, J. W., III, Ferre, R. A., Harr, J. E., Kosa, M. B., and Worland, S. T. (1999) Structure-based design, synthesis, and biological evaluation of irreversible human rhinovirus 3C protease inhibitors, *J. Med. Chem. 42*, 1213−1224.
19. Dragovich, P. S., Prins, T. J., Zhou, R., Fuhrman, S. A., Patick, A. K., Matthews, D. A., Ford, C. E., Meador, J. W., III, Ferre, R. A., and Worland, S. T. (1999) Structure-based design, synthesis, and biological evaluation of irreversible human rhinovirus 3C protease inhibitors. 3. Structure−activity studies of ketomethylene-containing peptidomimetics, *J. Med. Chem. 42*, 1203−1212.
20. Dragovich, P. S., Webber, S. E., Babine, R. E., Fuhrman, S. A., Patick, A. K., Matthews, D. A., Reich, S. H., Marakovits, J. T., Prins, T. J., Zhou, R., Tikhe, J., Littlefield, E. S., Bleckman, T. M., Wallace, M. B., Little, T. L., Ford, C. E., Meador, J. W., III, Ferre, R. A., Brown, E. L., Binford, S. L., DeLisle, D. M., and Worland, S. T. (1998) Structure-based design, synthesis, and biological evaluation of irreversible human rhinovirus 3C protease inhibitors. 2. Peptide structure−activity studies, *J. Med. Chem. 41*, 2819−2834.
21. Johnson, F. A., Lewis, S. D., and Shafer, J. A. (1981) Perturbations in the free energy and enthalpy of ionization of histidine-159 at the active site of papain as determined by fluorescence spectroscopy, *Biochemistry 20*, 52−58.
22. Dardenne, L. E., Werneck, A. S., de Oliveira, N. M., and Bisch, P. M. (2003) Electrostatic properties in the catalytic site of papain: a possible regulatory mechanism for the reactivity of the ion pair, *Proteins 52*, 236−253.
23. Brocklehurst, K., O'Driscoll, M., Kowlessur, D., Phillips, I. R., Templeton, W., Thomas, E. W., Topham, C. M., and Wharton, C. W. (1989) The interplay of electrostatic and binding interactions determining active centre chemistry and catalytic activity in actinidin and papain, *Biochem. J. 257*, 309−310.
24. Asboth, B., Stokum, E., Khan, I. U., and Polgar, L. (1985) Mechanism of action of cysteine proteinases: oxyanion binding site is not essential in the hydrolysis of specific substrate, *Biochemistry 24*, 606−609.
25. Sarkany, Z., Szeltner, Z., and Polgar, L. (2001) Thiolate-imidazolium ion pair is not an obligatory catalytic entity of

cysteine peptidases: the active site of picornain 3C, *Biochemistry 40*, 10601−10606.
26. Connolly, K. M., Smith, B. T., Pilpa, R., Ilangovan, U., Jung, M. E., and Clubb, R. T. (2003) Sortase from *Staphylococcus aureus* does not contain a thiolate-imidazolium ion pair in its active site, *J. Biol. Chem. 278*, 34061−34065.
27. Sarkany, Z., and Polgar, L. (2003) The unusual catalytic triad of poliovirus protease 3C, *Biochemistry 42*, 516−522.
28. Ziebuhr, J., Heusipp, G., and Siddell, S. G. (1997) Biosynthesis, purification, and characterization of the human coronavirus 229E 3C-like proteinase, *J. Virol. 71*, 3992−3997.
29. Fan, K., Wei, P., Feng, Q., Chen, S., Huang, C., Ma, L., Lai, B., Pei, J., Liu, Y., Chen, J., and Lai, L. (2004) Biosynthesis, purification and substrate specificity of SARS coronavirus 3C-like proteinase, *J. Biol. Chem. 279*, 1637−1642.
30. Hopkins, J. L., Betageri, R., Cohen, K. A., Emmanuel, M. J., Joseph, C. R., Bax, P. M., Pallai, P. V., and Skoog, M. T. (1991) Rhinovirus 3C protease catalyzes efficient cleavage of a fluorescein-labeled peptide affording a rapid and robust assay, *J. Biochem. Biophys. Methods 23*, 107−113.
31. Hata, S., Sato, T., Sorimachi, H., Ishiura, S., and Suzuki, K. (2000) A simple purification and fluorescent assay method of the poliovirus 3C protease searching for specific inhibitors, *J. Virol. Methods 84*, 117−126.
32. Wang, Q. M., Johnson, R. B., Cox, G. A., Villarreal, E. C., and Loncharich, R. J. (1997) A continuous colorimetric assay for rhinovirus-14 3C protease using peptide *p*-nitroanilides as substrates, *Anal. Biochem. 252*, 238−245.
33. Hegyi, A., Friebe, A., Gorbalenya, A. E., and Ziebuhr, J. (2002) Mutational analysis of the active centre of coronavirus 3C-like proteases, *J. Gen. Virol. 83*, 581−593.

BI036022Q