Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
The 3C-like proteinase of severe acute respiratory syndrome (SARS) coronavirus has been proposed to be a key target for structural-based drug design against SARS. In order to understand the active form and the substrate specificity of the enzyme, we have cloned, expressed, and purified SARS 3C-like proteinase. Analytic gel filtration shows a mixture of monomer and dimer at a protein concentration of 4 mg/ml and mostly monomer at 0.2 mg/ml, which correspond to the concentration used in the enzyme assays. The linear decrease of the enzymatic-specific activity with the decrease of enzyme concentration revealed that only the dimeric form is active and the dimeric interface could be targeted for structural-based drug design against SARS 3C-like proteinase. By using a high pressure liquid chromatography assay, SARS 3C-like proteinase was shown to cut the 11 peptides covering all of the 11 cleavage sites on the viral polyprotein with different efficiency. The two peptides corresponding to the two self-cleavage sites are the two with highest cleavage efficiency, whereas peptides with non-canonical residues at P2 or P1′ positions react slower. The P2 position of the substrates seems to favor large hydrophobic residues. Secondary structure studies for the peptide substrates revealed that substrates with more β-sheetlike structure tend to react fast. This study provides a basic understanding of the enzyme catalysis and a full substrate specificity spectrum for SARS 3C-like proteinase, which are helpful for structural-based inhibitor design against SARS and other coronaviruses.

The outbreak of a severe atypical pneumonia in early 2003 has caused 8422 cases and 916 related deaths. The World Health Organization has designated the illness as severe acute respiratory syndrome (SARS). A novel form of coronavirus has been identified as the major cause of SARS (1, 2). The genome of SARS coronavirus has been sequenced within a short period of time after confirmation of the virus (3, 4). Currently, 23 genome sequences of different variations of SARS coronavirus have been released at the NCBI web site (www.ncbi.nlm.nih.gov). 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, polyproteins 1a (~450 kDa) and 1ab (~750 kDa), which are conserved both in length and amino acid sequence to other coronavirus replicase proteins. Polyproteins 1a and 1ab are cleaved by the internally encoded 3C-like proteinase to release functional proteins necessary for virus replication. The SARS 3C-like proteinase is fully conserved among all of the released SARS coronavirus genome sequences and is highly homologous with other coronavirus 3C-like proteinase.

Two crystal structures of coronavirus 3C-like proteinase from transmissible gastroenteritis virus (TGEV) (5) and human coronavirus (hCoV) 229E have been solved (6). The structure of coronavirus 3C-like proteinase contains three domains. 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 3C-like proteinase shares the chymotrypsin fold part with the 3C proteinases from other viruses like rhinovirus called picornavirus (7, 8). The 3C proteinase of rhinovirus has been used as a target to develop drugs against the common cold (9–15). Because of the functional importance of SARS 3C-like proteinase in the viral life cycle, it 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 various groups (6, 16, 17), and the conformational flexibility of the substrate-binding site has been studied (17). Virtual screening of chemical compounds libraries has given possible inhibitors (16). An 8-mer peptide has been docked on the model of SARS 3C-like proteinase to study the possible interactions of the protein and the substrate (18).

Similar to other coronaviruses, a sequence analysis revealed 11 cleavage sites of the 3C-like proteinase on the SARS polyprotein. The substrate specificity of coronavirus 3C-like proteinase is determined mainly by the P1, P2, and P1′ positions (19). The P1 position has a well conserved Gln residue, and the P2 position has 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, and Met residues at P2 position. To study the substrate specificity of SARS 3C-like proteinase, we have cloned, expressed, and purified the protein and studied its activity toward 11 peptides covering the 11 cleavage sites on the virus polyprotein. Our results confirm that purified SARS 3C-like proteinase is active toward substrate peptides mapped from
the cleavage sites on the polyprotein and reveals the substrate requirement of the proteinase-binding site. This study helps to understand the mechanism of SARS-CoV polyprotein processing and provides clues for drug design.

**EXPERIMENTAL PROCEDURES**

Cloning of SARS-CoV 3C-like Proteinase—The reverse transcriptional mixture of SARS-CoV RNA from supernatant fluid of the virus-infected Vero cells using random primers was generously supplied by Dr. Y. Lu from Zhejiang Provincial Center for Disease Prevention and Control. For cloning of the cDNA of 3C-like proteinase of the virus, the first strand cDNA mixture was subjected to PCR amplification using a pair of specific primers, comprising 5′-AGT GGT GTT AAG AAA ATG GCA TTC CC-3′ and R108 (5′-TTG GAA GGT AAC AGA GC-3′) to amplify a 917-bp fragment containing full-length 3C-like proteinase coding sequence. The PCR products were purified by agarose gel electrophoresis and then cloned directly into pGEM T Easy vector (Promega, Madison, WI). The resultant sequence confirmed that the amplified fragment was the same as that of SARS-CoV 3C-like proteinase.

**Construction of Plasmid pET3 CLP-21x**—The gene of the SARS 3C-like proteinase was amplified by PCR from the cloning vector (pGEM T Easy) described above using primers CLP-NheI (5′-CAGTCTG-CATAAGGTTTATAGAAATGCTCC-3′) and 3CLP-XhoI (5′-CACTGGAAGTTAAAGTACACGAGG-3′). The PCR product was digested with NheI and XhoI and ligated with NdeI/XhoI-digested pET21a DNA. The resulting plasmid pET 3CLP-21x encodes a 35.1 kDa protein containing a C-terminal His6- tag.

Expression of 3C-like Proteinase—pET 3CLP-21x was transformed into Escherichia coli BL21 (DE3) cells. Cultures were grown at 37 °C in 1 liter of LB medium containing ampicillin (100 μg/ml) until the A600 reached 0.8 and then induced with 0.5 mM isopropyl-1-thio-β-D-galactopyranoside at 30 °C for 3 h. The cells were harvested by centrifugation at 5000 × g for 10 min. The pelleted cells were suspended in buffer A (40 mM Tris-HCl, pH 8.0, 100 mM NaCl, 10 mM imidazole, 7.5 mM 2-mercaptoethanol), at 2% of the original culture volume. After cell lysis by ultrasonic, the cell lysate was separated by centrifugation under 24,000 × g for 20 min. The filtrated supernatant was applied to a nickel-nitriotropic acid column (Qiagen) equilibrated by 50 mM of buffer A. After being washed in 100 mM of buffer A, the 3C-like proteinase was eluted with the gradient of 1–100% buffer B (40 mM Tris-HCl, pH 8.0, 100 mM NaCl, 250 mM imidazole, 7.5 mM 2-mercaptoethanol). The eluted enzyme was concentrated and loaded on a gel filtration column Sephacryl S-200 HR (Amersham Biosciences) equilibrated by 180 mM of buffer C (40 mM Tris-HCl, pH 8.0, 100 mM NaCl, 7.5 mM 2-mercaptoethanol). After elution with another 180 mM buffer C, we received over 95% purified 3C-like proteinase.

**Analytic Gel Filtration**—The aggregation state of the SARS 3C-like proteinase was analyzed using a Superdex 75 HR column (Amersham Biosciences) on AKTA fast protein liquid chromatography. Freshly purified protein was diluted to 4 and 0.2 mg/ml and equilibrated at room temperature. Column Sephacryl S-200 HR (Amersham Biosciences) equilibrated by 50% buffer A. As a result, the majority of 3C-like proteinase can be purified. The concentration of the enzyme varied from 1.07 to 17.1 μM because of the different cleavage activity to different substrates. A cell with a path length of 0.1 mm was used. Each spectrum was the average of four scans corrected by subtracting a spectrum of the (9-fluorenyl)methoxycarbonyl–butyl strategy (21). The cleavage of the peptide from Rink resin and removal of all of the side-chain protecting groups were achieved in trifluoroacetic acid solution. The crude peptide was purified by reversed-phase high performance liquid chromatography (RP-HPLC, LabPrep System, Gilson) on a Vydac C18 semi-preparative column (218TPS10, 10 by 250 mm, Vydac) with gradients of water/acetonitrile containing 0.1% trifluoroacetic acid. Peptide homogeneity and identity were analyzed by analytical HPLC and matrix-assisted laser desorption/ionization-time-of-flight mass spectrometry (MALDI-TOF MS), respectively. Other substrate peptides of HPLC purity from S02 to S11 were purchased from GL Biochemistry Ltd. (Shanghai, China).

**Peptide Cleavage**—The proteolysis activity of the SARS 3C-like proteinase was determined by peptide cleavage assay. Peptide S01 (Table 1) was incubated with the enzyme in Tris-HCl buffer, pH 7.3, at room temperature. The cleavage mixture was analyzed by RP-HPLC. To verify the cleavage site on the substrate peptide, the two products were purified by semi-preparative RP-HPLC using a 15-min 0–50% linear gradient of acetonitrile in 0.1% trifluoroacetic acid and lyophilized. The relative molecular weights of the products were identified by MALDI-TOF MS (BIFLEX III time-of-flight mass spectrometer, Bruker).

To determine the $k_{cat}/K_m$ for the substrate, 0.2 mM of substrate peptide was incubated with SARS 3C-like proteinase in 40 mM Tris-HCl buffer, pH 7.3. The concentration of the enzyme varied from 0.90 to 22.5 μM because of the different cleavage activity to different substrates. Reaction aliquots were removed at different times within 7 h and analyzed by RP-HPLC as described above. $k_{cat}/K_m$ was determined by plotting substrate peak area as Equation 1,

$$\ln P_A = C - \ln K_M + \ln \left( \frac{S_0}{k_{cat}} \right)$$

where $P_A$ is the peak area of the substrate peptide, $C$ is the total concentration of 3C-like proteinase, and $C$ is an experimental constant. $K_M$ and $k_{cat}$ of the proteinase for selected substrates were determined by incubation of the substrate peptide at different concentration varying from 2 to 0.1 mM with SARS 3C-like proteinase in 40 mM Tris-HCl buffer, pH 7.3, for 20 min and analyzed by RP-HPLC as described above. The concentration of the enzyme varied from 1.07 to 17.1 μM because of the different cleavage activity to different substrate. Peak areas were calculated by integration and converted to absolute units by using peptide standards. The reaction rate was calculated for all of the cleavage products of two experiments and averaged. $K_M$ and $k_{cat}$ were calculated by the Lineweaver-Burk plot.

**RESULTS AND DISCUSSION**

**Bio synthesis, Purification, and Secondary Structures of Recombinant SARS CoV 3CL Proteinase**—The C-terminal His tagged SARS 3C-like proteinase has been successfully expressed in E. coli and purified. Induction was first done at 37 °C. As a result, the majority of 3C-like proteinase can be found in the insoluble fraction of the cell lysate. Induction with isopropyl-1-thio-β-D-galactopyranoside was then done at 30 °C for 3 h. As a result, most of the 3C-like proteinase was found in the soluble fraction. The protein was purified by nickel column followed by gel filtration on a Sephacryl S-200 HR column (Fig. 1). Approximately 10 mg of purified protein can be obtained from 1-liter cells. The protein can be concentrated to 10 mg/ml in 50 mM Tris-HCl, 0.1 mM NaCl, and 1 mM dithiothreitol, pH 7.3.

The CD spectra of SARS 3C-like proteinase were shown in Fig. 2. Far-UV and near-UV CD spectra show that the purified protein has well defined secondary and tertiary structures. Far-UV CD spectrum (Fig. 2a) shows a positive peak at 196 nm and two negative peaks at 209 and 222 nm, respectively, which clearly indicates for a mixed α and β structure. Calculated secondary structure content shows 26% of α-helix and 25% of β-sheet, similar to those in the TGEV 3C-like proteinase crystallography.
tal structure (22 and 26%, respectively). Near-UV CD spectrum (Fig. 2b) shows a broad positive peak at ~280 nm and a small positive shoulder at 291 nm, indicating a well folded tertiary structure. Thermal melting of the protein by monitoring the CD signal at 218 nm (Fig. 2c) gives a sigmoid denaturation curve and a \( T_m \) of 61 °C, indicating a highly cooperative thermodenaturation.

Aggregation State and Enzyme Activity—Because 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 biological functional form of the protein (5, 6). 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 1–2 mg/ml (6). We have performed analytical gel filtration of SARS 3C-like proteinase at different concentrations using a Superdex 75 column (Fig. 3). At the concentration of 4 mg/ml, two peaks corresponding to the monomeric and dimeric form of the protein appeared while only one peak corresponding to the monomeric form was found at a lower concentration of 0.2 mg/ml. The dissociation constant of the dimer was estimated to be around 100 \( \mu \)M. In vitro peptide or protein cleavage assays of coronavirus 3C-like proteinases usually were performed with a protein concentration at micromolar level (22, 23). In this study, we also used a peptide cleavage assay with the SARS 3C-like proteinase at different concentrations not >22.5 \( \mu \)M that corresponds mainly to the monomer form of the enzyme. This raises an interesting question whether the minor amount of the dimer plays a major role in catalysis, although we are still not clear about the exact form of the protein in the cell under a molecular crowding environment.

To answer this question, the proteolysis activity of SARS CoV 3C-like proteinase at different enzyme concentrations was studied. Peptide S01 was used as substrate and was incubated with the proteinase in Tris-HCl buffer, pH 7.3, at room temperature. The cleavage mixture was analyzed by RP-HPLC. The peak area of the substrate S01 decreases as reaction time increases, whereas the area of two newly formed peaks increases. The products were collected and identified by MALDI-TOF MS with experimental \( (M + H)^+ \) of 593 and 618, respectively, that were identical to the theoretical \( (M + H)^+ \) of the C-terminal pentapeptide and N-terminal hexapeptide. This confirms that the substrate is cleaved by the SARS CoV 3C-like proteinase at the predicted fragile Gln-Ser peptide bond.

Fig. 1. Expression and purification of the SARS 3C-like proteinase. Samples taken at each step of purification were analyzed on a 12% SDS-polyacrylamide gel, and the protein was stained with Coomassie Brilliant Blue. Lane M, protein molecular mass marker; lanes 1 and 2, cell lysates from non-induced and isopropyl-\( \beta \)-D-galactopuranoside-induced E. coli BL21(DE3) containing the expression plasmid pET 3CLP-21x; lanes 3 and 4, the supernatant and the deposition separated from the cell lysate by centrifugation; lane 5, collected peak fractions from the nickel-nitrilotriacetic acid column; lane 6, collected peak fractions from the Sephacryl S-200 HR column.

Fig. 2. CD spectra of SARS 3C-like proteinase. The CD spectra of 3C-like proteinase were recorded in 40 mM Tris-HCl buffer, pH 8.0 at 20 °C. a, far-UV CD spectrum shows a positive peak at 196 nm and two negative peaks at 209 and 222 nm, respectively, indicating a mixed \( \alpha \) and \( \beta \) structure. b, near-UV CD spectrum shows a broad positive peak at ~280 nm and a small positive shoulder at 291 nm. c, thermal melting of the protein by monitoring the CD signal at 218 nm gives a sigmoid denaturation curve and a \( T_m \) of 61 °C.

The observed \( k_{cat}/K_m \) was determined as described under “Experimental Procedures” at different proteinase concentrations of 4.5 and 0.9 \( \mu \)M (shown in Fig. 4). The observed
1.4

is the active form of the proteinase. whereas the enzyme concentration increases, indicating the dimer form

\[ k \]

of TGEV 3C-like proteinase, residues P5 to P3 of the substrate

the enzyme and proteolysis. As shown in the crystal structure

Substrate Specificity of SARS 3C-like Proteinase

Most of the cleavage sites of coronavirus polyprotein have a conserved (Leu/Ile)-Gln ↓ (Ser, Ala, or Gly) core sequence (the cleavage site is indicated by ↓ ). However, the SARS coronavirus polyprotein has three noncanonical cleavage sites with Phe, Val, or Met in the P2 position and one noncanonical cleavage site with Asn in the P1’ position. Here, peptides S02, S03, and S07 are derived from the three noncanonical cleavage sites at P2 and peptide S05 is derived from the one noncanonical cleavage site at P1’. With the dipeptide sequence of Asn-Asn at age site at P1

Substrate Specificity—Eleven 11-mer peptides derived from the 11 cleavage sites of 3C-like proteinase upon the SARS polyprotein have been synthesized and tested for SARS 3C-like proteinase cleavage experiment (Table I). The proteolysis ac-

ity of expressed SARS CoV 3C-like proteinase at different pH was tested as described under “Experimental Procedures.” When peptide S01 was used as substrate as shown in Fig. 5, the protease had the highest activity at around pH 7 and decreased at another pH, resulting a bell-shaped curve. When pH decreased to 5, the proteolytic activity of the enzyme decreased to 10% because of the protonation of the catalytic His-41.

To investigate the relationship between the proteolysis activity and the secondary structure of the substrates, far-UV CD spectra were recorded for all of the peptides at 2 mM (for S09 the peptide concentration was 0.2 mM because of its low solubility) (see Fig. 6c). The secondary structure contents were calculated using VARSCL1 (20) with the CD data in the range of 184–260 nm for each peptide with the exception of S09 (which has very low solubility) and summarized in Table II. Although not obvious in the CD spectra, each peptide forms more or less sheet structures. The calculated sheet content reflects the formation of a β-strandlike extended conformation in equilibrium with other turnlike or random conformations. Fig. 6b shows an interesting tendency in which substrate peptide with small \( k_{cat}/K_m \) has relatively less content of sheet but more of other unordered structure, suggesting that the formation of β-strandlike extended structure is favored for binding to the enzyme and proteolysis. As shown in the crystal structure of TGEV 3C-like proteinase, residues P5 to P3 of the substrate form an antiparallel β-sheet with segment 164–167 of the long strand eI1 of the enzyme on one side and with segment 186–191 on the other side (6). The binding modes of substrate to different coronavirus 3C-like proteinases were predicted to be identical by comparing the crystal structure of the substrate-binding regions of the free proteinases of hCoV and SARS CoV and of TGEV proteinase in complex with a hexapeptidyl chloromethyl ketone inhibitor (6). Substrate peptide that tends to form an extended β-strandlike conformation would be the preferred choice to bind with the enzyme, resulting in a small \( K_m \) and a large \( k_{cat}/K_m \). It is also supported by a previous study (25) on 3C proteinase in which helix-desaturating residues were often found in close proximity to cleavage sites.

Of all of the 11 peptides tested, S01 and S02 were the most suitable substrates for SARS CoV 3C-like proteinase cleavage, which were derived from the N-terminal and C-terminal self-cleavage sites, respectively. \( K_m \) of S01 was determined as 1.15 ± 0.28 mM, which is three times larger than that of hCoV 229E 3C-like proteinase for a 15-mer substrate, 0.39 ± 0.07 mM. The relative large \( K_m \) was counteracted by a large \( k_{cat} \) 12.2 ± 2.9 min⁻¹, which is an order larger than that of the five other substrates whose \( k_{cat} \) and \( K_m \) values are also determined by the Lineweaver-Burk plot (Table I).

Most of the cleavage sites of coronavirus polyprotein have a conserved (Leu/Ile)-Gln ↓ (Ser, Ala, or Gly) core sequence (the cleavage site is indicated by ↓ ). However, the SARS coronavirus polyprotein has three noncanonical cleavage sites with Phe, Val, or Met in the P2 position and one noncanonical cleavage site with Asn in the P1’ position. Here, peptides S02, S03, and S07 are derived from the three noncanonical cleavage sites at P2 and peptide S05 is derived from the one noncanonical cleavage site at P1’. With the dipeptide sequence of Asn-Asn at age site at P1

Substrate Specificity—Eleven 11-mer peptides derived from the 11 cleavage sites of 3C-like proteinase upon the SARS polyprotein have been synthesized and tested for SARS 3C-like proteinase cleavage experiment (Table I). The proteolysis ac-

\[ k_{cat}/K_m \] increases linearly with the increase of the enzyme concentration. Using the estimated dissociation constant of \( \sim 100 \mu\text{M} \), we can deduce from the linear dependence that the monomeric form of SARS CoV 3C-like proteinase almost has no catalytic activity, whereas the \( k_{cat}/K_m \) of the dimeric form is \( 1.4 \times 10^3 \text{mM}^{-1}\text{min}^{-1} \). This result supports the previous prediction that the dimer of the 3C-like proteinase is the active form of the enzyme and also explains that the low in vitro activity of CoV 3C-like proteinases compared with other monomeric virus 3C proteinases is because of the low concentration of the active dimeric form under the assay conditions.

Substrate Specificity—Eleven 11-mer peptides derived from the 11 cleavage sites of 3C-like proteinase upon the SARS polyprotein have been synthesized and tested for SARS 3C-like proteinase cleavage experiment (Table I). The proteolysis ac-

\[ k_{cat}/K_m \] increases linearly with the increase of the enzyme concentration. Using the estimated dissociation constant of \( \sim 100 \mu\text{M} \), we can deduce from the linear dependence that the monomeric form of SARS CoV 3C-like proteinase almost has no catalytic activity, whereas the \( k_{cat}/K_m \) of the dimeric form is \( 1.4 \times 10^3 \text{mM}^{-1}\text{min}^{-1} \). This result supports the previous prediction that the dimer of the 3C-like proteinase is the active form of the enzyme and also explains that the low in vitro activity of CoV 3C-like proteinases compared with other monomeric virus 3C proteinases is because of the low concentration of the active dimeric form under the assay conditions.

Substrate Specificity—Eleven 11-mer peptides derived from the 11 cleavage sites of 3C-like proteinase upon the SARS polyprotein have been synthesized and tested for SARS 3C-like proteinase cleavage experiment (Table I). The proteolysis ac-

\[ k_{cat}/K_m \] increases linearly with the increase of the enzyme concentration. Using the estimated dissociation constant of \( \sim 100 \mu\text{M} \), we can deduce from the linear dependence that the monomeric form of SARS CoV 3C-like proteinase almost has no catalytic activity, whereas the \( k_{cat}/K_m \) of the dimeric form is \( 1.4 \times 10^3 \text{mM}^{-1}\text{min}^{-1} \). This result supports the previous prediction that the dimer of the 3C-like proteinase is the active form of the enzyme and also explains that the low in vitro activity of CoV 3C-like proteinases compared with other monomeric virus 3C proteinases is because of the low concentration of the active dimeric form under the assay conditions.

Substrate Specificity—Eleven 11-mer peptides derived from the 11 cleavage sites of 3C-like proteinase upon the SARS polyprotein have been synthesized and tested for SARS 3C-like proteinase cleavage experiment (Table I). The proteolysis ac-
substitution by a relative small hydrophobic amino acid Val decreased the cleavage activity, indicating an important role of pocket S2 in anti-SARS drug design.

On the other hand, peptide S07, which has a methionine in the P2 position, is the substrate with least cleavage activity among 11 substrates. The low cleavage activity of S07 may be caused by the mutant from Leu to Met, although it may also be affected by other reasons such as different secondary structures. As shown in the far-UV CD spectrum, peptide S07 may adopt an unusual helix-like, type II polyproline conformation, which has a strong negative peak at 200 nm and a positive peak at 230 nm. The formation of the polyproline helix will disturb the conformation transition to -strandlike-extended structure of the substrate peptide and disable the binding of the substrate to the enzyme. It is also supported by the factor that this peptide has the least content of sheet.

Hegyi and Ziebuhr (19) have studied the conservation of substrate specificities among coronavirus 3C-like proteinase from three coronavirus by using peptides corresponding to the four cleavage sites on the viral polyprotein. Our findings that the two peptides S01 (P1/P2) and S02 (P2/P3) corresponding to the two self-cleavage sites of the SARS 3C-like proteinase are the two most reactive ones and S05 (P5/P6) is less reactive are comparable to their results. This implies that SARS 3C-like proteinase follows the same substrate specificity rules governing all of the coronavirus 3C-like proteinase. As we have studied, the substrate specificities of SARS 3C-like proteinase using peptides covering all of the 11 cleavage sites try to correlate not only the conserved sequences but also secondary structures of

### TABLE I

| Peptide | Cleavage site | Sequence                   | $K_m$ (mM) | $k_{cat}$ (min$^{-1}$) | $k_{cat}/K_m$ (min$^{-1}$ mM$^{-1}$) | ($k_{cat}/K_m$)$_{rel}$ |
|---------|---------------|----------------------------|------------|------------------------|-------------------------------------|-------------------------|
| S01     | P1/P2         | TSAVLQ/SQFRK-NH$_2$         | 1.15 ± 0.28 | 12.2 ± 2.9             | 10.6 ± 0.6$^a$                      | 1.00                    |
| S02     | P2/P3         | SQVTQD/KEFKK                | 0.583 ± 0.086 | 2.55 ± 0.35             | 4.38 ± 0.26$^b$                      | 0.41                    |
| S03     | P3/P4         | KVATQV/SKMD                 | 0.353 ± 0.013$^b$ | 3.29 ± 1.07             | 0.556 ± 0.126$^b$                     | 0.05                    |
| S04     | P4/P5         | NRTAQI/ASID                 | 0.202 ± 0.003$^b$ | 1.68 ± 0.57             | 0.865 ± 0.028$^b$                     | 0.08                    |
| S05     | P5/P6         | SVKLOQ/NNELS                | 0.1076 ± 0.0022$^b$ | 0.976 ± 0.014$^b$         | 0.57 0.865                 | 0.02                    |
| S06     | P6/P7         | ATVRLQ/AGNAT                | 0.052 ± 0.002              | 1.57 ± 0.28             | 0.26 ± 0.21$^a$                      | 0.27                    |

$^a$ Determined by Lineweaver-Burk plot.
$^b$ Determined by fitting reaction kinetic data as in Equation 1 at low substrate concentration.
the substrate peptides. The full spectrum of substrate specificity for SARS 3C-like proteinase studied here can be extended to understand other coronavirus 3C-like proteinases.

In summary, the recombinant SARS 3C-like proteinase has been successfully cloned and expressed. The purified protein exists in a mixture of monomer and dimer at a concentration of 4 mg/mL but mostly monomer at 0.2 mg/mL. The specific activity of the enzyme decreases linearly with the decrease of enzyme concentration, implying that the active dimeric form is active and that the dimeric interface could be targeted for structural based drug design against SARS 3C-like proteinase. The enzyme can cut the 11 peptides covering all of the 11 cleavage sites on the viral polyprotein with different efficiency. Substrate specificities for SARS 3C-like proteinase studied here can be extended to other coronavirus 3C-like proteinases.

REFERENCES

1. Drosten, C., Gunther, S., Preiser, W., van der, W. S., Brodt, H. A., Becker, S., Rabanau, H., Panning, M., Kolenkova, L., Foucher, R. A., Berger, A., Burguiere, A. M., Cinatl, J., Rickmann, M., Ersou, N., Grywna, K., Kramme, S., Klenk, H. D., Osterhaus, A. D., Schmitz, H., and Doerr, H. W. (2003) Science 300, 1394–1399

2. Anand, K., Palm, G. J., Mesters, J. R., Siddell, S. G., Ziebuhr, J., and Hilenfeld, R. (2002) EMBO J. 21, 3213–3224

3. Anand, K., Ziebuhr, J., Wadhawan, P., Mesters, J. R., and Hilenfeld, R. (2003) Science 300, 1763–1767

4. Matthews, D. A., Smith, W. F., Ferre, R. A., Condon, B., Budahazi, G., Sisson, W., Vaissafra, J. E., Janson, C. A., McElroy, H. E., Gribskov, C. L., and Worland, S. (1994) Cell 77, 761–771

5. Schoppet, J., Guarnie, A., Bergmann, E., James, M., Sommergruber, W., Fita, I., and Skern, T. (1999) Virus Res. 62, 159–168

6. 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., Battugo, M. R., Gleeson, J. P., Sakata, S. K., Chen, L., Guzman, M. C., Meador, J. W., III, Ferre, R. A., and Worland, S. T. (2001) J. Med. Chem. 45, 1607–1623

7. 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., Hox, S., Meador, J. W., Ferre, R. A., and Worland, S. T. (2002) Bioorg. Med. Chem. Lett. 12, 733–738

8. Dragovich, P. S., Zhou, R., Webber, S. E., Prins, T. J., Kowk, 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) Bioorg. Med. Chem. Lett. 10, 45–48

9. Dragovich, P. S., Prins, T. J., Zhou, R., Marakovsky, J. T., Taktike, J., Fuhrman, S. A., Zalman, L. S., Patick, A. K., Matthews, D. A., Hendrickson, T. F., Kosa, M. B., Liu, B., Battugo, M. R., Gleeson, J. P., Sakata, S. K., Chen, L., Guzman, M. C., Meador, J. W., III, Ferre, R. A., and Worland, S. T. (1999) Bioorg. Med. Chem. Lett. 9, 2189–2194

10. Dragovich, P. S., Prins, T. J., Zhou, R., Marakovsky, J. T., Taktike, J., Fuhrman, S. A., Zalman, L. S., Patick, A. K., Matthews, D. A., Lee, C. A., Ford, C. E., Burke, B. J., Kejo, 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) J. Med. Chem. 42, 1203–1212

11. 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) J. Med. Chem. 42, 1212–1222

12. Dragovich, P. S., Prins, T. J., Zhou, R., Marakovsky, J. T., Taktike, J., Fuhrman, S. A., Zalman, L. S., Patick, A. K., Matthews, D. A., Lee, C. A., Ford, C. E., Burke, B. J., Kejo, 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) J. Med. Chem. 42, 1203–1212

13. Dragovich, P. S., Prins, T. J., Zhou, R., Marakovsky, J. T., Taktike, J., Fuhrman, S. A., Zalman, L. S., Patick, A. K., Matthews, D. A., Lee, C. A., Ford, C. E., Burke, B. J., Kato, 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) J. Med. Chem. 42, 1203–1212

14. 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) J. Med. Chem. 42, 1212–1222

15. Dragovich, P. S., Prins, T. J., Zhou, R., Fuhrman, S. A., Patick, A. K., Matthews, D. A., Reich, S. H., Marakovsky, J. T., Prins, T. J., Zhou, R., Taktike, 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., Battugo, M. R., DeLisle, D. M., and Worland, S. T. (1998) J. Med. Chem. 41, 2819–2834

16. Sun, X., Cui, S., Xue, Y., Luo, H., Cui, J., Luo, H. B., Chen, L. L., Li, G. W., Sun, T., Yu, C. Y., Yue, L. D., Duan, W. H., Shen, X. J., Qin, L., Shi, T. L., Li, X. Y., Chen, X. K., Luo, X. M., Shen, X., Shen, J. H., and Jiang, H. L. (2003) Acta Pharmacol. Sin. 24, 497–504

17. Liu, S., Pei, J., Chen, H., Zhu, X., Liu, Z., Ma, W., He, F., and Lai, L. (2003) Bioorg. Med. Chem. Lett. 13, 65–70

18. Chou, K. C., Wei, D. Q., and Zhong, W. Z. (2003) Biochem. Biophys. Res. Commun. 306, 148–151

19. Hegyi, A., and Ziebuhr, J. (2002) J. Gen. Virol. 83, 595–599

20. Johnson, W. C., J. Junior. (1988) Annu. Rev. Biophys. Biophys. Chem. 17, 145–166

21. Fields, G. B., and Noble, R. L. (1990) Int. J. Pept. Protein Res. 35, 161–214

22. Ziebuhr, J., and Siddell, S. G. (1995) J. Virol. 69, 4331–4338

23. Ziebuhr, J., Heusipp, G., and Siddell, S. G. (1997) J. Virol. 71, 3992–3997

24. Deleted in proof

25. Cordingly, M. G., Callahan, P. L., Sardana, V. V., Garsky, V. M., and Colonna, R. J. (1990) J. Biol. Chem. 265, 9062–9065