ISSN 1672-9145 **Acta Biochimica et Biophysica Sinica** 2005, **37**(12): 807–813 **CN** 31-1940/Q

Study on Substrate Specificity at Subsites for Severe Acute Respiratory Syndrome Coronavirus 3CL Protease

Yu-Fei SHAN[1] and Gen-Jun XU[1,2]*

[1] *Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China;*
[2] *College of Life Science, Zhejiang Sci-Tech University, Hangzhou 310018, China*

Abstract Autocleavage assay and peptide-based cleavage assay were used to study the substrate specificity of 3CL protease from the severe acute respiratory syndrome coronavirus. It was found that the recognition between the enzyme and its substrates involved many positions in the substrate, at least including residues from P4 to P2′. The deletion of either P4 or P2′ residue in the substrate would decrease its cleavage efficiency dramatically. In contrast to the previous suggestion that only small residues in substrate could be accommodated to the S1′ subsite, we have found that bulky residues such as Tyr and Trp were also acceptable. In addition, based on both peptide-based assay and autocleavage assay, Ile at the P1′ position could not be hydrolyzed, but the mutant L27A could hydrolyze the Ile peptide fragment. It suggested that there was a stereo hindrance between the S1′ subsite and the side chain of Ile in the substrate. All 20 amino acids except Pro could be the residue at the P2′ position in the substrate, but the cleavage efficiencies were clearly different. The specificity information of the enzyme is helpful for potent anti-virus inhibitor design and useful for other coronavirus studies.

Key words SARS coronavirus 3CL protease; substrate specificity; autocleavage; binding site; inhibitor design

The causative pathogen of severe acute respiratory syndrome (SARS) is a kind of positive strand RNA virus and belongs to the *Coronaviridae* family [1,2]. In the infected cells, the major part of its genome is directly expressed as large polyprotein (pp1a or pp1ab) precursors, which undergo a series of controlled proteolytic processing to generate functional viral proteins [3,4]. A 34 kDa virally-encoded protease named 3CL, also called the main protease, plays a crucial role in the cleavage cascade. The 3CL proteases, or 3C-like proteases, also exist in other viruses, many of which are infamous pathogens, and mediate the viral polyprotein processing in the same way [5].

The detailed information of substrate specificity is the theoretical base for inhibitor design [6], and great efforts have been made, in this regard, to throw light on the features of SARS-coronavirus (CoV) 3CL protease. Based on the homology and highly conserved substrate specificities [7], the cleavage sites of the 3CL protease in SARS-CoV polyprotein (pp1ab, approximately 750 kDa) were suggested to contain Q↓ (S, A, G, N) dipeptides (the cleavage site is indicated by ↓) [8]. The peptide-based cleavage assay *in vitro* confirmed those 11 putative cleavage sites [9]. The crystal structure resolved by Yang *et al*. provided an unexpected binding mode between the SARS-CoV 3CL protease and its substrate-analog inhibitor and it was used to explain the less stringent S2 subsite specificity of the enzyme [10]. However, the inhibitor used in crystal diffraction did not include the residues downstream of the scissile bond. It is insufficient for the design of a potent inhibitor for clinical trial.

We initiated this study to get more information about the substrate specificity of the SARS-CoV 3CL protease, especially on P1′ and P2′ positions. In peptide-based cleavage assay, it was found that many residues had effects on the cleavage efficiency and the deletion of residues

Received: August 23, 2005 Accepted: September 26, 2005
This work was supported by a grant from the National Natural Science Foundation of China (No. 39930060)
*Corresponding author: Tel, 86-21-54921257; Fax, 86-21-54921257; E-mail, gjxu@sibs.ac.cn

DOI: 10.1111/j.1745-7270.2005.00114.x

downstream of the P1′ position or upstream of the P3 position in the substrate would decrease its hydrolysis rate dramatically. It suggested P2′ and P4 residues may play important roles in the recognition between substrate and enzyme. A number of replacements at P1′, P2′ and P4 positions in substrates were made and examined by autocleavage assay. It was found that all residues except Pro were acceptable at the P2′ position in substrate and bulky residues could also be allowed at the P1′ position. We also found that Ile could not be the residue at the P1′ position in substrate, but it could be hydrolyzed by the enzyme mutant L27A (Leu27 of the enzyme was replaced by Ala). As for the P4 position, all eight residues tested were acceptable.

Materials and Methods

Cloning and expression of SARS-CoV 3CL protease

The gene was amplified by polymerase chain reaction (PCR) using forward primer NP2 and reverse primer CP2. All the primer sequences used in this study are listed in **Table 1**. The PCR product was purified and digested with *Nco*I and *Xho*I, then inserted into *Nco*I/*Xho*I cut plasmid pET-28a. The constructed plasmid (pET28a-3CLsm) encodes the SARS-CoV 3CL protease in which the first amino acid Ser was replaced by Met. pET28a-3CLsm was transformed into *Escherichia coli* BL21(DE3) and a single selected colony was grown in one liter of Luria Broth (LB) medium containing kanamycin (50 μg/ml). When the absorbance of culture at 600 nm reached 0.6, it was induced with 10 μM isopropyl β-*D*-thiogalactopyranoside (IPTG) for 6 h. Cells were harvested and the enzyme was purified with a metal-affinity column as described previously [11]. Finally the protease was made in buffer A [20 mM Tris-HCl, 200 mM NaCl, 1 mM EDTA, 1 mM dithiothreitol (DTT), 10% (*V*/*V*) glycerol, pH 7.5] and kept at –80 ºC as the stock solution (2 mg/ml; A_{280}=2.5).

Synthetic peptide substrates

A series of peptides, which were derived from the cleavage site of the SARS-CoV 3CL protease at its N-terminal, were prepared by the solid-phase method. The synthetic peptides were purified through a Beckman ultrasphere C-18 reverse-phase high performance liquid chromatography

Table 1 Oligonucleotide sequences of primers for amplification or site-directed mutagenesis

Primer	Sequence (5′→3′)	Polarity	Remark
NP1	ATGGCTAGCTCAATCACTTCTGCTGTTCTGCAGAGTGGTTT-TAGGAAAATGGCA	Forward	For N-terminal amplification
NP2	TATACCATGGGTTTTAGGAAAATGGCATTC	Forward	For N-terminal amplification
CP2	GGTGCTCGAGTTGGAAGGTAACACCAGAGCATTGTC	Reverse	For C-terminal amplification
Pri1′	ATGGCTAGCTCAATCACTTCTGCTGTTCTGCAG*NNN*GGTTT-TAGGAAAATGGCA	Forward	For mutations at P1′ position
Pri2′	ATGGCTAGCTCAATCACTTCTGCTGTTCTGCAGAGT*NNN*TT-TAGGAAAATGGCA	Forward	For mutations at P2′ position
Pri4	ATGGCTAGCTCAATCACTTCT*NNN*GTTCTGCAGAGTGGTTT-TAGGAAAATGGCA	Forward	For mutations at P4 position
P27A+	TGTGGAACTACAACTGCTAATGGATTGTGGTTG	Forward	For mutating
P27A–	CAACCACAATCCATTAGCAGTTGTAGTTCCACA	Reverse	Leu27 to Ala
P48A+	ATTTGCACAGCAGAAGCTATGCTTAATCCTAAC	Forward	For mutating
P48A–	GTTAGGATTAAGCATAGCTTCTGCTGTGCAAAT	Reverse	Asp48 to Ala
P48K+	ATTTGCACAGCAGAAAAGATGCTTAATCCTAAC	Forward	For mutating
P48K–	GTTAGGATTAAGCATCTTTTCTGCTGTGCAAAT	Reverse	Asp48 to Lys

The oligonucleotide sequences indicating mutated codons are underlined. Pri1′, five primers with different codons shown by *NNN*; the codons used for each mutant were: CGC (Arg), TGG (Trp), ATT (Ile), TTT (Phe), TAC (Tyr); Pri2′, 11 primers with different codons shown by *NNN*; the codons used for each mutant were: TTT (Phe), TGG (Trp), CAC (His), CCA (Pro), TAC (Tyr), TGT (Cys), GAC (Asp), ATG (Met), CAA (Gln), ACT (Thr), AGG (Arg); Pri4, eight primers with different codons shown by *NNN*; the codons used for each mutant were: TAC (Tyr), AGG (Arg), CCA (Pro), ATG (Met), TGT (Cys), CTG (Leu), ATT (Ile), GGT (Gly).

(RP-HPLC) column (4.6 mm×250 mm). The identity and homogeneity of peptides were confirmed by mass spectroscopy and RP-HPLC. The sequences of peptides are shown in **Table 2**.

Table 2 Relative severe acute respiratory syndrome coronavirus 3CL protease cleavage efficiencies of synthetic peptides with different sequences

Peptide	Sequence	Cleavage efficiency
Sub1	T S I T S A V L Q S G F R K M A	4.78
Sub2	T S A V L Q S G F R K M A	4.52
Sub3	S A V L Q S G F R K M A	4.49
Sub4	A V L Q S G F R K M A	3.61
Sub5	V L Q S G F R K M A	UD
Sub0	T S A V L Q S G F	1.00
Sub6	T S A V L Q S G	0.42
Sub7	T S A V L Q S	UD
Sub8	T S A V L Q S Y F	0.17
Sub9	T S A V L Q S W F	0.22
Sub10	T S A V L Q S P F	UC
Sub11	T S A V L Q I G F	UC
Sub12	T S A V L Q W G F	<0.05

UD, the cleavage efficiency for the peptides was too low to be determined in competitive cleavage reaction; UC, the peptides were not cleaved under the conditions provided in "Materials and Methods".

Peptide-based cleavage assay *in vitro*

Cleavage assays were performed in buffer B (20 mM Tris-HCl, pH 7.35, 200 mM NaCl, 1 mM DTT, and 1 mM EDTA) containing 200 μM of each peptide and 1 μM of enzyme. The reaction mixtures were incubated at 20 ºC for 0–16 h and quenched by trifluoroacetic acid (1% final concentration). The products were analyzed by RP-HPLC, using a 0%–75% linear gradient of acetonitrile containing 0.1% trifluoroacetic acid. The peaks were collected and identified by mass spectroscopy.

The cleavage efficiency of different peptides was compared by relative specificity constant $(k_{cat}/K_m)_{rel}$ according to Cordingley *et al*. [12]. Briefly, 200 μM of each peptide was incubated in buffer B with 1 μM enzyme in the presence of 200 μM peptide Sub0 for competitive cleavage reaction. The peak areas of products remained linear at the used time intervals (generally less than 10% substrate hydrolyzed) and the peak areas were integrated to calculate $(k_{cat}/K_m)_{rel}$.

Plasmids of proenzyme for autocleavage assay

To construct the plasmids encoding proenzyme for autocleavage assay, a pair of primers NP1 and CP2 was used. The PCR product was digested with *Nhe*I and *Xho*I, and then inserted into *Nhe*I/*Xho*I cut plasmid pET-28a. The constructed plasmid (pET28a-3CLsc) encodes a proenzyme with a 31-mer leading peptide (N_{31}-tag) at its N-terminal and a His_6-tag at C-terminal. There is an autocleavage site between N_{31}-tag and the first amino acid (Ser) of the protease. If autocleavage occurs, N_{31}-tag would be cut off, and the plasmid will produce a band with an apparent molecular weight of 35 kDa on sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). Otherwise, the molecular weight of the proenzyme would be approximately 39 kDa.

Site-directed mutagenesis of the enzyme

Site-directed mutagenesis at P4, P1′ and P2′ positions on N_{31}-tag was performed by a recombination PCR method (the position of amino acids in substrates named from N- to C-terminal as follows: -P4-P3-P2-P1↓P1′-P2′-P3′-P4′) [13]. Briefly, the gene of the SARS 3C-like protease was amplified by PCR using forward primers containing mutational codons and a reverse primer (CP2). The PCR products were purified and digested with *Nhe*I and *Xho*I, then inserted into *Nhe*I/*Xho*I cut plasmid pET-28a. Plasmids of mutants were transformed into BL21(DE3). Single colonies were selected to grow in LB medium and induced with 10 μM IPTG for 4–8 h when A_{600} got to 0.6. The cells were harvested and analyzed by 12% SDS-PAGE. Mutagenesis of Leu27 and Asp48 was performed by three-step PCR in a similar way.

Results

Expression and purification of SARS-CoV 3CL protease

As the first N-terminal amino acid of the enzyme is Ser, three strategies were used to express SARS-CoV 3CL protease in *E. coli*. The first strategy was replacing the first Ser with Met; the second was adding an additional Met prior to Ser; and the third was introducing a leading sequence to the N-terminal, which was cut off by proteases after expression. It was found that the specific activities of all the resulting enzymes were compatible, as reported previously. We had successfully over-expressed the SARS-CoV 3CL protease in *E. coli* by mutating the

first amino acid Ser to Met. The purification procedure was described previously by using a metal-nitrilotriacetic acid column followed by a gel filtration column (Sephacryl S-200; GE Healthcare, USA) [9,11,14]. The product showed a single band on SDS-PAGE. After determining the concentration, the enzyme was subsequently employed in peptide-based cleavage assays.

Peptide-based cleavage assay for core-part of substrates

In order to obtain information about how the length of substrate affects the enzyme catalysis, a group of peptides were synthesized. These peptides of different lengths were derived from the sequence (TSITSAVLQSGFRKMA), one of the cleavage sites for the SARS-CoV 3CL protease, which was previously proved to be an efficient substrate [11]. The capability of various peptides to compete with a 9-mer peptide (Sub0, P6–P3′) for cleavage by the 3CL protease was estimated by RP-HPLC. The $(k_{cat}/K_m)_{rel}$ for each peptide compared to the reference substrate (Sub0) is shown in **Fig. 1**.

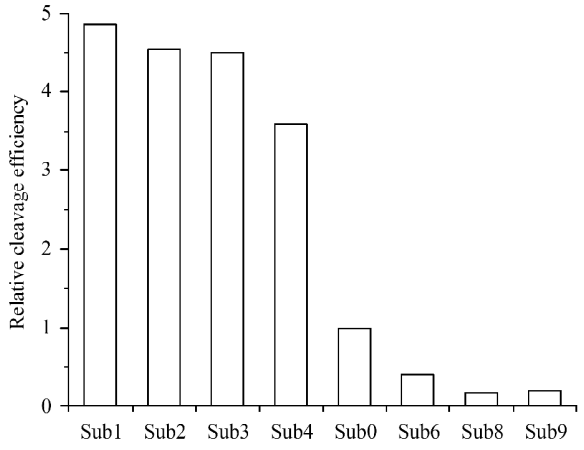

Fig. 1 Peptide-based competitive cleavage assay

In buffer B, 200 μM of each peptide was incubated with 1 μM enzyme in the presence of 200 μM peptide Sub0 for competitive cleavage reaction at 20 °C. Products were analyzed by RP-HPLC. The relative cleavage efficiency was the capability of various peptides to compete with a 9-mer peptide (Sub0, P6–P3′) for cleavage by the 3CL protease.

Results of competition assay showed that peptides Sub1 (P9–P7′), Sub2 (P6–P7′), Sub3 (P5–P7′) and Sub4 (P4–P7′) had similar $(k_{cat}/K_m)_{rel}$ values. It indicated that residues P5–P9 contributed little to the cleavage reaction. The $(k_{cat}/K_m)_{rel}$ value for Sub2 (P6–P7′) is about five times that for Sub0 (P6–P3′). The 8-mer peptide (Sub6, P6–P2′) had a $(k_{cat}/K_m)_{rel}$ value comparable to that of Sub0. It suggested that residues of P3′–P7′, taken as a whole, might contribute some energy to the binding between the enzyme and its substrate as the rate of hydrolysis enhanced. The cleavage rates for Sub7 (7-mer, P6–P1′) and Sub5 (10-mer, P3–P7′) were too low to be determined in the competitive cleavage assay. It meant that the deletion of residue P4 or P2′ would make it difficult for a substrate to be hydrolyzed by the SARS-CoV 3CL protease, suggesting that residues P4 and P2′ are two critical positions in the interaction between the enzyme and its substrate. P1' has been reported to play an important role in the recognition of the enzyme and the substrate. But systematic investigation of the P1′ position has not been reported in published work.

Requirements of P1′, P2′ and P4 positions in substrate

It was proposed that the P1′ position is one of the important determinants in substrate specificity for 3CL/3C protease [3,5,7,15]. P1′ positions of 11 known cleavage sites in the SARS-CoV 3CL protease are all occupied by amino acid residues with small side chains (Ser, Ala, Gly and Asn). Asn at the P1′ position is regarded as a noncanonical residue [8]. Consequently it was suggested in published reports that the SARS-CoV 3CL protease, similar to other 3CL proteases, had a relatively small S1′ subsite (required for P1′ residue binding) that may exclude bulky amino acid residues. In a previous study, we successfully constructed an autocleavage assay by introducing an N_{31}-tag to the N-terminal of the SARS-CoV 3CL protease [**Fig. 2(B)**]. We found that the expression product of a mutant proenzyme of C145S (Cys145 of the enzyme was replaced by Ser) appeared in two bands. One was 38.5 kDa; the other was 34.6 kDa [**Fig. 2(A)**, lane 8]. It was ascribed to an incomplete proteolysis of the mutant. Therefore, the mutant served as a marker for estimating molecular weights. The assay was simple, sensitive and reliable, so that it was employed to study the requirements of substrate for the SARS-CoV 3CL protease.

To verify the suggestion above, P1′ Ser in N_{31}-tag was mutated to a number of bulky amino acids (Tyr, Ile, Phe, Trp and Lys). Interestingly, all these mutants showed a 35 kDa band on SDS-PAGE except the P1′I mutant (SARS-CoV 3CL protease flanked an N_{31}-tag in which the P1′ Ser was replaced by Ile, the other mutants are designated similarly) (**Fig. 3**). It implied that most of the mutant proenzymes effectively performed autocleavage. That is to say, the S1′ subsite could accommodate bulky amino acids.

According to the peptide-based cleavage assay, the P2′ residue was more important than those of P3′ to P7′. But

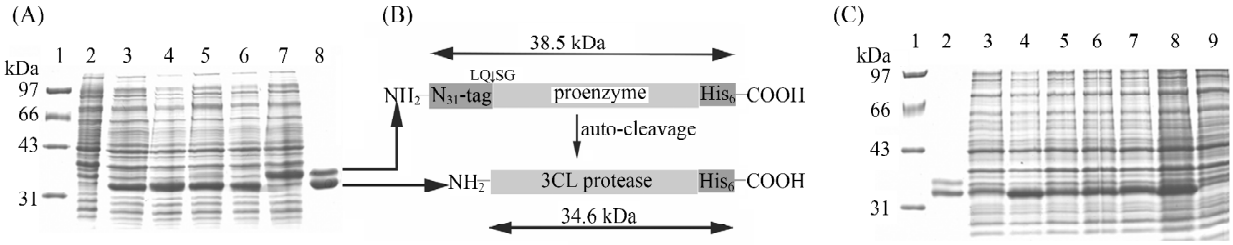

Fig. 2 Autocleavage assay of P1′ position mutants

Total cell lysates of isopropyl β-*D*-thiogalactopyranoside-induced *E. coli* samples were separated by 12% sodium dodecyl sulphate-polyacrylamide gel electrophoresis and stained with Coomassie brilliant blue. (A) 1, protein marker; 2, negative control; 3–7, mutants of P2′F, P2′H, P2′W, P2′Y and P2′P (from left to right); 8, positive control. (B) Schematic diagram of severe acute respiratory syndrome coronavirus 3CL protease autocleavage assay. (C) 1, protein marker; 2, negative control; 3–8, mutants of P2′C, P2′D, P2′M, P2′Q, P2′R and P2′T; 9, positive control.

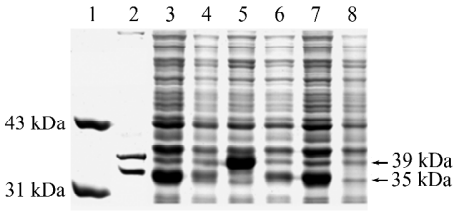

Fig. 3 Autocleavage assay of P2′ position mutants

Total cell lysates of isopropyl β-*D*-thiogalactopyranoside-induced *E. coli* samples were separated by 12% sodium dodecyl sulphate-polyacrylamide gel electrophoresis and stained with Coomassie brilliant blue. 1, protein marker; 2, a mutant ($C^{145}S$) has two expression products as an apparent molecular weight positive control; 3–7, mutants of P1′Y, P1′F, P1′I, P1′W, P1′K; 8, lysate of *E. coli* cell (containing plasmid pET-28a) as the negative control.

few reports supported the importance of the P2′ position to the substrate specificity for 3CL proteases and related 3C proteases. In viral protein pp1ab, nine amino acids (Gly, Lys, Ile, Asn, Ala, Val, Glu, Leu and Ser) were found at P2′ positions near the cleavage sites. So we replaced P2′ Gly with the other 11 residues (Cys, Asp, Phe, His, Met, Pro, Gln, Arg, Thr, Trp and Tyr). It was found that P2′ could accept a wide range of amino acid residues, except Pro as shown by autocleavage assay, in which P2′P resulted in only one band with an apparent molecular weight of 39 kDa [**Fig. 2(A)**, lane 6], whereas the others yielded a band of 35 kDa [**Fig. 2(A,C)**]. It clearly indicated that only Pro could not be allowed at the P2′ position.

The P4 position was also suggested as critical to the substrate specificity of 3CL proteases because of its limited variety, in that only four kinds of residues (Ala, Val, Pro and Thr) had been found in the cleavage sites for the SARS-CoV 3CL protease [9]. In this study, eight mutants (Tyr, Arg, Met, Pro, Cys, Leu, Ile and Gly) were constructed and tested by autocleavage assay. Results showed that all the mutants effectively performed autocleavage as shown by producing an apparent molecular weight of 35 kDa on SDS-PAGE (data not shown).

The region of the N-terminal in coronavirus 3CL proteases is essential [3,16]. The deletion of the N-terminal (residues 1–5) in the SARS-CoV 3CL protease eliminated most of its activity (unpublished data). In regard to the fact that the amino acids at P1′ and P2′ positions in N_{31}-tag were also the first two N-terminal residues in the enzyme, modification of any position may affect the enzyme activity. It was also reported previously that the *trans*-cleavage assay might have a different result from the *cis*-cleavage assay [3,11]. Both of them may consequently lead to an artificial conclusion. To exclude the possibilities, a parallel experiment of peptide-based cleavage assay should be conducted.

Five synthetic 9-mer peptides with a single-point substitution at P2′ in the substrate (**Table 2**) were afforded for *trans*-cleavage assay. Peptide Sub2 served as the positive control [**Fig. 4(A,B)**]. SARS-CoV 3CL protease hydrolyzed Sub8 (Tyr at P2′) and Sub9 (Trp at P2′) at a relatively low rate (**Fig. 1**). The hydrolysis rate for Sub12 (Trp at P1′) was even lower, and undetectable for Sub10 (Pro at P2′) even with a higher concentration of enzyme (10 μM) and a prolonged incubation time (48 h) employed [**Fig. 4(C,D)**].

Site-directed mutagenesis on Leu27 and Asp48

Based on the fact that the SARS-CoV 3CL protease could not hydrolyze the substrate with Ile at the P1′ position, it is reasonable to suggest a repulsion interaction may exist in Ile substrate. According to the suggestion of Anand *et al.*, the residues around S1′ in the SARS-CoV 3CL protease are comprised of Leu27, His41 and Asp48 [8]. In our previous study, we found the mutant of H41A was inactive, so that the mutation of His41 was not under consideration.

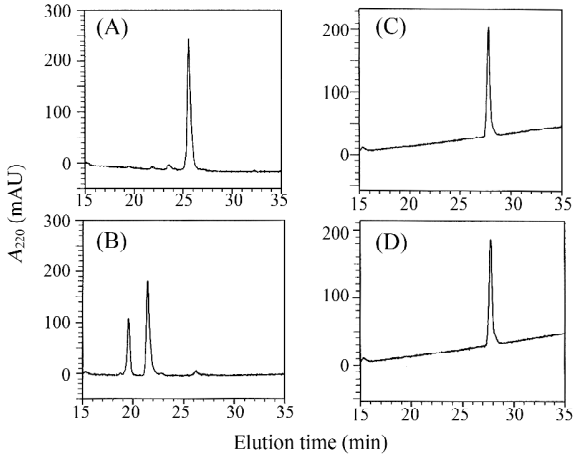

Fig. 4 Peptide-based cleavage assay *in vitro* for determining the effects of substitution at P2′ position

Peptides were cleaved by the severe acute respiratory syndrome coronavirus 3CL protease under conditions described in "Materials and Methods". Peptide Sub2 for 1 min (A) and 16 h (B); peptide Sub10 for 1 min (C) and 48 h (D).

We mutated the enzyme to L27A (Leu27 was replaced by Ala) and D48A (Asp48 was replace by Ala). Both of the mutants were active. In autocleavage assay, it was found that the proenzyme of L27A-P1′I (L27A mutant with Ile at the P1′ position of N_{31}-tag) effectively performed autocleavage (**Fig. 5**, lane 4), but L27A did not hydrolyze the substrate with Pro at the P2′ position, as shown by L27A-P2′P (L27A mutant with Pro at the P2′ position of N_{31}-tag) having only one band with an apparent molecular weight of 39 kDa on SDS-PAGE (**Fig. 5**, lane 5). Neither D48A nor D48R could cleave N_{31}-tag, if Ile presented at the P1′ position (**Fig. 5**, lanes 7 and 8). According to the

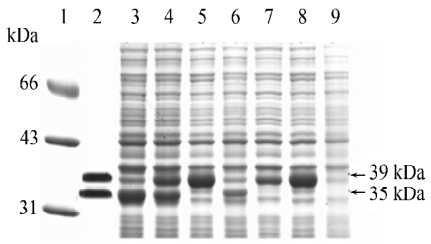

Fig. 5 Substitution effects of Leu27 and Asp48 as detected by autocleavage assay

Total cell lysates of isopropyl β-*D*-thiogalactopyranoside-induced *E. coli* were separated by 12% sodium dodecyl sulphate-polyacrylamide gel electrophoresis and stained with Coomassie brilliant blue. 1, protein marker; 2, positive control; 3–8, mutants of L27A, L27A-P1′I, L27A-P2′P, D48A, D48A-P1′I, D48K-P1′I; 9, negative control.

data of crystal diffraction, Asp48 is apart from the active site residues (His41 and Cys145), which may not be a part of the binding site [10]. No effect produced by the mutant was expected.

Discussion

The highly stringent substrate specificity of 3CL proteases is vital to all coronaviruses, as proteolytic cleavage must be accurate to release functional proteins from large polyprotein precursors step by step [3,17]. Otherwise, the biological function, which should take place at a certain phase in the viral life cycle, will be disturbed. The specificity is mainly realized by the recognition between the enzyme and a wide range of positions in the substrate. It was suggested that the positions P4–P3′ in the substrates of 3CL proteases all have a restricted variability [3]. The substrate binding sites of all 3CL proteases are conserved [3,6,7].

In this paper, firstly, we showed the substrate for the SARS-CoV 3CL protease had a core-part that enrolled residues P4–P2′ and all of the residues are indispensable for effective cleavage. The characteristics of amino acids at three positions on the core-part sequence were studied systematically by two parallel assays. We demonstrated that bulky amino acids such as Trp and Tyr could be also allowed at P1′ and P2′ positions in the substrate by peptide-based and autocleavage assays. However, it is evident that the presence of these two large residues at the P2′ position would decrease the efficiency of proteolysis (approximately 5-fold). Similarly, the cleavage rate for the peptide with Trp at the P1′ position was even lower.

Pro could take one of the *cis*- or *trans*-configurations, the predominant one existing in the peptide may not be in favor of the binding to the enzyme. Alternatively, Pro lacks a hydrogen atom at the peptide bond, which may give it different characteristics from other amino acids in properly binding to the enzyme. Both characteristics could explain why Pro could not be accepted at the P2′ position. The protease did not hydrolyze the substrates with Ile at the P1′ position. It may be explained in terms of steric hindrance, which was mainly produced by Leu27. This speculation was supported by the fact that the mutant enzyme L27A could hydrolyze the substrate with Ile at the P1′ position. The results also showed that residues P3′–P7′ were not essential, although they obviously affect the cleavage efficiency. This is consistent with a previous report [18]. This information should be helpful in the design of a potent 3CL protease inhibitor.

Acknowledgement

We thank Jian CHEN (College of Life Science, Zhejiang Sci-Tech University, Hangzhou, China) for operating the mass spectrometer.

References

1 Rota PA, Oberste MS, Monroe SS, Nix WA, Campagnoli R, Icenogle JP, Penaranda S *et al*. Characterization of a novel coronavirus associated with severe acute respiratory syndrome. Science 2003, 300: 1394–1399
2 Holmes KV, Enjuanes L. The SARS coronavirus: A postgenomic era. Science 2003, 300: 1377–1378
3 Ziebuhr J, Snijder EJ, Gorbalenya AE. Virus-encoded proteinases and proteolytic processing in the Nidovirales. J Gen Virol 2000, 81: 853–879
4 Anand K, Palm GJ, Mesters JR, Siddell SG, Ziebuhr J, Hilgenfeld R. Structure of coronavirus main proteinase reveals combination of a chymotrypsin fold with an extra α-helical domain. EMBO J 2002, 21: 3213–3224
5 Palmenberg AC. Proteolytic processing of picornaviral polyprotein. Annu Rev Microbiol 1990, 44: 603–623
6 Kiemer L, Lund O, Brunak S, Blom N. Coronavirus $3CL^{pro}$ proteinase cleavage sites: Possible relevance to SARS virus pathology. BMC Bioinformatics 2004, 5: 72
7 Hegyi A, Ziebuhr J. Conservation of substrate specificities among coronavirus main proteases. J Gen Virol 2002, 83: 595–599
8 Anand K, Ziebuhr J, Wadhwani P, Mesters JR, Hilgenfeld R. Coronavirus main proteinase ($3CL^{pro}$) structure: Basis for design of anti-SARS drugs. Science 2003, 300: 1763–1767
9 Fan K, Wei P, Feng Q, Chen S, Huang C, Ma L, Lai B *et al*. Biosynthesis, purification, and substrate specificity of severe acute respiratory syndrome coronavirus 3C-like proteinase. J Biol Chem 2004, 279: 1637–1642
10 Yang H, Yang M, Ding Y, Liu Y, Lou Z, Zhou Z, Sun L *et al*. The crystal structures of severe acute respiratory syndrome virus main protease and its complex with an inhibitor. Proc Natl Acad Sci USA 2003, 100: 13190–13195
11 Shan YF, Li SF, Xu GJ. A novel auto-cleavage assay for studying mutational effects on the active site of severe acute respiratory syndrome coronavirus 3C-like protease. Biochem Biophys Res Commun 2004, 324: 579–583
12 Cordingley MG, Callahan PL, Sardana VV, Garsky VM, Colonno RJ. Substrate requirements of human rhinovirus 3C protease for peptide cleavage *in vitro*. J Biol Chem 1990, 265: 9062–9065
13 Schechter I, Berger A. On the size of the active site in proteases. I. Papain. Biochem Biophys Res Commun 1967, 27: 157162
14 Sun H, Luo H, Yu C, Sun T, Chen J, Peng S, Qin J *et al*. Molecular cloning, expression, purification, and mass spectrometric characterization of 3C-like protease of SARS coronavirus. Protein Expr Purif 2003, 32: 302–308
15 Wirblich C, Sibilia M, Boniotti MB, Rossi C, Thiel HJ, Meyers G. 3C-like protease of rabbit hemorrhagic disease virus: Identification of cleavage sites in the ORF1 polyprotein and analysis of cleavage specificity. J Virol 1995, 69: 7159–7168
16 Hsu WC, Chang HC, Chou CY, Tsai PJ, Lin PI, Chang GG. Critical assessment of important regions in the subunit association and catalytic action of the severe acute respiratory syndrome coronavirus main protease. J Biol Chem 2005, 280: 22741–22748
17 Ziebuhr J, Heusipp G, Siddell SG. Biosynthesis, purification, and characterization of the human coronavirus 229E 3C-like proteinase. J Virol 1997, 71: 3992–3997
18 Shi J, Wei Z, Song J. Dissection study on the severe acute respiratory syndrome 3C-like protease reveals the critical role of the extra domain in dimerization of the enzyme: Defining the extra domain as a new target for design of highly specific protease inhibitors. J Biol Chem 2004, 279: 24765–24773

Edited by
Ming-Hua XU