ISSN: 2476-6917; Modares Journal of Biotechnology. 2020;11(2):145-154

Molecular Docking and Bioinformatics Study of Rare Codons in the *Lampyroidea maculata Luciferase* Gene

ARTICLE INFO

Article Type
Original Research

Authors
Mortazavi M.*[1] *PhD*,
Hosseinkhani S.[2] *PhD*,
Torkzadeh-Mahani M.[1] *PhD*,
Lotfi S.[1] *PhD*,
Emamzadeh R.[3] *PhD*

How to cite this article
Mortazavi M, Hosseinkhani S, Torkzadeh-Mahani M, Lotfi S, Emamzadeh R. Molecular Docking and Bioinformatics Study of Rare Codons in the *Lampyroidea maculata Luciferase* Gene. Modares Journal of Biotechnology. 2020;11(2):145-154.

[1]Biotechnology Department, Institute of Science and High Technology and Environmental Science, Graduate University of Advanced Technology, Kerman, Iran
[2]Biochemistry Department, Biosciences Faculty, Tarbiat Modares University, Tehran, Iran
[3]Biology Department, Sciences Faculty, University of Isfahan, Isfahan, Iran

***Correspondence**
Address: Graduate University of Advanced Technology, End of Haft-Baghe-Alavi Highway, Kerman, Iran. Postal code: 7631818356
Phone: +98 (34) 33776611
Fax: +98 (34) 33776617
m.mortazavi@kgt.ac.ir

Article History
Received: June 9, 2019
Accepted: December 28, 2019
ePublished: June 10, 2020

ABSTRACT

The bioluminescence process is a widespread phenomenon in nature. The luciferase enzymes are identified in some domains of life, but the luciferases from the Lampyridae family are considered for biological applications. The molecular cloning of a new type of Iranian firefly luciferase from *Lampyroidea maculata* was reported, previously. In this study, the rare codons of the Iranian insect *luciferase* gene were analyzed using the computational databases as ATGme, RACC, LaTcOm, and Sherlocc. Also, the structural modeling process of this enzyme was performed. Next, the status of these rare codons in this structural model was evaluatedusing SPDBV 4.10 and PyMOL 2.3.2 software. In the following, the substrate binding site in the enzyme's active site was studied using the AutoDock Vina 1.5.4. By molecular modeling, some rare codons were identified that may have a critical role in the structure and function of this luciferase. AutoDock Vina was used in the molecular docking that recognizes Asp531 that yield closely related to luciferin and AMP binding site. These bioinformatics analyzes play an important role in the design of new drugs.

Keywords *Lampyroidea maculate*; Luciferase; Rare Codon; Substrate Binding Site

CITATION LINKS

[1] Substrate-binding properties ... [2] Kinetics of the firefly luciferase ... [3] The chemi-and bioluminescence of firefly luciferin: An efficient chemical production of electronically excited ... [4] The origin, diversity, and structure ... [5] Molecular cloning, sequence analysis, and expression of a cDNA encoding the luciferase from the glow-worm ... [6] Surface charge modification increases firefly luciferase rigidity without alteration in bioluminescence ... [7] cDNA cloning, expression and homology modeling of ... [8] Molecular enigma of multicolor bioluminescence ... [9] Site-directed mutagenesis of firefly luciferase ... [10] Use of firefly luciferase in ATP-related assays of biomass ... [11] Codon choice and gene expression: Synonymous ... [12] Large-scale analysis of conserved rare codon clusters suggests an involvement in co-translational molecular recognition ... [13] Analysis of the distribution of functionally relevant ... [14] A part of codon bias in genes protects protein spatial structures from destabilization by random single point ... [15]ATGme: Open-source web application for rare codon identification and custom DNA sequence ... [16] LaTcOm: A web server for visualizing ... [17] Protein secondary structural types are ... [18] SWISS-MODEL and the Swiss-Pdb Viewer: An environment for comparative protein ... [19] I-TASSER server for protein 3D structure ... [20] Swiss-PDB viewer ... [21] The PyMOL molecular graphics ... [22] AutoDock Vina: Improving the speed and ... [23] LOMETS: A local meta-threading-server ... [24] Structural basis for the spectral difference in luciferase ... [25] Structural basis for the inhibition of firefly ... [26] What if: A molecular ... [27] PIC: Protein interactions ... [28] Autodock4 And ... [29] Open Babel: An open chemical ... [30] Co-variation of trna abundance and ... [31] Rare codons ... [32] Luc genes: Introduction of colour ... [33] Bioluminescent assays for high-throughput ... [34] Luminogenic enzyme substrates: The ... [35] Imaging of light emission from the expression ... [36] The generation of a conditional ... [37] Novel application of pH-sensitive ... [38] Bioinformatic identification of rare codon clusters (RCCs) in HBV genome ... [39] The maximum rate of gene ... [40] Generic algorithm to predict ... [41] Speeding with control: Codon ... [42] Effects of rare codon clusters ... [43] Synonymous codon substitutions ... [44] Recombinant protein ... [45] Ligplot: A program to generate schematic diagrams ... [46] Crystal structure of firefly luciferase in a second catalytic conformation ... [47] Molecular basis for the high-affinity binding ... [48] Molecular docking and rare codons evaluation in the luciola lateralis ... [49] Molecular basis for the high-affinity binding ...

Copyright© 2020, TMU Press. This open-access article is published under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License which permits Share (copy and redistribute the material in any medium or format) and Adapt (remix, transform, and build upon the material) under the Attribution-NonCommercial terms.

داکینگ مولکولی و مطالعه بیوانفورماتیک کدون‌های نادر ژن *لوسیفراز* ***Lampyroidea maculata***

مجتبی مرتضوی* PhD
گروه بیوتکنولوژی، پژوهشگاه علوم و تکنولوژی پیشرفته و علوم محیطی، دانشگاه تحصیلات تکمیلی صنعتی و فناوری پیشرفته، کرمان، ایران
سامان حسینخانی PhD
گروه بیوشیمی، دانشکده علوم زیستی، دانشگاه تربیت مدرس، تهران، ایران
مسعود ترک‌زاده‌ماهانی PhD
گروه بیوتکنولوژی، پژوهشگاه علوم و تکنولوژی پیشرفته و علوم محیطی، دانشگاه تحصیلات تکمیلی صنعتی و فناوری پیشرفته، کرمان، ایران
صفا لطفی PhD
گروه بیوتکنولوژی، پژوهشگاه علوم و تکنولوژی پیشرفته و علوم محیطی، دانشگاه تحصیلات تکمیلی صنعتی و فناوری پیشرفته، کرمان، ایران
رحمان امام‌زاده PhD
گروه زیست‌شناسی، دانشکده علوم، دانشگاه اصفهان، اصفهان، ایران

چکیده
فرآیند بیولومینسانس، یک پدیده گسترده در طبیعت است و آنزیم‌های لوسیفرازی در بخش‌های گسترده‌ای از حیات شناسایی شده‌اند. اما آنزیم لوسیفراز شناسایی‌شده در خانواده Lampyrid به‌دلیل ویژگی‌های برتر کاربردهای زیستی گسترده‌ای یافته است. به‌تازگی، کلونینگ یک ژن جدید از آنزیم لوسیفراز کرم شب‌تاب ایرانی با نام *Lampyroidea maculate* گزارش شده است. در این مطالعه، آنالیز کدون‌های نادر از ژن *لوسیفراز* حشره شب‌تاب ایرانی با استفاده از پایگاه‌های محاسباتی ATGme، RACC، LaTcOm و Sherlocc مورد بررسی قرار گرفت. همچنین، فرآیند مدل‌سازی ساختاری این آنزیم انجام شد. در ادامه، وضعیت این کدون‌های نادر در این مدل ساختاری به کمک نرم‌افزارهای SPDBV 4.10 و PyMOL 2.3.2 ارزیابی شد. جایگاه اتصال سوبسترا در دهانه فعال آنزیم با استفاده از نرم‌افزارهای داکینگ AutoDock Vina 1.5.4 بررسی شد. به کمک مدل‌سازی مولکولی، برخی از کدون‌های نادر شناسایی شدند که ممکن است نقش مهمی در ساختار و عملکرد این آنزیم داشته باشند. فرآیند داکینگ مولکولی به کمک AutoDock Vina انجام و Asp^{531} که در اتصال به لوسیفرین و AMP نقش دارد شناسایی شد. این آنالیزهای بیوانفورماتیکی نقش مهمی در طراحی داروهای جدید دارد.

کلیدواژه‌ها: *Lampyroidea maculata*، لوسیفراز، کدون نادر، جایگاه اتصال سوبسترا

تاریخ دریافت: ۱۳۹۸/۳/۱۹
تاریخ پذیرش: ۱۳۹۸/۱۰/۷
*نویسنده مسئول: m.mortazavi@kgt.ac.ir

مقدمه

آنزیم لوسیفراز حشره شب‌تاب (EC. ۱/۱۳/۱۲/۷)، مسئول در واکنش‌های بیولومینسانس است. این آنزیم واکنش اکسیداسیون لوسیفرین با اکسیژن مولکولی را در حضور ATP و Mg^{2+} کاتالیز می‌کند[1,2]. در این واکنش با تشکیل لوسیفریل- آدنیلات توسط لوسیفراز و واکنش آن با اکسیژن مولکولی، مولکول‌های PPi، CO، AMP و اکسی‌لوسیفرین جدا و نوری با طیف ناحیه زرد- سبز منتشر می‌شود[3]. قبلاً، ژن آنزیم لوسیفراز از گونه‌های مختلف حشره شب‌تاب در نقاط مختلف شناسایی و جدا شده و ویژگی‌های آنها نیز مطالعه شده است[4-6]. در مطالعات قبلی، دو توالی cDNA کدکننده ژن *لوسیفراز* از حشره‌های گونه ایرانی *Lampyris turkestanicus* و *Lampyroidea maculata*، (جنس Lampyrid) کلون شده‌اند[5, 7]. تاکنون مکانیزم واکنش بیولومینسانس به‌طور کامل شناسایی نشده است، اما برخی از رزیدوهای درگیر در واکنش‌های اتصال و کاتالیز شناسایی شده‌اند[8, 9]. با توجه به اینکه سنجش‌های مبتنی بر واکنش آنزیمی حشره شب‌تاب سریع، حساس و غیررادیواکتیو هستند، استفاده از این آنزیم به‌طور گسترده‌ای در زمینه‌های مختلف بیوتکنولوژی از جمله تشخیص ATP، تهیه کیت تشخیص سرطان و گزارشگرهای ژنتیکی کاربرد یافته است[10].

مطالعات نشان داده‌اند که برخی از کدون‌ها در بیان ژن‌ها با فراوانی بیشتری در بیان پروتئین‌های سلولی استفاده می‌شوند، در حالی که از تعدادی کدون‌ها تقریباً استفاده‌ای نمی‌شود که به آنها کدون‌های نادر گفته می‌شود[11]. علاوه بر این، برخی از گزارش‌ها نشان می‌دهند که در فرآیند ترجمه به‌دلیل پایین‌بودن میزان tRNAهای کدون‌های نادر، در موضع آنها، ریبوزوم تا زمانی که tRNA مربوطه فراهم شود توقف کوتاهی دارد[12]. مطالعات نشان می‌دهند که توقف‌های ریبوزومی در کدون‌های نادر در فرآیند فعالیت و تاخوردگی پروتئین‌ها نقش مهمی دارند[13]. ارزیابی کدون‌های نادر، می‌تواند بینش‌های جدیدی در حل مسایل مربوط به حوزه چالش‌های پروتئینی فراهم کند. توالی‌های نوکلئوتیدی کدکننده اسیدهای آمینه، علاوه بر انتقال اطلاعات ترتیب اسیدهای آمینه، حامل اطلاعات دیگری برای مکانیزم های مختلف مولکولی هستند[14]. مطالعات اخیر نشان می‌دهند که لایه‌های دیگری از اطلاعات علاوه بر کدینگ اسیدهای آمینه در درون توالی کدکننده پروتئین‌ها وجود دارند که در ظاهر دیده نمی‌شود، اما در میانجی‌گری فرآیند ترجمه نقش دارند[12]. تجزیه و تحلیل کدون‌ها همچنین به درک تعامل بین ویروس‌های RNAدار و پاسخ ایمنی میزبان کمک می‌کند. با توجه به این اطلاعات، در فرآیند مهندسی پروتئین آنزیم لوسیفرازی لازم است که دانش کاملی از موقعیت کدون‌های نادر این آنزیم در اختیار باشد. در فرآیندهایی همانند جهش‌زایی، بیان مناسب آنزیم لوسیفراز و جلوگیری از غیرفعال‌شدن لوسیفراز طراحی‌شده، توجه به موقعیت کدون‌های نادر با توجه با تاثیر آنها بر فرآیند فولدینگ صحیح آنزیم، بسیار حایز اهمیت است. مطالعه این اطلاعات پنهان در توالی کدونی می‌تواند فرآیند تکامل مولکولی موجودات را نشان دهد و همچنین بینشی روی تاریخچه ژن‌ها در ژنوم فراهم کند[12]. در مطالعه حاضر، بررسی‌های بیوانفورماتیکی کدون‌های نادر از ژن *لوسیفراز* حشره شب‌تاب ایرانی *L. maculata* (DQ137139) با استفاده از پایگاه‌های محاسباتی مختلف انجام شد[7, 15-17]. همچنین فرآیند مدل‌سازی ساختاری در پایگاه‌های مدل‌سازی Swiss Model و I-TASSER اجرا[18, 19] و ویژگی این کدون‌های نادر به کمک نرم‌افزارهای SPDBV 4.10 و PyMOL 2.3.2 مطالعه شد[20, 21]. در فرآیند شبیه‌سازی مولکولی

اتصال ATP و لوسیفرین به جایگاه فعال آنزیم از AutoDock Vina استفاده شد[22]. ارزیابی این اطلاعات پنهان (کدون‌های نادر) می‌تواند در نشان‌دادن نقش این کدون‌ها در ساختار و عملکرد لوسیفراز کرم شب‌تاب کمک کند.

مواد و روش‌ها

تشخیص خوشه کدون نادر در ژن و ساختار لوسیفراز

برای بررسی کدون‌های نادر، شماره Pfam آنزیم لوسیفراز خانواده Lampyrid در پایگاه داده Uniprot (http://www.uniprot.org) به دست آمد. Pfam یک پایگاه داده از خانواده پروتئین و با استفاده از مدل مارکوف تولید شده است. این شماره Pfam در برنامه محاسباتی کدون‌های نادر همانند Sherlocc آنالیز شد. توالی نوکلئوتیدی از لوسیفراز *L. maculata* از NCBI به دست آمد (https://www.ncbi.nlm.nih.gov/nuccore/DQ13713 9.1) و در نرم‌افزارهای تحت وب تشخیص کدون‌های نادر از جمله ATGme، LaTcOm و RACC مورد ارزیابی قرار گرفت. ابتدا به کمک جدول کدونی باکتری *E. coli* B [gbbct]: 11 CDS که از پایگاه داده‌ای جدول کدونی kazusa (http: //www.kazusa.or.jp) به دست آمد، کدون‌های نادر و بسیار نادر در ATGme شناسایی شدند. در ادامه، کدون آمینواسیدهای مشکل‌ساز در توالی اسیدآمینه از جمله آرژنین، لوسین، ایزولوسین و پرولین در پایگاه محاسباتی RACC ارزیابی و مشخص شدند. با توجه به ویژگی‌های خاص و تاثیرات این کدون‌های نادر، پایگاه محاسباتی آنها را به‌عنوان آمینواسیدهای مشکل‌ساز در نظر می‌گیرد. در نهایت با استفاده از سه الگوریتم تشخیص کدون‌های نادر به‌عنوان MSS، MINMAX و در پایگاه محاسباتی LaTcOm وضعیت خوشه‌های کدونی نادر لوسیفراز *L. maculata* مورد آنالیز قرار گرفت[16].

مطالعه کدون نادر در ساختار لوسیفراز

فرآیند مدل‌سازی آنزیم لوسیفراز در پایگاه‌های محاسباتی Swiss Model و همچنین پایگاه I-TASSER براساس هم‌ترازی چندگانه رشته‌ای انجام شد[18, 19, 23]. مدل‌های مناسب براساس شاخص‌های ویژه Confidence Score و Z-score انتخاب شدند. سپس وضعیت این کدون‌های نادر با استفاده از نرم‌افزارهای مشاهده ساختارها مانند PyMOL و SPDBV مورد ارزیابی قرار گرفت[18, 21]. در فرآیند مدل‌سازی، ساختار کریستال آنزیم‌های لوسیفرازی همانند لوسیفراز ژاپنی مقاوم به حرارت (PDB: 2D1S) و حشره شب‌تاب آمریکایی *Photinus pyralis* (PBD: 1ba3) به‌عنوان الگو در این پایگاه داده مدل‌سازی مولکولی استفاده شد[24, 25]. پارامترهای فیزیکوشیمیایی این مدل توسط سرور ProtParam Expasy (http://us.expasy.org/tools/ protparam.html) مورد مطالعه قرار گرفتند. همچنین میان‌کنش‌های هیدروژنی این لوسیفراز با استفاده از پایگاه‌های What IF و PIC محاسبه و بررسی شدند[26, 27].

داکینگ مولکولی با استفاده از **AutoDock Vina 1.5.4** برای بررسی جایگاه اتصال سوبستراهای AMP و لوسیفرین در ساختار مدل‌سازی‌شده لوسیفراز، فرآیند داکینگ بین آنزیم لوسیفراز، لوسیفرین و AMP مورد مطالعه و بررسی قرار گرفت. در AutoDock Vina 1.5.4، این مدل ساختاری سه بعدی آنزیم لوسیفراز با استفاده از نرم‌افزار MGL 1.5.6 به فرمت PDBQT به‌عنوان پذیرنده تبدیل شد[28, 22]. همچنین فرمت SDF سوبستراهای لوسیفرین و ATP از پایگاه داده PubChem (https://pubchem.ncbi.nlm.nih.gov) به دست آمد و به کمک نرم‌افزار Open Babel 2.3.1 به فرمت PDB و با استفاده از ابزار MGL 1.5.4 به فرمت PDBQT تبدیل شد[29]. فضای اتصال سوبسترا در داخل آنزیم مشابه با بستر محل اتصال سوبسترا به آنزیم لوسیفراز *P. pyralis* که ساختار کریستال آن تعیین شده است انتخاب شد[25]. فضای مربع جست‌وجو با مشخصات طول، عرض و ارتفاع برابر x= ۱۴ و y= ۱۴ و z= ۱۰ انتخاب شد. با استفاده از ابزار MGL 1.5.6، بهترین نتیجه داکینگ از اتصال لوسیفراز و لوسیفرین انتخاب و به فرمت PDBQT تبدیل شد. در ادامه، داکینگ ثانویه بین فرمت‌های PDBQT و AMP مورد مطالعه قرار گرفت. آزمایش اتصال به ارزش جامعیت ۲۵ انجام شد.

یافته‌ها

تشخیص خوشه کدون نادر در ژن و ساختار لوسیفراز

شماره Pfam آنزیم لوسیفراز *L. maculata* در پایگاه داده‌ای Uniprot به‌صورت Q1WLP6 شناسایی شد. این شماره در برنامه Sherlocc برای تشخیص خوشه کدون نادر بررسی شد[17] که نتایج نشان دادند که هیچ خوشه کدون نادری که قابل شناسایی باشد در توالی پروتئینی این خانواده آنزیمی لوسیفراز وجود ندارد.

در ادامه، توالی نوکلئوتیدی *لوسیفراز L. maculata* به‌منظور بررسی کدون‌های نادر آن در سرور ATGme (http://atgme.org) مورد مطالعه قرار گرفت. با استفاده از جدول کدونی *E.coli* B [gbbct]: 11 CDS، این توالی نوکلئوتیدی ارزیابی شد. نتایج نشان می‌دهند که میزان GC در توالی اصلی حدود ۳۸% و میزان AT حدود ۶۲% است. شکل ۱ کدون‌های نادر و بسیار نادر که به ترتیب با رنگ نارنجی و قرمز مشخص شده‌اند را نشان می‌دهد. برای درک بهتر این نتایج، کدون آمینواسیدهای مشکل‌ساز از جمله Arg، Ile، Leu و Pro در این توالی *لوسیفرازی* در سرور RACC مورد بررسی قرار گرفت. مطالعه حاضر نشان داد که این ژن دارای شش کدون نادر تنها از Arg، سه کدون نادر از Leu، سه کدون نادر از Ile و شش کدون نادر از Pro است (شکل ۱).

توالی نوکلئوتیدی لوسیفراز *L. maculata* در پایگاه محاسباتی LaTcOm تجزیه و تحلیل شد. این پایگاه از سه الگوریتم MINMAX، sliding window و MSS استفاده می‌کند. برای تجزیه و تحلیل داده‌ها از جدول کدونی *دانگ* به‌عنوان مقیاس مرجع استفاده شد (نمودار ۱) [30].

در این تحلیل، براساس معیارهای داخلی، خوشه‌های کدون‌های نادر مختلفی شناسایی شدند. سپس براساس اهمیت این کدون‌ها و با توجه به موقعیت ساختاری مهم، برخی از آنها انتخاب و در ساختار آنزیم لوسیفراز مطالعه شدند[31].

ATG GAA ATG GAA **AAG** GAG GAG AAT GTT GTA TAT GGC **CCT** CAG **CCA** TTC **TAC** **CCA** ATT GAA AAA **GGA** **TCC** GCA **GGA** ATT CAG TTG
M E M E K E E N V V Y G P Q P F Y P I E K G S A G I Q L
CAT **AAG** **TAC** ATG **CAT** CAA TAT GCC CAA CTT **GGA** GCA ATT GCT TTT **AGT** AAC GCC CTT **ACT** **GGA** GTG GAT ATT **TCT** **TAC** CAA GAA
H K Y M H Q Y A Q L G A I A F S N A L T G V D I S Y Q E
TAC TTT GAT ATT **TCA** **TGT** CGT TTA GCT GAG GCA ATG GAA AAC TAT GGT ATG AAA CCG GAA **GGA** CGT ATT GCT TTG **TGC** **AGT** GAA
Y F D I S C R L A E A M E N Y G M K P E G R I A L C S E
AAT **TGT** GAA GAA TTT TTC ATC **CCT** GTA CTT GCT GGT CTT **TAC** ATT GGG GTG GGT GTT GCA **CCT** **ACT** AAT GAA ATT **TAC** **ACT** TTA
N C E E F F I P V L A G L Y I G V G V A P T N E I Y T L
CGT GAA CTT AAT **CAC** **ACT** TTG GGC ATC GCA GAA **CCA** **ACT** ATT GTA TTC AGC **TCC** AAA AAA GGC TTA **CCT** AAA GTT TTA GAA GTG
R E L N H T L G I A E P T I V F S S K K G L P K V L E V
CAA AAA **ACA** GTT **ACA** **TGC** ATC AAA **ACA** ATT GTT ATT TTA AAT **AGT** AAA GTA AAC TTT **GGA** GGT **TAC** GAT **TGT** GTG GAA **ACT** TTT
Q K T V T C I K T I V I L N S K V N F G G Y D C V E T F
ATT **AAG** AAA AAT GTA GAA TTA GGT TTT CAA **CCA** **ACT** AGC TTT AAA **CCC** ATT GAT GTA **AAG** AAC CGT AAA GAA **CAC** GTT GCT TTG
I K K N V E L G F Q P T S F K P I D V K N R K E H V A L
ATT ATG AAT **TCT** **TCT** GGC **TCT** **ACT** GGT TTA **CCT** AAA GGT GTA CAA ATT **ACA** **CAC** GAA GCT **ACA** GTT **ACA** **AGA** TTC **TCA** **CAC** GCT
I M N S S G S T G L P K G V Q I T H E A T V T R F S H A
AAG GAT **CCA** ATT **TAC** **GGA** AAC CAA GTT **TCA** **CCT** GGT **ACT** GCT ATT TTA **ACT** GTC GTT CCG TTC **CAT** **CAT** **GGA** TTT GGC ATG TTT
K D P I Y G N Q V S P G T A I L T V V P F H H G F G M F
ACC **ACT** TTA **GGA** **TAC** TTT GCT **TGT** **GGA** **TAC** CGT ATT GTC ATG TTA **ACA** AAA TTC GAT GAA GAA **ATA** TTT TTG AAA **ACT** ATG CAA
T T L G Y F A C G Y R I V M L T K F D E E I F L K T M Q
GAT **TAC** AAA **TGT** ACC **AGT** GTT ATT CTT GTA **CCA** ACG TTA TTT GGT ATT CTC AAC **AAG** **AGT** GAA TTG ATC GAT **AAG** TTC GAT TTA
D Y K C T S V I L V P T L F G I L N K S E L I D K F D L
TCT AAT **CTA** **ACT** GAA ATT GCT **TCC** GGT **GGA** GCT **CCT** TTG GCC AAA GAA GTT GGT GAA GCA GTC GCT **AGA** **AGA** TTT AAT **CTA** **CCC**
S N L T E I A S G G A P L A K E V G E A V A R R F N L P
GGT GTC CGC CAG GGT **TAC** **GGA** TTA **ACA** GAA ACG **ACA** **TCT** GCA TTT ATT ATT **ACA** **CCA** GAG GGT GAT GAT AAA **CCT** **GGA** GCA **TCT**
G V R Q G Y G L T E T T S A F I I T P E G D D K P G A S
GGA AAA GTA GTA **CCC** TTA TTT AAA GTA AAA GTT ATT GAT CTT GAT **ACT** AAA AAA **ACT** TTG GGT GCC AAC **CGA** **CGA** **GGA** GAG ATT
G K V V P L F K V K V I D L D T K K T L G A N R R G E I
TGT GTA AAA GGT **CCC** **AGT** CTT ATG AAA GGG **TAC** **ATA** AAC AAT CCG GAA GCA **ACA** AAA GAA ATT ATT GAT GAA GAG GGT TGG ATG
C V K G P S L M K G Y I N N P E A T K E I I D E E G W M
CAC **ACT** **GGA** GAT ATT GGT TAT **TAC** GAC GAA GAC GAA **CAT** TTC TTC ATT GTA GAT CGT TTG AAA **TCA** TTA ATC AAA **TAC** AAA GGG
H T G D I G Y Y D E D E H F F I V D R L K S L I K Y K G
TAC CAG GTA **CCA** **CCC** GCT GAA CTG GAA **TCA** GTT CTT TTG CAA **CAT** **CCA** GAT **ATA** TTT GAT GCT GGT GTG GCT GGT GTC **CCC** GAT
Y Q V P P A E L E S V L L Q H P D I F D A G V A G V P D
CCT GAA GCT GGT GAA CTT **CCA** GGG GCT GTA GTT GTA ATG GAA AAA **GGA** AAA **ACT** ATG **ACT** GAA **AAG** GAA ATT GTG GAT TAT GTT
P E A G E L P G A V V V M E K G K T M T E K E I V D Y V
AAT **AGT** CAA GTA GTG AAC **CAC** AAA CGT CTG CGT GGT GGC GTT CGT TTT GTG GAT GAA GTA **CCT** AAA GGT **CTA** **ACT** GGT AAA ATT
N S Q V V N H K R L R G G V R F V D E V P K G L T G K I
GAT GCT AAA GTA ATT **AGA** GAG ATT CTT **AAG** AAA **CCA** CAA GCC **AAG** ATG
D A K V I R E I L K K P Q A K M

شکل ۱) الف) وضعیت کدون‌های نادر و بسیار نادر به‌صورت نارنجی و قرمز نشان داده شده است؛ ب) وضعیت کدون‌های نادر آمینواسیدهای آرژنین، ایزولوسین، لوسین و پرولین در توالی لوسیفراز *L. maculata* که با رنگ‌های قرمز و نارنجی نشان داده شده است.

الف

ب

ج

نمودار ۱) موقعیت خوشه‌های کدونی نادر در ژن *لوسیفراز* *L. maculata*؛ نتایج نشان می‌دهند که تحلیل این ژن در الگوریتم‌های MSS (الف)، minmax (ب) و sliding window (ج) میزان مختلفی از خوشه کدون نادر را محاسبه می‌کند که نشان‌دهنده قابلیت متنوع این الگوریتم‌ها در ارزیابی خوشه‌های کدونی نادر است.

اتصال سوبسترا- آنزیم

برای انجام مطالعه داکینگ سوبسترا- آنزیم، ساختار آنزیم لوسیفراز در سرور I-TSSAR مدل‌سازی شد. بهترین مدل از بین مدل‌های تولیدشده با مشخصات 0.96+-0.05 value of TM-Score, 1.76 value of overall C-score, and Exp. RMSD of 3.8+-2.6 انتخاب شد. خواص فیزیکوشیمیایی آنزیم لوسیفراز توسط پایگاه محاسباتی ProtParam آنالیز شد (جدول ۱).

مطالعه و مقایسه مدل لوسیفراز *L. maculata*. با ساختار کریستالوگرافی آنزیم لوسیفراز گونه *P. pyralis* (1ba3) نشان می‌دهد که جایگاه‌های فعال آنها با هم مشابه هستند[25]. به‌منظور مطالعه جایگاه اتصال لوسیفرین و AMP، مطالعات داکینگ با استفاده از نرم‌افزار AutoDock Vina 1.5.4 انجام شد[22]. مدل مولکولی لوسیفراز *L. maculata* به‌عنوان گیرنده و سوبستراهای لوسیفرین و AMP به‌عنوان یک لیگاند مولکول کوچک مورد ارزیابی قرار گرفتند. همان‌طور که گفته شد، فضای جست‌وجو با توجه به جایگاه فعال آنزیم لوسیفراز گونه *P. pyralis* (1ba3) طراحی شد. نتایج حاصل از داکینگ مولکولی با اندازه‌های مختلف از مربع جست‌وجو مورد بررسی قرار گرفت. نتایج داکینگ از لوسیفراز و لوسیفرین نشان می‌دهند که کمپلکس لوسیفراز- لوسیفرین دارای شبکه‌ای از میان‌کنش‌های غیرکووالان بین آنزیم و لوسیفرین است (شکل ۲).

در نهایت، کمپلکس آنزیم لوسیفراز- لوسیفرین به فرمت PDBQT تبدیل و به‌عنوان گیرنده در فرآیند داکینگ دوم با AMP به‌عنوان لیگاند در نظر گرفته شد. کمپلکس آنزیم لوسیفراز- لوسیفرین- AMP که از نتایج داکینگ به دست آمد، در شکل ۳ نشان داده شده است. طبق این شکل، شبکه‌ای از میان‌کنش‌های غیرکووالان مانند میان‌کنش هیدروژنی در این کمپلکس مشاهده می‌شود. همچنین چند میان‌کنش هیدروژنی می‌تواند بین این کمپلکس و آنزیم تشکیل شود. این کمپلکس نیز با آمینواسیدهای مختلف آنزیم لوسیفراز از طریق برخی فعل و انفعالات آب‌گریز و واندروالس در تعامل است.

جدول ۱) مطالعه خواص فیزیکوشیمیایی آنزیم لوسیفراز (آنالیزشده توسط پایگاه محاسباتی ProtParam)

شماره	ویژگی‌ها	لوسیفراز
۱	pH ایزوالکتریک تئوری (Theoretical pI)	۶/۰۲
۲	وزن مولکولی	۶۰۴۵۴/۸۱
۳	طول توالی	۵۴۸
۴	ضریب خاموشی(Extinction coefficients (M-1 cm-1at 260nm)*	۳۷۲۹۰-۳۶۷۹۰
۵	مجموع آسپارتیک‌اسید و گلوتامیک‌اسید	۶۸
۶	مجموع لیزین و آرژنین	۶۱
۷	شاخص ناپایداری	۳۱/۱۶
۸	میانگین آب‌گریزی (Grand average of hydropathicity)	-۰/۱۱۶
۹	شاخص آب‌دوستی	۹۰/۳۱

* محاسبه ضریب خاموشی در دو حالت انجام می‌شود. فرض بر این است که در حالت نخست (۳۷۲۹۰) هر دو سیستئین تشکیل سیستین و در حالت دوم (۳۶۷۹۰) هر دو سیستئین در حالت احیا هستند.

الف

ب

ج

شکل ۲) الف) نمایش فرآیند داکینگ و اتصال لوسیفرین به لوسیفراز انجام‌شده در PyMOL (رنگ آبی ساختار آنزیم لوسیفراز و بخش قرمز لوسیفرین را نشان می‌دهد.)؛ ب) نمایش میان‌کنش‌های قطبی به‌صورت خطوط فیروزه‌ای رنگی در نرم‌افزار LIGPLOT 2.1؛ ج) نمایش میان‌کنش‌های ناحیه اتصال لوسیفرین به لوسیفراز انجام‌شده در PyMOL (بخش سبز لوسیفرین را نشان می‌دهد.)

الف

ب

ج

شکل ۳) الف) نمایش فرآیند داکینگ و اتصال AMP به کمپلکس لوسیفراز-لوسیفرین انجام‌شده در PyMOL (رنگ آبی ساختار آنزیم لوسیفراز، بخش سبز لوسیفرین و بخش قرمز AMP را نشان می‌دهد.)؛ ب) نمایش میان‌کنش‌های قطبی به‌صورت خطوط فیروزه‌ای رنگی در نرم‌افزار LIGPLOT 2.1؛ ج) نمایش میان‌کنش‌های ناحیه اتصال AMP به کمپلکس لوسیفراز- لوسیفرین انجام‌شده در PyMOL (بخش قرمز AMP و بخش سبز لوسیفرین را نشان می‌دهد.)

مطالعه کدون‌های نادر در ساختار آنزیم لوسیفراز

مکان‌هایی از این کدون‌های نادر در مدل‌های ساختاری آنزیم لوسیفراز بررسی شدند. این مطالعه نشان می‌دهد که کدون‌های نادر در مناطق مختلف ساختار لوسیفراز واقع شده‌اند. این ساختار به همراه برخی کدون‌های نادر آرژنین در موقعیت‌های ۲۲۰، ۳۳۱، ۳۳۲، ۳۸۸، ۳۸۹ و ۵۳۸ و سوبستراهای لوسیفرین و AMP در شکل ۴ نشان داده شده‌اند. درواقع این شکل موقعیت مکان‌هایی از این کدون‌های نادر را در آنزیم لوسیفراز نشان می‌دهد. کدون‌های نادر Arg میان‌کنش‌های هیدروژنی با آمینواسیدهای دیگر آنزیم لوسیفراز برقرار می‌کنند و بررسی اولیه در محل این کدون‌های نشان‌دهنده نقش مهم این آمینواسید در فرآیند فولدینگ ساختار این آنزیم است.

شکل ۴) نمایش شماتیک ساختار آنزیم لوسیفراز همراه با لوسیفرین (سبز)، AMP (قرمز) و آمینواسید کدون نادر آرژنین (زرد)

بحث

امروزه در حوزه علوم زیستی، از واکنش لوسیفراز و نشر نور می‌توان برای شناسایی و مشاهده فرآیندهای بیولوژیکی استفاده کرد[32]. همچنین می‌توان ژن *لوسیفراز* را از طریق مهندسی ژنتیک به داخل ارگانیزم و یا سلول‌ها برای اهداف بیولوژیک منتقل کرد. در تحقیقات بیولوژیک، از ژن *لوسیفراز* معمولاً به‌عنوان یک گزارشگر برای ارزیابی فعالیت رونویسی در سلول[33]، تشخیص سطح ATP سلولی در سنجش میزان حیات سلولی[34]، تصویربرداری کامل حیوانی[35] و ردیابی تومورزایی و پاسخ تومور به درمان در مدل‌های حیوانی[36] استفاده می‌شود. با این حال عوامل محیطی و درمانی ممکن است در شدت نشر نور ایجادشده در این موارد اختلال ایجاد کنند. علاوه بر این، شدت سیگنال به عوامل مختلفی مانند pH داخل سلولی و میزان لوسیفراز تاخورده مناسب نیز بستگی دارد[37].

پیش از این، فرآیند کلونینگ مولکولی و بیان *لوسیفراز* *L. maculata* انجام شده است[7]. اگرچه تاکنون ابزارهای مولکولی مختلفی برای بیان پروتئین‌های نوترکیب در باکتری *E. coli* ایجاد شده‌اند، ارزیابی چالش‌های مشکل‌ساز جدید در بیان پروتئین

همانند کدون‌های نادر، بسیار حیاتی هستند. علاوه بر این ارزیابی محل اتصال سوبسترا نقش خاصی در مهندسی پروتئین برای جهش‌زایی هدف‌دار و بیان پروتئین دارد. از آنجا که هیچ مطالعه مشابهی روی این آنزیم وجود ندارد، بررسی جامعی از کدون‌های نادر و محل اتصال سوبسترا در سازه ژنی و پروتئینی این آنزیم انجام شد. پیش از این یک تجربه موفق از این نوع مطالعات روی ویروس هپاتیت به دست آمده است[38].

اگرچه این آنزیم لوسیفراز جدید را می‌توان برای کاربردهای بیولوژیک در نظر گرفت، اما برخی از موضوعات حل‌نشده آن همانند جایگاه اتصال سوبسترا وجود دارد. برای ارزیابی بهتر مکانیزم کاتالیستی، برخی از رویکردها همانند داکینگ مورد توجه قرار گرفته است. آنزیم لوسیفراز با بهره‌گیری از ATP و لوسیفرین به کمک اسیدهای آمینه خاصی باعث تولید نور می‌شود که نقش مهمی در این فرآیند دارند. در طراحی جهش به‌دلیل موقعیت‌های ساختاری، این آمینواسیدهای جایگاه فعال همراه با کدون‌های نادر باید از طریق مطالعه بیوانفورماتیک ارزیابی شوند. در این راستا موقعیت ساختاری و میان‌کنش‌های این آمینواسیدهای مهم در ژن و پروتئین لوسیفراز مورد مطالعه قرار گرفت. ساختار آنزیم لوسیفراز *L. maculata* تاکنون تعیین نشده است و ابتدا لازم است که یک مدل ساختاری اولیه از این آنزیم به کمک نرم‌افزارهای بیوانفورماتیکی و پایگاه‌های اطلاعاتی تهیه شود. برای این کار توالی ژن آنزیم لوسیفراز *Lampyroidea maculata* از پایگاه NCBI تهیه و در پایگاه اطلاعاتی Swiss model مدل‌سازی این آنزیم انجام شد.

به‌طور کلی، کدون‌ها توسط tRNAهای با فراوانی بالا خوانده و با بالاترین سرعت ترجمه می‌شوند تا صحت ترجمه تضمین شود[39]. از طرف دیگر، کدون‌های نادر توسط tRNAهای با فراوانی اندک ترجمه می‌شوند. این فراوانی متفاوت سرعت ترجمه را تغییر می‌دهد[40]. در ترجمه کدون‌های نادر، ریبوزوم‌ها مکث می‌کنند تا tRNA نادر فعال شود و اسیدآمینه بعدی را وارد مسیر ترجمه کند[41]. به همین ترتیب، چندین مطالعه که روی کدون‌های نادر و جایگزینی آنها با کدون‌های معمولی انجام شده‌اند، نشان دادند که کدون‌های نادر نقش ویژه‌ای در بیان، تاخوردگی و فعالیت آنزیم‌ها و پروتئین‌های مختلف دارند[6-8, 42-44]. این مکث در ترجمه ممکن است تاخوردگی مستقل برخی از مناطق از زنجیره های پلی‌پپتیدی را تضمین کند. به‌منظور بررسی این داده‌ها، کدون‌های نادر در این آنزیم بررسی شدند. همان‌طور که نشان داده شد، پایگاه محاسباتی Sherlocc قادر به شناسایی خوشه کدونی نادر در خانواده پروتئین لوسیفراز نبود (جدول ۱). با این حال پایگاه محاسباتی ATGme حدود ۱۳۱ کدون نادر و ۲۴ کدون بسیار نادر را شناسایی کرد. این نتایج در پایگاه محاسباتی RACC به کدون‌های نادر آرژنین، لوسین، ایزولوسین و پرولین خلاصه شد. در ادامه، با سه الگوریتم minmax، MSS و sliding window در پایگاه محاسباتی LaTcOm، تعداد مختلفی از خوشه کدونی نادر در ژن *لوسیفراز* شناسایی شد. سپس این کدون‌های نادر در ژن این آنزیم تعیین شدند و جایگاه این کدون‌ها در ساختار آنزیم مورد مطالعه و ارزیابی قرار گرفت.

در ادامه، شش کدون نادر از آرژنین در موقعیت‌های ۲۲۰، ۳۳۱، ۳۳۲، ۳۸۸، ۳۸۹ و ۵۳۸ انتخاب و در ساختار مدل مولکولی آنزیم لوسیفراز مورد مطالعه قرار گرفتند. تجزیه و تحلیل بیشتر نشان داد که این کدون‌های آرژنین با سایر آمینواسیدهای آنزیم میان‌کنش غیرکووالانسی برقرار می‌کنند (داده‌ها نشان داده نشده‌اند). این میان‌کنش‌های هیدروژنی، دومین‌های مختلف آنزیم لوسیفراز را با هم نگه می‌دارند و درنتیجه شکل‌گیری آنها ممکن است وقت‌گیر باشد. این کدون‌های نادر باعث کاهش سرعت تاخوردگی آنزیم در این موقعیت‌ها می‌شوند. با این حال، برای ارزیابی این ادعا شواهد تجربی مورد نیاز است.

مطالعات داکینگ مولکولی به کمک نرم‌افزار Vina AutoDock و مدل مولکولی لوسیفراز *L. maculata* به‌عنوان ماکرومولکول و AMP و لوسیفرین به‌عنوان لیگاند انجام شد[22]. ساختار کریستالوگرافی آنزیم لوسیفراز *P. pyralis* (1ba3) نشان می‌دهد که جایگاه فعال این آنزیم مشابه مدل از لوسیفراز *L. maculata* است و از این رو داکینگ براساس جایگاه محل اتصال این آنزیم انجام شد و نتایج آن با استفاده از نرم‌افزارهای PyMOL 2.3.2 و Ligplot 2.1 مورد مطالعه قرار گرفتند[25, 45]. بهترین مدل انتخاب و اسیدهای آمینه درگیر در اتصال مطالعه و گزارش شد. در ادامه موقعیت این کدون‌ها و جایگاه فعال نسبت به هم مطالعه و نتایج آن گزارش شد.

همان‌طور که در شکل‌های ۳ و ۴ مشاهده می‌شود، لوسیفرین و AMP چند میان‌کنش هیدروژنی با ساختار دیگر لوسیفراز تشکیل می‌دهند. این فرآیند اتصال با اندازه‌های مختلف فضای جست‌وجو انجام شد. فضای اتصال سوبستراهای لوسیفرین و AMP در داخل آنزیم مشابه با فضای اتصال سوبسترا به آنزیم لوسیفراز *P. pyralis* که ساختار کریستال آن تعیین شده است انتخاب شد[25]. با وجود اینکه موقعیت کلی محل قرارگیری سوبستراها شبیه به ساختار آنزیم لوسیفراز *P. pyralis* بود، اما همه آمینواسیدهای قابل پیش‌بینی در نتایج داکینگ پیدا نشدند. امر ممکن است به‌دلیل حذف یون Mg^{2+} از فرآیند داکینگ و همچنین چرخش ناحیه C-ترمینال و ایجاد دو کنفورماسیون در آنزیم لوسیفراز *P. pyralis* باشد[46]. در آنزیم لوسیفراز *P. pyralis* به‌دلیل چرخش ساختار بعد از اتصال سوبسترا، فضای نهایی انجام واکنش تا حدودی نسبت به فضای اولیه اتصال تغییر می‌کند. همچنین یون Mg^{2+} نقش مهمی در فعالیت کاتالیزوری دارد، اما به‌دلیل مشکل‌بودن داکینگ هم‌زمان سوبسترا و Mg^{2+} این یون از فرآیند داکینگ حذف شد. البته مطالعه مجدد داکینگ در مطالعات بعدی با یون Mg^{2+} انجام خواهد شد. نتایج بررسی‌های اولیه نشان می‌دهند که شش اسیدآمینه انتخاب‌شده دارای کدون نادر از جایگاه اتصال سوبسترای لوسیفرین و AMP دور هستند. با این حال، میان‌کنش‌های ایجادشده توسط

این آرژنین‌ها، با اتصال بخش‌های مختلف ساختار آنزیم، احتمالاً در ایجاد فضای مناسب ویژه جایگاه فعال از طریق میان‌کنش‌های از راه دور نقش بازی می‌کنند. مطالعات دیگری که روی اتصال لوسیفرین و آنزیم لوسیفراز انجام شد نشان داد که انرژی اتصال بین لوسیفرین و آنزیم لوسیفراز آزاد بیشتر از انرژی اتصال بین لوسیفرین و آنزیم لوسیفراز تثبیت‌شده است[47]. بنابراین در بررسی این میان‌کنش‌ها و طراحی جهش‌ها، باید این یافته‌ها را نیز مورد توجه قرار داد. همچنین در مطالعه‌ای دیگر، به کمک نرم‌افزار AutoDock Vina، اتصال بین سوبستراهای لوسیفرین و AMP و آنزیم لوسیفراز گونه *Luciola lateralis* به کمک ساختار لوسیفراز *Luciola cruciate* بررسی شد و نتایج نشان‌دهنده جایگاه اتصال مشابهی در این دو ساختار بود[24, 48]. مطالعه دیگری نشان داد که مدل اتصال لیگاندهای PTC124-AMP و DSLC در آنزیم لوسیفراز مشابه است[49]. در این مطالعه نیز طبق ساختارهای کریستال‌شده جایگاه‌های مشابهی در اتصال سوبستراها شناسایی شدند. تمامی این مطالعات نشان می‌دهند که نحوه اتصال و نتایج برنامه داکینگ AutoDock Vina دارای دقت مناسب و قابل اعتماد برای مطالعات آزمایشگاهی و استفاده از آنها در مهندسی پروتئین هستند.

کدون‌های نادر ممکن است نقش‌هایی در تشکیل ساختار مناسب آنزیم داشته باشند که برای بررسی نیاز به طراحی جهش‌های مناسب و آنالیزهای آزمایشگاهی دارد. یکی از بهترین روش‌ها برای مطالعه فعالیت لوسیفراز، ایجاد جهش‌های جدید براساس این نتایج است. در این مطالعه برخی از آمینواسیدهایی که ممکن است یک نقش حیاتی در جایگاه اتصال یا فرآیند تاخوردگی داشته باشند شناسایی شدند. همچنین مطالعه حاضر می‌تواند در طراحی بیوسنسورهای جدید در حوزه علوم زیستی مفید واقع شود.

تشکر و قدردانی: مطالعه حاضر در قالب طرح پژوهشی شماره ۹۶/۲۳۲/ص/۷ با استفاده از اعتبارات پژوهشی پژوهشگاه علوم و تکنولوژی پیشرفته و علوم محیطی، دانشگاه تحصیلات تکمیلی صنعتی و فناوری پیشرفته (کرمان؛ ایران) انجام شده است و بدین وسیله از این پژوهشگاه تشکر می‌شود.

تاییدیه‌اخلاقی: موردی گزارش نشده است.

تعارض منافع: هیچ‌گونه تعارض منافعی وجود ندارد.

سهم نویسندگان: مجتبی مرتضوی (نویسنده اول)، نگارنده مقدمه/روش‌شناس/پژوهشگر اصلی/تحلیلگر آماری/نگارنده بحث (۵۰%)؛ سامان حسینخانی (نویسنده دوم)، پژوهشگر کمکی/نگارنده بحث (۱۰%)؛ مسعود ترک‌زاده‌ماهانی (نویسنده سوم)، روش‌شناس/تحلیلگر آماری (۲۰%)؛ صفا لطفی (نویسنده چهارم، تحلیلگر آماری (۱۰%)؛ رحمان امام‌زاده (نویسنده پنجم)، روش‌شناس (۱۰%)

منابع مالی: مطالعه حاضر در قالب طرح پژوهشی شماره ۹۶/۲۳۲/ص/۷ با استفاده از اعتبارات پژوهشی پژوهشگاه علوم و تکنولوژی پیشرفته و علوم محیطی، دانشگاه تحصیلات تکمیلی صنعتی و فناوری پیشرفته (کرمان؛ ایران) انجام شده است.

منابع

1- Denburg JL, Lee RT, McElroy WD. Substrate-binding properties of firefly luciferase: I. Luciferin Bind Site. Arch Biochem Biophys. 1969;134(2):381-94.
2- DeLuca M, McElroy WD. Kinetics of the firefly luciferase catalyzed reactions. Biochemistry. 1974;13(5):921-5.
3- White EH, Rapaport E, Seliger HH, Hopkins TA. The chemi-and bioluminescence of firefly luciferin: An efficient chemical production of electronically excited states. Bioorganic Chem. 1971;1(1-2):92-122.
4- Viviani V. The origin, diversity, and structure function relationships of insect luciferases. Cell Mol Life Sci. 2002;59(11):1833-50.
5- Alipour BS, Hosseinkhani S, Nikkhah M, Naderi-Manesh H, Chaichi MJ, Kazempour Osaloo Sh. Molecular cloning, sequence analysis, and expression of a cDNA encoding the luciferase from the glow-worm, Lampyris turkestanicus. Biochem Biophys Res Commun. 2004;325(1):215-22.
6- Mortazavi M, Hosseinkhani S. Surface charge modification increases firefly luciferase rigidity without alteration in bioluminescence spectra. Enzyme Microb Technol. 2017;96:47-59.
7- Emamzadeh AR, Hosseinkhani S, Sadeghizadeh M, Nikkhah M, Chaichi MJ, Mortazavi M. cDNA cloning, expression and homology modeling of a luciferase from the firefly Lampyroidea maculata. BMB Rep. 2006;39(5):578-85.
8- Hosseinkhani S. Molecular enigma of multicolor bioluminescence of firefly luciferase. Cell Mol Life Sci. 2011;68(7):1167-82.
9- Branchini BR, Magyar RA, Murtiashaw MH, Anderson SM, Helgerson LC, Zimmer M. Site-directed mutagenesis of firefly luciferase active site amino acids: A proposed model for bioluminescence color. Biochemistry. 1999;38(40):13223-30.
10- Lundin A. Use of firefly luciferase in ATP-related assays of biomass, enzymes, and metabolites. Method Enzymol. 2000;305:346-70.
11- Dix DB, Thompson RC. Codon choice and gene expression: Synonymous codons differ in translational accuracy. Proc Natl Acad Sci. 1989;86(18):6888-92.
12- 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.
13- Widmann M, Clairo M, Dippon J, Pleiss J. Analysis of the distribution of functionally relevant rare codons. BMC Genom. 2008;9(1):207.
14- Kypr J. A part of codon bias in genes protects protein spatial structures from destabilization by random single point mutations. Biochem Biophys Res Commun. 1986;139(3):1094-7.
15- Daniel E, Onwukwe GU, Wierenga RK, Quaggin SE, Vainio SJ, Krause M. ATGme: Open-source web application for rare codon identification and custom DNA sequence optimization. BMC Bioinform. 2015;16(1):303.
16- Theodosiou A, Promponas VJ. LaTcOm: A web server for visualizing rare codon clusters in coding sequences. Bioinformatics. 2012;28(4):591-2.
17- Thanaraj T, Argos P. Protein secondary structural types are differentially coded on messenger RNA. Protein Sci. 1996;5(10):1973-83.
18- Guex N, Peitsch MC. SWISS-MODEL and the Swiss-Pdb Viewer: An environment for comparative protein modeling. Electrophoresis. 1997;18(15):2714-23.

19- Zhang Y. I-TASSER server for protein 3D structure prediction. BMC Bioinform. 2008;9(1):40.
20- Kaplan W, Littlejohn TG. Swiss-PDB viewer (deep view). Brief Bioinform. 2001;2(2):195-7.
21- DeLano WL. The PyMOL molecular graphics system. San Carlos: Delano Scientific; 2002.
22- Trott O, Olson AJ. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem. 2010;31(2):455-61.
23- Wu S, Zhang Y. LOMETS: A local meta-threading-server for protein structure prediction. Nucleic Acids Res. 2007;35(10):3375-82.
24- Nakatsu T, Ichiyama S, Hiratake J, Saldanha A, Kobashi N, Sakata K, et al. Structural basis for the spectral difference in luciferase bioluminescence. Nature. 2006;440(7082):372-6.
25- Franks NP, Jenkins A, Conti E, Lieb WR, Brick P. Structural basis for the inhibition of firefly luciferase by a general anesthetic. Biophys J. 1998;75(5):2205-11.
26- Vriend G. What if: A molecular modeling and drug design program. J Mol Gr. 1990;8(1):52-6.
27- Tina K, Bhadra R, Srinivasan N. PIC: Protein interactions calculator. Nucleic Acids Res. 2007;35(suppl-2):W473-6.
28- Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, et al. Autodock4 And Autodocktools4: Automated docking with selective receptor flexibility. J Comput Chem. 2009;30(16):2785-91.
29- O'Boyle NM, Banck M, James CA, Morley Ch, Vandermeersch T, Hutchison GR. Open Babel: An open chemical toolbox. J Cheminformatics. 2011;3(1):33.
30- Dong H, Nilsson L, Kurland CG. Co-variation of trna abundance and codon usage inescherichia coliat different growth rates. J Mol Biol. 1996;260(5):649-63.
31- Clarke TF, Clark PL. Rare codons cluster. PLoS One. 2008;3(10):e3412.
32- Wood KV. Luc genes: Introduction of colour into bioluminescence assays. J Biolumin Chemilumin. 1990;5(2):107-14.
33- Fan F, Wood KV. Bioluminescent assays for high-throughput screening. Assay Drug Dev Technol. 2007;5(1):127-36.
34- Meisenheimer PL, O'Brien MA, Cali JJ. Luminogenic enzyme substrates: The basis for a new paradigm in assay design. Cell Notes. 2008;22:10-4.
35- Greer LF, Szalay AA. Imaging of light emission from the expression of luciferases in living cells and organisms: A review. Lumin J Biol Chem Lumin. 2002;17(1):43-74.
36- Lyons SK, Meuwissen R, Krimpenfort P, Berns A. The generation of a conditional reporter that enables bioluminescence imaging of Cre/loxP-dependent tumorigenesis in mice. Cancer Res. 2003;63(21):7042-6.
37- Gabriel GV, Viviani VR. Novel application of pH-sensitive firefly luciferases as dual reporter genes for simultaneous ratiometric analysis of intracellular pH and gene expression/location. Photochem Photobiol Sci. 2014;13(12):1661-70.
38- 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. Hepat Mon. 2016;16(10):e39909.
39- Varenne S, Baty D, Verheij H, Shire D, Lazdunski C. The maximum rate of gene expression is dependent in the downstream context of unfavourable codons. Biochimie. 1989;71(11-12):1221-9.
40- Zhang G, Ignatova Z. Generic algorithm to predict the speed of translational elongation: Implications for protein biogenesis. PLoS One. 2009;4(4):e5036.
41- Novoa EM, De Pouplana LR. Speeding with control: Codon usage, tRNAs, and ribosomes. Trends Genet. 2012;28(11):574-81.
42- Kane JF. Effects of rare codon clusters on high-level expression of heterologous proteins in Escherichia coli. Curr Opinion Biotechnol 1995;6(5):494-500.
43- Komar AA, Lesnik T, Reiss C. Synonymous codon substitutions affect ribosome traffic and protein folding during in vitro translation. FEBS Lett. 1999;462(3):387-91.
44- Overton TW. Recombinant protein production in bacterial hosts. Drug Discov Today. 2014;19(5):590-601.
45- Wallace AC, Laskowski RA, Thornton JM. Ligplot: A program to generate schematic diagrams of protein-ligand interactions. Protein Eng. 1995;8(2):127-34.
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 D, Lovell S, Thorne N, Lea W, Maloney D, Shen M, et al. Molecular basis for the high-affinity binding and stabilization of firefly luciferase by PTC124. Proc Natl Acad Sci. 2011;107(11):4878-83.
48- Mortezavi M, Torkzadeh-Mahani M, Nezafat N, Malekpour A, Zarenezhad M, Hemmati R, et al. Molecular docking and rare codons evaluation in the luciola lateralis luciferase, an in silico study. Biomacromolecular J. 2017;3(1):48-59.
49- 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. Proc Natl Acad Sci. 2010;107(11):4878-83.