Trends in Peptide and Protein Sciences
Volume 8 (2023): e3

DOI: https://doi.org/10.22037/tpps.v8i1.41854

Original Article

Structural Characterization of a Novel Luciferase-Like-Monooxygenase from *Pseudomonas meliae* – An in-Silico Approach

Mohammad Rayhan[a], Mohd. Faijanur - Rob Siddiquee[b], Asif Shahriar[c,d*], Hossain Ahmed[e], Aar Rafi Mahmud[f], Muhammad Shaiful Alam[g], Muhammad Ramiz Uddin[h], Mrityunjoy Acharjee[c,i], Mst. Sharmin Sultana Shimu[j], Mohd. Shahir Shamsir[a], Talha Bin Emran[k]

[a]*Faculty of Science, Universiti Teknologi Malaysia, Skudai, Malaysia.*
[b]*Department of Biochemistry and Molecular Biology, University of Dhaka, Bangladesh.*
[c]*Department of Microbiology, Stamford University Bangladesh, Bangladesh.*
[d]*Department of Chemistry, Biochemistry and Molecular Biology, College of Science, The University of Texas Rio Grande Valley, USA.*
[e]*Department of Biotechnology and Genetic Engineering, The University of Development Alternative, Dhanmondi, Bangladesh.*
[f]*Department of Biochemistry and Molecular Biology, Mawlana Bhashani Science and Technology University, Bangladesh.*
[g]*Department of Pharmacy, University of Science and Technology Chittagong, Bangladesh.*
[h]*Department of Chemistry and Biochemistry, University of Oklahoma, Norman, USA.*
[i]*Department of Molecular and Developmental Biology, Graduate School of Science and Technology, Shizuoka University, Japan.*
[j]*Genetic Engineering and Biotechnology, University of Rajshahi, Bangladesh.*
[k]*Department of Pharmacy, BGC Trust University Bangladesh, Bangladesh.*

Article history:
Received: 16 April 2023
Accepted: 3 July 2023

Keywords:
Bioluminescence
Luciferase
Luciferase-like monooxygenase
Plant pathogens
Pseudomonas meliae

Mohammad Rayhan:
https://orcid.org/0009-0001-3845-461X
Asif Shahriar:
https://orcid.org/0000-0001-5529-8893

*Corresponding Author:
Email: abasifbl@gmail.com
(A. Shahriar)

HIGHLIGHTS

- Structural characterization of luciferase-like monooxygenase in *P. meliae* .
- Bioluminescence can be used to evaluate antimicrobial efficacy by releasing light emissions.
- Luciferase-like monooxygenase: a potential therapeutic candidate for clinical applications.

ABSTRACT

Luciferase is a well-known oxidative enzyme that produces bioluminescence. The *Pseudomonas meliae* is a plant pathogen that causes wood to rot on nectarine and peach and possesses a luciferase-like monooxygenase. After activation, it produces bioluminescence, and the pathogen's bioluminescence is a visual indicator of contaminated plants. The present study aims to model and characterize the luciferase-like monooxygenase protein in *P. meliae* for its similarity to well-established luciferase. In this study, the luciferase-like monooxygenase from *P. meliae* infects chinaberry plants has been first modeled and then, studied by comparing it with existing known luciferase. In addition, the similarities between uncharacterized luciferase from *P. meliae* and the template from *Geobacillus thermodenitrificans* were analyzed. The results suggest that the absence of bioluminescence in *P. meliae* could be critical for the production of the luciferin substrate and the catalytic activity of the enzyme due to the evolutionary mutation in positions 138 and 311. The active site remains identical except for two amino acids. Therefore, mutation of the residues 138 and 311 in *P. meliae* Luciferase-like monooxygenase may restore luciferase light-emitting ability.

Cite this article as: Rayhan, M., Siddiquee, M.F.R., Shahriar, A., Ahmed, H., Mahmud, A.R., Alam, M.S., Uddin, M.R., Acharjee, M., Shimu, M.S.S., Shamsir, M.S. and Emran, T.B., (2023). Structural characterization of a novel luciferase-like-monooxygenase from *Pseudomonas meliae* – An in-silico approach. *Trends Pept. Protein Sci.*, **8**: e3.

Introduction

The names of luciferase and luciferin were introduced by Emil Du Bois-Reymond in 1885 (Roura et al., 2013). Later on, in 1940, Green and McElroy extracted and purified luciferase protein. They isolated the enzyme and determined its structure (England et al., 2016). Emil Du Bois-Reymond used cold water to investigate the components of the click beetle's bioluminescence, producing luminescence in the laboratory. Experiments resulted in two extracted components. One was named "Luciferine", which was the enzyme responsible for the chemical reaction, and the other was named luciferase. The cloning of firefly luciferase (FLuc) in *Escherichia coli* by Marlene Deluca has made it a commonly used technique in various luciferase systems (Thorne et al., 2012; Pozzo et al., 2018). FLuc is responsible for the oxidative decarboxylation of luciferin in the presence of ATP, Mg^{2+}, and O_2, producing oxyluciferin (Feeney et al., 2016). Due to the longer-wavelength light emission, firefly luciferase can be utilized as a reporter gene in living cells and organisms, enhancing animal tissue penetration (Feeney et al., 2016; Pozzo et al., 2018).

Luciferase, specially sourced from firefly (*Photinus pyralis*), has been utilized as a reporter protein in different assay systems, including gene expression, and was applied in high-throughput screening for drug discovery (Inouye, 2010). Using bioluminescence as a visual cue to signal changes has been well established. This bioluminescence happens in nature in different green growths, microscopic organisms, parasites, and some oceanic creatures, such as jellyfish (Thorne et al., 2010). The luciferase gene is extracted and used to visualize various organisms' gene expressions as a reporter gene. The first luciferase protein was purified from fireflies in the 1940s (Kirkpatrick et al., 2019).

Firefly luciferase has an established application of luciferase-liken in reporting gene regulation and pharmaceutical screening. The *P. meliae*, is a plant pathogenic bacterium that causes wood rot on a nectarine (Fleiss and Sarkisyan, 2019). In this case, *P. meliae*, creates fireflies intriguing prospect of the pathogen's bioluminescence as a visual indicator of contaminated plants. If the pathogen's protein can be activated when the plant has been infected, its bioluminescent bacterial gall can identify affected plants. In this study, the luciferase like monooxygenase from *P. meliae* that infects chinaberry plants, was studied genetically for the first time (Aeini and Taghavi, 2014).

This study aims to examine the luciferase-like monooxygenase gene found in *P. meliae* for its similarity to well-established luciferase enzymes. The novel protein sequence was modeled and compared with known structures using bioinformatics tools; such as AlphaFold which predicts protein structures using deep learning and neural networks. Accurate predictions can aid in understanding protein function, drug design, disease research, and protein engineering. Sequence comparison and physiochemical analysis help to find the potentiality of two compared proteins to use as an aid in the research of proteomics. This study will further develop the luciferase from *P. meliae* as a reporter for gene expression.

Materials and Methods

Selection and retrieval of protein sequence

The selection and retrieval of the luciferase sequence of *P. meliae* are obtained from the Uniprot database. UniProt ID was found using the enzyme name with the pathogen name. A separate investigation was performed using the retrieved sequence in the protein databank (www.rcsb.org) to confirm the absence of a 3D structure.

Composition and physicochemical analysis

The retrieved sequence's primary structure was analyzed using the ProtParam suite of tools (https://web.expasy.org/protparam/). This tool is used to identify the atomic composition, formula, total number of atoms, half-life, estimated amino acid composition, molecular weight, instability index, pI value, hydrophobicity, and other information of the protein model (O'Malley et al., 2012).

Sequence alignment

Sequence alignment was performed to compare the retrieved protein sequence of *P. meliae* with another similar sequence from the UniProt database (https://www.uniprot.org), and UniProt ID (A0A0P9UTV8). Sequence alignment was also performed using ESPript 3.0 (http://espript.ibcp.fr/ESPript/ESPript/) to identify the conserved region between 2 sequences (Gouet et al., 2003).

Secondary structure analysis

The secondary structure obtained from the protein sequence is used to identify the beta-sheet, alpha helix, and amino acid sequence. The predicted secondary structure was also focused on a graphical sequence view on the possibility of perpetration of beta-sheet, alpha helix, and turns. SOPMA (Self-Optimized Prediction Method with Alignment) was used for secondary structure analysis (Geourjon and Deleage, 1995).

Tertiary structure prediction

a) Template selection

The most accurate computational approach for constructing an accurate structural model, AlphaFold, is widely used in many biological applications (Jumper et al., 2021). The 3D structure of a query protein helped template the protein's sequence alignment. For the 3D model, the AlphaFold (https://alphafold.ebi.ac.uk/) web tool was used. It has easily determined the gap between known protein sequences (Varadi et al., 2022).

b) Model validation

The protein 3D model built from AlphaFold was evaluated by ERRAT (https://servicesn.mbi.ucla.edu/ERRAT/). ERRAT evaluates model quality and non-bonded interactions between atoms (Sumitha et al., 2020). PROCHECK (https://www.ebi.ac.uk/thornton-srv/software/PROCHECK/) has been used to check a protein structure's stereochemical quality, creating several PostScript plot analyses (Hameduh et al., 2020). Verify 3D (https://servicesn.mbi.ucla.edu/Verify3D/) has been used for checking the exactness of the modeled 3D structure. ProSA (https://prosa.services.came.sbg.ac.at/prosa.php) server was also used to determine the quality of the 3D structure of the protein (Wiederstein and Sippl, 2007).

c) Structural comparison

The Alpha Fold Protein Structure Database server (Varadi et al., 2022) was used as a template with high similarity to the predicted 3D model of *P. meliae*. The Alkaline monooxygenase (LadA) from *Geobacillus thermodenitrificans* (strain NG80-2) is compared with the luciferase model. Both model and template are superimposed by UCSF ChimeraX software. The template 3B9O and the predicted structure A0A0P9UTV8 were compared and visualized using the UCSF ChimeraX software (Pettersen et al., 2021).

Results

Selection and retrieval of protein sequence

Sequence alignment comparison, the protein luciferase-like monooxygenase from *P. meliae* was chosen from UniProt ID A0A0P9UTV8. The sequence has no known tertiary structure in UniProt and a subsequent search in the protein databank (www.rcsb.org) confirmed that this sequence has no prior structure. The sequence consists of 466 dimer amino acids, which are retrieved from UniProt, and fits the expected length usually found in luciferase protein families.

Composition and physicochemical analysis

ProtParam calculated the amino acid composition of luciferase-like monooxygenase from *P. meliae* and alkaline monooxygenase (LadA) from *Geobacillus thermodenitrificans* (strain NG80-2) for comparison. Fig. 1A shows the highest differences between luciferase-like monooxygenase and alkane monooxygenase are alanine (A) (10.1% vs 6.4%) and leucine (L) (4.5% vs 7.0%). Alanine richness indicated more hydrophobicity of the model. All the other amino acids are nearly similar between model luciferase like monooxygenase and suggested template alkane monooxygenase.

Luciferase-like monooxygenase has 466 amino acid residues and a molecular weight of 52.9 kDa. Alkane monooxygenase has 440 amino acid residues and a molecular weight of 50.20 kDa. Both enzymes have a computed pI value below 7.0 and are acidic, with luciferase-like monooxygenase having a pI of 5.77 and alkane monooxygenase having 6.38. The high acidic residue with different amino acids suggests that the two enzymes may serve similar functions. Luciferase-like monooxygenase is less stable than alkane monooxygenase with an instability index of 38.14 and 29.06, respectively. The Grand average of hydropathicity (GRAVY) for luciferase-like monooxygenase is -0.336 and for alkane monooxygenase it is -0.466. GRAVY calculates the hydrophobicity of a protein or peptide.

Sequence alignment

Multiple sequence alignment presents useful algorithmic tools for pairwise alignment (Chatzou et al., 2016). The program of ESPript (Easy Sequencing in PostScript)

(Gouet et al., 2003) allows rapid visualization. A comparison was made between the luciferase-like monooxygenase from *P. meliae* and the alkane monooxygenase from *Geobacillus thermodenitrificans*. They share a similar sequence, but also have differences. The similarities between the two enzymes are 46.36%. The sequence includes alpha helix α (1-14) and beta-sheet β (1-13). Fig. 1B shows the alignment between the two enzymes.

Secondary structure analysis

The formation of secondary structure in proteins occurs through hydrogen bonds between the amino hydrogen and carboxyl oxygen atoms in the peptide backbone. Protein's secondary structure analysis involves examining the regular secondary structures (helices and strands), which are the essential building blocks. Besides, non-regular structures like loops also exist. Both regular and non-regular structures play a role in interfacing.

Comparison of the secondary structure using SOPMA showed that the percentage of helices, strands, and coils are almost similar (Fig. 1C-D). The template (39.09%%) and model (36.48%) both had dominant alpha helix presence. The model had more amino acids and a lower percentage of strands (11.8%) compared to the template (12.27%). The template had a lower percentage of coil region (48.64%) than the model (51.72%).

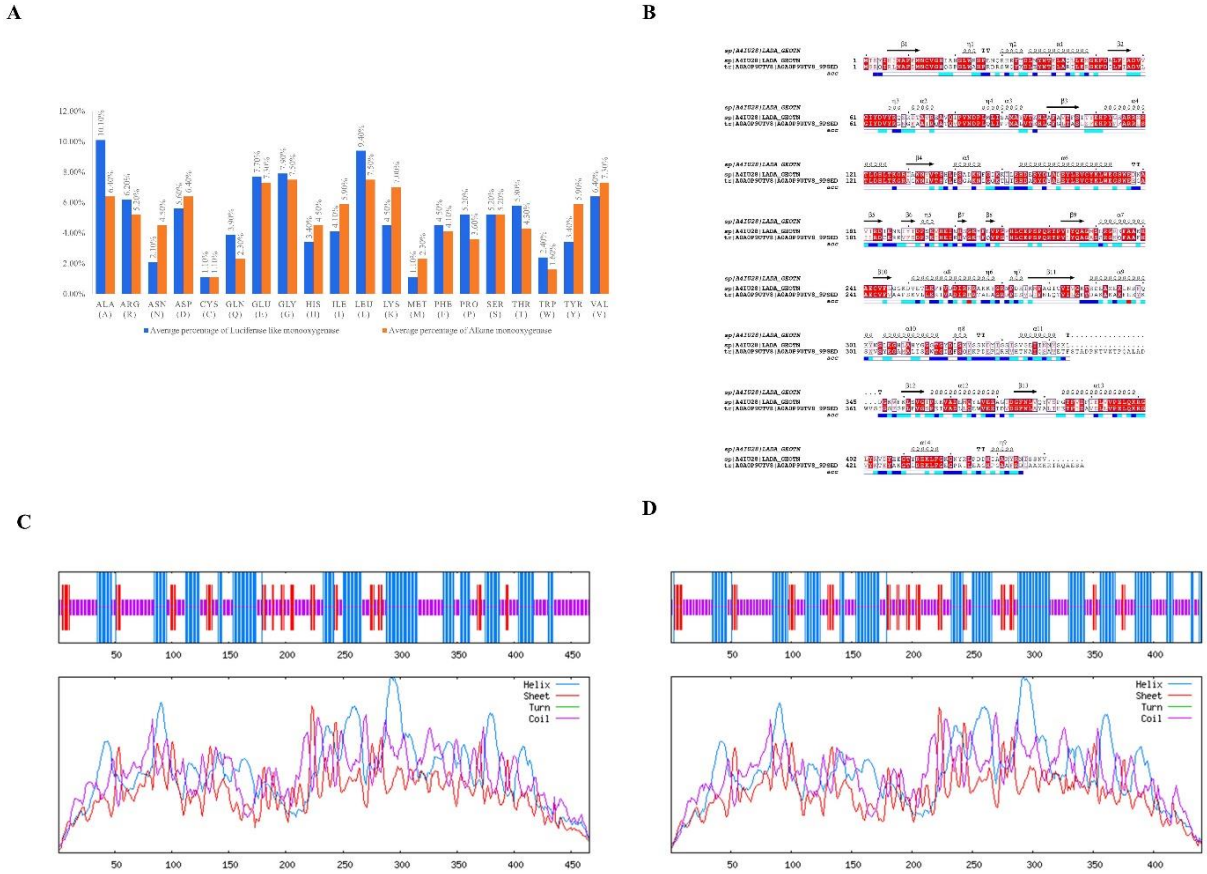

Figure 1: Physiochemical comparison and secondary structure analysis. A) Comparison of the physicochemical characterization in percentages of total residues from luciferase like monooxygenase (blue) and alkane monooxygenase (orange). B) The multiple sequence alignment of alkane monooxygenase *Geobacillus thermodenitrificans* (strain NG80-2) (top) and luciferase-like monooxygenase gene found in *P. meliae* (bottom). The red bands denote identical residues between the two sequences. The black boxes highlighted the active site of the template and predicted active site of the model. C) Secondary structure of template predicted by using SOPMA of network protein sequence analysis of luciferase like monooxygenase. D) Secondary structure of template predicted by using SOPMA of network protein sequence analysis of Alkane monooxygenase.

Template selection

The protein sequence submitted to the Alpha fold Protein structure database (Jumper et al., 2021) resulted in the selection of alkane monooxygenase from *Geobacillus thermodenitrificans* (strain NG80-2, PDB ID: 3B9O) and obtained the best score 52.86% with a high percentage of identity and low e-value 2.2e-56 (Table 1).

Tertiary structure prediction

Three-dimensional (3D) structure of luciferase-like monooxygenase from *P. meliae* (A0A0P9UTV8) as modeled by using Alpha fold software shown in Fig. 2A.

Structure validation

- *ERRAT2*

ERRAT software produces high-quality structures with an overall quality factor of around 95%. The predicted structure of luciferase-like monooxygenase had a lower value of 93.231. Two lines indicate the confidence to reject regions beyond the error values. The average overall quality factor for lower resolutions is around 91%. (Fig. 2B).

- *PROCHECK*

The Ramachandran plot (Fig. 2C) validates the structure of the AlphaFold template's luciferase-like monooxygenase protein prediction model. White regions indicate atoms in the polypeptide coming closer than their van der Waals radii. The plot score was 92.8%. Glycine is the only amino acid that can fit in regions where other amino acids cannot due to its lack of a side chain. The allowed regions are alpha-helical and beta-sheet conformations, shown in red, while the yellow areas show slightly shorter allowed regions (Table 2).

Table 2. Ramachandran plot validation percentage.

Evaluation of residues	Score
Residues in most favoured regions [A,B,L]	(374) 92.8%
Residues in additional allowed regions [a,b,l,p]	(27) 6.7%
Residues in generously allowed regions [~a,~b,~l,~p]	(0) 0.0%
Residues in disallowed regions	(2) 0.5%
Number of non-glycine and non-proline residues	(403) 100.0%
Number of end-residues (excl. Gly and Pro)	2
Number of glycine residues (shown as triangles)	37
Number of proline residues	24
Total number of residues	466

Table 1. Selected a template for modelling by Accession number, protein name, and organism

Enzyme	Server name	Protein Identification	Length	Identity	E-value
Luciferase-like monooxygenase	Alpha fold	Alkane monooxygenase	440	52.86	NR
		Dimethyl-sulfide monooxygenase	488	52.85	NR
		Riboflavin lyase	461	36.26	NR
	Phyre2	Alkane monooxygenase	440	53	NR
		Dimethyl-sulfide monooxygenase	488	53	NR
		Riboflavin lyase	461	36	NR
	HHpred	Alkane monooxygenase	440	53	2.2e-56
		Dimethyl-sulfide monooxygenase	488	53	3e-70
		Riboflavin Lyase	461	35	5.5e-51

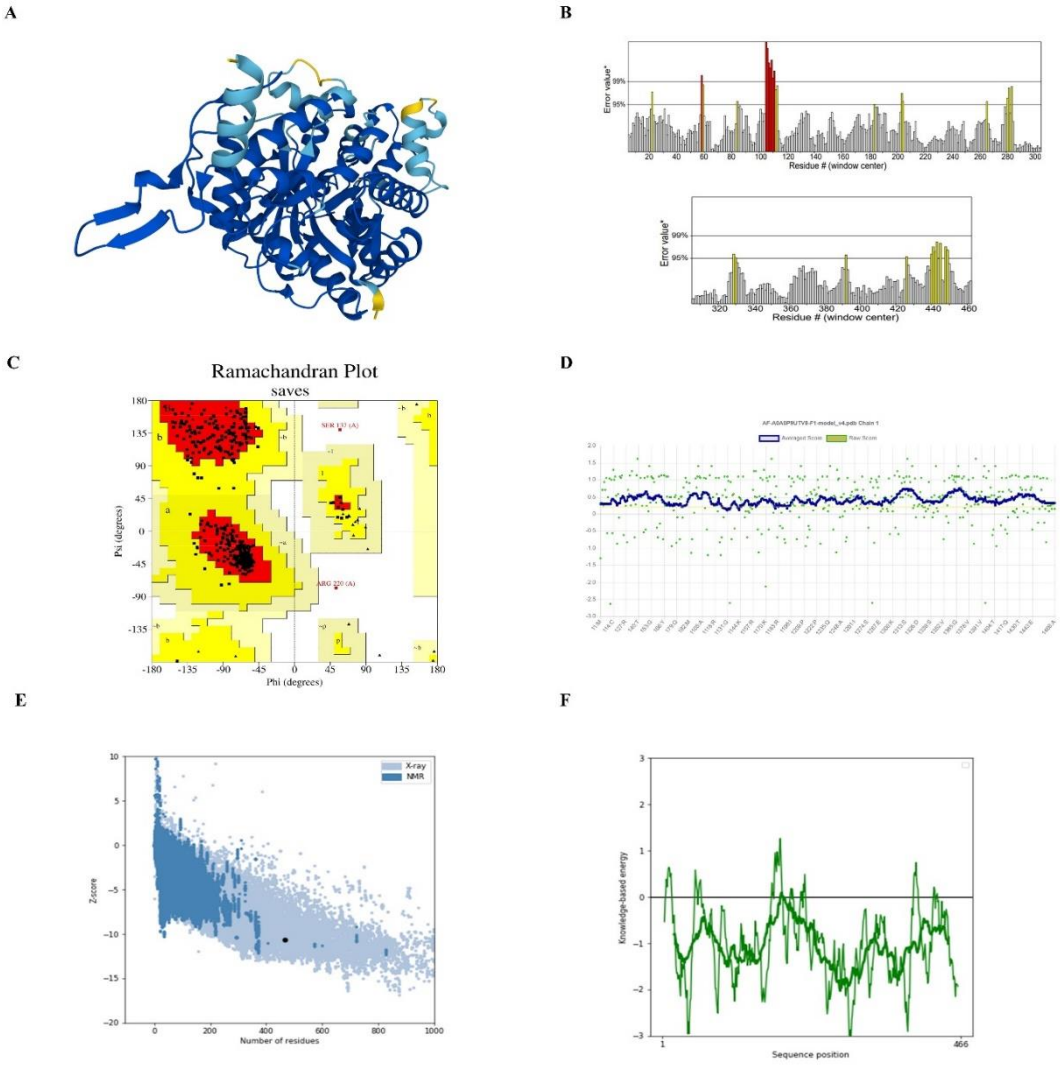

Figure 2: Selection and validation of 3D structures**.** A) 3D Structure of luciferase-like monooxygenase from *P. meliae* (A0A0P9UTV8) as modeled using AlphaFold Deepmind software. The head of the structure started from Methionine and ended of the amino acid is Alanine. B) Structure validation using ERRAT2. Two lines were drawn to indicate the possibilities to reject regions that exceed that error value. C) Model validation using Ramachandran plot to show the background of phi-psi probabilities. D) The compatibility of an atomic model (3D) with its own amino acid sequence (1D) by assigning a structural class-based Model validation using Verify 3D. E) ProSA server for Z-score (-10.67) of Model for overall quality for luciferase like monooxygenase. F) Local model quality assessment with ProSA server for luciferase like monooxygenase.

- *VERIFY 3D*

Verify 3D checks if the 3D model of a protein matches its amino acid sequence. A correct match with high scores means the model is valid. At least 80% of amino acids must score ≥ 0.2 for validation. The thermophile luciferase protein's 3D AlphaFold structure had 97.21% of residues with an average score of ≥ 0.2 (Fig. 2D).

- *ProSA*

ProSA (Protein Structure Analysis) shows the z-scores on overall model quality. It also helps to show local model quality by plotting energies as a function of amino acid sequence position (Fig. 2E-F).

Structural comparison

The overall structure and orientation of the modeled structure fit the shape and requirements of a monooxygenase. Fig. 3A shows the surface of the model structure luciferase-like monooxygenase and template alkane monooxygenase.

Active site prediction

The model protein has 46.36% similarity with template 3B9O. Template 3B9O was used as a reference to locate similar amino acids in the active site of the luciferase-like monooxygenase protein. Eight important amino acids were identified and aligned with the model protein. The two structures' superimposition has been highlighted using different colors (Fig. 3BI). The template and model luciferase-like-monooxygenase have similar and different amino acids highlighted in green and red. The active site is highlighted in blue (Fig. 3BII). The 3B9O residues are near the modeled luciferase residues, even though they are at a distance in the sequence. Both active sites are in a similar location within the luciferase structure.

Fig. 3C shows the specific position of the active site of template alkane monooxygenase and the predicted active site of luciferase-like-monooxygenase. The active site comparison revealed two differences between the template and model amino acid positions: at 138 (histidine replaced with tyrosine in the model) and 311 (histidine replaced with leucine in the model).

When generating a surface image of a molecule, the entire molecule is depicted rather than just individual residues. This surface image is like an outer layer of the molecule, which is stretched over an imaginary Van der Waals surface. As a result, only the external part of the molecule is visible, while the underlying molecular representation is not. Fig. 3DI and 3DII for examples of surface images of the active sites of model luciferase-like-monooxygenase from *P. meliae* and template alkane monooxygenase from *Geobacillus thermodenitrificans.*

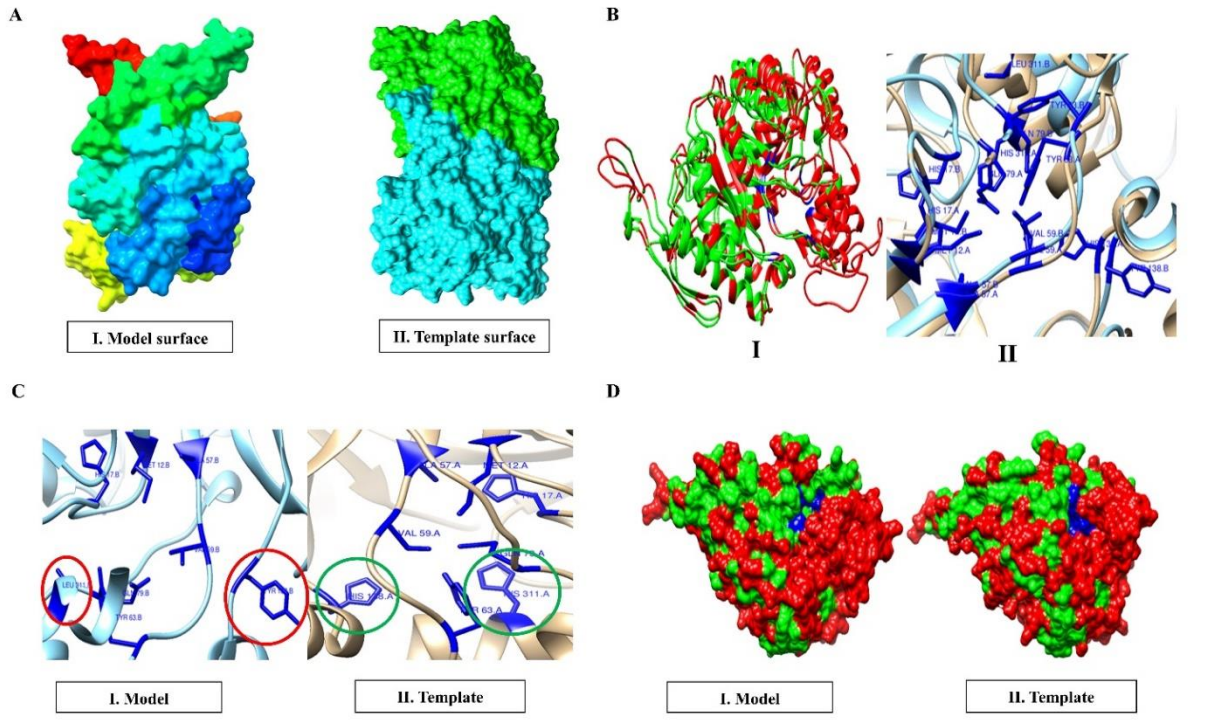

Figure 3: Structural comparison of structure**.** A) Surface structure of the model (I) Luciferase-like monooxygenase and (II) the template alkane monooxygenase. B) The superimpose structure of the template and model (I) the surface image of the active site (blue) between the model and template 3B9O (II). C) Active site residues of model luciferase-like-monooxygenase (I) and active site residues of template alkane monooxygenase (II). D) The probability of predicted active site (blue) of luciferase-like-monooxygenase (I) from *Pseudomonas meliae and* II) the surface image of active site (blue) template alkane monooxygenase from *Geobacillus thermodenitrificans* (strain NG80-2).

Overall structure analysis

a) Strands

Beta strand is an important secondary structure element, which consists of several beta-strands—stretched segments of the polypeptide chain combined by the network of hydrogen bonds between adjacent strands. Beta-strand is an essential mode of protein-protein interaction (Siepen et al., 2009). Strands superimpose between model luciferase-like monooxygenase (Fig. 4A) and template alkane monooxygenase. Blue highlights similarities, red differences. The same amino acids are in the red areas in Fig. 4A. Template and model strands differed in protein amount. The template has 12.27% strands, while the model has 11.8%. (Fig. 4B).

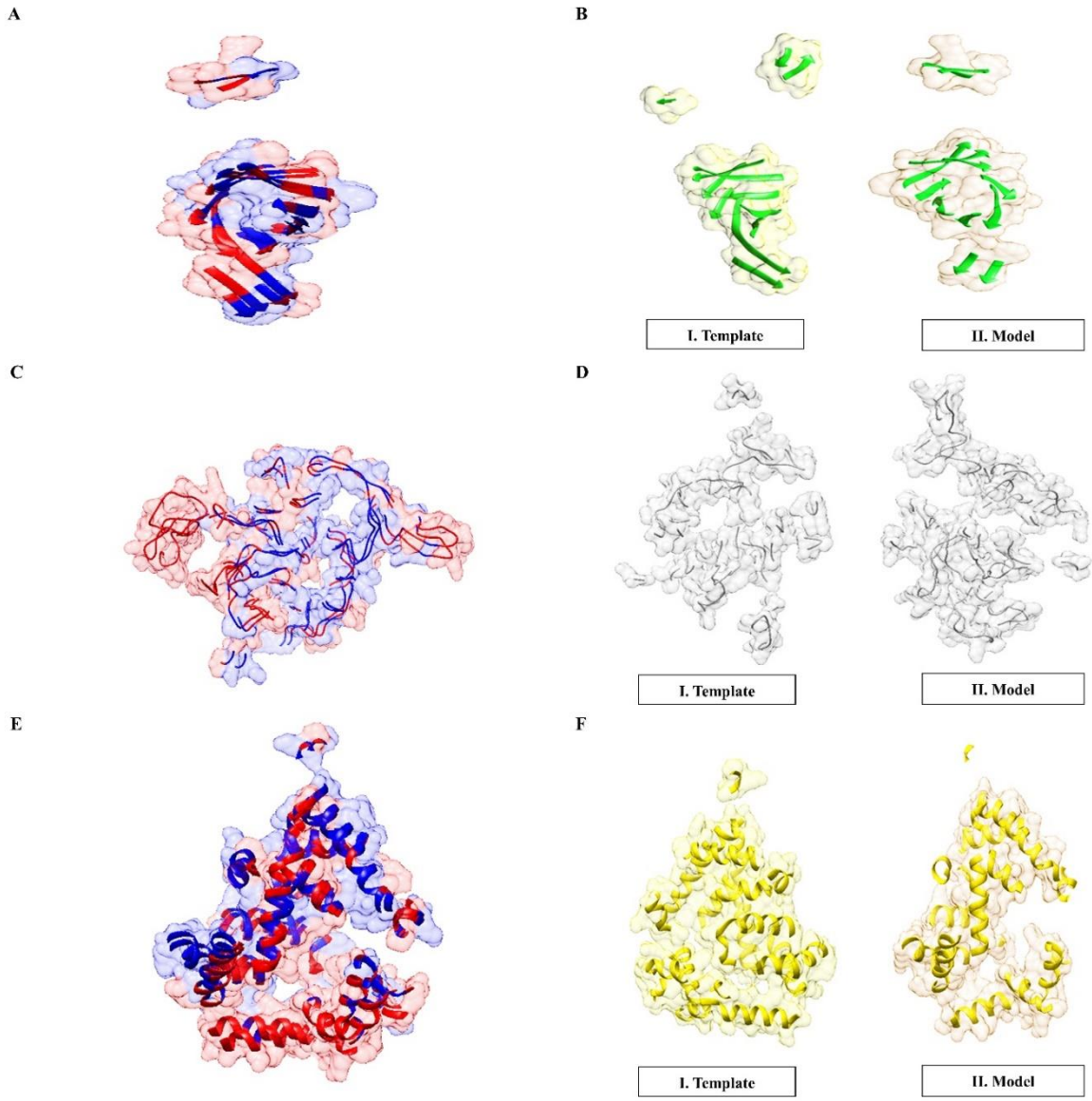

Figure 4: Structure Analysis of luciferase like monooxygenase of *P. meliae* and alkane monooxygenase *Geobacillus thermodenitrificans.* A) The superimposition of model and template strands. B) The strands of (I) Template alkane monooxygenase and (II) Model luciferase-like-monooxygenase. C) The superimposition of model and template loops identify the similarities between model and template. D) The loops of (I) Template alkane monooxygenase and (II) Model luciferase-like-monooxygenase. E) Helices superimposition of model luciferase-like monooxygenase and template alkane monooxygenase. F) The helices between (I) Template alkane monooxygenase and (II) model luciferase-like-monooxygenase.

The 7th and 224th model positions have Leucine (L), and the template has Isoleucine (I), the same alkyl group. On the other hand, at the 56th position, the model's amino acid is Isoleucine (I), and the template's amino acid is Leucine (L), the same group. At the 11th position, the model amino acids are aspartic acid (D), and the template amino acids are glutamic acid (E). Both are different amino acids but the same (carboxylic acid group). At the 103rd and 285th positions, the model's amino acids are Leucine (L) and template amino acids are Valine (V). Both are not similar, but they are from the same alkyl group. Howevr, for 246th position, it was reversed. The template's amino acid is Leucine (L) and the model's amino acid is Valine (V), both belong to the same alkyl group. At the 130th position, the template's amino acid is isoleucine (I), and the model's amino acid is alanine (A), both belong the same alkyl group. The template and model's amino acids are Valine (V) and Isoleucine (I), but both amino acids belong to the same alkyl group. And finally, at the 397th position, Alanine (A) for the model and Valine (V) for the template that both belong to the same alkyl group (Patel et al., 2021).

b) Loops

Loop regions are important in protein function and connect secondary structures. The flexible coil segments, known as loop regions, contribute to catalytic and ligand binding sites (Li, 2013). Protein function and the binding of unaligned regions in sequence alignments between alpha-helix and beta-sheet are facilitated by the loop region. The template structure comprises 48.64% coils or loops, while the model structure contains 51.72% coils, as shown in Fig. 4C. Fig. 4D presents the discrepancies between the template and the model structure of loops.

The loops have been highlighted by color. The blue color highlights the similarities between the template and the model. The red color highlights the differences, but some amino acids are from the same group. For example, Valine (V) and Leucine (L), which are from the same alkyl group, are 40, 400, and 421. Isoleucine (I) and Valine (V) are different amino acids, but they belong to the same alkyl group. They are at 80, 181, 189, 196, 275, and 391. Serine (S) and Threonine (T) are not similar, but they are from the same hydroxyl group and are 108, 191, 374, and 401. At 151 and 182, the template and model have Isoleucine (I) and Leucine (L) from the same alkyl group. Two different amino acids in the carboxylic acid at 178th and 425th are aspartic acid (D) and Glutamic acid (E). At the position of 340, the model has Alanine (A) and the template has Isoleucine (I), but they are from the same alkyl group. They have been shown in the same group because of their polarity. Alanine and Isoleucine are non-polar and part of aliphatic amino acids in nature.

c) Helices

Helices are an essential type of secondary structure element found in proteins (Fodje and Al-Karadaghi, 2002). Helices play a prominent role in genome maintenance, how the enzyme changes conformations, and transitions between different conformational states, regulating nucleic acid and reshaping structure (Ma et al., 2018). The number of helices in the template structure is 39.09%, compared to 36.48% helices in the model (Fig. 4F).

Blue highlights similarities, red highlights differences, but some red amino acids belong to the same group in the superimposed structure (Fig. 4E). At the 41st position, the model amino acid is aspartic acid (D), and the template amino acid is glutamic acid (E), which is different but different from the same group. At the 71st position, the model's amino acid is glutamic acid (E), and the template's amino acid is aspartic acid (D). At the 74th and 406th positions, the model has isoleucine (I), and the template has valine (V), and at the 409th position, the amino acids are reversed. The template has isoleucine (I), and the model has valine (V) but belongs to the same alkyl group. There are two different amino acids at 89, 318, 358, and 364, but they are from the same alkyl group, including isoleucine (I) and valine (V). In addition, the final similar group of amino acids includes valine (V) and leucine (L), at the position of 412.

The structural elucidation of the template and model has been highlighted with a different color. The beta-strands are highlighted by blue color, the loops are highlighted by grey color, and a shade of light brown highlights helices (Fig. 5A).

Binding site

In proteins, binding sites are small tertiary structure pockets where the ligands bind to them using weak forces (non-covalent bonding) (Patel et al., 2018). The binding sites of proteins play an important role in a wide range of applications, including molecular docking, drug design, structure identification, and comparison of functional sites (Guo and Wang, 2012). One of the receptor's most fundamental properties is the set of amino acids available for interactions with ligands

(Khazanov and Carlson, 2013). The binding site of the predicted model luciferase-like monooxygenase from *P. meliae* has been given the Uniprot ID. The binding side's position is Asp58 and Thr104 (Fig. 5BI). The template alkane monooxygenase from *Geobacillus thermodenitrificans (strain NG80-2)* has given the binding side at 58, 104, 158 245 (Fig. 5BII).

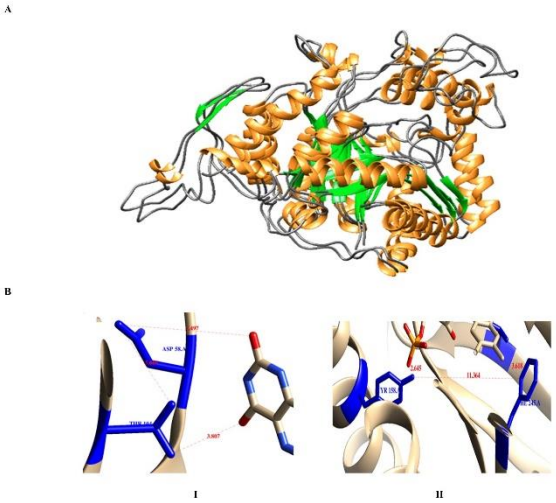

Figure 5: Binding site analysis of luciferase like mono-oxygenase of *P. meliae* and alkane monooxygenase *Geobacillus thermodenitrificans*. A) Different colors highlight the structures between the template and model: green for beta-strands, grey for loops, and light brown for helices. B) Binding residues of model luciferase-like monooxygenase are 104 and 58 with ligand (I). The binding residues of template alkane monooxygenase are 104, 58, 158, and 245 (II).

Discussion

In this study, the sequence of luciferase-like monooxygenase was retrieved from UniProt (Pettersen et al., 2021) (Organism: *P. meliae*, ID A0A0P9UTV8). There was no known structure in protein databases, therefore, through homology analysis some important features of this protein were revealed. In physicochemical comparison, we found maximum differences between luciferase-like monooxygenase and alkane monooxygenase in the percentage of alanine (A) and leucine (L), that is, 10.1% and 6.4%, respectively, while other amino acids were similar. Alanine indicates more hydrophobicity in the model (Lefèvre et al., 1997). The computed (pI) value of both enzymes is below 7.0. Luciferase-like monooxygenase has a pI value of 5.77, and alkane monooxygenase has 6.38, meaning both enzymes are acidic. The hydrophobicity nature also supports by the GRAVY value. Here, the GRAVY value of luciferase-like monooxygenase is -0.336, and for alkane monooxygenase, it is -0.466. An increasing positive score indicates a greater hydrophobicity (Kyte and Doolittle, 1982). The similarity in size and acidity suggests that the model may serve similar functions. In the instability index, luciferase-like monooxygenase scores relatively higher than alkane monooxygenase, 38.14 and 29.06, respectively. A less than 40 instability index score indicates more excellent stability.

The active sites of both proteins are located at a distance from each other along the protein sequence. Highlighting the 3D superimposed structure location revealed that both active sites exist at a similar location within the conformational structure of luciferase. In both enzymes, all active residues were similar except for positions 138 and 311. At position 138, the template amino acid has histidine, and the model amino acid has tyrosine. At position 311, the template amino acid has histidine; at the model's 311, the amino acid is leucine. The findings indicate that the lack of bioluminescence in *P. meliae* of luciferase-like-monooxygenase is due to an evolutionary mutation of an amino acid at positions 138 and 311. *Pseudomonas fluorescens*, for example, has been shown to react with light and emit luminescence when exposed to UV light (Scales et al., 2014; Galet et al., 2015). By measuring light emission, bioluminescent *P. aeruginosa* may be used to determine the antimicrobial efficacy of wound dressings (Dartnell et al., 2013). As a result, if the residues 138 and 311 are mutated to recover luciferase light-emitting capacity through future studies, *P. meliae* could have a few possible applications in the future. There is still space for luciferase-like-monooxygenase from *P. meliae* to be improved, activated, and repurposed as a disease marker. For potential biological applications, such as living cells and living tissue, fluorescence may be used to detect specific components of complex bio-molecular assemblies.

According to the results, the data and model obtained for the luciferase-like monooxygenase from *P. meliae* indicate that the protein of this plant pathogen possesses significant characteristics of light-emitting luciferase. The protein's structure, composition, and profile, including its acidity, suggest that it has not evolved significantly differently from other proteins in the light-emitting luciferase family. The active site remains the same, except for two amino acids: tyrosine 138 in the *P. meliae* model replaces histidine 138 in the template, and leucine 311 replaces histidine 311. In the template,

tyrosine is classified under the hydroxyl group while histidine is under the amino group. Both leucine and histidine amino acids have functional groups that consist of an α-amino group and a carboxylic group, given certain biological conditions. The luciferase-like monooxygenase protein has various properties that are comparable to the template alkane monooxygenase, including primary structure characteristics, amino acid sequences, binding sites, and predicted active sites. Both structures share key characteristics, such as a high presence of amino acid residues in the primary protein structure.

Conclusion

The findings indicate that the absence of bioluminescence in *P. meliae* may be due to the absence of a mutation of amino acid in the positions of 138 and 311, preventing light emission by luciferase-like-monooxygenase. Mutating the remaining residues at these positions may uncover the light-emitting ability of luciferase and create opportunities for further development, activation, and repurposing of the enzyme from *P. meliae* as a disease marker. Bioluminescent *P. aeruginosa* has been used to assess antimicrobial efficacy by releasing light emissions. Fluorescence can also be used to detect complex bio-molecular assemblies in living cells and tissues for future biological applications.

Ethics Statement

This study did not included any human and animal studies.

Competing Interests

The authors declare that they have no competing interests.

Funding

No funding was reported.

Authors Contribution

Mohammad Rayhan and Mohd. Faijanur - Rob Siddiquee conceived the study, designed the experiments, and wrote the manuscript. Mohammad Rayhan, Mohd. Faijanur - Rob Siddiquee, Asif Shahriar, Muhammad Shaiful Alam, Hossain Ahmed, Aar Rafi Mahmud, Mst. Sharmin Sultana Shimu and Mohd. Shahir Shamsir performed the dry-lab experiments. *Muhammad Ramiz Uddin* and Mrityunjoy Acharjee were performed the English Editing and Proofreading. Mohd. Shahir Shamsir designed and planned the studies, and supervised the experiments. Mohd. Shahir Shamsir and Talha Bin Emran supervised the research. All authors approved the final version of the manuscript.

References

Aeini, M. and S.M. Taghavi, (2014). ″Genotypic characteristics of the causal agent of chinaberry gall.″ *Archives of Phytopathology and Plant Protection*, **47**(12): 1466–1474. DOI: https://doi.org/10.1080/03235408.2013.845997.

Chatzou, M., Magis, C., Chang, J.M., Kemena, C., Bussotti, G., Erb, I. and C. Notredame, (2016). ″Multiple sequence alignment modeling: Methods and applications. ″ *Briefings in Bioinformatics*, **17**(6): 1009–1023. DOI: https://doi.org/10.1093/BIB/BBV099.

Dartnell, L.R., Roberts, T.A., Moore, G., Ward, J.M. and J.P. Muller, (2013). ″Fluorescence characterization of clinically important bacteria.″ *PLoS ONE*, **8**(9): e75270. DOI: https://doi.org/10.1371/journal.pone.0075270.

England, C.G., Ehlerding, E.B. and W. Cai, (2016). ″NanoLuc: A small luciferase is brightening up the field of bioluminescence.″ *Bioconjugate Chemistry*, **27**(5): 1175–1187. DOI: https://doi.org/10.1021/acs.bioconjchem.6b00112.

Feeney, K.A., Putker, M., Brancaccio, M. and J.S. O'Neill, (2016). ″In-depth characterization of firefly luciferase as a reporter of circadian gene expression in mammalian cells.″ *Journal of Biological Rhythms*, **31**(6): 540–550. DOI: https://doi.org/10.1177/0748730416668898.

Fleiss, A. and K.S. Sarkisyan, (2019). ″A brief review of bioluminescent systems.″ *Current Genetics*, **65**(4): 877–882. DOI: https://doi.org/10.1007/s00294-019-00951-5.

Fodje, M.N. and S. Al-Karadaghi, (2002). ″Occurrence, conformational features and amino acid propensities for the π-helix.″ *Protein Engineering*, **15**(5): 353–358. DOI: https://doi.org/10.1093/protein/15.5.353.

Galet, J., Deveau, A., Hôtel, L., Frey-Klett, P., Leblond, P. and B. Aigle, (2015). ″*Pseudomonas fluorescens* pirates both ferrioxamine and ferricoelichelin siderophores from *Streptomyces ambofaciens*.″ *Applied and Environmental Microbiology*, **81**(9): 3132–3141. DOI: https://doi.org/10.1128/AEM.03520-14.

Geourjon, C. and G. Deleage, (1995). ″SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments.″ *Bioinformatics*, **11**(6): 681-684. DOI: https://doi.org/10.1093/bioinformatics/11.6.681.

Gouet, P., Robert, X. and E. Courcelle, (2003). ″ESPript/ENDscript: Extracting and rendering sequence and 3D information from atomic structures of proteins.″ *Nucleic Acids Research*, **31**(13): 3320–3323. DOI: https://doi.org/10.1093/nar/gkg556.

Guo, F., & Wang, L. (2012). ″Computing the protein binding sites.″ *BMC Bioinformatics*, **13** (Suppl 10): 25–26. DOI: https://doi.org/10.1186/1471-2105-13-S10-S2.

Hameduh, T., Haddad, Y., Adam, V. and Z. Heger, (2020). ″Homology modeling in the time of collective and artificial intelligence.″ *Computational and Structural Biotechnology Journal*, **18:** 3494–3506. DOI: https://doi.org/10.1016/j.csbj.2020.11.007.

Inouye, S. (2010). ″Firefly luciferase: An adenylate-forming enzyme for multicatalytic functions.″ *Cellular and Molecular Life Sciences*, **67**(3): 387–404. DOI: https://doi.org/10.1007/s00018-009-0170-8.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A.W., Kavukcuoglu, K., Kohli, P. and D. Hassabis, (2021). ″Highly accurate protein structure prediction with AlphaFold.″ *Nature*, **596**(7873): 583–589. DOI: https://doi.org/10.1038/s41586-021-03819-2.

Khazanov, N.A. and H.A. Carlson, (2013). ″Exploring the composition of protein-ligand binding sites on a large scale.″ *PLoS Computational Biology*, **9**(11): e1003321. DOI: https://doi.org/10.1371/journal.pcbi.1003321.

Kirkpatrick, A., Xu, T., Ripp, S., Sayler, G. and D. Close, (2019). ″Biotechnological advances in luciferase enzymes.″ In: *Bioluminescence - Analytical Applications and Basic Biology*, Suzuki, H. (Ed), IntechOpen, pp. 1-23. DOI: https://doi.org/10.5772/intechopen.85313.

Kyte, J. and R.F. Doolittle, (1982). ″A simple method for displaying the hydropathic character of a protein.″ *Journal of Molecular Biology*, **157**(1): 105-132. DOI: https://doi.org/10.1016/0022-2836(82)90515-0.

Lefèvre, F., Rémy, M.H. and J.M. Masson, (1997). ″Alanine-stretch scanning mutagenesis: A simple and efficient method to probe protein structure and function.″ *Nucleic Acids Research*, **25**(2): 447–448. DOI: https://doi.org/10.1093/nar/25.2.447.

Li, Y. (2013). ″Conformational sampling in template-free protein loop structure modeling: An overview.″ *Computational and Structural Biotechnology Journal*, **5**(6): e201302003. DOI: https://doi.org/10.5936/csbj.201302003.

Ma, W., Whitley, K. D., Chemla, Y. R., Luthey-Schulten, Z. and K. Schulten, (2018). ″Free-energy simulations reveal molecular mechanism for functional switch of a DNA helicase.″ *ELife*, **7**: 1–21. DOI: https://doi.org/10.7554/eLife.34186.

O'Malley, C.J., Montague, G.A., Martin, E.B., Liddell, J.M., Kara, B. and N.J. Titchener-Hooker, (2012). ″Utilisation of key descriptors from protein sequence data to aid bioprocess route selection.″ *Food and Bioproducts Processing*, **90**(4): 755–761. DOI: https://doi.org/10.1016/j.fbp.2012.01.005.

Patel, C.N., Georrge, J.J., Modi, K.M., Narechania, M.B., Patel, D.P., Gonzalez, F.J. and H.A. Pandya, (2018). ″Pharmacophore-based virtual screening of catechol-o-methyltransferase (COMT) inhibitors to combat Alzheimer's disease.″ *Journal of Biomolecular Structure and Dynamics*, **36**(15): 3938-3957. DOI: https://doi.org/10.1080/07391102.2017.1404931.

Patel, C. N., Jani, S. P., Jaiswal, D. G., Kumar, S. P., Mangukia, N., Parmar, R. M., Rawal, R. M. and H.A. Pandya, (2021). ″Identification of antiviral phytochemicals as a potential SARS-CoV-2 main protease (Mpro) inhibitor using docking and molecular dynamics simulations.″ *Scientific Reports*, **11**(1): 1–13. DOI: https://doi.org/10.1038/s41598-021-99165-4.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I., Morris, J.H. and T.E. Ferrin, (2021). ″UCSF ChimeraX: Structure visualization for researchers, educators, and developers.″ *Protein Science*, **30**(1): 70–82. DOI: https://doi.org/10.1002/pro.3943.

Pozzo, T., Akter, F., Nomura, Y., Louie, A.Y. and Y. Yokobayashi, (2018). ″Firefly luciferase mutant with enhanced activity and thermostability.″ *ACS Omega*, **3**(3): 2628–2633. DOI: https://doi.org/10.1021/acsomega.7b02068.

Roura, S., Gálvez-Montón, C. and A. Bayes-Genis, (2013). ″Bioluminescence imaging: A shining future for cardiac regeneration.″ *Journal of Cellular and Molecular Medicine*, **17**(6): 693–703. DOI: https://doi.org/10.1111/jcmm.12018.

Scales, B.S., Dickson, R.P., Lipuma, J.J. and G.B. Huffnagle, (2014). ″Microbiology, genomics, and clinical significance of the Pseudomonas fluorescens species complex, an unappreciated colonizer of humans.″ *Clinical Microbiology Reviews*, **27**(4): 927–948. DOI: https://doi.org/10.1128/CMR.00044-14.

Siepen, J.A., Radford, S.E. and D.R. Westhead, (2009). ″β Edge strands in protein structure prediction and aggregation.″ *Protein Science*, **12**(10): 2348–2359. DOI: https://doi.org/10.1110/ps.03234503.

Sumitha, A., Devi, P.B., Hari, S. and R. Dhanasekaran, (2020). ″COVID-19 - In silico structure prediction and molecular docking studies with doxycycline and quinine.″ *Biomedical and Pharmacology Journal*, **13**(3): 1185–1193. DOI: https://doi.org/10.13005/bpj/1986.

Thorne, N., Inglese, J. and D.S. Auld, (2010). ″Illuminating insights into firefly luciferase and other bioluminescent reporters used in chemical Biology.″ *Chemistry and Biology*, **17**(6): 646–657. DOI: https://doi.org/10.1016/j.chembiol.2010.05.012.

Thorne, N., Shen, M., Lea, W.A., Simeonov, A., Lovell, S., Auld, D.S. and J. Inglese, (2012). ″Firefly luciferase in chemical biology: A compendium of inhibitors, mechanistic evaluation of chemotypes, and suggested use as a reporter.″ *Chemistry and Biology*, **19**(8): 1060–1072. DOI: https://doi.org/10.1016/j.chembiol.2012.07.015.

Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., Yuan, D., Stroe, O., Wood, G., Laydon, A., Zídek, A., Green, T., Tunyasuvunakool, K., Petersen, S., Jumper, J., Clancy, E., Green, R., Vora, A., Lutfi, M., Figurnov, M., Cowie, A., Hobbs, N., Kohli, P., Kleywegt, G., Birney, E., Hassabis, D. and S. Velankar, (2022). ″AlphaFold Protein Structure Database: Massively expanding the structural coverage of protein-sequence space with high-accuracy models.″ *Nucleic Acids Research*, **50**(D1): D439–D444. https://doi.org/10.1093/nar/gkab1061.

Wiederstein, M. and M.J. Sippl, (2007). ″ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins.″ *Nucleic Acids Research*, **35**(suppl_2): W407-W410. DOI: https://doi.org/10.1093/nar/gkm290.

