Gaziosmanpasa University
Graduate School of Natural and Applied Sciences
Journal of New Results in Science

ISSN: 1304-7981 **http://jnrs.gop.edu.tr**

Received: 24.08.2012
Accepted: 10.09.2012

Editors-in-Chief: Naim Çağman
Area Editor: Emel Turgut

Codon Optimization and Cloning of Bovine Chymosin Gene into pTOLT Expression Plasmid of *Escherichia coli*

Yakup ULUSU[1]* and İsa GÖKÇE[2]

Abstract

Chymosin is a very important industrial enzyme that commonly used in cheese manufacture. Bovine chymosin is an aspartic protease which is extracted from abomasum of suckling calves. In this study, bovine chymosin gene was first optimized and then cloned into pTOLT *E.coli* expression system for production of chymosin. Protein engineering of chymosin has also been attempted. A number of companies are now producing recombinant chymosin for commercial use in cheese manufacture.

Keywords: *pTOLT, codon optimization, chymosin*

1. Introduction

Chymosin (rennin, EC 3.4.23.4) is an aspartic protease that is produced in the abomasum of suckling calves [1]. Chymosin is used as the milk clotting agent rennet in cheese manufacture. 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 [2]. The coagulants used for cheese manufacturing come from those sources: microorganisms, plants, animals and recombinant protein expression. Milk coagulants from different sources have different characteristics, which determine their applications. Microbial milk-clotting enzymes come from the culture supernatant of microorganisms such as *Bacillus subtilis, Mucor miehei* [3]. However, the content of specific milk-clotting enzymes in the microbial rennet is low, and the use of these enzyme preparations in cheese production can result in the product having a bitter flavor and low yield of curd [4]. The coagulants from plants are used for making special cheeses in certain areas of the world. For example, coagulants from flowers (*Cynara scolymus*) or fig tree extract (*Ficus carica*) are used for the production of traditional cheeses in northern and southern Algeria [5]. Animal coagulants are extracted from the abomasums of unweaned ruminants such as cows, goats, pigs and sheep. Among the all milk-

[1] *Corresponding Author*, Department of Biology, University of Gaziosmanpasa, 60250 Tokat, Turkey. (e- mail: yakupulusu@yahoo.com)
[2] Department of Bioengineering, University of Gaziosmanpasa, 60250 Tokat, Turkey. (e- mail: isa_gokce@yahoo.co.uk)

* This article is partially produced from the Yakup ULUSU's PhD thesis.

clotting enzymes, calf chymosin is the most effective enzyme for cheesemaking process, because cheese made with this enzyme has a unique texture and flavor.

The traditional method for the production of calf chymosin requires slaughtering of large numbers of calves annually. Therefore, the cheese industry has been seeking novel enzyme sources as an alternative to calf rennet. Recombinant calf chymosin has been found to be an effective alternative since it gives several advantages over microbial and plant rennets.

Chymosin is synthesized *in vivo* as preprochymosin. It has 365 amino acids at first (Figure 1). The 16 amino acids hydrophobic leader pre-sequence is a signal sequence, which is important in secretion of chymosin across the cell membranes [6]. A 42 amino acid pro-sequence follows to do that. It has been known that chymosin is secreted as an inactive zymogen called prochymosin, having a molecular weight of 40,777 Da whose inactive state is maintained by the N-terminal propeptide [7]. 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 pseudochymosin (337 amino acids, observed at pH around 2). Both chymosin and pseudochymosin show milk clotting activity.

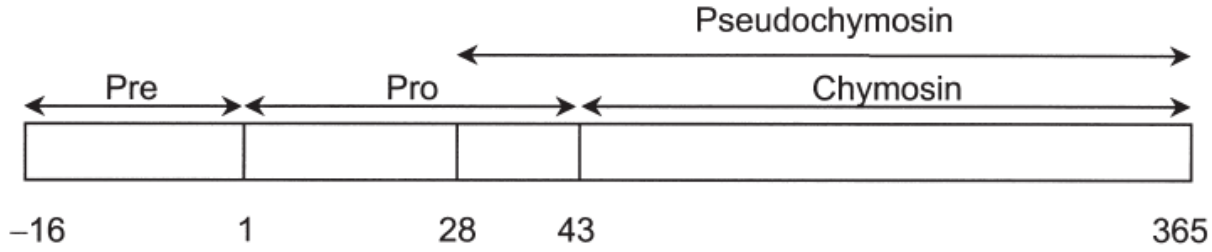

Figure 1. Chymosin protein structure [8]

This enzyme contains two aspartic acid residues at the active site, Asp32-Asp215, allowing to specifically break down the Phe105-Met106 peptide bond of milk *κ-casein* to form insoluble *para κ-casein* [8]. Coagulation is a two-stage process; the primary phase (proteolytic stage) involves the enzymatic production of paracasein (insoluble) and soluble glycomacropeptides (soluble). The secondary phase (coagulation stage) involves the precipitation or gelation of paracasein by Ca^{2++} at a temperature higher than 20 ^{0}C [9].

Enzymolysis of *κ-casein* fraction with chymosin at pH 6.7 in the default of calcium causes precipitation of para-κ-caseinate fraction. Action of chymosin is dependent on the addition of calcium in the mixture. When calcium is absent, the para-κ-casein fraction, which by itself is insoluble, interacts with calcium sensitive caseins to keep from precipitating. When calcium ions are present, the calcium sensitive caseinate fraction along with insoluble para-κ-casein fraction forms a clot [9, 10]. Chymosin acts to start milk clotting by cleavage of κ-casein between Phe 105 and Met 106 bond [11, 12]. This bond is much more sensitive to acid proteases than other peptide bonds in the milk protein system. The unique sensitivity of the Phe-Met bond has aroused considerable interest. Neither the di-peptide HPhe-MetOH, nor tri- or tetra-peptides containing a Phe-Met bond are hydrolyzed. However, this bond is hydrolyzed in the penta-peptide, HSer-Ser-Phe-Met-Ala-OH [9, 13, 14]. The length of the peptide and the sequence around the sectile bond are important determinants of enzyme–substrate interaction [9].

2. Codon Optimization of Bovine Chymosin Gene

Escherichia coli has been widely used as a host cell for expressing recombinant proteins because it has many attractive features, including well-characterized genetics, rapid growth and the availability of numerous vectors [15-16].

Not all 61 mRNA codons are used equally. The so-called major codons are those that occur in highly expressed proteins, whereas the minor or rare codons tend to be in genes expressed at a low level. Which of the codons are the rare ones depends strongly on the organism.

Usually, the frequency of the codon usage reflects the abundance of their cognate tRNAs. Therefore, when the codon usage of your target protein differs significantly from the average codon usage of the expression host, this could cause problems during expression. The following problems are often encountered:

- Decreased mRNA stability (by slowing down translation).
- Premature termination of transcription and/or translation, which leads to a variety of truncated protein products.
- Frameshifts, deletions and mis-incorporations (*e.g.* lysine for arginine).
- Inhibition of protein synthesis and cell growth.

As a consequence, the observed levels of expression are often too low or there will be no expression at all [17]. Especially in cases where rare codons are present at the 5'-end of the mRNA or in clusters expression levels are low and truncated protein products are found [18].

The expression levels of a recombinant protein can be improved by codon optimisation that is replacing rare codons with more favourable codons throughout the whole gene. Codon optimization is a useful technique to maximize the protein expression in host organism by increasing the translational efficiency of gene of interest by transforming DNA sequence of one species into DNA sequence of another species. Like human sequence to bacteria or yeast sequences, etc. [19]. In this process amino acid will be the same, but codon of low frequency of an amino acid will be replaced with codon of high frequency. For example, suppose in one species amino acid arginine codon CGG has low frequency and in desired species arginine codon CGC has high frequency, so CGG will be replaced with CGC [20].

3. Results and Discussion

3.1 Codon Optimization of Bovine Chymosin Gene for *E.coli* expression

In this work, rare codons in the DNA sequence of bovine chymosin gene were identified using codon usage frequency of *E. Coli* K12 organism, and these low frequency codons were replaced with higher frequency codons [20]. Rare codons for Methionine AGG, AGA, CGG, CGA were replaced with CGT and low frequency codons for Leucine CTA and also for Isoleucine ATA were replaced with CTG and AUU. DNA alignment for codon optimised and wild type chymosin gene were given below. The DNA alignment was taken from BLAST (Basic Local alignment tool) [21].

```
Optimz 1    ATGGGGGAGGTGGCCAGCGTGCCCCTGACCAACTACCTGGATAGTCAGTACTTTGGGAAG  60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Wild   1    ATGGGGGAGGTGGCCAGCGTGCCCCTGACCAACTACCTGGATAGTCAGTACTTTGGGAAG  60

Optimz 61   ATCTACCTCGGGACCCCGCCCCAGGAGTTCACCGTGCTGTTTGACACTGGCTCCTCTGAC  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Wild   61   ATCTACCTCGGGACCCCGCCCCAGGAGTTCACCGTGCTGTTTGACACTGGCTCCTCTGAC  120

Optimz 121  TTCTGGGTACCCTCTATCTACTGCAAGAGCAATGCCTGCAAAAACCACCAGCGCTTCGAC  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Wild   121  TTCTGGGTACCCTCTATCTACTGCAAGAGCAATGCCTGCAAAAACCACCAGCGCTTCGAC  180

Optimz 181  CCGCGTAAGTCGTCCACCTTCCAGAACCTGGGCAAGCCCCTGTCTATCCACTACGGGACA  240
            ||| | ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Wild   181  CCGAGAAAGTCGTCCACCTTCCAGAACCTGGGCAAGCCCCTGTCTATCCACTACGGGACA  240

Optimz 241  GGCAGCATGCAGGGCATCCTAGGCTATGACACCGTCACTGTCTCCAACATTGTGGACATC  300
            |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Wild   241  GGCAGCATGCAGGGCATCCTGGGCTATGACACCGTCACTGTCTCCAACATTGTGGACATC  300

Optimz 301  CAGCAGACAGTAGGCCTGAGCACCCAGGAGCCCGGGGACGTCTTCACCTATGCCGAATTC  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Wild   301  CAGCAGACAGTAGGCCTGAGCACCCAGGAGCCCGGGGACGTCTTCACCTATGCCGAATTC  360

Optimz 361  GACGGGATCCTGGGGATGGCCTACCCCTCGCTCGCCTCAGAGTACTCGATACCCGTGTTT  420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Wild   361  GACGGGATCCTGGGGATGGCCTACCCCTCGCTCGCCTCAGAGTACTCGATACCCGTGTTT  420

Optimz 421  GACAACATGATGAACCGTCACCTGGTGGCCCAAGACCTGTTCTCGGTTTACATGGACCGT  480
            ||||||||||||||| | ||||||||||||||||||||||||||||||||||||| |
Wild   421  GACAACATGATGAACAGGCACCTGGTGGCCCAAGACCTGTTCTCGGTTTACATGGACAGG  480

Optimz 481  AATGGCCAGGAGAGCATGCTCACGCTGGGGGCCATCAACCCGTCCTACTACACAGGGTCC  540
            |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Wild   481  AATGGCCAGGAGAGCATGCTCACGCTGGGGGCCATCGACCCGTCCTACTACACAGGGTCC  540

Optimz 541  CTGCACTGGGTGCCCGTGACAGTGCAGCAGTACTGGCAGTTCACTGTGGACAGTGTCACC  600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Wild   541  CTGCACTGGGTGCCCGTGACAGTGCAGCAGTACTGGCAGTTCACTGTGGACAGTGTCACC  600

Optimz 601  ATCAGCGGTGTGGTTGTGGCCTGTGAGGGTGGCTGTCAGGCCATCTTGGACACGGGCACC  660
            |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Wild   601  ATCAGCGGTGTGGTTGTGGCCTGTGAGGGTGGCTGTCAGGCCATCCTGGACACGGGCACC  660

Optimz 661  TCCAAGCTGGTCGGGCCCAGCAGCGACATCCTCAACATCCAGCAGGCCATTGGAGCCACA  720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Wild   661  TCCAAGCTGGTCGGGCCCAGCAGCGACATCCTCAACATCCAGCAGGCCATTGGAGCCACA  720

Optimz 721  CAGAACCAGTACGGTGAGTTTGACATCGACTGCGACAACCTGAGCTACATGCCCACTGTG  780
            |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Wild   721  CAGAACCAGTACGATGAGTTTGACATCGACTGCGACAACCTGAGCTACATGCCCACTGTG  780

Optimz 781  GTCTTTGAGATCAATGGCAAAATGTACCCACTGACCCCCTCCGCCTATACCAGCCAAGAC  840
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Wild   781  GTCTTTGAGATCAATGGCAAAATGTACCCACTGACCCCCTCCGCCTATACCAGCCAGGAC  840

Optimz 841  CAGGGCTTCTGTACCAGTGGCTTCCAGAGTGAAAATCATTCCCAGAAATGGATCCTGGGG  900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Wild   841  CAGGGCTTCTGTACCAGTGGCTTCCAGAGTGAAAATCATTCCCAGAAATGGATCCTGGGG  900

Optimz 901  GATGTTTTCATCCGTGAGTATTACAGCGTCTTTGACCGTGCCAACAACCTCGTGGGGCTG  960
            |||||||||||||| |||||||||||||||||||| | ||||||||||||||||||||||
Wild   901  GATGTTTTCATCCGAGAGTATTACAGCGTCTTTGACAGGGCCAACAACCTCGTGGGGCTG  960

Optimz 961  GCCAAAGCCATC  972
            ||||||||||||
Wild   961  GCCAAAGCCATC  972

Length=972 Score = 1718 bits (930),  Expect = 0.0
Identities = 958/972 (99%), Gaps = 0/972 (0%) Strand=Plus/Plus
```

Accession	Max score	Total score	Query coverage	E value	Max identity
24813	1718	1718	97%	0.0	99%

Table 1: Sequences producing significant alignments

3. 2. Cloning of Bovine Chymosin Gene into pTOLT Plasmid Vector

DNA fragment encoding for bovine chymosin was amplified using two oligonucleotides (Table 2). Template used for PCR reaction was pIDTSMART-ISAKMZN plasmid vector. Chymosin gene was introduced into pTOLT plasmid using *Not* I and *Mlu* I restriction sites. Final plasmid named as a pTOLTCHY (Figure 2). Restriction fragment analysis, analytical colony PCR (Figure 3) and DNA sequencing of this plasmid indicated that chymosin encoding DNA fragment was correctly inserted. This final construct was used initially to transform *E. coli* DH5α cells with ampiciline selection. Successful transformants were selected on the basis of mini prep restriction digest analysis or analytical PCR, and a subsequent plasmid preparation was used for DNA sequencing and to transform *E. coli* BL21 (DE3) cells.

Primers	Sequence
Chy NotI tolt sense	TTTTTTTTTGCGGCCGCATGGGGGAGGT **29 bp**
Chy MluI tolt reverse	TTTTTACGCGTTCAGATGGCTTTGGCCAG **30 bp**

Table 2. Oligonucleotides used for PCR amplification of cyhmosin gene fragment

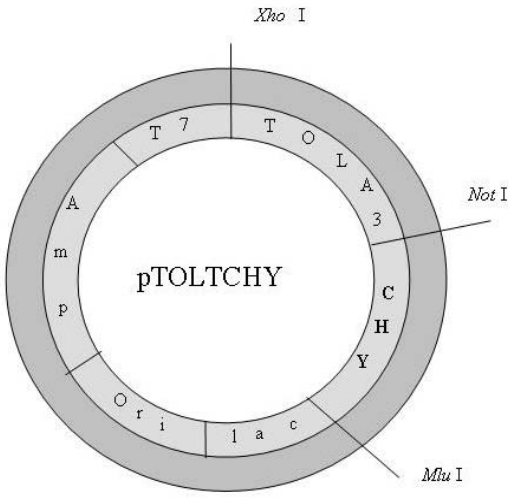

Figure 2. Circular plasmid map of the pTOLCHY construct to produce Tol-A-III and Chymosin fusion.

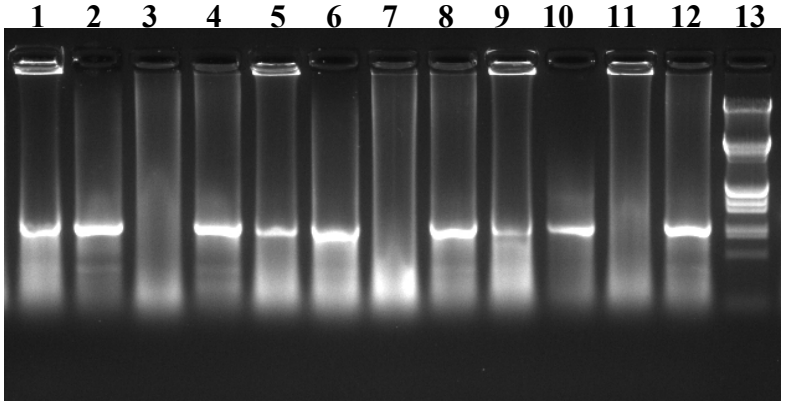

Figure 2. 1% Agarose gel demonstrating the analytical PCR reactions for the chymosin gene. Positive 972 bp product indicates that chymosin gene fragment correctly inserted to pTolT system. 1-12 PCR reactions, 13 λ-EcoR I /Hind III DNA marker. 9 out of 12 samples had right chymosin gene insert.

Sometimes the levels of protein expression in *E. coli* are too low despite the use of strong viral transcriptional and translational signals (T7 promoter); therefore, the rare codon optimisation of the chymosin gene approach was used in this work to optimize chymosin expression levels. We have replaced the codons that have been associated with translation problems in *E. coli* with more favourable codons throughout the whole chymosin gene.

We also successfully cloned the bovine chymosin gene into the pTOLT system using restriction enzymes and ligation processes. Constructed plasmid in this study is designed for bovine chymosin protein expression with the fusion gene under the control of a T7 promoter and an N-terminal 6 Histidine tag to facilitate protein purification using an affinity chromotagraphy column.

References

[1] A. Kumar, s. Grover, J. Sharma, V.K. Batish, Chymosin and other milk coagulants: Source and biotechnological interventions. Crit Rev Biotechnol, 2010, Volume:30, pp: 243-258.

[2] B., Foltman, A review on prorennin and rennin, Compt Rend Trav Lab Carlsberg, 1966, Volume: 35, pp: 143-231.

[3] X.P. Jang, M.L. Yin, P. Chen, Q. Yang, Constitutive expression, purification and characterization of bovine prochymosin in *Pichia pastoris* GS115. World J Microbiol Biotechnol, 2012, DOI:10.1007/s11274-012-1012-7.

[4] F.L. Davis, B.A. Law, Advances in the microbiology and biochemistry of cheese and fermented milk. 1984, Elsevier Applied Science, London.

[5] A. Nounai, E. Dako, A. Morsli, N. Belhamiche, S. Belbraouet, M.M. Bellal, A. Dadie, Characterization of the purified coagulant extracted derived from artichoke flowers (*Cynara scolymus*) and from the fig latex (*Ficus carica*) in light of their use in the manufacture of traditional cheeses in Algeria. 2009, Journal of Food Technol, Volume: 7, pp: 20-29.

[6] D.F. Steiner, P.S. Quinn, S.J. Chan, J. Marsh, H.S. Tager, Processing mechanisms in the biosynthesis of proteins. Ann NY Acad Sci, 1980, Volume: 343, pp: 1-16.

[7] B. Foltman, V.B. Pedersen, M. Jacobsen, D. Kauffman, G. Wijbrandt, The complete amino acid sequence of prochymosin. 1977, Volume: 74, pp: 2321-2314.

[8] A.K., Mohanty, U.K., Mukhopadhyay, S., Grover, V.K., Batish, Bovine chymosin: Production by rRNA technology and application in cheese manufacture. Biotechnology Advances, 1999, Volume: 17, pp: 205-217.

[9] P.F. Fox, Rennets and their action in cheese manufacture and ripening. Biotechnol Appl Biochem 1988, Volume: 10 pp: 522–35.

[10] M.P. Mathur, R.D. Dutta, Rennet from living calves. Dairy Guide 1983, Volume: 5(12) pp: 39–47.

[11] B. Foltman, General and molecular aspects of rennets. In: Fox PF, editor. Cheese: Chemistry, Physics and Microbiology, 1987, Volume: 1 pp: 37–68.

[12] N.M.C. Kaye, P. Jolles, The involvement of one of the three histidine residues of cow k-casein in the chymosin initiated milk clotting process. 1978, Biochim Biophys Acta, Volume:40, pp: 536:329.

[13] R.D. Hill, The nature of the rennin sensitive bond in casein and its possible relation to sensitive bonds in other proteins. 1968, Biochem Biophys Res Commun, Volume: 33 pp:659–63.

[14] R.D. Hill, Synthetic peptide and ester substrates for rennin. 1969. J Dairy Res, Volume: 36 pp: 409–15.

[15] S. Jana, J.K. Deb, Strategies for efficient production of heterologous protein in Escherichia coli. Appl Microbiol Biotechnol. 2005, Volume: 67 pp: 289–298.

[16] T. Makino, G. Skretas, G: Georgiou, Strain engineering for improved expression of recombinant proteins in bacteria. Microb Cell Fact. 2011, 10:32

[17] Looman AC, Bodlaender J, Comstock LJ, et al. Influence of the codon following the AUG initiation codon on the expression of a modified lacZ gene in Escherichia coli. EMBO Journal. 1987;6(8):2489–2492.

[18] (http://www.embl.de/pepcore/pepcore_services/protein_expression/ecoli/optimisation _expression_levels/).

[19] D. Baev, X.W. Lil, M. Edgerton, Genetically engineered human salivary histatin genes are functional in Candida albicans: development of a new system for studying histatin candidacidal activity, Microbiol-Sgm. 2001, Volume: 147, pp. 3323-3334.

[20] http://www.kazusa.or.jp/codon/cgi-bin/showcodon.cgi?species=83333, 13.08.2012.

[21] http://blast.ncbi.nlm.nih.gov, 13.08.2012.