

# Cloning and sequence analysis of preprochymosin cDNA of Xinong Saanen Milk Goat

Baojin Yang<sup>1,2,\*</sup>, Xueqing Wang<sup>1</sup>, Jun Luo<sup>1</sup>, Guifen Tang<sup>2</sup>, Wengang Li<sup>2</sup>, Lijuan Zhang<sup>1</sup>

<sup>1</sup>Northwest Agriculture and Forestry University, Yangling, Shaanxi 712100, China; <sup>2</sup>Zhengzhou College of Animal Husbandry Engineering, Zhengzhou, Henan 450011, China

Received May 14, 2007

## Abstract

A cDNA coding for Saanen goat preprochymosin was synthesized from total RNA of two abomasums of suckling-goat by RT-PCR with two oligonucleotide probes corresponding to the 5'- and 3'- ends of the coding region for *Ovis aries* (Spain goat) preprochymosin gene. The objective gene was ligated to pMD18-T Vectors and electric-transformed to *E.coli* ElectroMAX DH18B. The sequencing result indicated that the cDNA was a gene coded for chymosin B and had 1292 nucleotides which coded 381 amino acid residues. The sequence of amino acid included a predicted signal peptide region (16 amino acids) and a prochymosin region (42 amino acids). Comparison of the Xinong Saanen Goat preprochymosin sequence with the other sequences of the *Capra hircus* (Spain goat), *Ovis aries* (sheep) and *Bos taurus* (cattle) revealed 99.44%, 98.76%, 95.12% and 99.21%, 98.42%, 93.7% identity at the nucleotide and amino acid level. [Life Science Journal. 2007; 4(3): 64 – 69] (ISSN: 1097 – 8135).

**Keywords:** Xinong Saanen Goat; chymosin; cloning; RT-PCR

## 1 Introduction

Chymosin (rennin; EC 3.4.23.4) belongs to the aspartic proteinase family which is widely distributed in many organisms and tissues with different physiological and functional properties. It is used extensively in cheese production because it cleaves casein in a specific manner, at the Phe<sup>105</sup> – Met<sup>106</sup> bond<sup>[1,2]</sup>. This enzyme possesses a very high milk clotting activity with a low proteolytic activity and this makes chymosin particularly suitable for the manufacture of cheese<sup>[3,4]</sup>. The traditional source of chymosin for cheese-making is calf rennet. But the increasing world production of cheese, coupled with a decline in the number of slaughtered calves, has stimulated a search for alternative sources of chymosin. Several rennet substitutes have been developed including bovine rennet from adult cows, fungus proteinases and other proteolytic enzymes. However, they have a much greater level of non-specific proteolytic activity and higher thermostability which more completely degrade the milk proteins to peptides,

leading to a reduction in yield and poor flavour development in some types of cheese. Consequently, numerous other methods to produce chymosin are attempted.

Chymosin produced using recombinant DNA technology represents one of the first successful applications of recombinant DNA technology in the food industry. The first commercial application of agricultural biotechnology approved by the FDA in 1990 was the development of fermentation-derived chymosin<sup>[5]</sup>. A number of companies are now producing recombinant chymosin for commercial use in cheese manufacture in the world<sup>[6]</sup>. There is only a few papers which report the study of recombinant calf chymosin<sup>[7,8]</sup> but no report on the molecular cloning and expression of goat prochymosin cDNA in China. Because of increasing production of goat milk in China in recent years, more and more milk production plants want to produce goat milk cheese to satisfy the increasing demands of people. Recombinant goat chymosin has the characteristics compatible with the natural goat chymosin and is the best source of milk-clotting factor for goat milk cheese production. So the development of recombinant goat chymosin has important significance in goat cheese-making in China.

\*Corresponding author. Tel: 86-371-65775160; Email: baojiny@sina.com

In this study we attempt to clone the goat preprochymosin gene using RT-PCR strategy and lay a foundation in chymosin gene expression in future.

## 2 Materials and Methods

### 2.1 Materials

Two suckling kid-goats (5 days old) from Xinong Saanen Goat Plant of North-west Agriculture and Forest University were selected. Two goat abomasum tissues were collected in the lab, washed by DECP water, wrapped in the sterilized silver paper, and then frozen in liquid nitrogen quickly.

SuperScript™ Plasmid System with Gateway Technology for cDNA Synthesis and Cloning kit, Trizol RNA isolation kit and ElectroMAX DH18B Cells were purchased from Invitrogen Co. USA. TaKaRa Ex taq, pMD18-T Vector, Agarose Gel DNA Purification Kit Ver, T4 DNA ligase, restriction endonuclease (*Nde* I, *Eco*R I, *Stu* I, *Pst* I and *Sca* I) and  $\lambda$ -*Eco*T14 I digest were provided by TaKaRa Co. Dalian, China. GoldView™ nucleic acid stain was provided by Shanghai Saibaisheng Gene Co. China. Sanbo plasmid isolation kit was supplied by Beijing Sanbo Co. China.

### 2.2 Primers design

Oligonucleotide primers were designed by the Primer5.0 software based on the published corresponding sequence of *Ovis aries* (sheep) preprochymosin cDNA reported in GeneBank (accession No. X53037). The sequence of primer with a sense restriction endonuclease site of *Nde* I was: 5'-GGAATTCATATGCCAGATC-CAAGATGAG-3'. The sequence of antisense primer with a restriction endonuclease site of *Eco*R I was: 5'-CC-GGAATTCGCAGAAAGACAACATTTA-3'. In order to identify the amplified products right or not, one pair of inner primers was designed too. One inner primer was 5'-GCAAGCCCCTGTCTATCCGCTATGG-3', another was 5'-TGAGAATCATCTGTCTGGAAAC-3'. All primers were synthesized in TaKaRa Co. Dalian, China.

### 2.3 Total RNA isolation and RT-PCR

Total RNA was isolated from the frozen abomasums of two suckling kid-goats using the Trizol RNA isolation kit. The first strand cDNA was synthesized, antisense primer as primer, using SuperScript™ Plasmid System with Gateway Technology for cDNA Synthesis and Cloning kit following the manufacturer's recommended protocol. The second strand cDNA was synthesized by PCR using the RT-PCR product as template and antisense and sense primer as primer. PCR was performed in a 25  $\mu$ l reaction

mixture in an automatic thermal cycle. The PCR mixture contained 2.5  $\mu$ l of 10  $\times$  EX-taq buffer, 1  $\mu$ l of RT-PCR product, 1  $\mu$ l of dNTP (10 mM), 1  $\mu$ l of 0.25  $\mu$ mol/L oligonucleotide sense primers, 1  $\mu$ l of 0.25  $\mu$ mol/L oligonucleotide antisense primers, 18  $\mu$ l of MilliQ H<sub>2</sub>O, 0.5  $\mu$ l Taq DNA polymerase. The parameters for PCR were as follows: 95 °C for 5 minutes, 1 cycle; 94 °C for 30 seconds, 60 °C for 45 seconds, 72 °C for 1 minute, 35 cycles; 72 °C for 10 minutes, 1 cycle. The amplified products (5  $\mu$ l) were observed by electrophoresis on 10 g/L agarose gel containing 5  $\mu$ l/100 ml of GoldView™ nucleic acid stain in TBE buffer. The PCR product was visualized under UV light and photographed.

Using the last PCR product as template, PCR was performed once again but using inner primers. The amplified product of inner primers was electrophoresed and photographed.

According to restriction endonuclease sites of *Ovis aries* preprochymosin cDNA reported in GeneBank, the objective gene was digested at 37 °C/2 h by *Stu* I, *Pst* I and *Sca* I respectively. The digestive products were electrophoresed and photographed too.

### 2.4 Construction of recombinant plasmids and sequence determination

PCR products of normal primers and pMD18-T Vectors were digested by *Nde* I and *Eco*R I respectively according to the protocol and then ligated by T4 DNA ligase at 16 °C/16 h. The recombinant plasmid was electro-transformed into *E. coli* ElectroMAX DH18B Cells in BIO-RAD MicroPulser (BIO-RAD Co. USA) at 16.6 kv/cm and 4 ms. The *E. coli* containing the recombinant plasmid was cultured in SOC culture medium for 1 h, inoculated into anti-AMP LB plate and then cultivated 16 h. The single white colony was selected, inoculated into LB medium containing 100 mg/L of ampicillin and cultured at 200 r/min at 37 °C overnight.

In order to judge the result right or not, the recombinant plasmid was isolated using plasmid isolation kit, digested by *Nde* I and *Eco*R I. The products were electrophoresed and photographed again.

The sequence determination of objective fragment was performed by Dalian TaKaRa Co. China.

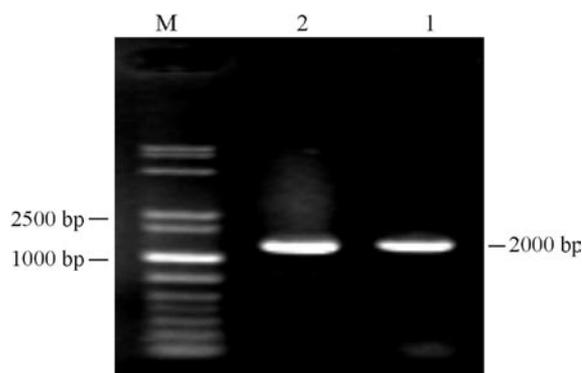
In order to judge the result right or not, the recombinant plasmid was isolated using plasmid isolation kit, digested by *Nde* I and *Eco*R I. The products were electrophoresed and photographed again.

The sequence determination of objective fragment was performed by Dalian TaKaRa Co. China.

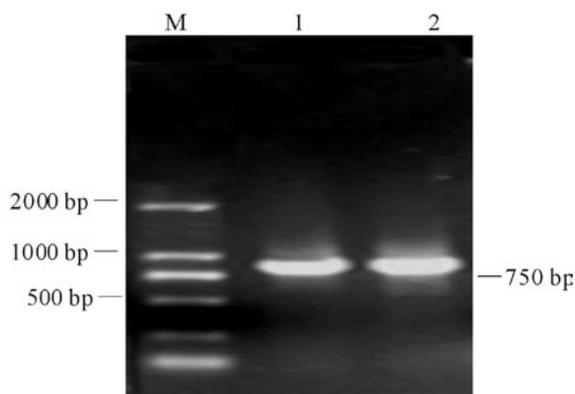
## 3 Results

### 3.1 The result of PCR amplification

The electrophoretic picture of objective gene amplified by RT-PCR from the above normal primers (deduced from lamb preprochymosin sequences, accession No. X53037) was observed in Figure 1. The target was approximately 1.3 kb. The PCR product of inner primers was photographed in Figure 2. The length of fragment was about 870 bp. The picture of all digested products of *Stu* I, *Pst* I and *Sca* I were showed in Figure 3. The length of fragment digested by *Stu* I were about 800 and 500 bp, by *Pst* I about 540, 350, 170 bp and less, and by *Sca* I about 750, 550 and 150 bp. All the results of this three pictures revealed that the size of objective cDNA fragment amplified by PCR was compatible with the size expected.



**Figure 1.** The result of preprochymosin gene using RT-PCR. Lane M: DNA Marker DL15,000; Lane 1 and Lane 2: the products of objective gene by RT-PCR.

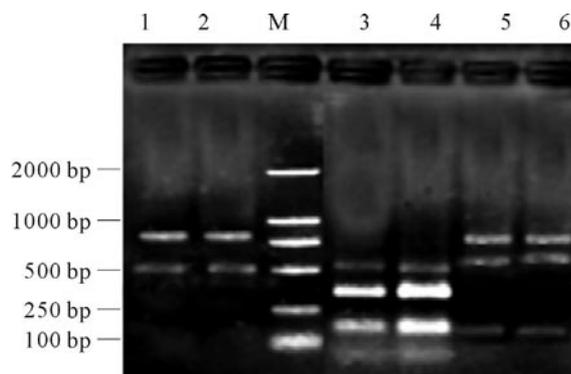


**Figure 2.** Identification of objective gene by PCR using inner primer. Lane M: DNA Marker DL 2000; Lane 1 and Lane 2: amplified products of inner primer.

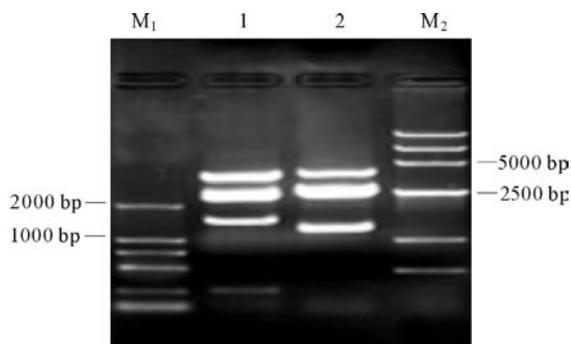
### 3.2 Identification of recombinant plasmids

Recombinant plasmid was isolated from transformed *E. coli*, digested by *Nde* I and *Eco*R I. The digestive product was visualized on 10 g/L agarose gel and showed in

Figure 4. The recombinant plasmid (about 4 kb) was digested to about 2.7 kb and 1.3 kb DNA fragment which was the objective gene. It demonstrated that the recombinant plasmid containing objective gene had been constructed successfully. The target gene could be amplified also from the recombinant plasmid by PCR.



**Figure 3.** The digestive result of restriction endonuclease. Lane M: DNA Marker DL2000; Lane 1 and Lane 2: the products digested by *Stu* I; Lane 3 and Lane 4: the products digested by *Pst* I; Lane 5 and Lane 6: the products digested by *Sca* I.



**Figure 4.** Identification of recombinant plasmid digested by *Nde* I and *Eco*R I. Lane M<sub>1</sub>: DL2000; Lane M<sub>2</sub>: DL15,000; Lane 1 and Lane 2: the product of recombinant plasmid digested by *Nde* I and *Eco*R I.

### 3.3 Sequence analysis

Sequencing result and deduced amino acid sequence were showed in Figure 5. The Xinong Sannen goat (*Capra hircus*) preprochymosin gene consists of 1292 base pairs. The sequencing result of Sannen goat preprochymosin was published in the GenBank (the accession number: EF199763).

It has been proved that there are two allelic forms of calf chymosin designated as chymosins A and B, respectively, differing by a single amino acid substitution (Asp/Gly) at position 243<sup>[9]</sup>. There was a glycine residue at position 243. Therefore, the Xinong goat prochymosin described

here should be the B form.

Comparison with other published aspartic proteinase (porcine pepsin A) sequence<sup>[10]</sup> indicated that an open reading frame beginning at the ATG codon predicted a protein of 381 amino acids residues with a molecular mass of 42.1 kD. This protein had an N-terminal leader sequence and a proenzyme region of 16 and 42 amino acids, respectively. Using the computer program Signal

P3.0<sup>[11]</sup>, a cleavage site for the removal of the signal sequence was predicted to occur between Gly<sup>-1</sup> and Ala<sup>1p</sup>. There were 323 amino acids in enzyme, including six cysteine residues which were very important for enzyme configuration stability because they could form three disulfide bridges at position 45 – 50, 206 – 210, 249 – 282. In addition, the two aspartate residues, Asp<sup>32</sup> and Asp<sup>215</sup> that function as catalytic residues in aspartic proteinases<sup>[12]</sup>

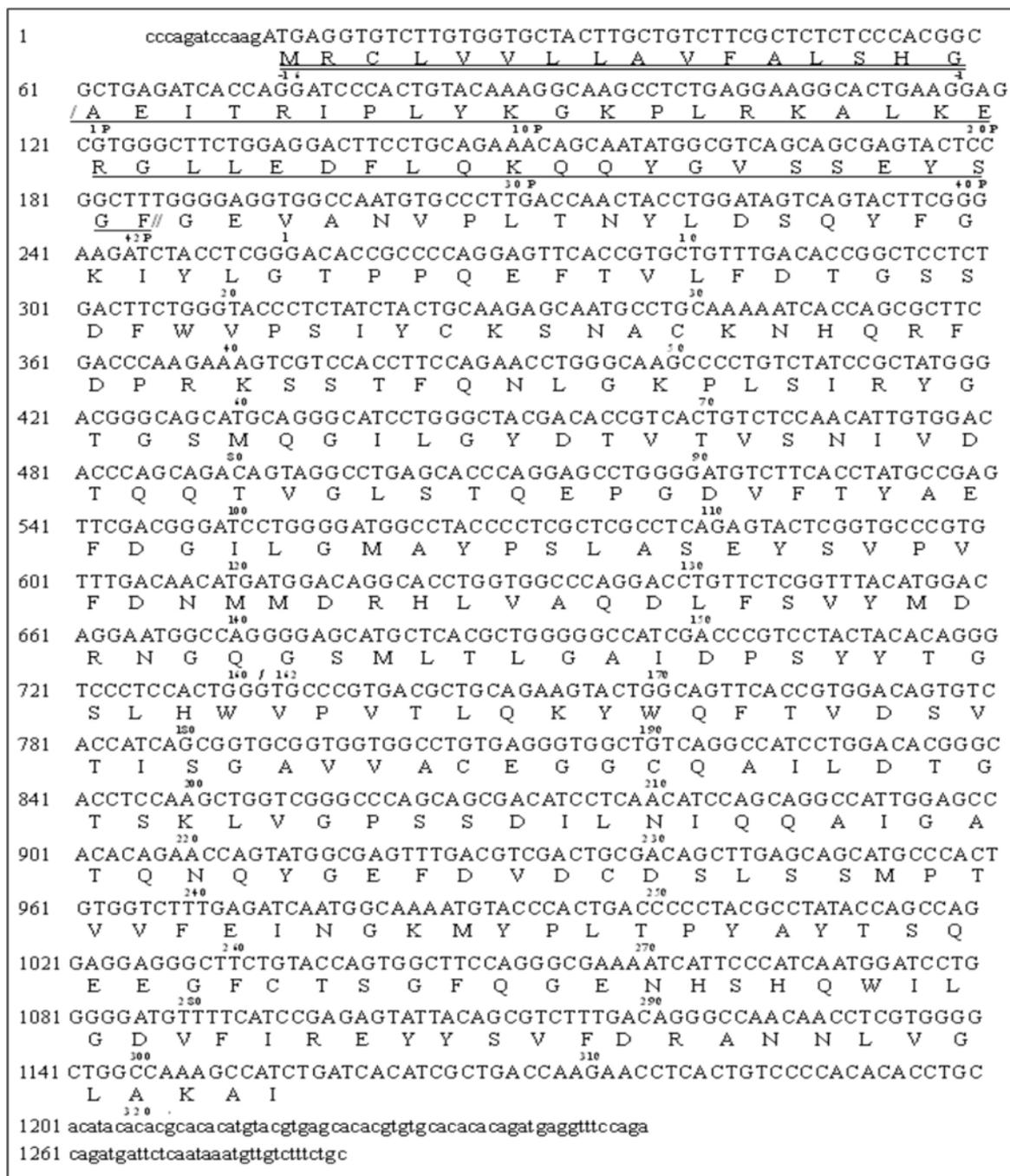


Figure 5. The sequence of nucleotide and putative amino acids of Xinong Saanen Goat preprochymosin.

were present in Saanen goat chymosin.

### 3.4 Homology analysis

NCBI BLASTn analysis indicated that the homologies of Xinong Saanen Goat (*Capra hircus*) preprochymosin nucleotide sequences compared with other published *Capra hircus* (Spain goat), *Ovis aries* (sheep) and *Bos taurus* (cattle) preprochymosin gene sequences were 99.44, 98.76 and 95.12% respectively. The objective gene had 7, 16 and 63 nucleic acid site differences with other three sequence.

Comparison of the deduced amino acid sequence of Saanen goat preprochymosin with its counterparts from

other mammals by NCBI BLASTp revealed a high sequence identity (Figure 6). The homologies of Xinong Saanen goat (*Capra hircus*) preprochymosin amino acid compared with that of other the *Capra hircus* (Spain goat), *Ovis aries* (sheep) and *Bos taurus* (cattle)'s were 99.21%, 98.42% and 93.70% respectively. The Xinong Saanen goat and Spain goat polypeptides differed in only three amino acid residues in the chymosin segment; goat and sheep polypeptides differed in six residues: one in the signal peptide and other five in the enzyme region; while goat and calf polypeptides differed in 24 residues: one in the signal peptide, five in the prochymosin region and 18 in the chymosin region.

B:	MRCLVLLAVFALSQGEITRIPLYKKGKSLRKALKEHGLLEDFLQKQQYGISSKYSYSGFGE	60
S:	MRCLVLLAVFALSQGAETRIPLYKKGKPLRKALKERGLLEDFLQKQQYGVSSSEYSGFGE	60
G:	MRCLVLLAVFALSHGAEITRIPLYKKGKPLRKALKERGLLEDFLQKQQYGVSSSEYSGFGE	60
X:	MRCLVLLAVFALSHGAEITRIPLYKKGKPLRKALKERGLLEDFLQKQQYGVSSSEYSGFGE	60
B:	VA <u>S</u> VPLTNYLDSQYFGKIYLGTPPQEFTVLFDTGSSDFWVPSIYCKSNACKNHQRFDPRK	120
S:	VA <u>S</u> VPLTNYLDSQYFGKIYLGTPPQEFTVLFDTGSSDFWVPSIYCKSNACKNHQRFDPRK	120
G:	VA <u>S</u> VPLTNYLDSQYFGKIYLGTPPQEFTVLFDTGSSDFWVPSIYCKSNACKNHQRFDPRK	120
X:	VANVPLTNYLDSQYFGKIYLGTPPQEFTVLFDTGSSDFWVPSIYCKSNACKNHQRFDPRK	120
B:	SSTFQNLGKPLSI <u>H</u> YGTGSMQGILGYDTVTVSNIVDIQQTVGLSTQEPGDVFTYAEFDGI	180
S:	SSTFQNLGKPLSIRYGTGSMQGILGYDTVTVSNIVDIQQTVGLSTQEPGDVFTYAEFDGI	180
G:	SSTFQNLGKPLSIRYGTGSMQGILGYDTVTVSNIVDTQQTVGLSTQEPGDVFTYAEFDGI	180
X:	SSTFQNLGKPLSIRYGTGSMQGILGYDTVTVSNIVDTQQTVGLSTQEPGDVFTYAEFDGI	180
B:	LGMAYPSLASEYSI <u>P</u> VFDNMMN <u>R</u> HLVAQDLFSVYMDRNGQ <u>E</u> SMLTLGAIDPSYYTGS <sub>L</sub> HW	240
S:	LGMAYPSLASEYSVPVFDNMMDR <u>R</u> LVAQDLFSVYMDR <u>S</u> GQGSMLTLGAIDPSYYTGS <sub>L</sub> HW	240
G:	LGMAYPSLASEYSVPVFDNMMDR <u>R</u> LVAQDLFSVYMDRNGQGSMLTLGAIDPSYYTGS <sub>L</sub> HW	240
X:	LGMAYPSLASEYSVPVFDNMMDRHLVAQDLFSVYMDRNGQGSMLTLGAIDPSYYTGS <sub>L</sub> HW	240
B:	VPVT <u>Y</u> QQYWQFTVDSVTISG <u>V</u> VVACEGGCQAILDTGTSKLVGPSSDILNIQQAIGATQNQ	300
S:	VPVT <u>L</u> QKYWQFTVDSVTISGAVVACEGGCQAILDTGTSKLVGPSSDILNIQQAIGATQNQ	300
G:	VPVT <u>L</u> QKYWQFTVDSVTISGAVVACEGGCQAILDTGTSKLVGPSSDILNIQQAIGATQNQ	300
X:	VPVT <u>L</u> QKYWQFTVDSVTISGAVVACEGGCQAILDTGTSKLVGPSSDILNIQQAIGATQNQ	300
B:	YGEFD <u>I</u> DCD <u>N</u> LSYMPVVFEINGKMYPLTP <u>S</u> AYTSQ <u>D</u> OGFCTSGFQ <u>S</u> ENHS <u>Q</u> KWILGDVF	360
S:	YGEFD <u>I</u> DCDSLSSMPTVVFEINGKMYPLTPYAYTSQEEGFCTSGFQGENHSHQWILGDVF	360
G:	YGEFD <u>I</u> DCDSLSSMPTVVFEINGKMYPLTPYAYTSQEEGFCTSGFQGENHSHQWILGDVF	360
X:	YGEFDVDCDSLSSMPTVVFEINGKMYPLTPYAYTSQEEGFCTSGFQGENHSHQWILGDVF	360
B:	IREYYSVFD <u>R</u> ANNLVGLAKAI	381
S:	IREYYSVFD <u>R</u> ANNLVGLAKAI	381
G:	IREYYSVFD <u>R</u> ANNLVGLAKAI	381
X:	IREYYSVFD <u>R</u> ANNLVGLAKAI	381

**Figure 6.** Comparison of amino acids of different animal preprochymosin. B: bovine; S: sheep; G: Spain goat; X: Xinong Saanen goat.

These analysis indicated that the Saanen goat preprochymosin gene sequence and its putative amino acid sequence were quite conservative and might be a potential enzyme candidate for milk-clotting enzyme development.

## 4 Discussion

Chymosin is synthesized *in vivo* as preprochymosin. It is characterized as a protein of 365 amino acids. The 16 amino acids hydrophobic leader presequence<sup>[13]</sup> is a signal sequence which is important in secretion of chymosin across the cell membranes. This is followed by a 42 amino acid pro sequence. It has long been known that chymosin is secreted as an inactive zymogen called prochymosin<sup>[3]</sup>, having a molecular weight of 40,777 Da whose inactive state is maintained by the N-terminal propeptide<sup>[9]</sup>. At acidic pH, the precursor undergoes autocatalytic activation to chymosin (35,600 Da molecular weight, 323 amino acids, observed at pH around 5.0) or pseudo-chymosin (337 amino acids, observed at pH around 2). Both chymosin and pseudo-chymosin show milk clotting activity<sup>[14]</sup>.

In this study, the Xinong Saanen Goat preprochymosin cDNA was cloned using an RT-PCR strategy with total RNA from the abomasums of two suckling kid-goats. The cDNA is predicted to code for a protein of 381 amino acids with an N-terminal leader sequence and a proenzyme region of 16 and 42 amino acids, respectively. The disulfide bond positions and enzyme catalytic residues have no change. The homologies of the nucleotide and putative amino acid sequences of the Saanen preprochymosin gene compared with the other published preprochymosin gene sequences were as high as 99.44% – 95.12% and

99.21% – 93.7%, respectively. These data indicate that the mutation level of Saanen goat preprochymosin gene is within the range reported by GenBank, and suggest that the Saanen goat preprochymosin gene is excellent and has a potentiality to develop recombinant chymosin.

## References

1. Bingham EW. Action of Remin on K-casein. Dairy Sci J 1975; 58: 13 – 8.
2. Braun I, Kunath H. Investigations on biochemical properties of milk-clotting enzymes. Die Nahrung J 1988; 32(4): 375 – 81.
3. Foltman B. A review on prorennin and rennin. Compt Rend Trav Lab Carlsberg 1966; 35: 143 – 231.
4. Beppu T. The cloning and expression of chymosin (rennin) genes in microorganisms. Trends Biotechnol 1983; 1: 85 – 9.
5. Flamm EL. How FDA approved chymosin: a case history. Biotechnology (NY) 1991; 9(4): 349 – 51.
6. Mohanty AK, Mukhopadhyay UK, Grover S, et al. Bovine chymosin: production by rDNA technology and application in cheese manufacture. Biotechnol Adv 1999; 17 (2 – 3): 205 – 17.
7. Zhang Y, Zhou W, Liu N, et al. Expression of calf prochymosin gene in *Escherichia coli*. Biotechnol 1991; 7(3): 195 – 200.
8. Wang G, Liu NJ, Yang KY. Regulation of the expression of calf pro-chymosin gene in *Escherichia coli*. Chinese Journal of Biotechnology 1994; 10(2): 162 – 7.
9. Foltman B, Pedersen VB, Jacobsen H, et al. The complete amino acid sequence of prochymosin. J Proc Natl Acad Sci 1977; 74: 2312 – 24.
10. Tang J, Sepulveda P, Marciszyn JrJ, et al. Amino-acid sequence of porcine pepsin. Proc Natl Acad Sci USA 1973; 70: 3437 – 9.
11. Bendtsen JD, Nielsen H, von Heijne J, et al. Improved prediction of signal peptides: Signal P 3.0. J Mol Biol 2004; 340: 783 – 95.
12. Kageyama T. Pepsinogens, progastricsins, and prochymosins: structure, function, evolution, and development. Cell Mol Life Sci 2002; 59: 288 – 306.
13. Steiner DF, Quinn PS, Chan SJ, et al. Processing mechanisms in the biosynthesis of proteins. Ann NY Acad Sci 1980; 343: 1 – 16.
14. Pedersen VB, Christensen KA, Foltman B. Investigations on the activation of bovine prochymosin. Eur J Biochem 1979; 94: 573 – 80.