# The N-Finger Peptide of SARS-CoV 3CL$^{pro}$ is Essential for the Enzyme Activity

Sun-Hua Zhang[1,2]*, Ming-Fang Ji[3], Xiao-Xia Shao[2]

([1]*Department of Biology, School of Life Science, East China Normal University, Shanghai 200062, China;* [2]*Institute of Protein Research, Tongji University, Shanghai 200092, China;* [3]*Scino Pharm Kunshan Biochemical Ltd. Co, Kunshan 215300, China*)

**Abstract** The genes encoding 3CL protease (3CL$^{pro}$) of SARS virus and its mutant 3CL$^{pro}$($\Delta_{1-7}$) were constructed and expressed in *E.coli* system. The 3CL$^{pro}$ showed a proteolytic activity on the fluorogenic substrate containing the consensus auto-cleavage site of 3CL$^{pro}$. However, the 3CL$^{pro}$($\Delta_{1-7}$) mutant showed no enzymatic activity at the same conditions as wild type did, indicating that the N-finger peptide of 3CL$^{pro}$ is essential for the enzymatic activity though it is not the enzyme reactive center. We also synthesized the N-finger peptide by solid-phase peptide synthesis method and this peptide showed partially inhibitory effect on the SARS 3CL protease.

**Key words** SARS-coronavirus; 3CL protease; N-finger peptide

Atypical pneumonia, with overall fatality rates of about 10%, characteristic of high fever, headache, and nonproductive cough emerged in southern China in late 2002[1,2] and then rapidly spread to other countries. The disease was later named the severe acute respiratory syndrome (SARS) by WHO, and a novel coronavirus, termed SARS-CoV, was rapidly identified as the pathogenic agent[3,4]. Subsequently, its genome sequence has been elucidated[5]. It is generally believed that SARS-CoV has evolved from an animal coronavirus that recently crossed the species barrier[6].

Coronaviruses are enveloped, plus-strand RNA viruses that feature the largest RNA genomes currently known[7]. Through ribosomal frame shift switching, the replicase gene of SARS coronavirus is translated into two polyproteins with overlapping sequences[8]. Replicative polyproteins 1a and 1ab are essential for virus RNA synthesis. The central and C-terminal regions of these polyproteins are cleaved into functional proteins by 3CL$^{pro}$, a protease with a chymotrypsin-like fold. 3CL$^{pro}$ encoded and located at the central part of replicase gene is auto-processed at the help of hydrophobic proteins encoded also by virus gene, and is released from the polyprotein[9]. Because of its distant relationship with the 3C proteases of picornaviruses, this protease is named 3C-like protease (3CL$^{pro}$)[9]. 3CL$^{pro}$ plays a pivotal role in coronavirus polyprotein processing and also releases the key replicative functions of the virus, such as RdRp and helicase; therefore, it is also called the coronavirus main protease[9]. Due to its functional importance in the viral life cycle, 3CL$^{pro}$ has been made an attractive target for developing drugs directly against the new disease.

Recently, Yang *et al.*[10] solved the crystal structure of 3CL$^{pro}$. This protease employs a catalytic Cys-His dyad and has a three-domain structure. The first two domains form a chymotrypsin-like fold, which is responsible for the catalytic reaction. The third domain is α-helical unique to the coronavirus 3CL proteases, which contributes to the dimerization of 3CL$^{pro}$, thus switching the enzyme from the inactive form (monomer) to the active form (dimer)[11]. The N-terminal residues 1–7 (N-finger) play an important role in the dimerization and formation of the active site of 3CL$^{pro}$[10,12]. Here we report clone and expressed 3CL$^{pro}$ and its mutant 3CL$^{pro}$ ($\Delta_{1-7}$). The wild type exhibited an enzyme activity while the mutant didn't suggesting that the N-finger is essential

Received: April 9, 2005 Accepted: July 14, 2005

This work was supported by the Major State Basic Research Development Program of China (973 Program) (No.2003CB514107)

*Corresponding author. Tel: 86-21-65988805, Fax: 86-21-65988403, E-mail: zhangsh0108@126.com

for maintenance of enzyme activity. In addition the synthesis of the N-finger peptide by solid-phase peptide synthesis method and its inhibitory effect on the SARS 3CL protease was described. These results might provide a clue to design a specific inhibitor on $3CL^{pro}$ targeting out of its active site.

## 1 Materials and Methods

### 1.1 Materials

The plasmid pGEX-$M^{pro}$, which contains the DNA sequence of SARS $3CL^{pro}$, was a kind gift from Prof. Zhi-He Rao. Restriction enzymes *Bam*HI and *Xho*I were purchased from New England Biolabs (NEB); *Pfu* DNA polymerase from MBI Fermentas; T4 DNA ligase from TaKaRa; IPTG and DTT from Sigma; the GST affinity matrix from Amersham-Pharmacia Biotech; enterokinase light chain from New England Biolabs (NEB); protein quantitative analysis kit from Shenergy Biocolor Company (China). All other chemicals in this study were of analytical or higher grade.

### 1.2 Construction of the GST-fusion plasmids of pGEX-6P-1-$3CL^{pro}$ and pGEX-6P-1-$3CL^{pro}$ ($\Delta_{1-7}$)

The GST-rhinovirus protease was used to release $3CL^{pro}$ from the fusion protein transformed from the original plasmid provided by Prof. Rao. This cleavage leaves 5 extra amino acid residues (Gly-Pro-Leu-Gly-Ser) upstream from the N-terminal of $3CL^{pro}$. In order to remove these 5 extra residues, the expression plasmid pGEX-6P-1-$3CL^{pro}$ was reconstructed, and the enterokinase cleavage site DDDDK was introduced (GAT GAC GAT GAC AAG upstream from the 5' end of $3CL^{pro}$). A forward primer of 5'-CGG GAT CCG ATG ACG ATG ACA AGA GTG GTT TTA GGA AAA TGG CAT TC-3' and a reverse primer of 5'-CCG CTC GAG TTA TTG GAA GGT AAC ACC AGA GCA T-3' were designed to amplify the DNA fragment encoding the entire $3CL^{pro}$. The PCR products encoding $3CL^{pro}$ without 5 extra amino acid residues were cloned into the pGEX-6P-1 vector (Pharmacia) and sequenced using *Bam*HI/*Xho*I restriction sites.

The plasmid pGEX-6P-1-$3CL^{pro}$ ($\Delta_{1-7}$) with 1–21 nucleotide deleted at the 5' end was also constructed to produce the N-finger peptide truncated $3CL^{pro}$. The forward primer was 5'-CGG GAT CCG ATG ACG ATG ACA AGT TCC CGT CAG GCA AAG TTG AAG GT -3', and the reverse primer 5'- CCG CTC GAG TTA TTG GAA GGT AAC ACC AGG CAT -3', the template was pGEX-$M^{pro}$. The enterokinase cleavage site gene was also inserted upstream of the $3CL^{pro}$ gene.

### 1.3 Expression and purification of recombinant proteins $3CL^{pro}$ and $3CL^{pro}$ ($\Delta_{1-7}$)

The expression constructs were transformed into the *Escherichia coli* strain BL21 to over-express the GST-fusion proteins. Briefly, the cells were cultured at 37 ℃ until the absorbance at 600 nm reached 1.2. Then 0.5 mmol/L IPTG was added into the cell culture to induce the foreign protein expression at 30 ℃ for 3 h. The cells was then spanned down and sonicated in 1 × PBS (140 mmol/L NaCl, 2.7 mmol/L KCl, 10 mmol/L $Na_2HPO_4$, 1.8 mmol/L $KH_2PO_4$, pH 7.3), 1 mmol/L DTT, 1% Triton X-100, 100 μg/ml PMSF to release GST-proteins. The supernatant was subsequently loaded on a GST-affinity column pre-equilibrated with 1 × PBS and washed with 20 mmol/L Tris-HCl buffer (pH 8.0). Then the column was eluted with 15 mmo/L reduced glutathione in 20 mmo/L Tris-HCl buffer (pH 8.0). The GST-proteins solution was added with urea and guanidine hydrochloride to a final concentration of 1 mmo/L, and then cleaved with enterokinase for 16 h at room temperature to release $3CL^{pro}$ or its mutant. The mixture was dialyzed and reloaded on the glutathione Sepharose beads pre-equilibrated with 1 × PBS to remove GST. The purified proteins $3CL^{pro}$ and its mutant were then dialyzed, lyophilized, and then stored at –80 ℃

### 1.4 Enzyme activity assay with a fluorogenic substrate

The enzyme activity of the full-length $3CL^{pro}$ and $3CL^{pro}$($\Delta_{1-7}$) was measured by continuous kinetic assays using the fluorogenic substrate containing a consensus auto-cleavage site of $3CL^{pro}$ (Dabcyl-KTSAVL**Q**SGFRKME-Edans, the cleaved peptide bond is bolded) kindly provided by Prof. Po-Huang Liang. The released fluorescence was monitored at 37 ℃ in the period of 120 s on the fluorescence spectrophotometer[13,14] ( F-2500, Hithachi, Japan), using an excitation and an emission

wavelength of 355 nm and 538 nm, respectively. The final concentration of substrate dissolved in 20 mmo/L Bis-Tris (pH 7.0) containing 2 mmo/L DTT was 480 nmo/L. The final concentration of enzyme added was 34 nmo/L.

### 1.5 Peptide synthesis

The N-finger peptide: $H_2N$-Ser-Gly-Phe-Arg-Lys-Met-Ala-COOH were prepared by solid-phase peptide synthesis and purified by high-performance liquid chromatography (HPLC) on a reversed-phase C8 column[15] (10 mm × 250 mm, Beckman, USA). Identity and homogeneity of the peptides were confirmed by mass spectrometry and analytical reversed-phase chromatography.

### 1.6 N-finger peptide partial inhibits on SARS 3CL protease

The partially inhibitory effect of N-finger was measured in a reaction mixture containing 34 nmol/L SARS 3CL protease, 480 nmol/L fluorogenic substrate in a buffer of 20 mmol/L Bis-Tris (pH 7.0), 2 mmol/L DTT in the presence of various concentrations of N-finger peptide, which ranged from 50 μmol/L to 300 μmol/L. The fluorescence change resulted from the reaction was followed with time using the fluorescence spectrophotometer (F-2500, Hithachi, Japan). Protease (final concentration of 34 nmol/L) was incubated for 15 min at 37 ℃ with the N-finger peptide (final concentration of 50 μmol/L to 300 μmol/L) and the reaction was initiated by adding substrate to the desired final concentration (34 nmol/L for experiments at a constant substrate concentration).

## 2 Results

### 2.1 Expression and purification of the recombinant $3CL^{pro}$ and its mutant

To express a fusion protein containing the complete $3CL^{pro}$ ($\Delta_{1-7}$), the bacterial expression vector pGEX-6P-1-$3CL^{pro}$ ($\Delta_{1-7}$) was constructed as described in Materials and Methods. This construct encodes a fusion protein that is essentially composed of the N-terminal glutathione S transferase fused to $3CL^{pro}$($\Delta_{1-7}$). The expression construct was transformed into *E.coli* strain B21. After induction, the cells were sonicated, the GST-$3CL^{pro}$ ($\Delta_{1-7}$) fusion protein was bound to the GST-affinity column from the supernatant of the lysate, and eluted by the reduced glutathione (Fig.1, Fig.2). The final yield of fusion protein was about 10 mg/L culture. The fusion protein was then cleaved with enterokinase, and the GST was removed by using GST-affinity chromatography. The Coomassie blue-stained SDS-PAGE revealed that one major band of the recombinant $3CL^{pro}$($\Delta_{1-7}$) protein was in the unbound fraction (Fig.3). The purified $3CL^{pro}$ ($\Delta_{1-7}$) did not contain extra 5 amino acid residues as

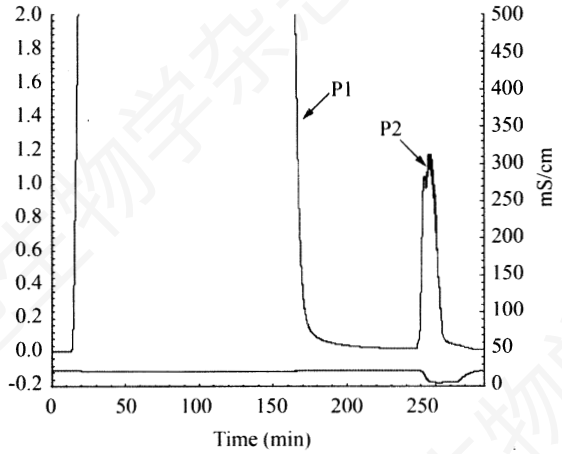


**Fig.1 Purification of GST-$3CL^{pro}$ ($\Delta_{1-7}$) fusion protein by GST affinity chromatography**

P1 peak represents the impurity proteins of the lysate flowing through the GST column; P2 peak represents the eluted GST-$3CL^{pro}$ ($\Delta_{1-7}$) fusion protein by the reduced glutathione.

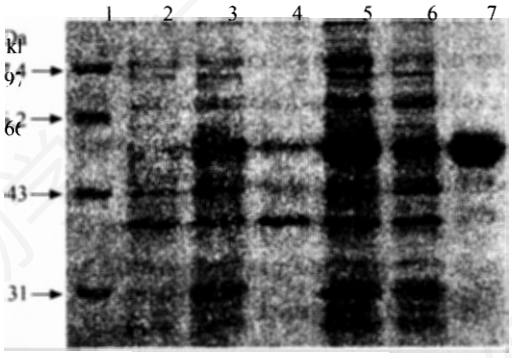


**Fig.2 SDS-electrophoresis of the GST-$3CL^{pro}$ ($\Delta_{1-7}$) fusion protein**

GST-$3CL^{pro}$ ($\Delta_{1-7}$) samples taken at each step of purification were analyzed on a 15% SDS-PAGE gel, and the protein was stained with Coomassie Brilliant Blue. Lane 1, marker of standard proteins; lanes 2 and 3, the cell lysates from the non-induced and IPTG induced *E. coli* BL21 (DE3); lanes 4 and 5, the supernatant and precipitate separated by centrifugation from the cell lysate; lane 6, unabsorbed proteins eluted through the GST column; lane 7, the fractions eluted by reduced glutathione from the GST column.

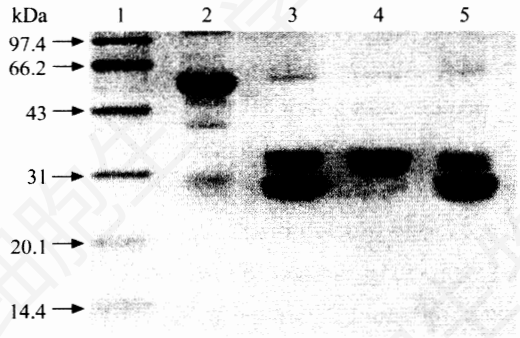


**Fig.3 SDS-PAGE analysis of the fusion protein GST-3CL$^{pro}$ ($\Delta_{1-7}$) and the free 3CL$^{pro}$ ($\Delta_{1-7}$)**

Lane 1, marker of the standard proteins; Lane 2, the fusion protein prior the cleavage by entrokinase; Lane 3, proteins cleaved by enterokinase for 16 h; Lane 4, the unbound 3CL$^{pro}$ ($\Delta_{1-7}$) protein from the GST affinity column.eluted by the reduced glutathione; Lane 5, the unbound protein eluted by the reduced glutathione. 3CL$^{pro}$ ($\Delta_{1-7}$) protein from the GST affinity column.

obtained in the case of using the expression plasmid pGEX-M$^{pro}$. The purified 3CL$^{pro}$($\Delta_{1-7}$) was dialyzed and lyophilized and stored at –80 ℃.

The 3CL$^{pro}$ was obtained by the same way as 3CL$^{pro}$ ($\Delta_{1-7}$). The purified 3CL$^{pro}$ did not contain extra 5 amino acid residues as obtained in the case of using the expression plasmid pGEX-M$^{pro}$. SDS-PAGE revealed that one major band of recombinant 3CL$^{pro}$ protein was in the unbond fraction, Data not show.

### 2.2 The enzyme activity of 3CL protease and its mutant

The fluorogenic substrate was used to measure the protease activity of the recombinant 3CL$^{pro}$. The increased fluorescence was monitored as shown in Fig.4. In contrast, under the same conditions, there was no increased fluorescence for the 3CL$^{pro}$($\Delta_{1-7}$), indicating that 3CL$^{pro}$ ($\Delta_{1-7}$) did not have any enzyme activity.

### 2.3 N-finger peptide partially inhibitory activity on SARS 3CL protease

To determine the inhibitory activity of the N-finger peptide, the concentration of the N-finger peptide was converted to molar units. This assay was completed in a short period of time (120 s) using fluorogenic substrate Dabcyl-KTSAVLQSGFRKME-Edans on a fluorescence spectrophotometer.

Fig.5 shows that N-finger peptide exhibits partially competitive inhibitory effect on SARS 3CL protease by continuous fluorescence-based assay.

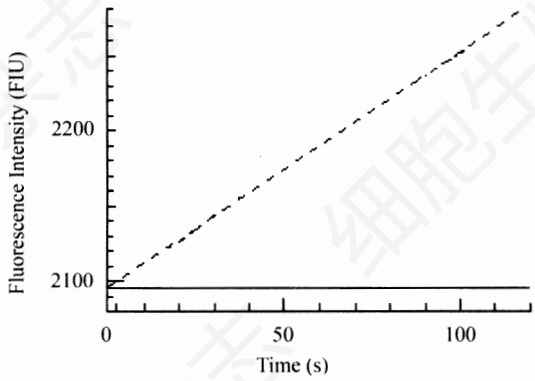


**Fig.4 The enzyme activity of 3CL$^{pro}$ and 3CL$^{pro}$($\Delta_{1-7}$) measured on fluorogenic substrate**

The enzymatic reaction was carried out at 37 ℃ in 20 mmol/L Bis-tris (pH 7.0) monitored on fluorescence spectrophotometer. Both the final concentration of 3CLpro and 3CLpro($\Delta_{1-7}$) were 34 nmol/L, and the concentration of fluorogenic substrate was 480 nmol/L. EM: 538 nmol/L, EX: 355 nmol/L. The dotted and black lines indicate the activities of 3CLpro and 3CLpro($\Delta_{1-7}$), respectively.

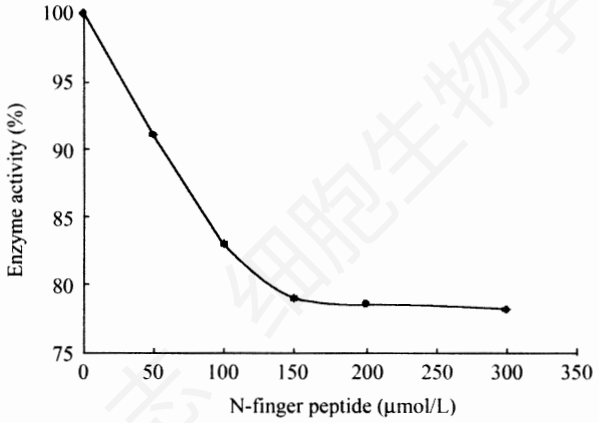


**Fig.5 Inhibitory action of N-finger peptide on 3CL$^{pro}$ activity**

The inhibitory effect of N-finger was measured at 37 ℃ in 20 mmol/L Bis-tris (pH 7.0), 2 mmol/L DDT and monitored on fluorescence spectrophotometer. The final concentrations of 3CL$^{pro}$ and fluorogenic subsrate were 34 nmol/L and 480 nmol/L, respectively. The concentrations of N-finger peptide ranged from 50 μmol/L to 300 μmol/L. The 3CL$^{pro}$ was incubated for 15 min at 37 ℃ with the N-finger peptide prior to adding substrate to the desired final concentration. EM: 538 nm, EX: 355 nm.

## 3 Discussion

The structure of the 3CL$^{pro}$ indicated that the active enzyme is in the form of dimer. The interaction between the domains III of two promoters stabilizes the formation of dimer[11]. The residues 1–7 of one monomer insert between the domain II and III of the other[10] and thus it may be critically involved in stabilizing the conformation of the enzyme active site. As expected, mutant

deleted the N-finger peptide was not active under the same conditions as the wild type did. The result is similar to that the mutant M$^{pro}$($\Delta_{1-7}$) of TGEV Mpro show only 0.3% activity against the native[12]. The result implies that the N-finger fragment is an important element of 3CL$^{pro}$. We also synthesized the N-finger peptide as above, and the hepta-peptide show partially competitive inhibitory action on the 3CL$^{pro}$. The free hepta-peptide, due to its small size, may easily get access to the N-terminal domain of the 3CL$^{pro}$, competing with the N-terminal 1-7 residues in its native structure. Since the competitor exists in large quantity, the dynamic balance between the competitors and the native amino acid residues would greatly favor the hepta-peptide. The free hepta-peptide interact with this N-terminal pocket through hydrogen bond and/or salt bond in similar way as the native residues and thus disturbs the conformation of the active site. The hepta-peptide may be effective drug against SARS virus after being optimized.

**Acknowledgements** We thank Prof. Zhi-He Rao for providing plasmid pGEX-Mpro and Prof. Po-Huang Liang for providing fluorogenic substrate. We thank Prof. Cheng-Wu Chi, Da-Fu Cui, You-Shang Zhang for discussion and advice.

## References

[1] Peiris JS *et al. N Engl J Med*, 2003, **349**: 2431
[2] Zhao Z *et al. J Med Microbiol*, 2003, **52**: 715
[3] Ksiazek TG *et al. N Engl J Med*, 2003, **348**: 1953
[4] Rota PA *et al. Science*, 2003, **300**: 1394
[5] Marra MA *et al. Science*, 2003, **300**: 1399
[6] Guan Y *et al. Science*, 2003, **302**: 276
[7] Lai MM *et al. Adv Virus Res*, 1997, **48**: 1
[8] Thiel V *et al. J Gen Virol*, 2003, **84**: 2305
[9] Ziebuhr J *et al. J Gen Virol*, 2000, **81**: 853
[10] Yang H *et al. Proc Natl Acad Sci USA*, 2003, **100**: 13190
[11] Shi J *et al. J Biol Chem*, 2004, **279**: 24765
[12] Anand K *et al. EMBO J*, 2002, **21**: 3213
[13] Wu CY *et al. Proc Natl Acad Sci USA*, 2004, **101**: 10012
[14] Kuo CJ *et al. Biochem Biophys Res Commun*, 2004, **318**: 862 (*Erratum in: Biochem Biophys Res Commum*, 2004, **320**: 623)
[15] King DS *et al. Int J Pept Protein Res*, 1990, **36**: 255

# SARS 冠状病毒蛋白酶 3CL$^{pro}$ N-finger 多肽是酶活性所必需

张笋华[1,2]* 姬明放[3] 邵晓霞[2]
([1]华东师范大学生命科学学院生物系，上海 200062；[2]同济大学蛋白质研究所，上海 200092；
[3]神隆(昆山)生化科技有限公司，昆山 215300)

**摘要** 构建了 SARS 冠状病毒主要蛋白酶 3CL$^{pro}$ 及其缺失 N 端七肽即 N-finger 的突变体 3CL$^{pro}$($\Delta_{1-7}$)融合表达质粒，并于大肠杆菌表达系统中表达。得到的融合蛋白经肠激酶酶切，亲和层析，最终得到纯化的 3CL$^{pro}$ 及其突变体 3CL$^{pro}$($\Delta_{1-7}$)。酶活性实验显示，去除 N-finger 多肽的突变体 3CL$^{pro}$($\Delta_{1-7}$)丧失了 3CL$^{pro}$ 所具有的对含有其自动水解位点的荧光底物的水解活性，表明处于非酶活性中心的 N-finger 多肽是酶活性所必需的。利用固相合成法，进一步合成了 N-finger 七肽。酶抑制实验表明 N-finger 七肽表现了对 3CL$^{pro}$ 部分竞争抑制活性。

**关键词** SARS 冠状病毒；3CL$^{pro}$ 蛋白酶；N-finger 多肽

收稿日期：2005-04-09 接受日期：2005-07-14
国家重点基础研究发展规划项目(973 计划) (No.2003CB514107)
* 通讯作者。Tel: 021-65988805, Fax: 021-65988403, E-mail: zhangsh0108@126.com