

## A novel auto-cleavage assay for studying mutational effects on the active site of severe acute respiratory syndrome coronavirus 3C-like protease<sup>☆</sup>

Yu-Fei Shan<sup>a</sup>, Shou-Feng Li<sup>a</sup>, Gen-Jun Xu<sup>a,b,\*</sup>

<sup>a</sup> Institute of Biochemistry and Cell biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China

<sup>b</sup> College of Life Science, Zhejiang University of Sciences, Hangzhou 310018, China

Received 19 August 2004

Available online 2 October 2004

### Abstract

The 3C-like protease (3CL<sup>PRO</sup>) of severe acute respiratory syndrome (SARS) has been proposed as an attractive target for drug design. His<sup>41</sup> and Cys<sup>145</sup> were essential for the active site as the principal catalytic residues. In this study, we mutated the two sites, expressed four resulting mutants in *Escherichia coli* and characterized. All mutants showed undetectable activity in *trans*-cleavage assay. In addition, we introduced a 31-mer peptide containing an auto-cleavage site to the N-terminal of the proteases and found the peptide could be cleaved efficiently by 3CLsc itself, but, among the four mutants, only the mutant Cys<sup>145</sup> → Ser showed residual activity as detected by the auto-cleavage assay. The data supported the proposition unequivocally that SARS-CoV 3CL<sup>PRO</sup> was a member of serine proteases involving His<sup>41</sup> and Cys<sup>145</sup> residues at the active site. The auto-cleavage assay also provided a sensitive and reliable compensation to the traditional *trans*-cleavage assay.

© 2004 Elsevier Inc. All rights reserved.

**Keywords:** SARS; 3CL protease; Active site; Auto-cleavage assay; N<sub>31</sub>-tag

Severe acute respiratory syndrome (SARS) broke out in southern China in 2003 and spread rapidly. The pathogen of SARS was identified as a novel coronavirus (CoV), which belongs to *Coronaviridae* family [1,2]. Coronavirus is a kind of positive-stranded RNA virus characterized with an exceptionally large genome size (up to 31 kb) known to date. SARS coronavirus (SARS-CoV) replicase gene encodes two overlapping large proteins, polyprotein 1a (pp1a, ~450 kDa) and polyprotein 1ab (pp1ab, ~750 kDa), which mediate the functions needed for transcription and replication of the virus. pp1a and pp1ab are cleaved extensively

by the specific proteases to release functional proteins and peptides. One of the proteases encoded by the virus cleaves at least 11 conservative sites on pp1ab and plays a main role in proteolytic processing [3]. It is named main protease (M<sup>PRO</sup>) and also called 3C-like protease (3CL<sup>PRO</sup>). Ascribing to its pivotal role in virus life circle, SARS-CoV 3CL<sup>PRO</sup> has been suggested as a promising target for antiviral drug design [3].

The structure of the 33.8 kDa protease of SARS-CoV has three domains. And the active site is predicted to involve a catalytic dyad of His and Cys, which is similar to those of main proteases in other coronaviruses. There is a controversy about the coronavirus main proteases whether they all belong to the class of serine proteases because they have an “additional” domain at C-terminal. In previous studies, replacements of Cys with Ser in some of CoV main proteases produced inactive

<sup>☆</sup> This work was supported by a grant from the National Natural Science Foundation of China (No 39930060).

\* Corresponding author. Fax: +86 21 54921257.

E-mail address: [gjxu@sibs.ac.cn](mailto:gjxu@sibs.ac.cn) (G.-J. Xu).

enzymes as detected by *trans*-cleavage assay [4,5]. There were a few reports that Cys-to-Ser mutants showed residual activities either in a *cis*-cleavage assay [6] or in a *trans*-cleavage assay with an extremely high enzyme concentration [7]. Because no enzyme was detected simultaneously in both *trans*- and *cis*-cleavage assays, it is difficult to draw an unequivocal answer to the discrepancy about the role of Ser.

In the present study, we mutated the two residues at active site and detected the activities of mutants in both *trans*- and a newly developed *cis*-cleavage assays. The results showed that Cys<sup>145</sup> → Ser mutant had a residual activity only in the *cis*-cleavage assay. It suggested that our assay exploiting the auto-processing capability of 3CL<sup>Pro</sup> has advantages over the traditional peptide-based activity detection system, and can be employed in the substrate specificity study or other relevant researches.

## Materials and methods

**Construction of pET28a-3CLsm and pET28a-3CLsc.** The gene of SARS 3C-like protease was a gift from Dr. Li and was amplified by PCR using primers 3CLsm and 3CLr (see Table 1 for the sequences of primers). The PCR product was digested with *Nco*I and *Xho*I, and inserted into *Nco*I/*Xho*I cut plasmid pET-28a. The plasmid, pET28a-3CLsm, encodes a 34.7 kDa (3CLsm) protein flanked at a C-terminal His<sub>6</sub>-tag and with the first amino acid Ser replaced by Met. The PCR product of a mutant using primers 3CLsc and 3CLr was digested with *Nhe*I and *Xho*I, and inserted into *Nhe*I/*Xho*I cut plasmid pET-28a. The resulting plasmid, pET28a-3CLsc, encodes a protein (3CLsc) containing a 31-mer peptide (MGSSHHHHHSSGLVPRGSHMAS SITS AVLQ; N<sub>31</sub>-tag) flanked at N-terminal and a C-terminal His<sub>6</sub>-tag. An auto-cleavage site was located between N<sub>31</sub>-tag and the N-terminal of the protein. This 31-mer sequence was also introduced to four mutants using the identical method as pET28a-3CLsc.

**Expression and purification.** pET28a-3CLsm and pET28a-3CLsc were transformed, respectively, into *Escherichia coli* BL21(DE3) cells. The transformed cells were grown in 1 liter LB medium containing kanamycin (50 µg/ml) at 22 °C until A<sub>600</sub> reached to 0.6–0.8 and then

induced with 10 µM isopropyl-1-thio-β-D-galactopyranoside for 6 h. The cells were harvested by centrifugation (5000g for 10 min) and suspended into 40 ml buffer A (20 mM Tris-HCl, pH 8.0, 200 mM NaCl, and 2 mM of 2-mercaptoethanol). Lysis of the cells was achieved by ultrasonic and the suspension was centrifuged at 25,000g for 20 min. Supernatant was applied to a buffer A equilibrated cobalt-nitrilotriacetic acid column (4 ml beads). The column was washed with 100 ml buffer A and then with buffer B (20 mM Tris-HCl, pH 8.0, 200 mM NaCl, and 10 mM imidazole) until A<sub>280</sub> was less than 0.02. The target protein was eluted with 20 ml buffer C (20 mM Tris-HCl, pH 8.0, 200 mM NaCl, and 2 mM of 2-mercaptoethanol) containing 200 mM imidazole. The recovered protein was concentrated and loaded on a gel filtration column (Sephacryl S-200HR), which was equilibrated with 250 ml buffer C. The eluted enzyme was concentrated and stored at –20 °C in buffer C containing 50% glycerol. All mutants were performed in a similar way.

**Peptide synthesis and trans-cleavage assay.** A 16-mer peptide, P439 (TSITSAVLQ ↓ SGFRKMA; the arrow indicates cleavage site), was synthesized by solid-phase method. The sequence was taken from the processing site of 3CL<sup>Pro</sup>. The synthetic peptide was purified through a C-18 reverse-phase HPLC column and eluted with acetonitrile/water (linear gradient 0–80% containing 0.1% trifluoroacetic acid). Identity and homogeneity were confirmed by mass spectrometry (matrix-assisted laser desorption/ionization time-of-flight, MALDI-TOF). An 11-mer peptide, P440 (TSAVLP ↓ SGFRK), was performed in a similar way.

The cleavage assay was carried out in 40 µl buffer D (20 mM Tris-HCl, pH 7.35, containing 400 µM P439, 200 mM NaCl, 1 mM DTT, and 1 mM EDTA) and the reaction was initiated by adding 2 µg enzyme. The reaction mixtures were incubated at 20 °C for 0–16 h and then 40 µl of 2% trifluoroacetic acid was added to stop the reaction. The products were characterized by reverse-phase HPLC using a 5–65% linear gradient of acetonitrile/water containing 0.1% trifluoroacetic acid. The peaks were estimated by integrating the area and identified by mass spectroscopy.

**Site-directed mutagenesis.** Site-directed mutagenesis was performed by a recombination PCR method. The nucleotide sequences of primers are given in Table 1. Mutants were also constructed into pET-28a plasmid and expressed in BL21 (DE3) cells. The homogeneities of the purified mutant proteins were characterized by SDS-PAGE.

**Auto-cleavage assay in vitro.** Two milligrams of purified SARS-CoV 3CL protease of C<sup>145</sup>S mutant was incubated in 1 ml buffer D at 20 °C for 0–36 h. Aliquot of 50 µl was taken out every 9 h, mixed with 50 µl of 2× SDS-loading buffer, and analyzed by SDS-PAGE on 12% polyacrylamide gel.

Table 1  
Primers used for the amplification or mutagenesis of SARS-CoV 3CL<sup>Pro</sup>

Primer	Oligonucleotide sequence (5' → 3')	Polarity	Remark
3CLsc	ATGGCTAGCTCAATCACTTCTGCTGTTCTG CAGAGTGGTTTTAGGAAAATGGCA	Forward	Introduce N <sub>31</sub> -tag
3CLsm	TATACCATGGGTTTTAGGAAAATGGCATTC	Forward	Amplification
3CLr	GGTGCTCGAGTTGGAAGGTAACACCAGA GCATTGTC	Reverse	Introduce His-tag
PH1	TACTGTCCTAGGGCTGTCATTTGCACAGC	Forward	His <sup>41</sup> to Ala
PH2	AATGACAGCCTAGGACAGTATACTGTGTC	Reverse	
PH3	TACTGTCCTAGGTGGGTCATTTGCACAGC	Forward	His <sup>41</sup> to Trp
PH4	AATGACCCACCTAGGACAGTATACTGTGTC	Reverse	
PC1	AATGGATCCGCTGGTAGTGTGGTTTTA	Forward	Cys <sup>145</sup> to Ala
PC2	ACTACCAGCGGATCCATTAAGGAAAG	Reverse	
PC3	AATGGATCCAGTGGTAGTGTGGTTTTA	Forward	Cys <sup>145</sup> to Ser
PC4	ACTACCAGTGGATCCATTAAGGAAAG	Reverse	

The oligonucleotide sequences indicating mutant codons are underlined.

## Results and discussion

### Expression, purification, and proteolytic activity of SARS-CoV 3CL<sup>pro</sup>

3CLsm was successfully over-expressed in *E. coli*. Most of the target protein was found in the soluble fraction of the cell lysate. About 20 mg of 3CLsm protein was recovered from 1 liter LB media. When 3CLsc was expressed in *E. coli*, the product had a similar apparent molecular mass to that of 3CLsm (see Fig. 1). It indicated that an auto-cleavage had occurred during 3CLsc expression.

The activity of the recombinant 3CLsm was analyzed in peptide-based *trans*-cleavage assay in which 3CLsm (2  $\mu$ g) was incubated with 400  $\mu$ M peptide P439 in 40  $\mu$ l buffer D. Analysis of the reaction products by reverse-phase HPLC and mass spectrum revealed that P439 could be hydrolyzed. As shown in Fig. 2A, more than half of the substrate was digested in 20 min.

### Mutation analysis of the catalytic site of SARS-CoV 3CL<sup>pro</sup>

It has been reported that His<sup>41</sup> and Cys<sup>145</sup> might play a key role in the catalytic process for coronavirus main proteases [3]. We mutated these two residues of the enzyme and obtained four mutants, H<sup>41</sup>A, H<sup>41</sup>W, C<sup>145</sup>A, and C<sup>145</sup>S. The mutants were characterized by SDS-PAGE and the proteolytic activities were measured in *trans*-cleavage assay. But, different from 3CLsm, the activities of H<sup>41</sup>A, H<sup>41</sup>W, C<sup>145</sup>A, and C<sup>145</sup>S were all under the detection limitation in *trans*-cleavage assay,

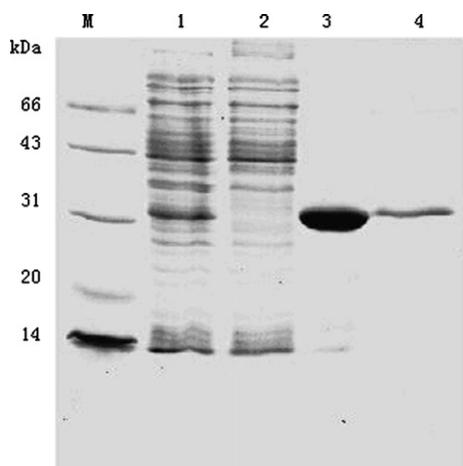


Fig. 1. Bacterial expression and purification of recombinant SARS-CoV 3CL<sup>pro</sup>. Samples were analyzed by SDS-PAGE on 12% polyacrylamide gel and stained with Coomassie brilliant blue; M, protein molecular mass markers (kDa); 1, total cell lysate from IPTG-induced *E. coli* BL21 (pET28a-3CLsm); 2, cell lysate pass through Cobalt-affinity column (pET28a-3CLsm); 3, target 3CL<sup>pro</sup> eluted from Sephacryl S-200HR column (3CLsm); and 4, pooled peak fractions from the Sephacryl S-200HR column (3CLsc).

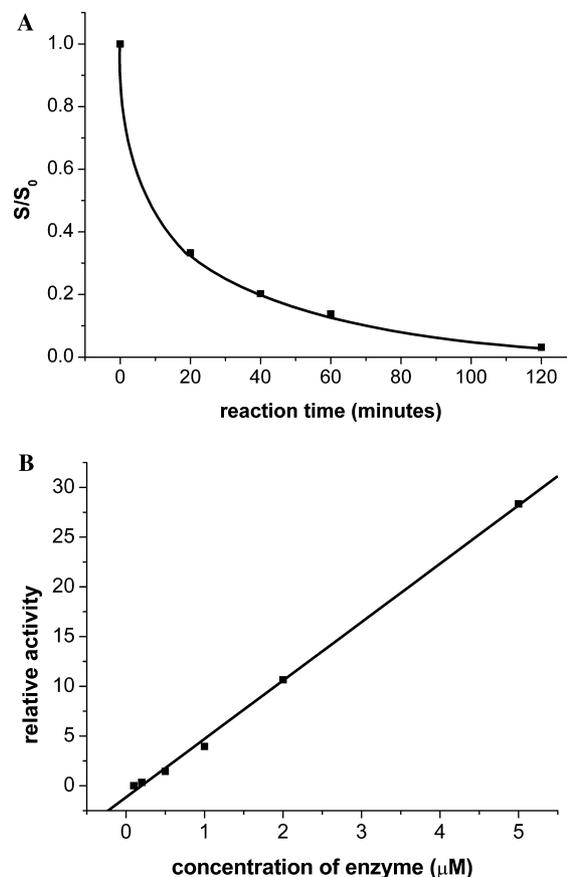


Fig. 2. Peptide P439 cleaved by 3CLsm. Shown in (A) is the time course of 400  $\mu$ M P439 cleaved by 2  $\mu$ g 3CLsm in 40  $\mu$ l cleavage buffer. Shown in (B) are the cleavage activities of recombinant SARS-CoV 3CL<sup>pro</sup> at different final enzyme concentrations implying strongly dependence of specific activity with enzyme concentration.

even, when the incubation time was prolonged to 16 h. As shown in Fig. 2B, the activity of 3CL<sup>pro</sup> has a great dependence on enzyme concentration because the active form of enzyme was suggested to be a dimer. We had increased the concentration of the mutants up to 15  $\mu$ M, but still no activity was detected. Even higher enzyme concentration could not be used because it will cause aggregation and precipitation at room temperature within a few hours. The great difference in *trans*-cleavage assay between 3CLsm and the mutant enzymes strongly supports a crucial catalytic function of His<sup>41</sup> and Cys<sup>145</sup>, and is consistent with the published data of other 3CL/3C proteases [8].

Tibbles et al. [6] reported that a mutant with the substitution of the active site Cys with Ser in avian infectious bronchitis virus 3CL<sup>pro</sup> produced an enzyme with a residual activity in *cis*-cleavage assay. In contrast with their finding, although all 3CL proteases have conservative catalytic site characteristics, no activity could be detected for the corresponding Cys-to-Ser mutants of the mouse hepatitis virus and human coronavirus strain 229E 3CL<sup>pro</sup> in *trans*-cleavage assay [4,5]. One

explanation for this discrepancy is that *cis*-cleavage assay has high sensitivity over the *trans*-cleavage one [8]. In our study, Cys-to-Ser substitution also produced an enzyme with undetectable activity in *trans*-cleavage assay. So we decided to test the cleavage activity of this mutant in the novel *cis*-cleavage assay.

#### Auto-cleavage assay of SARS-CoV 3CL<sup>pro</sup> mutants

As mentioned above, the 31-mer peptide (N<sub>31</sub>-tag) introduced to the N-terminal of 3CLsc could be cleaved efficiently, so it might be applied in *cis*-cleavage assay. Therefore, the N<sub>31</sub>-tag was introduced to four mutants. After expression, three of the mutants, H<sup>41</sup>A, H<sup>41</sup>W, and C<sup>145</sup>A, yielded a higher apparent molecular mass than that of 3CLsc (see Fig. 3). It suggested that the three mutants did not show enzyme activity. The results could serve as a negative control; it excluded the possibility that proteolytic process was caused by the expression system. The cleavage should be performed by SARS-CoV 3CL<sup>pro</sup> itself. Interestingly, for C<sup>145</sup>S mutant, there were two bands visualized in SDS-PAGE (see Fig. 3, lanes 4 and 6). The apparent molecular mass showed that the upper band retained the N<sub>31</sub>-tag and the lower one did not. It suggested that substitution Cys<sup>145</sup> with Ser in SARS-CoV 3CL<sup>pro</sup> produced an enzyme with a very low activity, despite the fact that the mutant did not show activity in *trans*-cleavage assay.

To verify the fact, we purified C<sup>145</sup>S (with the N<sub>31</sub>-tag) and incubated 2 mg of the proteins in 1 ml buffer D at 20 °C for 0–36 h. The reaction mixture was analyzed by SDS-PAGE on a 12% polyacrylamide gel. As shown in Fig. 4, the band with N<sub>31</sub>-tag disappeared gradually, but the reaction was not completed even after

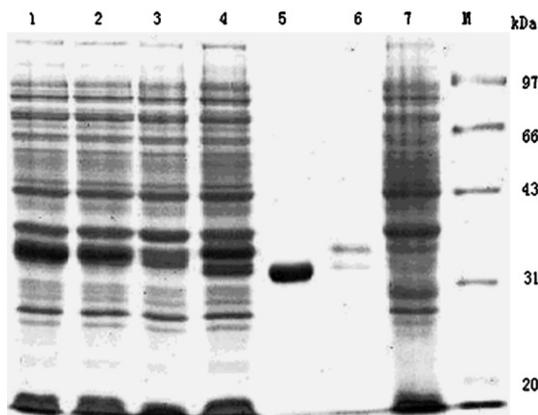


Fig. 3. Expression of SARS-CoV 3CL<sup>pro</sup> mutants containing N<sub>31</sub>-tag in *E. coli* BL21 (DE3). Total cell lysates of IPTG-induced *E. coli* samples were analyzed by SDS-PAGE on 12% polyacrylamide gel and stained with Coomassie brilliant blue; 1, lysate of mutant H<sup>41</sup>A; 2, lysate of mutant H<sup>41</sup>W; 3, lysate of mutant C<sup>145</sup>A; 4, lysate of mutant C<sup>145</sup>S; 5, purified 3CL<sup>pro</sup> (3CLsc) as a control; 6, purified 3CL<sup>pro</sup> (C<sup>145</sup>S); 7, total cell lysate of *E. coli* (3CLsc) without being induced; and M, protein molecular mass markers (kDa).

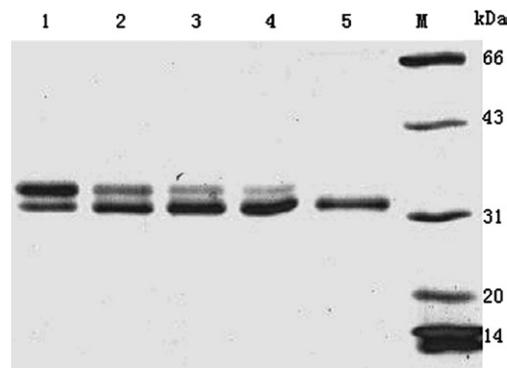


Fig. 4. Time course of SARS-CoV 3CL<sup>pro</sup> C<sup>145</sup>S in auto-cleavage assay. The purified 3CL<sup>pro</sup> C<sup>145</sup>S was incubated in buffer D for 0–36 h at 20 °C as described in Materials and methods, the reaction products were analyzed by SDS-PAGE on 12% polyacrylamide gel and stained with Coomassie brilliant blue. 1, reacted for 0 h; 2, reacted for 9 h; 3, reacted for 18 h; 4, reacted for 36 h; 5, 3CL-sc protease; and M, protein molecular mass markers (kDa).

36 h incubation. Under the same conditions, 6 μg purified C<sup>145</sup>S protein was incubated with 400 μM peptide P439 in 40 μl buffer D at 20 °C for *trans*-cleavage peptide assay, still no protease activity could be detected.

The cleavage sites of 3CL<sup>pro</sup> commonly contain a dipeptide (LQ) sequence, and the amino acid Gln is reported as an “absolutely” conservative residue [9]. Fan et al. [10] reported a peptide (TSAVLQ↓SGFRK) could be cleaved by SARS-CoV 3CL protease. We synthesized the peptide P440 (TSAVLP↓SGFRK) in which the Gln was replaced by Pro and performed *trans*-cleavage assay at 10 μM 3CLsm. No enzyme activity was detectable within 24 h. We also introduced LP dipeptide to the N<sub>31</sub>-tag at N-terminal and tested auto-cleavage assay. It had been found that, even though at a slow rate, the corresponding cleavage site could be processed (data not shown). It implied that the novel auto-cleavage assay is more sensitive.

In the study of Ziebuhr et al. [5], the final concentration of enzyme was 1 μM which is the same in the study of Seybert et al. [4]. Comparing with our data, those two 3CL proteases (C<sup>145</sup>S) probably also have residual activities in a *cis*-cleavage assay. Huang et al. [7] reported the very low cleavage activity of Cys-to-Ser mutant of 3CL protease in a peptide-based assay by using a synthetic *p*-nitroanilide compound (colorimetric substrate) and three synthetic peptides (S10, S06, and S09), in which the activity for substrate S09 was too low to be determined. It should be noted that the enzyme concentration used in their assay was extremely high (76.9 μM; about 2.6 mg/ml). Under such a high concentration of enzyme, unexpected non-specific side reactions could not be completely excluded. In addition, in case of the synthetic colorimetric substrate, it might not involve the specificities contributed by downstream subsites. And Shi et al. [11] reported the amino acids at C-terminal to the cleavage

site also have extensive interactions with the 3CL protease. For these two reasons, the explicit conclusion could not be drawn unequivocally.

The novel assay is simple, sensitive, and reliable, and might be extended to other relevant protease studies. In addition, the results also support that the SARS-CoV 3CL<sup>pro</sup> belongs to serine protease of the chymotrypsin family and clarified the discrepancy of previous researches.

## References

- [1] K.V. Holmes, L. Enjuanes, The SARS coronavirus: a postgenomic era, *Science* 300 (2003) 1377–1378.
- [2] P.A. Rota, M.S. Oberste, S.S. Monroe, W.A. Nix, R. Campagnoli, J.P. Icenogle, S. Penaranda, B. Bankamp, K. Maher, M.H. Chen, S. Tong, A. Tamin, L. Lowe, M. Frace, J.L. DeRisi, Q. Chen, D. Wang, D.D. Erdman, T.C. Peret, C. Burns, T.G. Ksiazek, P.E. Rollin, A. Sanchez, S. Liffick, B. Holloway, J. Limor, K. McCaustland, M. Olsen-Rasmussen, R. Fouchier, S. Gunther, A.D. Osterhaus, C. Drosten, M.A. Pallansch, L.J. Anderson, W.J. Bellini, Characterization of a novel coronavirus associated with severe acute respiratory syndrome, *Science* 300 (2003) 1394–1399.
- [3] K. Anand, J. Ziebuhr, P. Wadhvani, J.R. Mesters, R. Hilgenfeld, Coronavirus main proteinase (3CL<sup>pro</sup>) structure: basis for design of anti-SARS drugs, *Science* 300 (2003) 1763–1767.
- [4] A. Seybert, J. Ziebuhr, S.G. Siddell, Expression and characterization of a recombinant murine coronavirus 3C-like proteinase, *J. Gen. Virol.* 78 (1997) 71–75.
- [5] J. Ziebuhr, G. Heusipp, S.G. Siddell, Biosynthesis, purification, and characterization of the human coronavirus 229E 3C-like proteinase, *J. Virol.* 71 (1997) 3992–3997.
- [6] K.W. Tibbles, I. Brierley, D. Cavanagh, T.D. Brown, Characterization in vitro of an autocatalytic processing activity associated with the predicted 3C-like proteinase domain of the coronavirus avian infectious bronchitis virus, *J. Virol.* 70 (1996) 1923–1930.
- [7] C. Huang, P. Wei, K. Fan, Y. Liu, L. Lai, 3C-like Proteinase from SARS coronavirus catalyzes substrate hydrolysis by a general base mechanism, *Biochemistry* 43 (2004) 4568–4574.
- [8] J. Ziebuhr, E.J. Snijder, A.E. Gorbalenya, Virus-encoded proteinases and proteolytic processing in the Nidovirales, *J. Gen. Virol.* 81 (2000) 853–879.
- [9] H. Yang, M. Yang, Y. Ding, Y. Liu, Z. Lou, Z. Zhou, L. Sun, L. Mo, S. Ye, H. Pang, G.F. Gao, K. Anand, M. Bartlam, R. Hilgenfeld, Z. Rao, The crystal structures of severe acute respiratory syndrome virus main protease and its complex with an inhibitor, *Proc. Natl. Acad. Sci. USA* 100 (2003) 13190–13195.
- [10] K. Fan, P. Wei, Q. Feng, S. Chen, C. Huang, L. Ma, B. Lai, J. Pei, Y. Liu, J. Chen, L. Lai, Biosynthesis, purification, and substrate specificity of severe acute respiratory syndrome coronavirus 3C-like proteinase, *J. Biol. Chem.* 279 (2004) 1637–1642.
- [11] J. Shi, Z. Wei, J. Song, Dissection study on the severe acute respiratory syndrome 3C-like protease reveals the critical role of the extra domain in dimerization of the enzyme: defining the extra domain as a new target for design of highly specific protease inhibitors, *J. Biol. Chem.* 279 (2004) 24765–24773.