Biomacromol. J., Vol. 7, No. 3, 126-137, December 2021.

Biomacromolecular Journal
www.bmmj.org

Bioinformatics Aanalysis of Lampyridae Family Luciferases by Homology Modeling and Substrate Docking Studies

M. Mortazavi[a,*], M. Torkzadeh-Mahani[a], A. Riahi-Madvar[b], M. Rahimi[a], M. Maleki[a] and S. Lotfi[a]

[a]*Department of Biotechnology, Institute of Science and High Technology and Environmental Sciences, Graduate University of Advanced Technology, Kerman, Iran*

[b]*Department of Molecular and Cell Biology, Faculty of Basic Sciences, Kosar University of Bojnord, Bojnord, Iran*

(Received 9 June 2022, Accepted 3 November 2022)

ABSTRACT

Codon usage and rare codons have mixed results on the protein structure and function. An increasing amount of data is shown that replacing the rare codons with frequently synonymous ones has diverse results as a decrease in a protein's specific activity, changing the folding pathway, and reducing protein solubility. In this study, we investigated the situation of codon usage of the Lampyridae family luciferases using computational databases. For this, the codon feature of these luciferases was studied, bioinformatically. Also, *in silico* analyses of this enzyme were conducted by structural modeling on the I-TASSER web server. The status of these rare codons in these structural models was studied using SPDBV and PyMOL software. Finally, the binding site properties were studied using the AutoDock Vina. Using molecular modeling, two rare codons (Arg533 and Arg536) were analyzed that may have a critical role in the structure and function of these luciferases. AutoDock Vina was used in molecular docking that recognizes some residues that yield closely related to luciferyl-adenylate binding sites. These analyses created a new understanding of the sequence and structure of these luciferases, and our findings can be used in some fields of clinical and industrial biotechnology. This bioinformatics analysis plays an essential role in the design of new drugs.

Keywords: Luciferase, Codon bias, Rare codon, Lampyridae, Molecular docking

INTRODUCTION

Today, many of the luciferase enzymes have been identified in diverse families of life as the Lampyridae family. This family is called fireflies for their conspicuous use of bioluminescence [1]. Lampyridae has three subfamilies, including Luciolinae, Lampyrinae, and Photurinae [2], that utilize similar luciferase enzymes in the bioluminescence reaction [3]. Firefly luciferases catalyze the oxidation of luciferin in the presence of Mg^{2+}, ATP, and O_2 that produce the oxyluciferin in an excited state, which decays to the ground state by the emission of a photon [4]. The luciferase-based assay has been widely studied and applied in biological sciences such as in pyrosequencing, *in vivo* imaging, ATP assay, gene reporters, and luciferase-based split biosensors [5,6].

Previous studies have shown that synonymous codons are not used with the same frequency in organisms [7]. Other reports demonstrate that at the translational level, the rate of protein elongation is specified by the properties of mRNA, tRNA abundances, and codon usage [8,9]. Commonly, the desired codons are translated at higher rates and are read by more abundant tRNAs [10]. On the other hand, rare codons are taken by lowly abundant tRNAs and this creates changes in the translation speed [11]. Other studies show that rare codons have a special role in protein expression, folding, and enzyme activity [12]. In the translation of rare codons, the ribosomes pause until the rarely activated tRNA brings the next amino acid [13]. This mediates the local kinetics of translation [14] and by mutation of the slow-translating

*Corresponding author. E-mail: mortezavimm@gmail.com

messenger to a fast-translating messenger, the folding yield was reduced [15].

The luciferase enzymes suffer from low turnover numbers, high K_m for the substrate ATP, and inactivation at high temperatures [16]. So far, many extensive protein-engineering studies have been conducted to overcome the problems associated with the luciferase enzyme [17]. Although synonymous mutations are generally selected as a neutral base, it has been found that changing the composition and order of the codon (Codon Usage) can have great effects on protein expression and function [18]. According to these findings, the situation of codon usage and rare codon in the luciferases of the Lampyridae family is being studied and done (data is being published). For this, the nucleotide sequences were retrieved from the NCBI (https://www.ncbi.nlm.nih.gov/) and some of the rare codons were identified in these enzymes.

Arg has special properties as its guanidinium group participates in the ionic and hydrogen interactions, which leads to an unexpected increase in thermal stability, and for this reason, further studies have been conducted on the Arg codons [19]. The Arg d-guanido moiety has a decreased chemical reactivity and provides high surface area for charged interactions and has wide tendency to take part in salt bridge interaction [20]. Arg is encoded by six different codons and two of the six Arg codons, AGA AGG, are in low abundance [21]. The Arg synonymous mutation has been studied and show that rare arginine codons AGA and AGG affect the heterologous expression of proteins in Eschericha coli [22]. Due to the fact that we had previously introduced several Arg for mutagenesis and the C-terminal region of the enzyme had not been introduced for the purpose of mutagenesis, these two regions were selected. Our chosen location within the sequence was based on the assessment of internal structure and conservation during evolution, and this region showed that it is conserved and has few changes during evolution in different species of this family.

By *in silico* analysis, two Arg codons (R^{533} and R^{536}) were bioinformatically analyzed. For this, the structures of these luciferases were retrieved or modeled in the Swiss model and I-TASSER [23] web server. In the following, with the help of Swiss PDB Viewer software [24] and PyMOL Molecular Graphics System [25], the location and situation of these Arg rare codons were analyzed in the structure of these luciferases. These results may help in better recognition of enzyme activity, functional development of bioluminescence assay, and a new understanding of the molecular evolution of the Lampyridae family.

MATERIAL AND METHODS

Gene and Amino Acid Sequences

For bioinformatics analysis, the nucleotide and amino acid reference sequences of Lampyridae luciferases and their features were retrieved from GenBank (http://ncbi.nlm.nih.gov/) and UniProtKB (http://www.expasy.org/uniprot) Databases.

Sequence Alignment

The protein reference sequences of similar species were retrieved from the NCBI database. MSA of these amino acid sequences was performed by the ClustalW2 program (https://www.ebi.ac.uk/Tools/msa/clustalo/) and MEGA 7 [26]. This program was used to identify the identity and similarity of these sequences, as for the creation of a phylogenetic tree, the MSA is a critical step [27].

Bioinformatics Studies

For *in silico* studies, the crystal structures of these luciferases were retrieved from PDB or modeled on the I-TASSER web server [23] and the Swiss model [28]. I-TASSER web server generated a total of five most suitable models of luciferases based on multiple-threading alignments by LOMETS [29]. The models with the best "Confidence Score" and Z-score were chosen. The best models were visualized by Swiss PDB [30) and PyMOL [25] viewers. Hydrogen bonds were also calculated by WHAT IF [31] and PIC web servers [32]. The physicochemical parameters of these models, such as molecular weight and instability index, were calculated on the Expasy ProtParam server [33].

Molecular Docking

Molecular docking was conducted at AutoDock Vina (version 1.1.2) [34]. 3D molecular models of luciferases were used as a target. The SDF format of luciferyl-adenylate was obtained from PubChem (https://pubchem.ncbi.nlm.nih.gov/) and converted to PDB format by Open Babel (version 2.3.1) [35]. Finally, using

MGL tools (version 1.5.4), the PDB format of luciferyl-adenylate was converted to PDBQT format [36]. Molecular docking was adjusted using different box sizes between luciferases and luciferyl-adenylate. The best-docked conformation was visualized using the PyMOL [25] and Ligplot [37].

RESULTS

Luciferases Sequences

The amino acid sequences of these luciferases were from NCBI. Table 1 shows the list of species names and features of the luciferase enzyme gene along with accession numbers.

Multiple Sequence Alignment and Phylogenetic Tree

The amino acid sequence references of these luciferases are retrieved and aligned in the Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/). The MSA was saved in the clustal_num format and analyzed with the Jalview software [38] (Fig. 1).

Table 1. The Protein Properties of Luciferase in the Lampyridae Family

ORGANISM	Protein gene bank	
Pyrearinus termitilluminans	AOC83873	528-532
Photinus pyralis	AAA29795.1	
Pyrocoelia miyako	AAC37254.1	
Pyrocoelia rufa	AAG45439.1	
Photuris pensylvanica15	BAA05005.1	
Photuris pensylvanica16	BAA05006.1	
Hotaria unmunsana	AAM00429.1	
Luciola mingrelica	AAB26932.1	
Cratomorphus distinctus	AAV32457.1	
Luciola lateralis	BAL46510.1	
Luciola cruciata	AAA29135.1	
Lampyris noctiluca	AAR20794.1	
Photuris pennsylvanica40	AAB60897.1	
Hotaria parvula	AAC37253.1	
Lampyris turkestanicus	AAU85360.1	

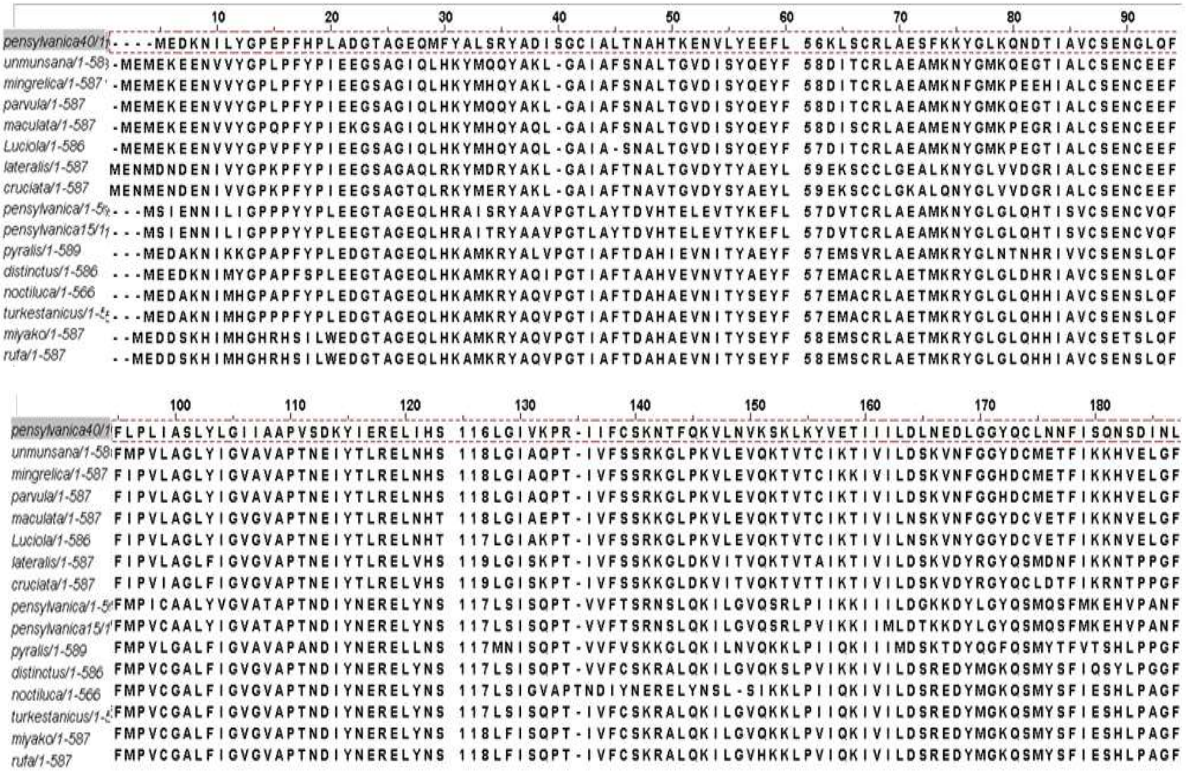

Fig. 1. The analysis of the MSA file with the Jalview software.

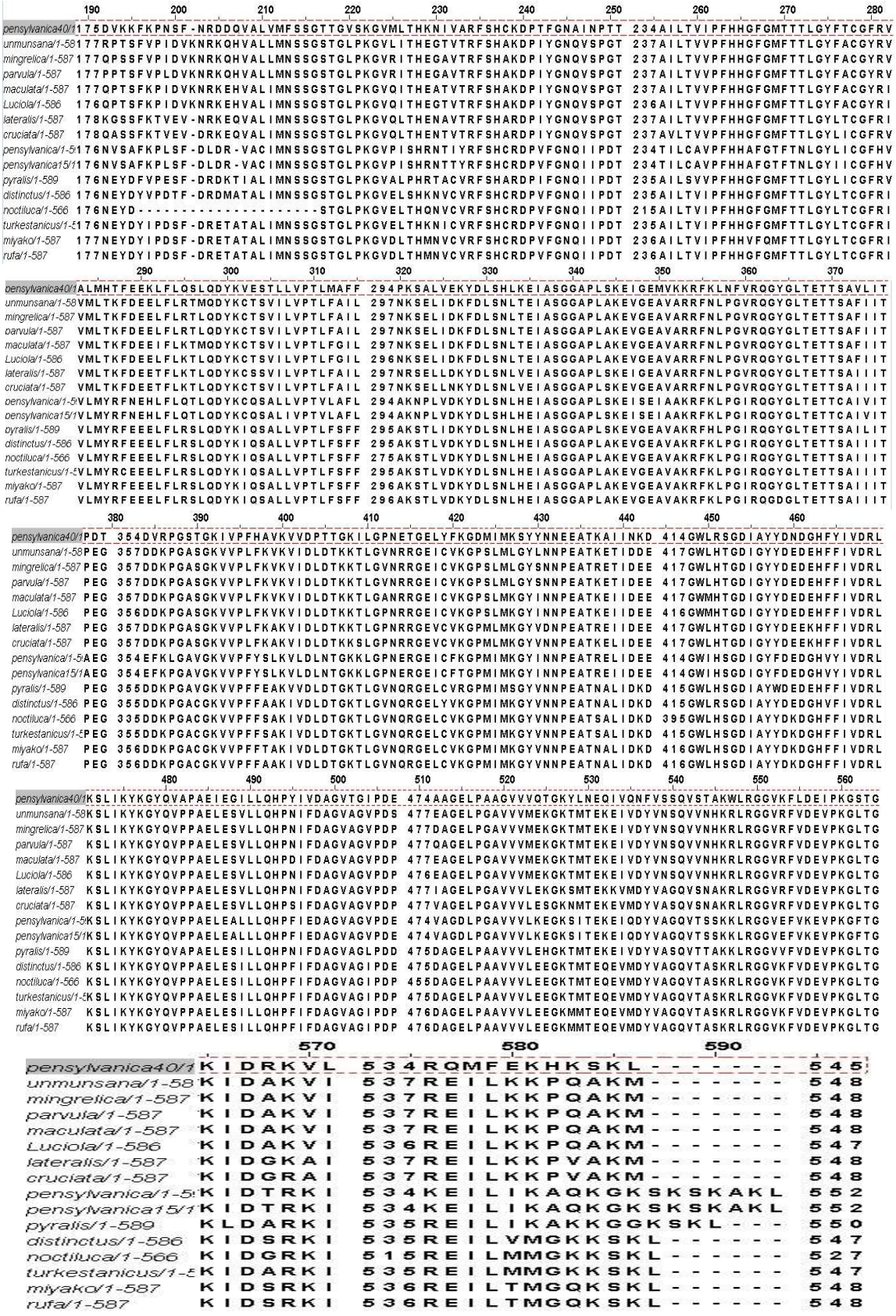

Fig. 1. Continued.

This analysis shows that the frequency of a number of rare codons has been high during evolution. But other rare codons have low frequency and some are repeated only once. Based on some special properties of Arg, the Arg residues were evaluated, and finally, the two residues of R533 and R536 were selected for further analysis.

Evolutionary analyses were conducted in MEGA7 [26] and the evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 1.73339343 is shown (Fig. 2). The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. Evolutionary distances were computed using the Poisson correction method and are in the units of the number of amino acid substitutions per site.

Bioinformatics Studies

Based on the input information, some of the rare codons in these genes were identified. For a better analysis of these rare codons, the 3D structures of these luciferases were retrieved from the PDB or modeled in the I-TASSER (Iterative Threading ASSEmbly Refinement) [39]. For further analysis, the situation of these residues (R533 and R536) was precisely studied in the structure of these luciferases. Based on the results of the modeling, these two residues have been located in the C-terminal domain of luciferase. Analyzing the 3D model of luciferase showed that these three residues establish an extensive network of hydrogen bonds with other residues (Fig. 3). A structural review of these rare codons and hydrogen interactions indicated that these residues may have roles in the propern folding of these luciferases. The non-covalent interactions between these residues were calculated by WHAT IF [31] and PIC Web servers [40]. Docking simulation studies were conducted in AutoDock Vina [34]. The crystal structure of luciferases was treated as a receptor, whereas luciferyl-adenylate was used as a small molecule ligand. The luciferase-luciferin complex obtained from docking results is shown in Fig. 3.

In the following, we conducted the cross-validation using AlphaFold2 [41] and RoseTTAFold [42] and then compared their results with those of I-TASSER. For this, we performed the modeling process on the luciferase enzymes from the

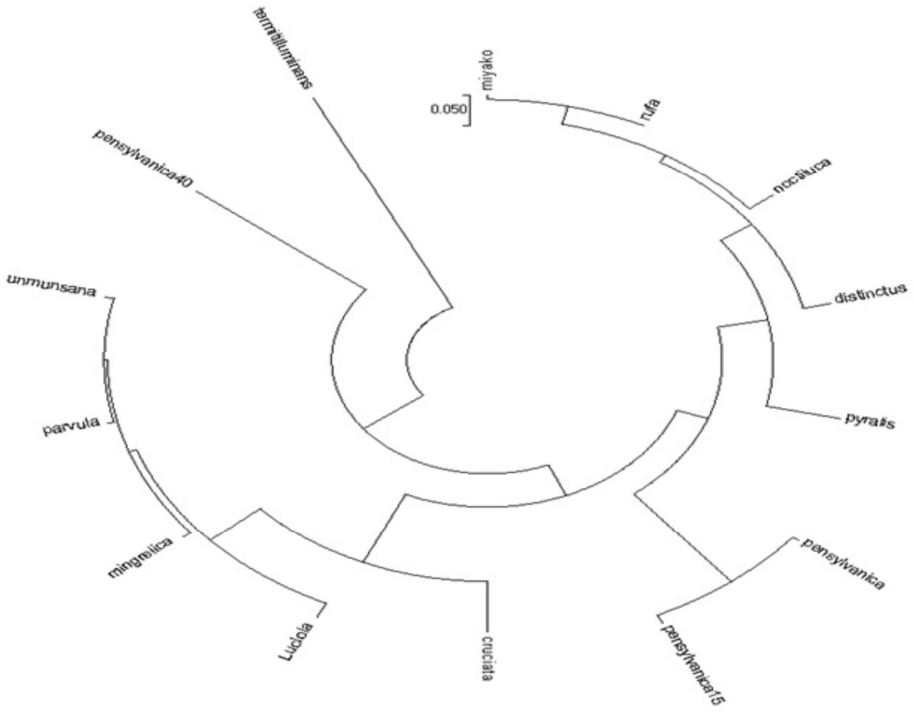

Fig. 2. Evolutionary analyses. The analysis involved 14 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 519 positions in the final dataset.

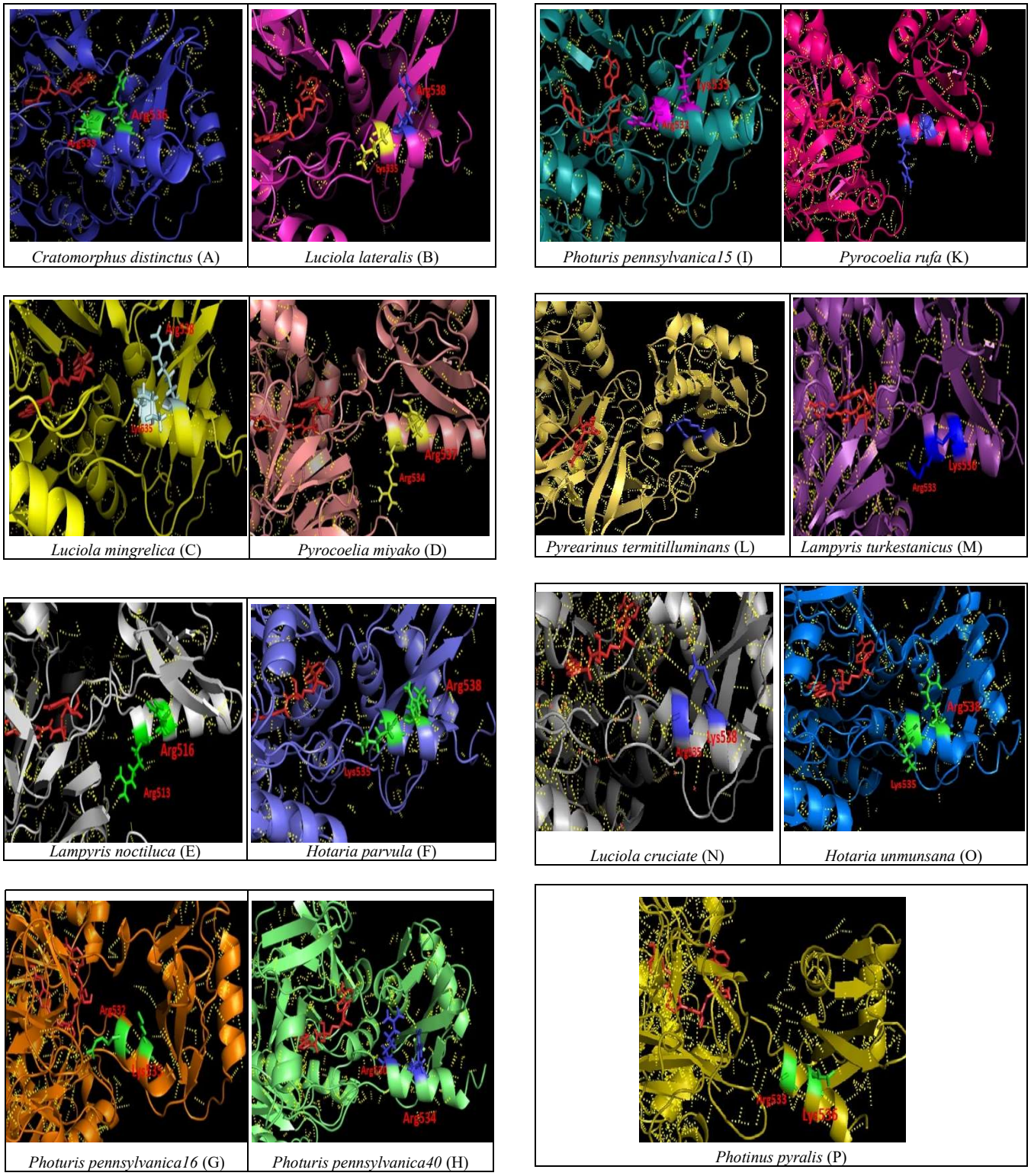

Fig. 3. A-P) PyMOL diagram of docking situation of luciferyl-adenylate into the luciferase. PyMOL diagram showing the interaction of luciferyl-adenylate with luciferase (Red stick: luciferyl-adenylate). Polar interactions are shown as yellow color lines. The relative rareness of the Arg codons in these positions is shown.

Luciola mingrelica in the AlphaFold2 (Fig. 4) and compared the results with the models from the I-TASSER by the superimposition of the models.

I-TASSER was ranked as the No 1 server for protein structure prediction in CASP9 [43]. I-TASSAR Web Server generated five models for *Luciola mingrelica*. The best model showed a 1.76 value for the overall C-score, Exp. RMSD was 3.8 ± 2.6, and 0.96 ± 0.05 value of TM-Score. C-score is typically in the range of [-5,2], where a C-score of higher value signifies a model with high confidence and vice-versa. A TM-score > 0.5 indicates a model of correct topology and a TM-score < 0.17 means a random similarity. The results show that there is a high similarity between the created models. AlphaFold produces a per-residue estimate of its confidence on a scale from 0-100. This confidence measure is called pLDDT and corresponds to the model's predicted score on the lDDT-Cα metric. Regions with pLDDT > 90 are expected to be modeled with high accuracy. The pLDDT is 90.8 (ptmscore 0.856) which shows the high accuracy of the model.

In the following, we performed the modeling process on the luciferase enzymes from Cratomorphus distinctus in the RoseTTAFold (Fig. 5) and compared the results with the models from the I-TASSER.

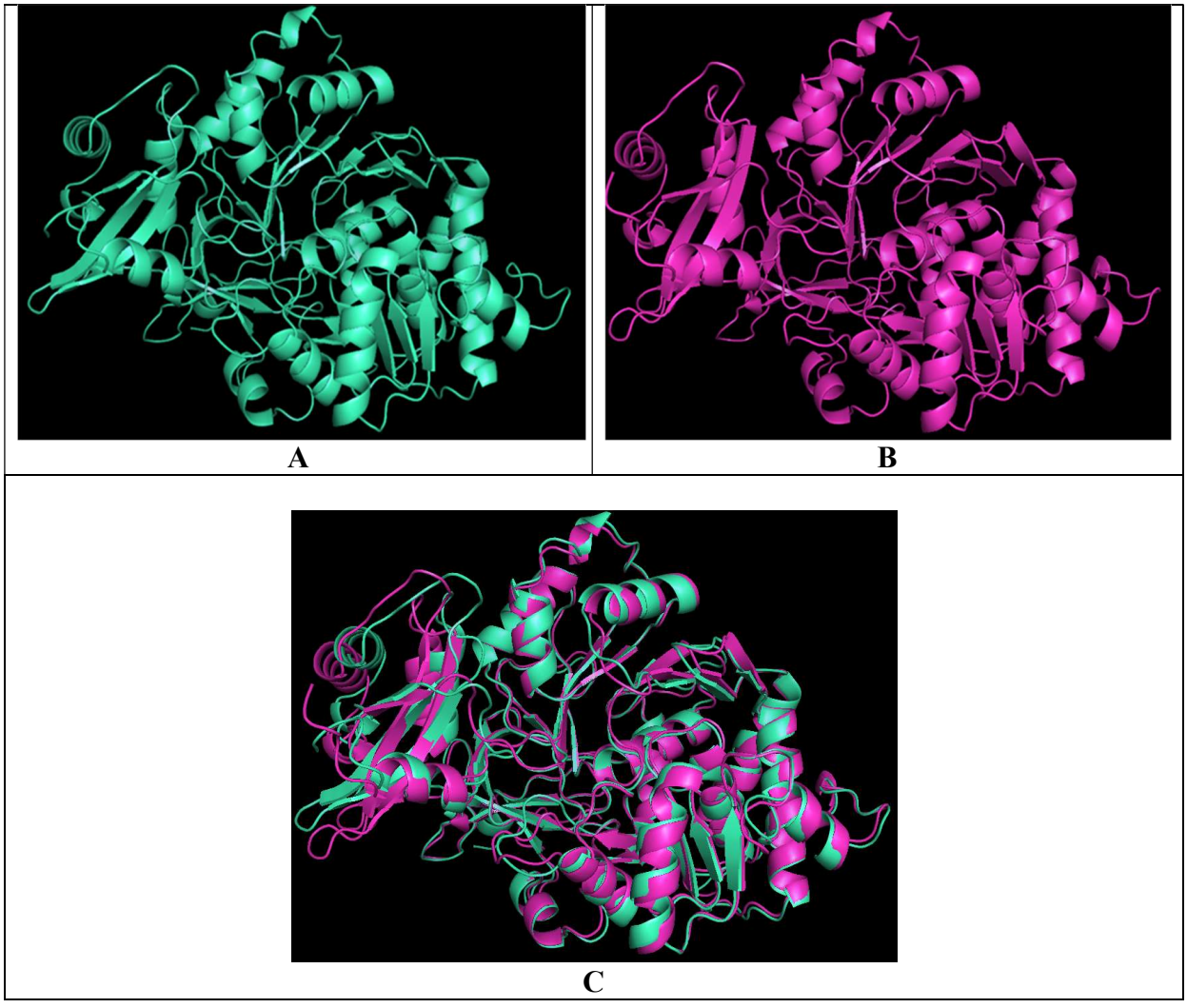

Fig. 4. Structures modeled of *Luciola mingrelica* in the I-TASSER (A), AlphaFold2 (B), and Superimposition of these models in the AlphaFold2 and I-TASSER (C).

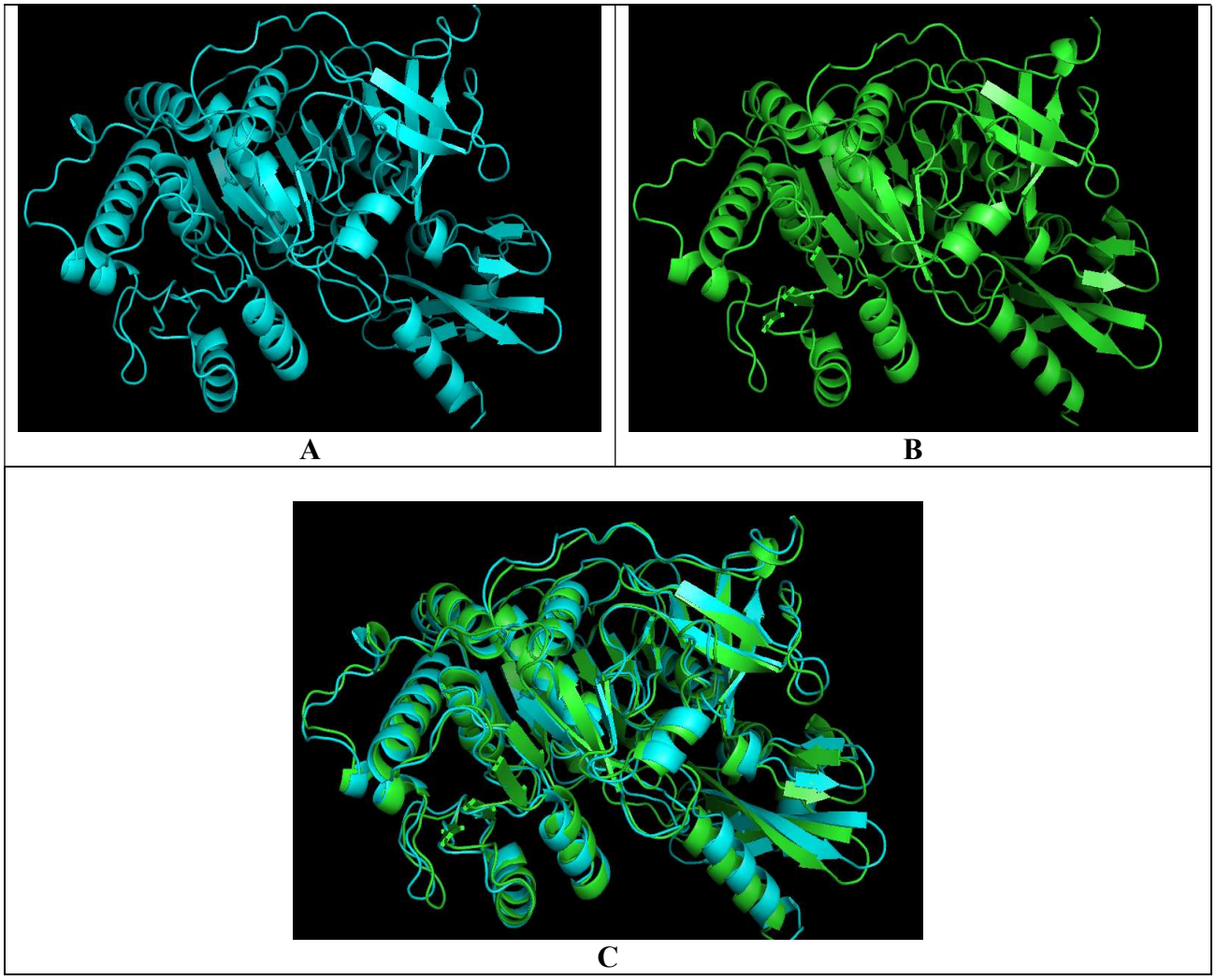

Fig. 5. Structures modeled of *Cratomorphus distinctus* in the I-TASSER (A), RoseTTAFold (B), and superimposition of these models in the RoseTTAFold and I-TASSER (C).

I-TASSAR Web Server generated five models for *Cratomorphus distinctus*. The best model showed a 1.71 value of overall C-score, Exp. RMSD was 3.9 ± 2.7, and 0.95 ± 0.05 value of TM-Score. In the RoseTTAFold, the 0.0 < TM-score < 0.17 shows the random structural similarity and 0.5 < TM-score < 1.00 shows in about the same fold [44]. RMSD, TM-score of the luciferase model from *Cratomorphus distinctus* are 1.831 and 0.9606 (d0 = 8.24), respectively. Based on the TM-score, these models from RoseTTAFold and I-TASSER have the similar fold.

There is a diversity of non-covalent interactions in the enzyme-substrate complex and some hydrogen bonds can be formed between the luciferyl-adenylate and luciferase as reported previously [45]. Using the crystal structure *P. pyralis* luciferase in the adenylate-forming conformation bound to DLSA [46] and firefly luciferase (FLuc) in complex with PTC124-AMP [47], the situation of these residues in relation to the active site was investigated. This study shows that these rare codons do not connect directly with the substrate, but they are located in the vicinity of the substrate. To investigate the importance of the protein folding rate in these areas, by the %MinMax algorithm, the relative rareness of the Arg codons in the native and mutant mRNA sequence of luciferase was calculated (data not shown).

DISCUSSION

Although genome diversity in codon usage has been studied [48], it remains unknown why particular codons are utilized rarely in protein-coding genes [49]. However, modifying codon usage has been found to have remarkable results [50]. Other studies that replaced rare codons with frequently synonymous ones have diverse results, including a decrease in a protein's specific activity [51], a change in substrate specificity [48,52], an alteration in the folding pathways [53], and a decrease in protein solubility [54]. By changing the ribosome translation rate, presumably through the constitution of new hydrogen bonds or changes of previous interactions, the structural rigidity has changed that confirmed by structural analysis. Furthermore, synonymous mutations in the methylobacterium extorquens reduce enzyme activity in comparison to the wild-type. These variations could be the result of altered co-translational protein folding and/or mistranslation [49]. These findings reveal that synonymous mutations have a visible effect on enzyme activity, which can meet the biological and nanotechnological needs of enzymes. Data show that we must reconsider our ideas regarding synonymous mutations.

In this regard, the situation of rare codons in the luciferase of the Lampyridae family was studied to obtain new insights into evolutionary relationships. In the Lampyridae family, more than 2000 species have been described [55]. Although extensive studies have been conducted on luciferases, there have been some unresolved issues [56]. Previously, we have conducted some *in silico* analyses of rare codons in different proteins [57-60]. In this regard, by *in silico* analysis, these rare codons were evaluated structurally in the luciferases of the Lampyridae family. In this study, some parameters of codon usage were analyzed. After preliminary analysis, it was found that despite the high similarity of the nucleotide sequences and based on evolutionary relationships, the *Pyrearinus termitilluminans* have some fundamental differences from the other luciferases.

In the following, the rare codons of the Arg were identified and showed that the rare Arg codon has the highest frequency of these nucleotide sequences. Consequently, based on Arg properties [61,62], two Arg residues (Arg^{533} and Arg^{336}) were selected for further analysis. Structural analysis shows that these Arg residues are located near the active site and have a wide network of non-covalent interactions where different parts of the structures are held together. The establishment of these bonds may be very important for the regulation of the folding rate in the luciferase structure. *Pyrearinus termitilluminans, Photinus pyralis, Pyrocoelia miyako, Pyrocoelia rufa, Luciola cruciata, Photuris pennsylvanica40, Hotaria parvula,* and *Lampyris turkestanicus* have two rare codons of *Arg.* On the other hand, *Photuris pensylvanica15, Photuris pensylvanica16, Hotaria unmunsana, Luciola mingrelica, Cratomorphus distinctus,* and *Luciola lateralis* have one rare codon of Arg. *Lampyris noctiluca* has no rare codons in this area. Bioinformatics investigation reveals that variations in ribosome translation rate may affect the enzyme's folding process. Our findings show that the ribosome translation rate in the Arg533 and Arg533 regions of these luciferases are different in that some have two stages of deceleration in the ribosome translation rate, and some have one stage of deceleration and one luciferase does not have a slowing down stage of deceleration. These results show that the interaction situation of this α-helix has a critical role in the final folding of luciferases.

The structural rigidity may be altered by modifying the ribosome translation rate, probably due to the formation of new hydrogen bonds or a change in an existing contact. However, our study indicates that these rare codons might have an impact on enzyme activity and structure. Furthermore, in many enzyme investigations that combine site-directed mutagenesis and codon optimization (synonymous mutations), the reported modifications may be attributable to optimized codons rather than mutations. Overall, the tRNA population can explain variations in protein characteristics [63], and additional research is needed to interpret these findings. Our results show that although the ribosome translation rate in this location has changed, the situation of interactions has not been affected substantially by the proper folding of the enzyme.

ACKNOWLEDGMENTS

We gratefully acknowledge the research funding provided for this project (No. 96003716) by Iran National Science Foundation (INSF) and Graduate University of Advanced Technology, Kerman, Iran (7/s/96/2557).

REFERENCES

[1] Hackett K, Whitcomb R, Tully J, Lloyd J, Anderson J, Clark T, et al. Lampyridae (Coleoptera): A plethora of mollicute associations. Microbial ecology. 1992;23(2):181-93.

[2] Day JC, Goodall TI, Bailey MJ. The evolution of the adenylate-forming protein family in beetles: multiple luciferase gene paralogues in fireflies and glow-worms. Molecular phylogenetics and evolution. 2009;50(1):93-101.

[3] Inouye S. Firefly luciferase: an adenylate-forming enzyme for multicatalytic functions. Cellular and molecular life sciences. 2010;67(3):387-404.

[4] Deluca M, McElroy W. [1] Purification and properties of firefly luciferase. Methods in enzymology. 57: Elsevier; 1978. p. 3-15.

[5] Ataei F, Torkzadeh-Mahani M, Hosseinkhani S. A novel luminescent biosensor for rapid monitoring of IP3 by split-luciferase complementary assay. Biosensors and Bioelectronics. 2013;41:642-8.

[6] Roda A, Guardigli M, Michelini E, Mirasoli M. Bioluminescence in analytical chemistry and in vivo imaging. TrAC Trends in Analytical Chemistry. 2009;28(3):307-22.

[7] 7.Ikemura T. Codon usage and tRNA content in unicellular and multicellular organisms. Molecular biology and evolution. 1985;2(1):13-34.

[8] Gorgoni B, Ciandrini L, McFarland MR, Romano MC, Stansfield I. Identification of the mRNA targets of tRNA-specific regulation using genome-wide simulation of translation. Nucleic acids research. 2016;44(19):9231-44.

[9] Kemp AJ, Betney R, Ciandrini L, Schwenger AC, Romano MC, Stansfield I. A yeast tRNA mutant that causes pseudohyphal growth exhibits reduced rates of CAG codon translation. Molecular microbiology. 2013;87(2):284-300.

[10] Varenne S, Buc J, Lloubes R, Lazdunski C. Translation is a non-uniform process: effect of tRNA availability on the rate of elongation of nascent polypeptide chains. Journal of molecular biology. 1984;180(3):549-76.

[11] Zhang G, Ignatova Z. Generic algorithm to predict the speed of translational elongation: implications for protein biogenesis. PloS one. 2009;4(4):e5036.

[12] Overton TW. Recombinant protein production in bacterial hosts. Drug discovery today. 2014;19(5):590-601.

[13] Novoa EM, de Pouplana LR. Speeding with control: codon usage, tRNAs, and ribosomes. Trends in Genetics. 2012;28(11):574-81.

[14] Chartier M, Gaudreault F, Najmanovich R. Large-scale analysis of conserved rare codon clusters suggests an involvement in co-translational molecular recognition events. Bioinformatics. 2012;28(11):1438-45.

[15] O'brien EP, Vendruscolo M, Dobson CM. Prediction of variable translation rate effects on cotranslational protein folding. Nature communications. 2012;3:868.

[16] White PJ, SQUIRRELL DJ, ARNAUD P, LOWE CR, MURRAY JA. Improved thermostability of the North American firefly luciferase: saturation mutagenesis at position 354. Biochemical Journal. 1996;319(2):343-50.

[17] Tisi L, White P, Squirrell D, Murphy M, Lowe C, Murray J. Development of a thermostable firefly luciferase. Analytica Chimica Acta. 2002;457(1):115-23.

[18] Thanaraj T, Argos P. Protein secondary structural types are differentially coded on messenger RNA. Protein science. 1996;5(10):1973-83.

[19] Tanaka Y, Tsumoto K, Yasutake Y, Umetsu M, Yao M, Fukada H, et al. How oligomerization contributes to the thermostability of an archaeon protein protein L-isoaspartyl-O-methyltransferase from Sulfolobus tokodaii. Journal of Biological Chemistry. 2004;279(31):32957-67.

[20] Mrabet NT, Van den Broeck A, Van den Brande I, Stanssens P, Laroche Y, Lambeir AM, et al. Arginine residues as stabilizing elements in proteins. 1992;31(8):2239-53.

[21] Kane JFJCoib. Effects of rare codon clusters on high-level expression of heterologous proteins in Escherichia coli. 1995;6(5):494-500.

[22] Aguirre-López B, Cabrera N, de Gómez-Puyou MT, Perez-Montfort R, Gómez-Puyou AJBr. The importance of arginine codons AGA and AGG for the expression in E. coli of triosephosphate isomerase from seven different species. 2017;13:42-8.

[23] Zhang Y. I-TASSER server for protein 3D structure prediction. BMC bioinformatics. 2008;9(1):40.

[24] Kaplan W, Littlejohn TG. Swiss-PDB viewer (deep view). Briefings in bioinformatics. 2001;2(2):195-7.

[25] DeLano WL. The PyMOL molecular graphics system. 2002.

[26] Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Molecular biology and evolution. 2016;33(7):1870-4.

[27] Baxevanis AD, Ouellette BF. Bioinformatics: a practical guide to the analysis of genes and proteins: John Wiley & Sons; 2004.

[28] Schwede T, Kopp J, Guex N, Peitsch MC. SWISS-MODEL: an automated protein homology-modeling server. Nucleic acids research. 2003;31(13):3381-5.

[29] Wu S, Zhang Y. LOMETS: a local meta-threading-server for protein structure prediction. Nucleic acids research. 2007;35(10):3375-82.

[30] Guex N, Peitsch M. Swiss-PdbViewer: a fast and easy-to-use PDB viewer for Macintosh and PC. Protein Data Bank Quaterly Newsletter. 1996;77(7).

[31] Vriend G. WHAT IF: a molecular modeling and drug design program. Journal of molecular graphics. 1990;8(1):52-6.

[32] Tina K, Bhadra R, Srinivasan N. Nucleic Acids Res 35. Web Server issue) W473–W476. 2007.

[33] Gasteiger E, Hoogland C, Gattiker A, Wilkins MR, Appel RD, Bairoch A. Protein identification and analysis tools on the ExPASy server. The proteomics protocols handbook: Springer; 2005. p. 571-607.

[34] Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. Journal of computational chemistry. 2010;31(2):455-61.

[35] OLBoyle NM, Banck M, James CA, Morley C, Vandermeersch T, Hutchison GR. Open Babel: An open chemical toolbox. J Cheminf. 2011;3:33.

[36] Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, et al. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. Journal of computational chemistry. 2009;30(16):2785-91.

[37] Wallace AC, Laskowski RA, Thornton JM. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. Protein engineering. 1995;8(2):127-34.

[38] Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. Bioinformatics. 2009;25(9):1189-91.

[39] Conti E, Franks NP, Brick P. Crystal structure of firefly luciferase throws light on a superfamily of adenylate-forming enzymes. Structure. 1996;4(3):287-98.

[40] Tina K, Bhadra R, Srinivasan N. PIC: protein interactions calculator. Nucleic acids research. 2007;35(suppl_2):W473-W6.

[41] Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Tunyasuvunakool K, et al. AlphaFold 2. 2020.

[42] Baek M, DiMaio F, Anishchenko I, Dauparas J, Ovchinnikov S, Lee GR, et al. Accurate prediction of protein structures and interactions using a three-track neural network. 2021;373(6557):871-6.

[43] Zhang YJBb. I-TASSER server for protein 3D structure prediction. 2008;9(1):1-8.

[44] Zhang Y, Skolnick JJPS, Function,, Bioinformatics. Scoring function for automated assessment of protein structure template quality. 2004;57(4):702-10.

[45] Franks N, Jenkins A, Conti E, Lieb W, Brick P. Structural basis for the inhibition of firefly luciferase by a general anesthetic. Biophysical journal. 1998;75(5):2205-11.

[46] Sundlov JA, Fontaine DM, Southworth TL, Branchini BR, Gulick AM. Crystal structure of firefly luciferase in a second catalytic conformation supports a domain alternation mechanism. Biochemistry. 2012;51(33):6493-5.

[47] Auld DS, Lovell S, Thorne N, Lea WA, Maloney DJ, Shen M, et al. Molecular basis for the high-affinity binding and stabilization of firefly luciferase by PTC124. Proceedings of the National Academy of Sciences. 2010;107(11):4878-83.

[48] Kimchi-Sarfaty C, Oh JM, Kim I-W, Sauna ZE, Calcagno AM, Ambudkar SV, et al. A" silent" polymorphism in the MDR1 gene changes substrate specificity. Science. 2007;315(5811):525-8.

[49] Agashe D, Martinez-Gomez NC, Drummond DA, Marx

CJJMb, evolution. Good codons, bad transcript: large reductions in gene expression and fitness arising from synonymous mutations in a key enzyme. 2013;30(3):549-60.

[50] Plotkin JB, Kudla GJNRG. Synonymous but not the same: the causes and consequences of codon bias. 2011;12(1):32-42.

[51] Yu C-H, Dang Y, Zhou Z, Wu C, Zhao F, Sachs MS, et al. Codon usage influences the local rate of translation elongation to regulate co-translational protein folding. Molecular cell. 2015;59(5):744-54.

[52] Gow JM, Hodges LM, Chinn LW, Kroetz DL. Substrate-dependent effects of human ABCB1 coding polymorphisms. Journal of Pharmacology and Experimental Therapeutics. 2008;325(2):435-42.

[53] Tsai C-J, Sauna ZE, Kimchi-Sarfaty C, Ambudkar SV, Gottesman MM, Nussinov R. Synonymous mutations and ribosome stalling can lead to altered folding pathways and distinct minima. Journal of molecular biology. 2008;383(2):281-91.

[54] Nørholm MH, Light S, Virkki MT, Elofsson A, von Heijne G, Daley DO. Manipulating the genetic code for membrane protein production: what have we learnt so far? Biochimica et Biophysica Acta (BBA)-Biomembranes. 2012;1818(4):1091-6.

[55] Alekseev VI. The beetles (Insecta: Coleoptera) of Baltic amber: the checklist of described species and preliminary analysis of biodiversity. Zoology and Ecology. 2013;23(1):5-12.

[56] Viviani VR. The origin, diversity, and structure function relationships of insect luciferases. Cellular and Molecular Life Sciences CMLS. 2002;59(11):1833-50.

[57] Fattahi M, Malekpour A, Mortazavi M, Safarpour A, Naseri N. The Characteristics of Rare Codon Clusters in the Genome and Proteins of Hepatitis C Virus; a Bioinformatics Look. Middle East journal of digestive diseases. 2014;6(4):214.

[58] Gulick AM. Conformational dynamics in the Acyl-CoA synthetases, adenylation domains of non-ribosomal peptide synthetases, and firefly luciferase. ACS chemical biology. 2009;4(10):811-27.

[59] Mortazavi M, Zarenezhad M, Gholamzadeh S, Alavian SM, Ghorbani M, Dehghani R, et al. Bioinformatic Identification of Rare Codon Clusters (RCCs) in HBV Genome and Evaluation of RCCs in Proteins Structure of Hepatitis B Virus. Hepatitis monthly. 2016;16(10).

[60] Kargar F, Mortazavi M, Savardashtaki A, Hosseinkhani S, Mahani MT, Ghasemi Y. Genomic and protein structure analysis of the luciferase from the Iranian bioluminescent beetle, Luciola sp. International journal of biological macromolecules. 2019;124:689-98.

[61] Zhou X-X, Wang Y-B, Pan Y-J, Li W-F. Differences in amino acids composition and coupling patterns between mesophilic and thermophilic proteins. Amino acids. 2008;34(1):25-33.

[62] Mrabet NT, Van den Broeck A, Van den Brande I, Stanssens P, Laroche Y, Lambeir AM, et al. Arginine residues as stabilizing elements in proteins. Biochemistry. 1992;31(8):2239-53.

[63] Cabrera-Cabrera F, Marín M, Fernández-Calero T, Ehrlich R. Silent Polymorphisms: Can the tRNA Population Explain Changes in Protein Properties? Life (2075-1729). 2016;6(1).