ORIGINAL ARTICLE

Optimization of Col H Gene Encoding *Clostridium histolyticum* Collagenase to Express in *Escherichia coli*

Seyed Mohammad Amin Mahdian[1], PhD student; Abbasali Raz[2], PhD; Navid Dinparast Djadid[2], PhD; Hamzeh Alipour[3], PhD

[1]Department of Pharmaceutical nanotechnology, Faculty of pharmacy, Tehran University of Medical Sciences, Tehran, Iran
[2]Malaria and Vector Research Group [MVRG], Biotechnology Research Center [BRC], Pasteur Institute of Iran, Tehran, Iran
[3]Research Center for Health Sciences, Institute of Health, Department of Biology and Control of Disease Vectors, School of Health, Shiraz University of Medical Sciences, Shiraz, Iran

Correspondence:
Hamzeh Alipour, PhD;
Research Center for Health Sciences, Institute of Health, Department of Vector Biology and Control of Diseases, School of Health, Shiraz University of Medical Sciences, Shiraz, Iran
Tel: +98 71 37251001
Email: alipoorh@sums.ac.ir
Received: 30 October 2021
Revised: 11 November 2021
Accepted: 12 December 2021

Abstract

Background: A given amino acid sequence can be encoded by a huge number of different nucleic acid sequences. These sequences, however, have proved not to be equally useful. The choice of sequence can significantly impact the expression of an encoded protein. Given the importance of protein-coding sequence and promising industrial and medicinal applications of *Clostridium histolyticum* collagenase, this study examined the codon optimization of the Col *H* gene so as to enhance collagenase expression in *Escherichia coli* (*E. coli*).

Methods: This is an experimental study in which the CDS of Col H gene was optimized according to the codon usage of E. coli, using Gene Designer software (DNA 2.0).

Results: The results revealed that relative frequency of codon usage in Col H gene was adapted to the most preferred triplets in *E. coli* in such a way that codon usage bias in *E. coli* was enhanced after codon optimization. The higher level of collagenase expression was more likely the result of substituting rare codons with optimal codons.

Conclusion: The findings of this study suggest that codon optimization provides a theoretical improvement in Col H gene expression in *E. coli*. Nevertheless, experimental research is needed to confirm the improvement.

Please cite this article as: Mahdian SMA, Raz AA, Dinparast Djadid N, Alipour H. Optimization of Col H Gene Encoding *Clostridium histolyticum* Collagenase to Express in *Escherichia coli*. J Health Sci Surveillance Sys. 2022;10(1):126-133.

Keywords: Codon, *Clostridium histolyticum*, Recombinant proteins, *Escherichia coli*

Introduction

Collagens existing in a large number of cell types are the integral component of animal tissues such as the skin, tendons, and cartilage as well as the organic constituent of the bones, teeth, and cornea. In fact, it accounts for about 25 to 33% of the total protein in mammalian organs.[1] Collagen in the connective tissues of nearly all organs is as insoluble fibers. It is embeded in the mucopolysaccharides and protein of the extracellular matrix and plays a vital role in the strength of tissues.[2] A change in its production or degradation has been shown to result in a variety of diseases. In such cases, proteolytic enzymes provide a useful clinical way to the treatment of collagen-centered disorders.[3] On the other hand, making use of enzymes as a drug enjoys two advantages which differentiate them from all other types of drugs.[4-6] To start with, enzymes frequently bind and act on their substrates with a high affinity. In the second place, enzymes are catalytic and transform target molecules into certain products. Thus, they are more specific and potent drugs than small molecules that can accomplish therapeutic biochemistry in the body. These features have given rise to the development of many enzymatic drugs for a large variety of disorders.[5] A tight triple helical structure that makes up collagen causes its resistance to most proteases; all the same, collagenases can specifically degrade collagen.[6] Bacterial collagenases, in addition, have been shown to display broader substrate specificity than vertebrate

collagenases.[7] Collagenases derived from *Clostridium histolyticum*, namely Col G and Col H, are a case in point since these are capable of easily digesting collagens,[8, 9] no matter what their size and type are.[9, 10] Recently, making use of Clostridium collagenases has attracted a great deal of researchers' interest as a non-invasive therapeutic procedure.[3, 11] As such, these collagenases have been examined for the treatment of Dupuytren's disease,[3, 12, 13] Peyronie's Disease,[14, 15] Herniated Lumbar Disk,[3, 4, 16] retained placenta,[17, 20] wound healing,[21, 22] the debridement of burn,[23, 24] and in the preparation of pancreatic islet cells for transplantation.[25-27] Nowadays, culturing C. Histolyticum and subsequent purifying all of the produced bacterial proteins is a widely used method of producing collagenases for clinical applications.[28, 29] Nevertheless, the isolation and extraction of enzyme from natural sources due to the low expression levels and the intracellular localization of the enzyme, have technical problems and increases the costs of production. For commercial success, decreasing the cost of production is necessary which, in turn, depends on the expression level of the enzyme and the purification costs. As such, there has been growing interest in producing enzyme by recombinant methods.[30] Prokaryotic expression systems, especially E. coli, are one of the most common systems for the industrial production of proteins of therapeutic or commercial applications. E. coli has advantages including growth on inexpensive media, rapid biomass accumulation, convenient genetic manipulation, and simple scale-up.[31] Nonetheless, the production of heterologous protein in the organism may be decreased by codon bias phenomenon, in which proteins of interest contain codons that are rarely used in E. coli.[32, 33] It has revealed that the presence of rare codons results in decreasing translation speed and inducing translational errors.[34, 35] The rarest codons in E. coli include Arginine (AGG, AGA, CGG, CGA), GGG), Isoleucine (ATA) Leucine (CTA, TTA), Proline (CCC), Serine (TCG, TCA, AGT), and Therionine (ACA). In addition to rare codons, GC content can affect the expression levels. GC-rich mRNAs can contribute to forming powerful secondary structures and, especially in bacteria, such a powerful structure near ribosome-binding site obstructs the translation initiation. On the other hand, GC-poor mRNAs cannot fold strongly and frequently carry sequence elements limiting expression. For instance, low GC content has been seen to restrict the expression of Plasmodium falciparum genes in E. coli. Such mRNAs seem to be the targets for RNase E cleaving AU-rich sequence.[34-36] Codon optimization has been considered to be a common strategy to improve the efficiency and accuracy of translation.[37, 38] Codon optimization, indeed, is to alter rare codons in target gene so as to adapt them to the codon usage of specific expression host. Recently, a huge number of studies have reported an increase in the expression levels by codon optimization.[39, 40] Here, we explored codon optimization of Col H gene to express in E. coli.

Methods

Codon Optimization of Col H

The protein sequence of Col H was taken from UniProt database, and 40 amino acids of putative signal (MKRKCLSKRLMLAIT MATIFTVNSTLPIYAAVDKNNATAA) were removed in order to generate a mature enzyme. Then, the coding region of mature Col H gene was optimized according to the codon usage of E. coli, using Gene Designer software (DNA 2.0). This software using proprietary algorithms replaces rare codons and eliminates problematic mRNA structures and repetitive sequences.

Gene Sequence Analyses

The sequence analysis of native and optimized Col H gene was performed using online software involving Rare Codon Analysis Tool and Sequence Analysis which are available on websites www.genscript.com and www.bioinformatics.org, respectively.

Results

Codon oOptimization of Col H

The native gene used tandem rare codons such as AGA AGG which have been shown to greatly affect heterologous expression in E. coli. These effects include ribosome pausing and cotranslational cleavage of mRNA, ribosomal frame shifting or amino acid misincorporation.[32-34] Codon optimization substituted such codons and adapted the frequency of codon usage to the most preferred triplets in E. coli, so that its codon usage bias was enhanced. Moreover, the sequence of native (GenBank accession number D29981) and codon-optimized genes was aligned, indicating that codon optimization did not alter the amino acid sequence and that 666 out of 982 codons (67.82%) were substituted.

Percentage of Non-optimal and Optimal Codons Before and After Codon Optimization

The native and optimized codon sequences of Col H enzyme are compared using Sequence Analysis software. The gene coding collagenase contains 2946 bp which encodes a protein with 982 amino acids Before codon optimization, the numbers of rare codon GGA and GGG (codons encoding glycine) were 38 and 6, respectively. On the other hand, the codons GGC and GGT that are optimal for glycine were 3 and 26, respectively. After codon optimization, the rare codons reached 0, and the optimal codons GGC and GGT increased to 37 and 36, respectively. Moreover, the rare codons AGA and AGG (coding arginine) that were 23 and 5 fell to zero after optimization. In contrast, the codons CGT and CGC which are optimal for arginine increased from 2 and 1 to 20 and

Table 1: The results of the codon analysis of Collagenase H gene before and after optimization

Before codon optimization					After codon optimization				
Amino Acid	Codon	Number	/1000	Fraction	Amino Acid	Codon	Number	/1000	Fraction
Ala	GCG	1.00	1.02	0.02	Ala	GCG	14.00	14.26	0.33
Ala	GCA	22.00	22.40	0.52	Ala	GCA	13.00	13.24	0.31
Ala	GCT	17.00	17.31	0.40	Ala	GCT	8.00	8.15	0.19
Ala	GCC	2.00	2.04	0.05	Ala	GCC	7.00	7.13	0.17
Arg	AGG	5.00	5.09	0.16	Arg	AGG	0.00	0.00	0.00
Arg	AGA	23.00	23.42	0.74	Arg	AGA	0.00	0.00	0.00
Arg	CGG	0.00	0.00	0.00	Arg	CGG	0.00	0.00	0.00
Arg	CGA	0.00	0.00	0.00	Arg	CGA	0.00	0.00	0.00
Arg	CGT	2.00	2.04	0.06	Arg	CGT	20.00	20.37	0.65
Arg	CGC	1.00	1.02	0.03	Arg	CGC	11.00	11.20	0.35
Asn	AAT	57.00	58.04	0.83	Asn	AAT	10.00	10.18	0.14
Asn	AAC	12.00	12.22	0.17	Asn	AAC	59.00	60.08	0.86
Asp	GAT	59.00	60.08	0.83	Asp	GAT	35.00	35.64	0.49
Asp	GAC	12.00	12.22	0.17	Asp	GAC	36.00	36.66	0.51
Cys	TGT	3.00	3.05	1.00	Cys	TGT	1.00	1.02	0.33
Cys	TGC	0.00	0.00	0.00	Cys	TGC	2.00	2.04	0.67
Gln	CAG	3.00	3.05	0.13	Gln	CAG	16.00	16.29	0.70
Gln	CAA	20.00	20.37	0.87	Gln	CAA	7.00	7.13	0.30
Glu	GAG	10.00	10.18	0.14	Glu	GAG	19.00	19.35	0.27
Glu	GAA	60.00	61.10	0.86	Glu	GAA	51.00	51.93	0.73
Gly	GGG	6.00	6.11	0.08	Gly	GGG	0.00	0.00	0.00
Gly	GGA	38.00	38.70	0.52	Gly	GGA	0.00	0.00	0.00
Gly	GGT	26.00	26.48	0.36	Gly	GGT	36.00	36.66	0.49
Gly	GGC	3.00	3.05	0.04	Gly	GGC	37.00	37.68	0.51
His	CAT	16.00	16.29	0.94	His	CAT	5.00	5.09	0.29
His	CAC	1.00	1.02	0.06	His	CAC	12.00	12.22	0.71
Ile	ATA	41.00	41.75	0.71	Ile	ATA	0.00	0.00	0.00
Ile	ATT	11.00	11.20	0.19	Ile	ATT	23.00	23.42	0.40
Ile	ATC	6.00	6.11	0.10	Ile	ATC	35.00	35.64	0.60
Leu	TTG	4.00	4.07	0.06	Leu	TTG	0.00	0.00	0.00
Leu	TTA	42.00	42.77	0.65	Leu	TTA	0.00	0.00	0.00
Leu	CTG	0.00	0.00	0.00	Leu	CTG	65.00	66.19	1.00
Leu	CTA	4.00	4.07	0.06	Leu	CTA	0.00	0.00	0.00
Leu	CTT	14.00	14.26	0.22	Leu	CTT	0.00	0.00	0.00
Leu	CTC	1.00	1.02	0.02	Leu	CTC	0.00	0.00	0.00
Lys	AAG	32.00	32.59	0.35	Lys	AAG	25.00	25.46	0.27
Lys	AAA	59.00	60.08	0.65	Lys	AAA	66.00	67.21	0.73
Met	ATG	19.00	19.35	1.00	Met	ATG	19.00	19.35	1.00
Phe	TTT	24.00	24.44	0.65	Phe	TTT	12.00	12.22	0.32
Phe	TTC	13.00	13.24	0.35	Phe	TTC	25.00	25.46	0.68
Pro	CCG	1.00	1.02	0.03	Pro	CCG	33.00	33.60	0.82
Pro	CCA	23.00	23.42	0.57	Pro	CCA	6.00	6.11	0.15
Pro	CCT	15.00	15.27	0.38	Pro	CCT	1.00	1.02	0.03
Pro	CCC	1.00	1.02	0.03	Pro	CCC	0.00	0.00	0.00
Ser	AGT	25.00	25.46	0.36	Ser	AGT	0.00	0.00	0.00
Ser	AGC	5.00	5.09	0.07	Ser	AGC	17.00	17.31	0.25
Ser	TCG	0.00	0.00	0.00	Ser	TCG	0.00	0.00	0.00
Ser	TCA	16.00	16.29	0.23	Ser	TCA	0.00	0.00	0.00
Ser	TCT	21.00	21.38	0.30	Ser	TCT	34.00	34.62	0.49
Ser	TCC	2.00	2.04	0.03	Ser	TCC	18.00	18.33	0.26
Thr	ACG	1.00	1.02	0.02	Thr	ACG	6.00	6.11	0.10
Thr	ACA	23.00	23.42	0.37	Thr	ACA	0.00	0.00	0.00
Thr	ACT	34.00	34.62	0.55	Thr	ACT	16.00	16.29	0.26
Thr	ACC	4.00	4.07	0.06	Thr	ACC	40.00	40.73	0.65
Trp	TGG	9.00	9.16	1.00	Trp	TGG	9.00	9.16	1.00
Tyr	TAT	64.00	65.17	0.83	Tyr	TAT	23.00	23.42	0.30
Tyr	TAC	13.00	13.24	0.17	Tyr	TAC	54.00	54.99	0.70
Val	GTG	2.00	2.04	0.04	Val	GTG	19.00	19.35	0.34
Val	GTA	31.00	31.57	0.55	Val	GTA	11.00	11.20	0.20
Val	GTT	23.00	23.42	0.41	Val	GTT	19.00	19.35	0.34
Val	GTC	0.00	0.00	0.00	Val	GTC	7.00	7.13	0.13

11, respectively, after optimization. Additionally, the number of rare codons for leucine, i.e. TTA and CTA, decreased from 42 and 4 to zero. On the contrary, codon optimization raised the optimal codon CTG to 65. Furthermore, the codons AGT and TCA that are rare for serine diminished to zero; in contrast, increase in the numbers in the optimal codons TCT, TCC, and AGC emerged from codon optimization. In addition, after codon optimization, the rare codon ACA encoding threonine declined from 23 to 0, and the optimal codon ACC increased from 4 to 40. As far as isoleucine amino acid is concerned, the rare codon ATA reduced from 41 to 0 and the optimal codons ATC and ATT rose to 35 and 23, respectively. As for proline, the rare codon CCC decreased from 1 to 0, and optimal codon CCG increased from 1 to 33. As to other amino acids, it has also been shown that codons are biased towards more frequent codons via codon optimization. As an example, less frequent codon AAT that encode asparagine reduced from 57 to 10;on the contrary, more frequent codon AAC increased from 12 to 59 (Table 1).

Codon Adaption Index

Codon Adaption Index (CAI) assesses the extent of bias in favor of codons which are involved in highly expressed genes. The levels of protein expression and CAI are known to be correlated. A CAI of >0.8 is regarded to be optimum for expression in the expression host of interest. As shown in Figure 1, the CAI value for original gene was 0.62; however, codon optimization increased the CAI of the coding sequence to 0.84 that is ideal for expression in E. coli.

Frequency of Optimal Codon

The simplest method for measuring species-specific codon usage bias is the frequency of optimal codons (Fop):

$$\text{Fop} = \frac{Xop}{Xop + Xnon}$$

Where Xop and Xnon are the number of optimal and non-optimal codons in a gene, respectively. Codons excluded from the calculation contain stop codons and codons for tryptophan and methionine. Optimal codons for E. coli were originally determined based on the availability of tRNA and the nature of the codon-anticodon interplay. These codons are thought to be translationally optimal and are more frequently involved in genes expressed highly than lowly expressed ones.[41] Non-optimal codon content is currently known to limit the expression of heterologous proteins owing to restricting available cognate tRNAs in the expression host.[42] As can be seen in Figure 2,

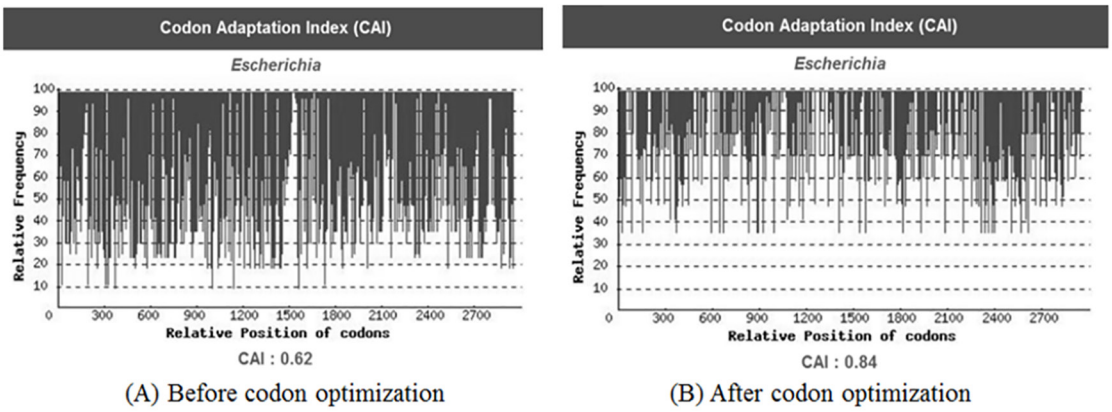

Figure 1: The distribution of codon usage frequency along the coding region of Collagenase H. The distribution of codon usage frequency along the coding region of Collagenase H. (A) and (B) depict the coding region of Col H before and after codon optimization, respectively

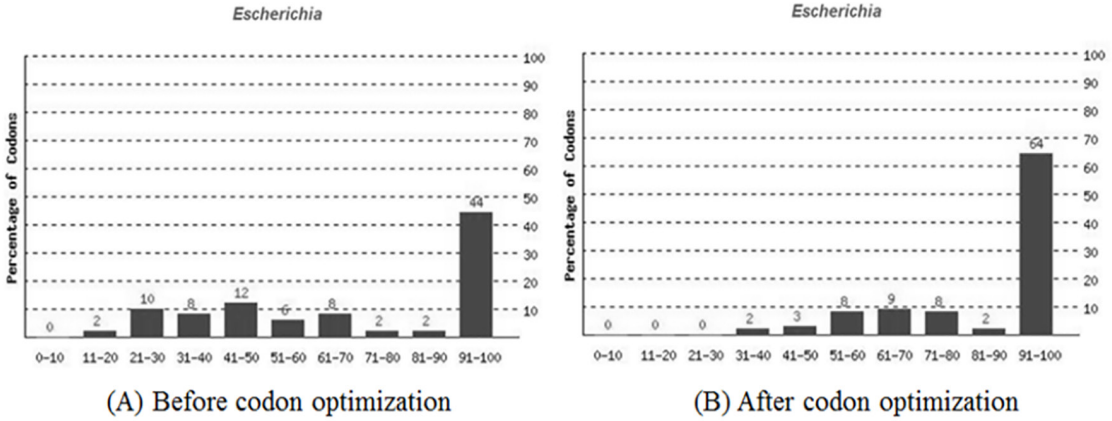

Figure 2: The percentage distribution of codons on the basis of their qualification. The value of 100 is assigned to the codon with the highest usage frequency for a specific amino acid in the expression organism of interest. The values of < 30 are assigned to codons which hinder the expression efficiency and accuracy. (A) And (B) illustrate the percentage of Collagenase H codons before and after codon optimization, respectively

44% of Col H codons showed the value of 100 before codon optimization, while they increased to 64% with codon optimization. Additionally, the percentage of rare codons (codons with values lower than 30) was 12, which decreased to 0 following codon optimization.

GC Content Adjustment

Original gene encoding Col H showed a GC low content, resulting in low expression. However, codon optimization increased it from 30.99% to 47.46%.

Discussion

The codon contains three nucleotides that encode a specific amino acid in protein synthesis. 64 codons make up the structure of proteins, but only 20 amino acids lead to the multiplicity of the genetic code. For example, leucine is one of the amino acids encoded by six different codons, including UUG, CUC, CUA, CUG, UUA, and CUU, while the amino acid cysteine is encoded by only two codons, UGU and UGC. In total, 2 amino acids are encoded from 20 types of amino acids with 1 codon, 9 amino acids with 2 codons, 1 amino acid with 3 codons, 5 amino acids with 4 codons, and finally 3 amino acids with 6 codons, which generally lead to 61 basic codons in nature. The TAA, TAG, and TGA codons are stop codons and are involved in protein termination. Degradation of codon leads to many protein encoding methods; in nature, not all possible encodings and outcome sequences are seen equally. However, in a particular organism some similar codons can be used more than others to encode a special amino acid. This is called codon bias or codon usage bias.[43]

Two measures are available for quantifying the efficiency of translation. The first type evaluates the codon bias of genes and CAI is a good case in this regard. Another kind is based on the availability of tRNA at each codon along the gene and Fop is a good case in this regard. A privileged status of the Fop over the CAI is that it reduces the need to identify a set of highly expressed genes as a reference. On the contrary, it only needs the recognition of all tRNA genes in the genome and their classification in accordance with their anti-codon.[43] Here, the gene designer stand-alone software applied to optimize codons considers both methods. Furthermore, GC content in the Col H gene was balanced using the software. As such, better results are expected to be obtained using the software, as reported elsewhere.[44, 45] The findings from this study suggest that codon optimization provides a theoretical improvement in Col H gene expression in E. coli. However, experimental research is needed to confirm the improvement. This is because although codon bias greatly impacts gene expression, it is not the only contributing factor. The selection of expression vectors and transcriptional promoters is also imperative.[46] In addition, a sequence motif in the vicinity of the initiation AUG and mRNA stability at the 5′terminus were shown to play a role in the gene expression. The competence of heterologous enzyme and protein production in *E. coli* can be reduced by biased codon usage. Commonly used approaches to overcome this problem include targeted mutagenesis to remove rare codons or add rare codon tRNAs to specific cell lines. Recently, advances in cost-effective synthetic gene production technology have made it a viable option.[45] Nikola et al. have shown that expression of the natural genes in assemblage with rare codon tRNAs imitated the behavior of artificial genes in the native strain.[46] The tendency is that the heterologous gene expression of some enzyme and proteins in bacterial host can be modified by fixing codon preference, However, this effect can mostly be amplified by the introduction of rare codon tRNAs into the bacterial host cell.[47] A 2020 study on codon usage and phenotypic divergence of SARS-CoV-2 genes has shown that the higher divergence observed for the last three genes could be a significant barrier to the development of antiviral drugs against SARS-CoV-2.[48] There are multiple studies of developed heterologous gene expression that were partially or completely replaced with synthetic DNA.[49, 50] In rotation, genes expression including rare codons can be delivered by co-expressing rare related tRNA genes[51] plasmid vectors encoding several tRNA genes are attainable commercially.[52]

Conclusion

Recombinant collagenase production has many applications in medicine, pharmacy, food industry, and health. Researchers use different expression vectors to express the genes. The codons need to be optimized to increase the yield of the protein produced. The results of this study are useful for companies and researchers producing recombinant proteins.

Ethics approval and consent to participate: Not applicable.

Consent for publication: Not applicable.

Availability of data and material: All data is available upon request.

Funding: No funding

Competing interests

SMAM conducted experiments, wrote, and edited the figures of the first draft. HA and AR performed molecular biology and experimental set up to produce the particles. All authors contributed to the design, analysis, and manuscript preparation. All authors read and approved the final manuscript

Acknowledgement

We would like to thank Dr. Ramin Rahmany for reviewing and editing the manuscript.

Conflicts of interest: None declared.

References

1 Bretaud S, Guillon E, Karppinen S-M, Pihlajaniemi T, Ruggiero F. Collagen XV, a multifaceted multiplexin present across tissues and species. Matrix Biology Plus. 2020:100023.

2 Ng HW, Zhang Y, Naffa R, Prabakar S. Monitoring the Degradation of Collagen Hydrogels by Collagenase Clostridium histolyticum. Gels. 2020;6 4:46.

3 Alipour H, Raz A, Zakeri S, Djadid ND. Therapeutic applications of collagenase (metalloproteases): A review. Asian Pacific Journal of Tropical Biomedicine. 2016;6 11:975-81.

4 Duarte AS, Correia A, Esteves AC. Bacterial collagenases–a review. Critical Reviews in Microbiology. 2016;42 1:106-26.

5 Danial S, Natsir H, Dali S, Leliani L. Production and Characterization of Collagenase From Bacillus sp. 6-2 Isolated From Fish Liquid Waste. Jurnal Akta Kimia Indonesia (Indonesia Chimica Acta). 2019;12 1:58-66.

6 Shouroki FK, Neghab M, Mozdarani H, Alipour H, Yousefinejad S, Fardid R. Genotoxicity of inhalational anesthetics and its relationship with the polymorphisms of GSTT1, GSTM1, and GSTP1 genes. Environmental Science and Pollution Research. 2019;26(4):3530-41.

7 Alipour H, Raz A, Zakeri S, Djadid ND. Molecular characterization of matrix metalloproteinase-1 (MMP-1) in Lucilia sericata larvae for potential therapeutic applications. Electronic Journal of Biotechnology. 2017;29:47-56.

8 Alipour H, Shahriari-Namadi M, Ebrahimi S, Moemenbellah-Fard MD. Wound healing potential: evaluation of molecular profiling and amplification of Lucilia sericata angiopoietin-1 mRNA mid-part. BMC Research Notes. 2020;13(1):1-7.

9 Alipour H, Raz A, Djadid ND, Zakeri S. Expression of a New Recombinant Collagenase Protein of Lucilia Sericata in SF9 Insect Cell as a Potential Method for Wound Healing. Iranian Journal of Biotechnology. 2019;17(4):e2429.

10 Mills SA, Gelbard MK. Sixty years in the making: collagenase Clostridium histolyticum, from benchtop to FDA approval and beyond. World Journal of Urology. 2020;38 2:269-77.

11 Ohbayashi N, Yamagata N, Goto M, Watanabe K, Yamagata Y, Murayama K. Enhancement of the structural stability of full-length clostridial collagenase by calcium ions. Applied and environmental microbiology. 2012;78 16:5839-44.

12 Preet Kaur S, Azmi W. Cost Effective Production of a Novel Collagenase from a Non-Pathogenic Isolate Bacillus tequilensis. Current Biotechnology. 2013;2 1:17-22.

13 Preet Kaur S, Azmi W. The Association of Collagenase with Human Diseases and its Therapeutic Potential in Overcoming them. Current Biotechnology. 2013;2 1:10-6.

14 Hurst LC, Badalamente MA, Kulovitz K, Relevo M. Collagenase injection: journey from bench to current advanced clinical use. Dupuytren Disease and Related Diseases-The Cutting Edge: Springer; 2017. p. 115-25.

15 Badalamente MA, Hurst LC. Development of collagenase treatment for Dupuytren disease. Hand clinics. 2018;34 3:345-9.

16 Capece M, Arcaniolo D, Manfredi C, Palmieri A, De Sio M, Verze P, et al. Second cycle of intralesional Collagenase Clostridium histolyticum for Peyronie's disease using the modified shortened protocol: Results from a retrospective analysis. Andrologia. 2020;52 3:e13527.

17 Cwikla DJ, Yafi FA. Intralesional collagenase Clostridium histolyticum in the management of Peyronie's disease: current best practice. Therapeutic Advances in Urology. 2018;10 4:139-53.

18 Dolor A, Szoka Jr FC. Digesting a path forward: the utility of collagenase tumor treatment for improved drug delivery. Molecular pharmaceutics. 2018;15 6:2069-83.

19 Abou-Aiana R, Hammad R, Gabr M, Amer A, Ahmadi E, Alharoon A. EFFECT OF THE METHOD OF POSTPARTUM ADMINISTRATION OF OXYTOCIN OR PGF2AON THE DROP OF RETAINED PLACENTA IN COWS AND SUBSEQUENT REPRODUCTIVE AND REPRODUCTIVE PERFORMANCE. Kafrelsheikh Veterinary Medical Journal. 2019;17 1:18-35.

20 Fitzpatrick J, Richardson C, Klaber I, Richardson MD. Clostridium histolyticum (AA4500) for the Treatment of Adhesive Capsulitis of the Shoulder: A Randomised Double-Blind, Placebo-Controlled Study for the Safety and Efficacy of Collagenase–Single Site Report. Drug Design, Development and Therapy. 2020;14:2707.

21 Badalamente MA, Wang E: Methods for non-surgical treatment of carpal tunnel syndrome. Google Patents; 2018.

22 Karahan N, Ozdemir G, Kolukısa D, Duman S, Arslanoğlu F, Çetin M. Can collagenase be used in the treatment of adhesive capsulitis? Medical Principles and Practice. 2020;29 2:174-80.

23 Sheets AR, Demidova-Rice TN, Shi L, Ronfard V, Grover KV, Herman IM. Identification and characterization of novel matrix-derived bioactive peptides: a role for collagenase from Santyl® ointment in post-debridement wound healing? PloS one. 2016;11 7:e0159598.

24 Abood A, Salman AM, El-Hakim AE, Abdel-Aty AM, Hashem AM. Purification and characterization of a new thermophilic collagenase from Nocardiopsis dassonvillei NRC2aza and its application in wound healing. International journal of biological macromolecules. 2018;116:801-10.

25 Pham CH, Collier ZJ, Fang M, Howell A, Gillenwater TJ. The role of collagenase ointment in acute burns: a systematic review and meta-analysis. Journal of wound care. 2019;28 Sup2:S9-S15.

26 Tanaka K, Okitsu T, Teramura N, Iijima K, Hayashida O, Teramae H, et al. Recombinant collagenase from Grimontia hollisae as a tissue dissociation enzyme for isolating primary cells. Scientific reports. 2020;10 1:1-14.

27 Corrales N, Park S, Lau H, Xu I, Luong C, Rodriguez S, et al. Comparison of Islet Characterization from Use of Standard Crude Collagenase to GMP-Grade Collagenase Enzyme Blends in Preweaned Porcine Islet Isolations. Cell Transplantation. 2020;29:0963689720977835.

28 Spiers RM, Marzi J, Brauchle EM, Cross SE, Vaughan RH, Bateman PA, et al. Donor age significantly influences the Raman spectroscopic biomolecular fingerprint of human pancreatic extracellular matrix proteins following collagenase-based digestion. Acta biomaterialia. 2019;99:269-83.

29 Campora S, Alberte PS, Bruno C, Ghersi G. Isolation of adult rat cardiomyocytes using recombinant collagenases. Chemical Engineering Transactions. 2018;64:25-30.

30 Трухин ВП, Москвичев БВ, Жиренкина ЕН, Зайцева ЕС, Уварова АВ, Анисенкова ОА, et al. Способ получения окрашенного коллагена и определения активности коллагеназы. 2019.

31 Katrolia P, Liu X, Li G, Kopparapu NK. Enhanced Properties and Lactose Hydrolysis Efficiencies of Food-Grade β-Galactosidases Immobilized on Various Supports: a Comparative Approach. Applied biochemistry and biotechnology. 2019;188 2:410-23.

32 Aziz NM, AL-Hajaj MA. Cloning, Expression, Purification, and Bio-functional Analysis of Human Interferon-γ.

33 Burgess-Brown NA, Mahajan P, Strain-Damerell C, Fernandez-Cid A, Gileadi O, Gräslund S. Screening and Production of Recombinant Human Proteins: Protein Production in E. coli. Structural Genomics: Springer. p. 45-66.

34 Chakraborty S, Barbhuiya PA, Mazumder GA, Deb B, Uddin A. Compositional features and codon usage pattern of TP63 gene. Computational biology and chemistry. 2019;83:107119.

35 Mordstein C, Savisaar R, Young RS, Bazile J, Talmane L, Luft J, et al. Codon Usage and Splicing Jointly Influence mRNA Localization. Cell Systems. 2020.

36 Mordstein C, Savisaar R, Young RS, Bazile J, Talmane L, Luft J, et al. Splicing buffers suboptimal codon usage in human cells. BioRxiv. 2019:527440.

37 Lamolle G, Fontenla S, Rijo G, Tort JF, Smircich P. Compositional analysis of flatworm genomes shows strong codon usage biases across all classes. Frontiers in genetics. 2019;10:771.

38 Liao X, Lin W, Chen N, Li L, Huang D, Zhong B, et al. Overexpression of the regulatory subunit of protein kinase A increases heterologous protein expression in Pichia pastoris. Biotechnology Letters. 2020;42 12:2685-92.

39 Frumkin I, Yofe I, Bar-Ziv R, Gurvich Y, Lu Y-Y, Voichek Y, et al. Evolution of intron splicing towards optimized gene expression is based on various Cis-and Trans-molecular mechanisms. PLoS biology. 2019;17 8:e3000423.

40 Chu D, Wei L. Trade-off between cost and efficiency during mRNA translation is largely driven by natural selection in angiosperms. Plant Systematics and Evolution. 2020;306 6:1-12.

41 Wang L, Lu H, Tu Z. Genome codon bias analysis of dengue virus type 1. African Journal of Microbiology Research. 2020;14 12:648-56.

42 Nakashima Y, Yano M, Kobayashi Y, Moriyama S, Sasaki H, Toyama T, et al. Endostatin gene therapy on murine lung metastases model utilizing cationic vector-mediated intravenous gene delivery. Gene therapy. 2003;10 2:123-30.

43 Angov E. Codon usage: nature's roadmap to expression and folding of proteins. Biotechnology journal. 2011;6 6:650-9.

44 Şen A, Kargar K, Akgün E, Ç Pınar M, Codon optimization: a mathematical programing approach, *Bioinformatics*, Volume 36, Issue 13, July 2020, Pages 4012–4020

45 Reboul P, Pelletier J-P, Tardif G, Cloutier J-M, Martel-Pelletier J. The new collagenase, collagenase-3, is expressed and synthesized by human chondrocytes but not by synoviocytes. A role in osteoarthritis. The Journal of clinical investigation. 1996;97 9:2011-9.

46 Slusarewicz P, Zhu K, Kirking B, Toungate J, Hedman T. Optimization of protein crosslinking formulations for the treatment of degenerative disc disease. Spine. 2011;36 1:E7.

47 Nicola A. Burgess-Brown, Sujata Sharma, Frank Sobott, Christoph Loenarz, Udo Oppermann, Opher Gileadi, Codon optimization can improve expression of human genes in Escherichia coli: A multi-gene study, Protein Expression and Purification, Volume 59, Issue 1, 2008, Pages 94-102

48 Dilucca, M.; Forcelloni, S.; Georgakilas, A.G.; Giansanti, A.; Pavlopoulou, A. Codon Usage and Phenotypic Divergences of SARS-CoV-2 Genes. *Viruses* 2020, *12*, 498. https://doi.org/10.3390/v12050498

49 Mojzita D, Rantasalo A, Jäntti J. Gene expression

engineering in fungi. Current opinion in biotechnology. 2019;59:141-9

50 Zhang JJ, Tang X, Moore BS. Genetic platforms for heterologous expression of microbial natural products. Natural product reports. 2019;36(9):1313-32

51 Kaur J, Kumar A, Kaur J. Strategies for optimization of heterologous protein expression in E. coli: Roadblocks and reinforcements. International Journal of Biological Macromolecules. 2018;106:803-22

52 Song L, Ouedraogo J-P, Kolbusz M, Nguyen TTM, Tsang A. Efficient genome editing using tRNA promoter-driven CRISPR/Cas9 gRNA in Aspergillus niger. PloS one. 2018;13(8):e0202868