Biotechnology Journal International

24(5): 49-73, 2020; Article no.BJI.60696
ISSN: 2456-7051
(Past name: British Biotechnology Journal, Past ISSN: 2231–2927, NLM ID: 101616695)

Structural Characterization of Novel Luciferase from Bioluminescence Fungi *Verticillium longisporum*

A. A. Yarima[1], S. M. Sambo[2], M. D. Kwairanga[1], K. N. Sharbat[3], Z. Arifullah[4], A. Y. Fardami[5*], H. Y. Ismail[6] and Z. Rayhan[7]

[1]Department of Biology, Federal College of Education (Technical) Gombe, Gombe State, Nigeria.
[2]Department of Chemistry, Federal College of Education (Technical) Gombe, Gombe State, Nigeria.
[3]Department of Biology, Faculty of Science, Shaikh Zayed University, Khost, Afghanistan.
[4]Department of Biology, Faculty of Science, Nangarhar University, Darunta Jalalabad, Afghanistan.
[5]Department of Microbiology, Usmanu Danfodiyo University, Sokoto, Nigeria.
[6]Department of Microbiology, Faculty of Science, University of Maiduguri, Nigeria.
[7]Department of Bioscience, Faculty of Science, Universiti Teknologi, Malaysia.

Authors' contributions

This work was carried out in collaboration among all authors. Author AAY designed the study, performed the experiments and wrote the first draft of the manuscript. Authors SMS, KNS and RZ wrote the protocol and managed the analyses of the study. Authors MDK, AYF and HYI managed the literature searches and publication process. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/BJI/2020/v24i530117
Editor(s):
(1) Dr. Ghousia Begum, CSIR-Indian Institute of Chemical Technology, India.
(2) Chung-Jen Chiang, China Medical University, Taiwan.
Reviewers:
(1) Ekambaram Rajasekaran, V. S. B. Engineering College, India.
(2) Kulvinder Kochar Kaur, Dr. Kulvinder Kaur Centre for Human Reproduction, India.
Complete Peer review History: http://www.sdiarticle4.com/review-history/60696

Original Research Article

Received 28 June 2020
Accepted 05 September 2020
Published 16 September 2020

ABSTRACT

Luciferase is an enzyme that catalyses a reaction to produce a visible light using an oxidative process, a chemical reaction that is typically referred to as bioluminescent. Insects, bacterial origin or microorganism of marine nature were considered as the mainly sources of discovered luciferase. The protein was commercialized for biomedical and biotechnological use as reporter gene. The first discovered wild form of luciferase originally from *Photinu spyralis* (firefly). Hence, there is need for both exploration and examination of novel luciferase to be expanded to new sources such as fungal which may likely be exploited to serve commercial purposes and applications. In this study, a novel uncharacterized luciferase protein from a fungal species *Verticillium longisporum*, was modelled and

**Corresponding author: E-mail: aminufy@gmail.com;*

analysed using bioinformatic tools. The modelled 3D structure is of high quality with a PROCHECK score of 99.5%, ERRAT2 value of 91.01%, and Verify3D score of91.01%, showing that the conformational structure is acceptable. The result showed that the fungal luciferase enzyme share major characteristics with luciferase representative from various fungal and bacterial species. There is only a slight difference in the two nucleotide bindings in *V. longisporum* with a D/E substitution of D with E and S/T substitution. The difference of the two nucleotides binding from the two proteins may be related to the evolutionary trends. Other differences include increased number of hydrophobic and polar amino acid groups than aromatic and aliphatic ones, as well as more coils and loops with less strands. The distance between the ligand and the binding site that houses Asp 64 and Thr 110 from template proteins (Riboflavin lyaseRcaE) and Asp 543 and Thr 589 from model luciferase is similar. The only difference occurred in the *V. longisporum*; protein oxidoreductase activities acts on paired donors, incorporate or reduce molecular oxygen, while in the template protein oxidoreductase activities act on single donors with incorporation of molecular oxygen. This study on fungal sourced luciferase present a unique opportunity away from the more well established bacterial and insect based luciferase.

Keywords: Luciferase; bioluminescent; Verticillium longisporum; protein-oxidoreductase; bioinformatic.

1. INTRODUCTION

The luciferase enzyme and luciferin substrate are the essential elements in the diverse reaction of producing light. Luciferin substrates are oxidized in the presence of molecular oxygen to convert chemical energy into light, using ATP and Mg^{2+} by the luciferase enzymes [1]. Luciferase simply refers to an enzyme that catalyses reaction that produces visible light. The emission of light begins with the formation of a product in an electronically excited state and the emission of a photon of light occurred as a result of the return of the excited state to the ground state. Luciferases are greatly varied, and catalysing different types of reactions using a wide variety of substrates [2].

Catalytic conversion of chemical energy into light causes the bioluminescence phenomenon in living organisms. This complex process is very common among some terrestrial animals, but is predominantly more common in species inhabiting marine environments [3,4]. In many luciferin identified in species of lampyridae, the luciferins are of simpler single component, while firefly and bacterial luciferin which is two-component system are comprised of flavin mononucleotide and a fatty aldehyde are more complex [1].

In this regard, research started with the firefly *Photinuspyralis* commonly available in the North American and the focus was understanding how fireflies emits light [2]. The availability of luciferase crystal structure of *Luciolacruciata* and *Photinuspyralis* has significantly advanced the understanding of key structure-function relationships that account for the efficient enzyme-catalysed emission of light in the firefly [5,6]. Luciferase has the potential to be widely used in molecular biology and bioanalytical systems as a reporter molecule due to the high quantum yield of the bioluminescence, availability of stable mutant forms of the enzyme with prescribed spectral characteristics and abundance of bacterial expression systems suitable for production of recombinant proteins in limitless quantities [7]. Currently, the application of firefly based bioluminescence system is limited to medical and pharmaceutical methods, such as in vivo imaging to visualize tumours, monitoring gene expression and its regulation [8].

Currently, all luciferases are classified as oxidoreductases, where the incorporation of molecular oxygen occurs using a single donors action. However, the diversity of luciferases are apparent as it comes from many protein families that are often unrelated. The only unifying feature is the luciferase-luciferin reactions to date have been shown to require molecular oxygen at some stage.

Hence, luciferase varieties discovered are actually in reality limited to only 22 species with a majority of it from the marine sources organisms. These are *Diaphus gigas, Oplophorus gracilirostris, Lingulodinium polyedra, Noctiluca scintillans, Pyrocystis fusiformis, Pyrocystis lunula, Pyrocystis noctiluca, Clytiagregaria, Obelialongissima, Renillareniformis, Apogonellioti, Cypridina noctiluca, Porichthys porosissimus, Vargulahil gendorfii, Lampyristur kestanicus, Luciola cruciata, Photinus pyralis, Amydetes viviani, Aliivibrio fischeri, Photorhabdus luminescens, Vibrio harveyi and Watasenia scintillans.* From all these species,

only four species; *Lampyris turkestanicus* (Iranian firefly), *Luciola cruciata* (Japanese firefly), *Photinus pyralis* (American firefly), *Amydetes vivianii* (Brazilian firefly) are insects. As insects are relatively visible and easy to study, it created a historical preference which has resulted to given much attraction to either insects or marine organisms. Therefore, there is a need to explore luciferases from organisms other than those established in existing databases. In this study, a novel sequence of fungal luciferase from *Verticillium longisporum* studied. Analysis of this Luciferase sequence would be the first study on luciferases taken from fungal species [9]. Various prior studies on the characterisation has been conducted on luciferases from various organisms but so far, has biased towards insects and of bacterial origin. There is a need to expand exploration and examine novel luciferases from new sources that may be exploited into potential commercial purposes and applications as fungi are easy to cultivate.

2. MATERIALS AND METHODS

2.1 Protein Sequence Analysis

Primary structure analysis was done using ProtParam tool https://web.expasy.org/protparam/ [10]. This step was carried out to determine the amino acid composition, the estimated molecular weight, pI value, atomic composition, and hydrophobicity for all ten proteins.

2.2 Multiple Sequence Alignment

Three multiple sequence alignments were performed using ClustalW2 tool (https://www.ebi.ac.uk/Tools/msa/clustalo/) to identify the conserved regions and comparatively analyse the different sequences [11]. The first multiple sequence alignment was performed between the five bacteria proteins and the second between the five fungi proteins and the third between all eleven proteins including model and the template.

2.3 Secondary Structure Analysis

The prediction of the uncharacterized protein of A0A0G4L2F7 from *Verticillium longisporum* secondary structure was conducted using the PSIPRED Protein Sequence Analysis software http://bioinf.cs.ucl.ac.uk/psipred/ [12]. The percentages of alpha helix, beta sheets, turns and coils were observed for further comparison and analysis. Also, PSIPRED was used in calculating the percentages of different groups of amino acids residues in an alpha helix, beta sheets, turns and coils. Since the A0A3B6UEK8 protein has a known structure, the information about its secondary structure was obtained from PDBsum database http://www.ebi.ac.uk/pdbsum [13].

2.4 Tertiary Structure Analysis

2.4.1 Homology modelling

The amino acid sequence of the uncharacterized protein A0A0G4L219 from *Verticillium longisporum* was submitted to SWISS-MODEL program (https://swissmodel.expasy.org/) to obtain a recommended template for homology modelling [14]. The templates were chosen based on sequence identity and E-value. The models generated by Swiss-Model server were assessed based on QMEAN scoring function and CQME score.

2.4.2 Homology model validation

The 3D model of luciferase proteins provided by SWISS-MODEL were validated using ERRAT http://services.mbi.ucla.edu/ERRAT [15], to identify, analyse, and calculate the residues with error values, while PROCHECK http://services.mbi.ucla.edu/PROCHECK/ [16], and Verify3D http://services.mbi.ucla.edu/Verify_3D/ [17] were used to check the accuracy and validation of the modelled 3D structures.

2.4.3 Structural comparison and visualisation

The SWISS-MODEL server (https://swissmodel.expasy.org/) was used a modelling server to model the 3D structure of A0A0G4L219 from *Verticillium longisporum*. Both the protein template structure and protein modelled result were superimposed, compared and viewed by using Chimera software downloaded from https://www.cgl.ucsf.edu/chimera/.

2.5 Phylogenetic Study

The evolutionary relationship of the bioluminescent fungi *Verticillium longisporum* A0A0G4L2F7 protein was studied using phylogenetic tree analysis. The NCBI Blast P service was used to search for a complete protein sequence. A set of 6 proteins sequences including the template were selected and aligned

Table 1. List of software and databases used for data collection and analysis

Name	Description	Link	References
UniProt	Protein sequence and functional information database.	http://www.uniprot.org/	[19]
GenBank	The National Institutes of Health (NIH) genetic sequence database.	http://www.ncbi.nlm.nih.gov/genbank/	[20]
RCSDB PDB	Information database that consists experimentally-determined structures of proteins, nucleic acids and complex assemblies.	http://www.rcsb.org/pdb/home/home.do	[21]
NPS@ (network protein sequence analysis)	Interactive Web server dedicated to protein sequence analysis.	https://npsa-prabi.ibcp.fr	[21]
PSI-BLAST	The program compares nucleotide or protein sequences to sequence databases.	https://blast.ncbi.nlm.nih.gov/Blast.cgi	[22]
ProtParam	The computational tool of various physical and chemical parameters for a given protein.	http://web.expasy.org/protparam/	[10]
PSIPRED	Protein Sequence Analysis Workbench used to predict secondary structure of proteins.	http://bioinf.cs.ucl.ac.uk/psipred/	[12]
PDBsum	Generate a full set of structural analysis.	http://www.ebi.ac.uk/thornton-srv/databases/pdbsum	[13]
Clustal W	Server use for multiple sequence alignments.	(https://www.ebi.ac.uk/Tools/msa/clustalo/)	[11]
GOR IV	Web server for protein secondary structure prediction.	https://npsa-prabi.ibcp.fr	[23]
Mega X	Software to analyze the molecular evolution of protein and DNA sequences.	http://www.megasoftware.net/	[18]
SWISS-MODEL	Structural bioinformatics webserver for protein 3D structures prediction	https://swissmodel.expasy.org/	[14]
HHpred	Structural bioinformatics webserver for protein 3D structures prediction	https://toolkit.tuebingen.mpg.de	[11]
ERRAT	Verify protein structures determined by crystallography.	http://services.mbi.ucla.edu/ERRAT/	[15]
PROCHECK	Checks the stereochemical quality of a protein structure.	http://www.ebi.ac.uk/thornton-srv/software/PROCHECK/	[16]
Verify 3D	Analyze the compatibility of an atomic model (3D) with its own amino acid sequence (1D).	http://services.mbi.ucla.edu/Verify_3D/	[24]

Fig. 1. Shows a portion of align sequencesin green colour that indicate the binding sites from Uniprot alignment

using Clustal W [14]. The phylogenetic tree was constructed using neighbor-joining method from MEGA X program [18].

2.6 Molecular Function (MF) Evaluation

Gene ontology of A0A0G4L2F7 Bac_luciferase domain-containing protein a luciferase from *Verticillium longisporum* and that of template A0A3B6UEK8 Riboflavin domain-containing protein from *Herbiconiux* was obtained from http://www.uniprot.org/ and http://toolkit.tubeingen.mpg.de.

2.7 Tools Used

The software packages and databases used in this study are summarised in Table 1.

3. RESULTS AND DISCUSSION

3.1 Primary Structure Analysis

3.1.1 Protein sequence retrieval

The sequences of the amino acids of model and template proteins were obtained from Uniprot

Table 2. Accession number, protein name, gene name and organism for the ten proteins

ID	Protein Name	Gene Name	Organism
A0A0G4L2F7	Bac_luciferase domain-containing protein	BN1708_011680	*Verticillium longisporum*
A0A3B6UEK8	Riboflavin Lyase	RcaE	*Herbiconiux* sp
A0A428X76	Bac_luciferase domain-containing protein	CEP53_006386	*Fusarium* sp
A0A2N3NFJ8	Bac_luciferase domain-containing protein	jhhlp_002928	*Lomentospora prolificans*
A0A428X76	Bac_luciferase domain-containing protein	CEP52_001034	*Fusarium* sp.
A0A428QMM5	Bac_luciferase domain-containing protein	CEP54_003681	*Fusarium* sp.
A0A4Z0Z2C8	Bac_luciferase domain-containing protein	E0Z10_g6014	*Xylaria hypoxylon*
P23146	Alkanalmonooxygenase alpha chain	luxA	*Photorhabdus luminescens*
P07740	Alkanalmonooxygenase alpha chain	luxA	*Vibrio harveyi*
P29238	Alkanalmonooxygenase alpha chain	luxA	*Photobacterium leiognathi*
P19908	Alkanalmonooxygenase beta chain	luxB	*Aliivibrio fischeri*

Table 3. Physiochemical characteristicsof Bac_luciferase domain from the A0A0G4L2F7 model and A0A3B6UEK8 template proteins

Characteristics	Model Protein	Template Protein
Number of amino acids	875	362
Molecular weight	97147.11 kDa	39596.26 kDa
Theoretical Pi	7.10	4.77
Total number of negatively charged residues (Asp + Glu):	90	53
Total number of positively charged residues (Arg + Lys)	89	30
Atomic composition	13734	5481
Chemical formula	$C_{4443}H_{6852}N_{1174}O_{1253}S_{12}$	$C_{1758}H_{2688}N_{484}O_{540}S_{11}$
Aliphatic index	92.13	79.59
Average of hydropathicity	-0.078	-0.242
The length of Bac_luciferase domain amino acid	875 (76-950)	362 (27-388)

database http://www.uniprot.org/ in FASTA format [19]. These sequences were used all over the analysis in determining the secondary structure, multiple sequence alignments, and modelling of the 3 dimensional structures. Similar proteins from various species were chosen and run alignment using Uniprot alignment from Uniprot database. Among all ten proteins only one protein possess crystallised 3 dimensional structure Alkanal monooxygenase alpha chain **(**P07740**)** in addition with the template Riboflavin RcaE that are having 3D crystallised structures (see Table 2 and Fig. 1).

3.1.2 Physicochemical characterization

Primary structural analysis was done using ProtParam tool (https://web.expasy.org/protparam/), in which the model and template proteins information analysis obtained in a tabular form (10). The table comprised of molecular weight, theoretical isoelectric point (Pi) value, amino acid composition, total number of positively charged and negatively charged residue, atomic composition, chemical formula, estimated half-life, aliphatic index and hydropathicity value of the model protein (see Table 3).

The primary sequence of A0A0G4L2F7 model protein is composed of 956 amino acids residues while A0A3B6UEK8 template protein is composed with 461 amino acids residues. The physiochemical parameters of A0A0G4L2F7 model protein, and A0A3B6UEK8 template protein were computed using the ProtParam tool (Table 4). The analysis revealed that A0A0G4L2F7 is more than two times bigger than A0A3B6UEK8 with the calculated molecular weights of 97147.11 and 39596.26 respectively. The computed isoelectric point (Pi) value for the model enzyme is 7.10, while that of template is 4.77 respectively. The finding of this study shows that model enzyme contain almost balance charged amino acid residues 90 negatively charged (-R) and 89 positively charged (+R) as indicate in the (Table 4). In the case of template enzyme negatively charged residues (-R) are more in number when compared with the positively charged amino acid residues (+R) in which they are having 53 and 30 respectively, and the isoelectric point (Pi) values is 4.77 as indicated in (Table 3) also. As indicated in the (Table 4) the length of Bac_luciferase domain amino acid residues from model protein are 875, while that of template protein are 362 amino acid residues respectively. The Bac_luciferase domain from the A0A0G4L2F7 model protein comprised of 91.53% of the total enzyme amino acid residues, while A0A3B6UEK8 template protein which Riboflavin Rcae comprised of 78.52% accordingly. This shows that both model

Table 4. Composition of amino acidresidues and percentage of model luciferase protein and template protein

Amino acids	Number of amino acids in the model	Amino acids percentage %	Number of amino acids in the template	Amino acids percentage %
Ala (A)	78	8.9	46	12.7
Arg (R)	48	5.5	21	5.8
Asn (N)	17	1.9	6	1.7
Asp (D)	49	5.6	28	7.7
Cys (C)	3	0.3	4	1.1
Gln (Q)	28	3.2	13	3.6
Glu (E)	41	4.7	25	6.9
Gly (G)	67	7.7	30	8.3
His (H)	23	2.6	11	3.0
Ile (I)	68	7.8	13	3.6
Leu (L)	83	9.5	29	8.0
Lys (K)	41	4.7	9	2.5
Met (M)	9	1.0	7	1.9
Phe (F)	45	5.1	13	3.6
Pro (P)	52	5.9	19	5.2
Ser (S)	67	7.7	20	5.5
Thr (T)	58	6.6	21	5.8
Trp (W)	23	2.6	9	2.5
Tyr (Y)	27	3.1	11	3.0
Val (V)	48	5.5	27	7.5

Table 5. Shows the composition of various amino acid groups

Amino acid composition By Groups	Bac_luciferase for *V. longisporum*Model Percentage %	Riboflavin Rcae 5w48 Template Percentage %
Acidic D, E	10.3%	14.6%
Aliphatic I, L, V	22.8%	19.1%
Aromatic H, F, W, Y	13.4%	12.1%
Basic R, H, K	12.8%	11.3%
Charged R, D, E, H, K	23.1%	25.9%
Hydrophobic A, C, F, I, L, M, V, W, Y	43.8%	43.9%
Polar R, N, D, E, Q, H, K, S, T	42.5%	42.5%
Big E, F, H, I, K, L, M, Q, R, W, Y	49.8%	44.4%
Small A, C, D, G, N, P, S, T, V	46.6%	55.5%
Tiny A, C, G, S	24.6%	27.6%

and template Bac_luciferase domain within the enzyme are having higher number of amino acid residues compared with the remaining amino acid residues outside the domain.

3.1.3 The Composition of amino acid groups for the model and template

The research established the percentage composition of amino acid residues from model which was obtained from (https://web.expasy.org/protparam/) as indicated in (Table 5) as follows: 8.9% Alanine (A), 5.5% Arginine (R), 1.9% Asparagine (N), 0.3% Cysteine (C), 3.2% Glutamine (Q), 4.7% Glutamic acid (E), 7.7% Glycine (G), 2.6% Histidine (H), 7.8% Isoleucine (I), 9.5% Leucine (L), 4.7% Lysine (K), 1.0% Methionine (M), 5.1% Phenilalanine (F), 5.9% Proline (P), 7.7% Serine (S), 6.6% Threonine (T), 2.6% Tryptophan (W), 3.1% Tyrosine (Y), and 5.5% Valine respectively. The template protein structures comprise the following Amino acid residues 12.7% Alanine (A), 5.8% Arginine (R), 1.7% Asparagine (N), 1.1% Cysteine (C), 3.6% Glutamine (Q), 6.9% Glutamic acid (E), 8.3% Glycine (G), 3.0% Histidine (H), 3.6% Isoleucine (I), 8.0% Leucine (L), 2.5% Lysine (K), 1.9% Methionine (M), 3.6% Phenilalanine (F), 5.2% Proline (P), 5.5% Serine (S), 5.8% Threonine (T), 2.5% Tryptophan (W), 3.0% Tyrosine (Y), and 7.5% Valine respectively (see Table 5 and Fig. 8) (10).

3.1.4 The composition of amino acid groups

As both model and template proteins are made up of helices, strands, coils, and turns, it is outward that the constituent proteins are distributed across these features to form the unit's entirety. Their distribution configuration is in such a way that different types of amino acids characterize each of the constituent structures in different percentage. The distribution in percentage of the amino acid residues for model is: acidic possess (10.56 aliphatic (22.8%), aromatic (13.4%), basic (12.8%), hydrophobic (43.8%), polar (42.5%), Big (49.8%), small (46.6%), tiny (24.6%) respectively. The distribution in percentage of the amino acid residues for template protein is: acidic possess (14.6%), aliphatic (19.1%), aromatic (12.1%), basic (11.3%), hydrophobic (43.9%), polar (42.5 %), Big (44.4%), small (55.5%), and tiny (27.6%) respectively (see Table 5 and Fig. 2). Taking a look on the trends established by the two proteins, it is apparent that there are more number of amino acid residues in the model than in the template base on the constituent structures (see Table 5).

According to [18], the values of each category of amino acids in each region can be represented diagrammatically using the alphabetical letters that represent the areas in which each amino acid group falls. The diagram provides the same information contained in the data used to create a snip and sketch of the individual compositions of each material in a single unit of Bac_luciferase protein. Even though the percentages presented by [18] are not the same with the collected data, they show the same trends characterizing each region within the protein unit (see Figs. 10,11, and 12).

Besides the chat representation of the individual percentages of amino acids residues indicating the composition of each groups in the model and template, the term 'sheet region, is used which is the equivalent to the strand region presented in the data collected. It is apparent that the two sources of information are different primarily regarding the methods of data presentation

adopted. However, the implications of the given statistics are consistent across the two sources as they communicate the same patterns of amino acid distribution.

3.2 Secondary Structural Analysis

By using the PSIPRED Protein Sequence Analysis software for the prediction of the secondary structure of the protein as carried out it shows that, the number of possible amino acid residues from the model protein to form helix structure are 442, equals to 46.23%, while 104 amino acid residues form beta sheet (10.88%) and 410 amino acids residue form turns (42.89%) as shown in Fig. 2.

By using the PSIPRED Protein Sequence Analysis software for the prediction of the secondary structure of the protein as carried out it shows that, the number of possible amino acids that forms a helix equals to 39.05%, while 58 amino acid residues form beta sheet (12.58%) and 223 amino acids residue form turns (48.37%) as shown in Fig. 3.

3.2.1 The percentages of the constituent structures

The finding of this research discovered that each protein unit consists of helices, strands, coils, turns, and loops. The percentages of helices, strands, and loops sections constitute 33%, 19%, and 48% in the model Bac_luciferase protein from *V. longisporum* respectively. On the other hand, the percentage compositions by helices, stands, and sheet sections, according to the available data from the template Riboflavin RcaE are 39%, 12%, and 49%, respectively (see Table 6 and Figs. 4 and 3.7).

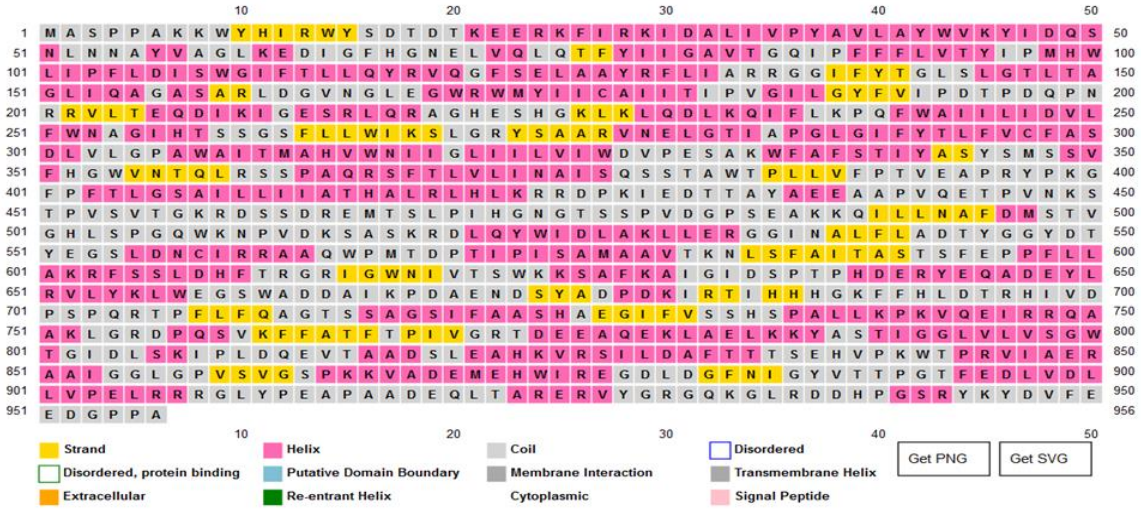


Fig. 2. Summary of helices, strands, coils, turns and loops Model

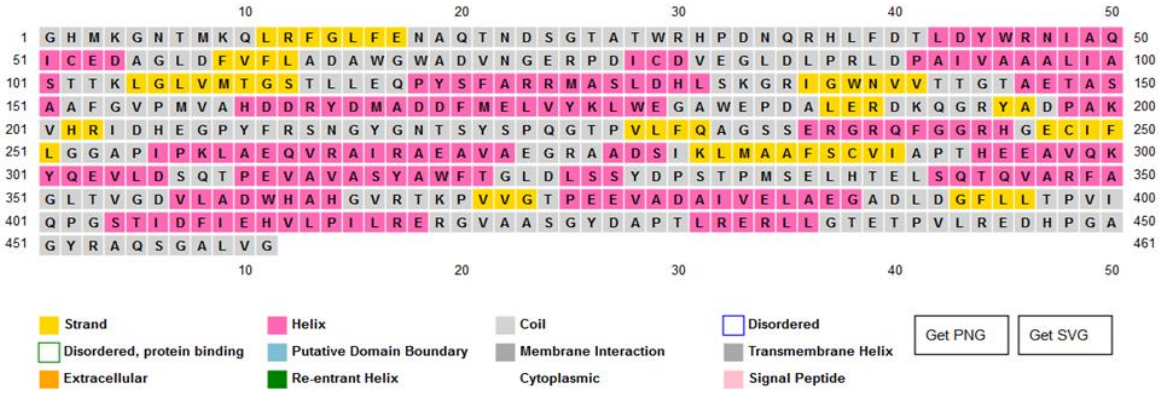
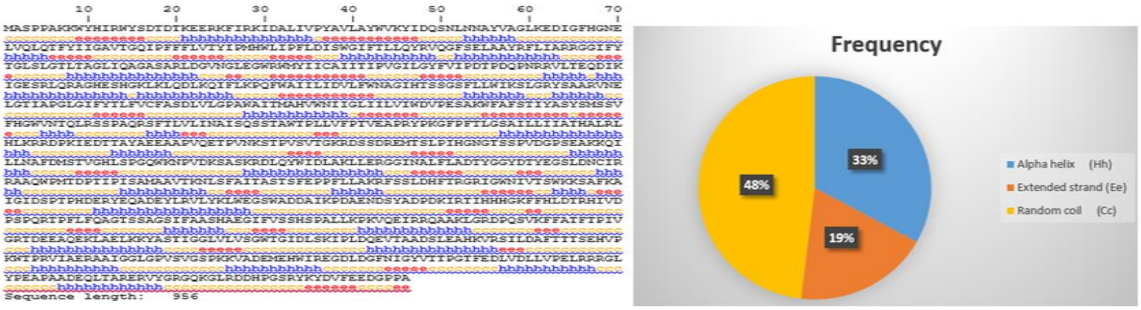


Fig. 3. Summary of helices, strands, coils, turns and loops Template

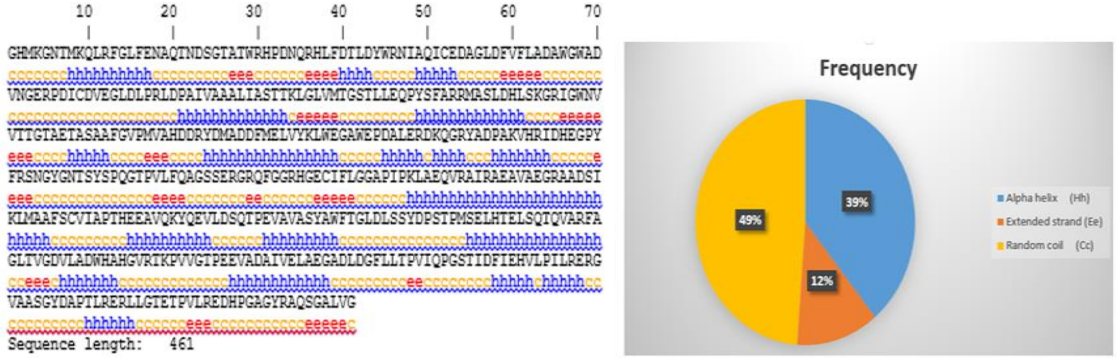
Table 6. The percentages of helices, strand, coils, turns, and loops in Model and template proteins

Types of Secondary structures	Model	Template
Helices	33%	39%
Strands	19%	12%
Coils, turns, and loops	48%	49%



(a) (b)

Fig. 4. Summary of alpha helix, extended strand and random coil of the Model protein from GOR IV



(a) (b)

Fig. 5. Summary of alpha helix, extended strand and random coil of the Template protein from GOR IV

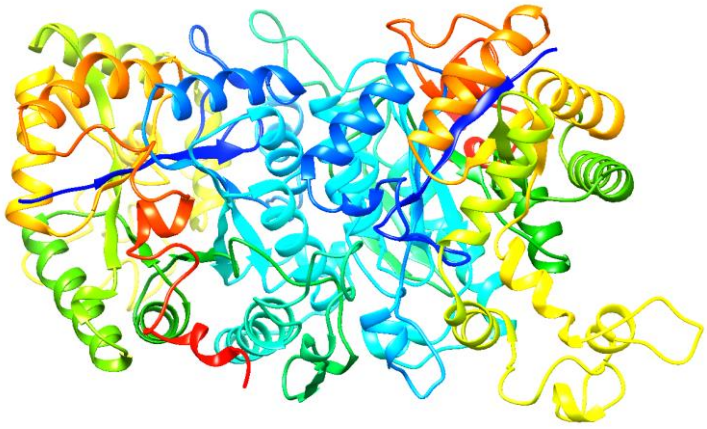


Fig. 6. 3D structure of Model protein Bac_luciferase (A0A0G4L2F7) from *V. longisporum*Ribon as modelled using the SWISS-MODEL database

Table 7. A0A0G4L2F7's top four proposed templates from two server

Enzymes name	Server name	Template	Protein name	Length	Identity	E-value
A0A0G4L2F7	HHpred	6AK1	Dimethyl-sulfide monooxygenase	488	51.4%	4.4e-57
		3B9O	Alkane monooxygenase	440	49.1%	2.3e-46
		5W4Z	Riboflavin Lyase	461	51%	1.9e-40
		5XKC	Dibenzothiophene desulfurization	453	50.4	7.2e-44
	Swiss Model	5W48	Riboflavin Lyase	457	33.48%	NA
		6LRI	Hexachlorobenzene oxidative dehalogenase	451	36.55%	NA
		6ASK	Putative monooxygenaseMoxC	434	31.07%	NA
		3SDO	Nitrilotriacetatemonooxygenase	434	29.98%	NA

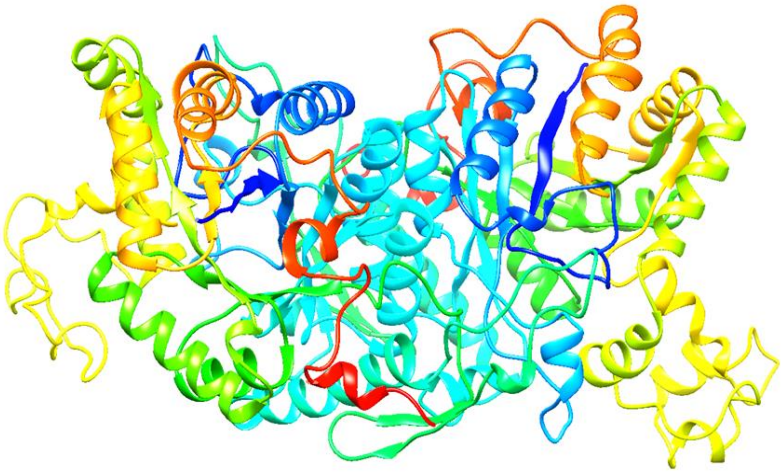


Fig. 7. 3D structure of Template protein template Riboflavin LyaseRcaE from *E. coli* with PDB ID 5w48.1.A as obtain from SWISS-MODEL software

By using the GOR IV software [25], the secondary structure of both A0A0G4L2F7 model and A0A3B6UEK8 template proteins were measured as indicated in (Figs. 2, 5 and 6). The entire amino acid residues of both model and template proteins consists of alpha helix (Hh), extended strand (Ee) and random coil (Cc), in which turn was dominant going by the percentage comprising (48%) and (49%) from both the model and template proteins, while helix consist of (33%) and (39) and also Beta sheets possess (19%) and (12%) accordingly (see Table 6, Fig. 4a and b and also 7).

The general framework of such protein in a single unit form is made up of strands, helices, coils, and turns. Therefore in each of these features establishes a specific percentage of the entire unit under consideration. The evidence as obtained from data-based indicates that a single unit of the Bac_luciferase protein from *V. longisporum* consists of coils, turns, and loops 48% and helices 33% while strands constitute the remaining 19% (see Table 6, and Fig. 4). This data indicates that coils, turns, and loops are the dominant structures in each unit of this model protein. It also reveals that helices constitute the second largest component of each composition. The finding is in conformity with the existing literature.

3.3 Tertiary Structure

3.3.1 Identification of the template

In order to have a well understanding of the protein structure, identifying the 3-dimensional (3D) structure of A0A0G4L2F7 enzymes is critical. The primary sequences of A0A0G4L2F7 were submitted to various protein alignment search tools to determine the suitable template for the 3D model building. The results from HHpred [26] and Swiss Model [27] servers were evaluated as represented in Table 7 shows the results of all two servers, the identity and e-values is within the acceptable range, result obtained from Swiss model database the identity is 33.48%, while the one from HHpred is 51.3% and e-values is 1.9e-40. The template and model proteins are both homodimers.

3.3.2 Homology modelling

The SWISS-MODEL database modeller used in predicting the 3D structure of Model protein Bac_luciferase from *V. longisporum*. The modelled structure (Table 7) showed 33.48% percentage identity to the template Riboflavin LyaseRcaE from *E. coli* with PDB ID 5w48.1.A.

3.4 Homology Modelling Validation

3.4.1 ERRAT2

ERRAT analyses the statistics of non-bonded interactions in between different type of atoms and plots the value of the error function against the position of a nine-residue sliding spaces obtained from a database http://services.mbi.ucla.edu/ERRAT. The different atoms types are distributed non-randomly with respect to each other within the proteins. Protein structure can be proved by distinguishing in between correctly and incorrectly determined sections of protein structures based on characteristic interaction of atoms. From ERRAT software, a good high-

resolution structure generally produces overall quality factor valuing around 95. The overall quality factor of predicted Bac_luciferase (A0A0G4L2F7) from *V. longisporum* from ERRAT2 analysis is 91.01.

3.4.2 Procheck

The prediction of the model Bac_luciferase protein from *Verticilliumlongisporum*a fungi specie from a SWISS model Database using a Ramachandran plot (Fig. 9) shows the validation of the template as shown in (Table 7). From the Table 7, the residues in most favoured region account for 90.6%, residues in additional allowed regions are 8.3%, and residues in generously allowed region 0.6% which give general total of residues that fall within allowed to 99.5% , while the amino acid residues which are not in allowed region are 0.5%.

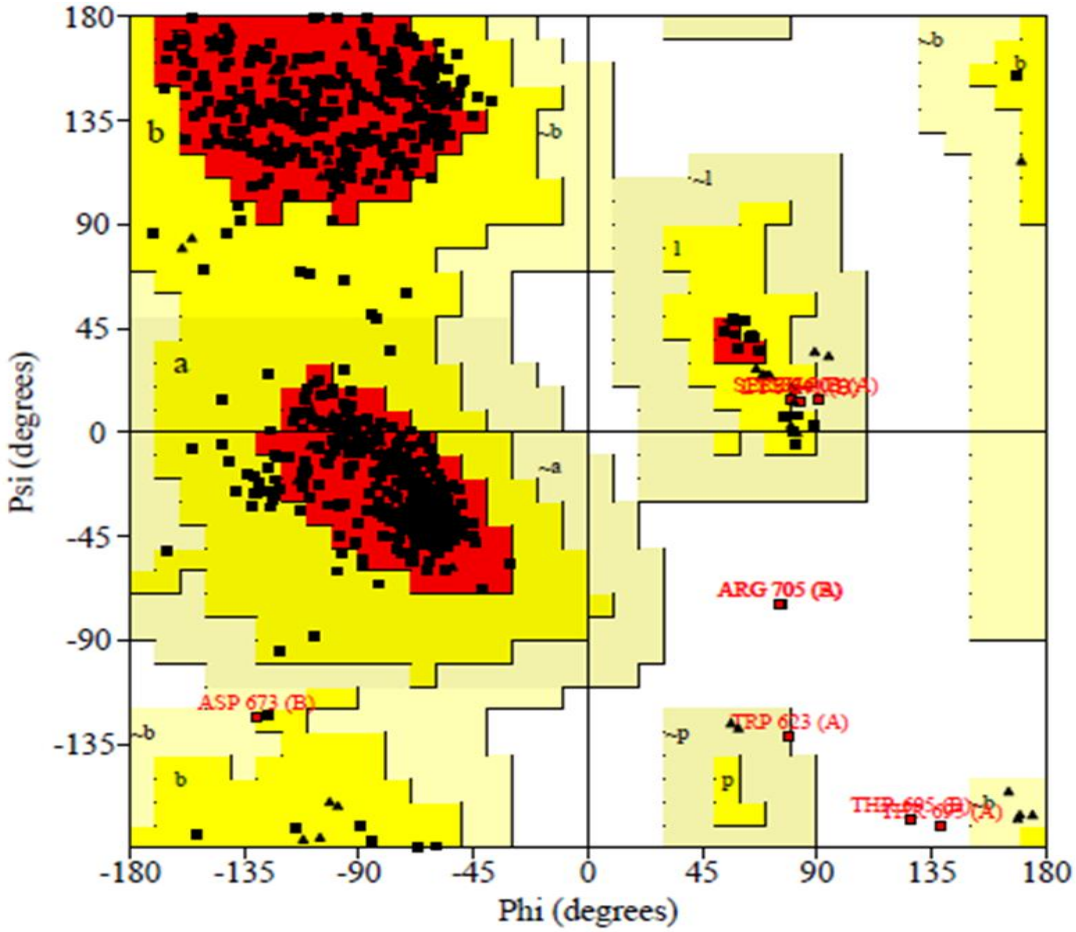


Fig. 8. Model validation using Ramachandran Plot [28]

Table 8. Ramachandran plot validation percentage of model luciferase protein

Evaluation of residues	Score
Residues in most favoured regions [A,B,L]	(720) 90.6%
Residues in additional allowed regions [a,b,l,p]	(66) 8.3%
Residues in generously allowed regions [~a,~b,~l,~p]	(5) 0.6%
Residues in disallowed regions	(4) 0.5%
Number of non-glycine and non-proline residues	(795) 100.0%
Number of end-residues (excl. Gly and Pro)	4
Number of glycine residues (shown as triangles)	68
Number of proline residues	56
Total number of residues	923

3.4.3 Verify3D

Verify3D program determines the compatibility of the atomic coordinates of the 3D model with its own amino acid sequence (1D). Three-dimensional profiles computed from correct protein structures match their own sequences with high scores. The model structure attain a pass from validation if at least 80% of the amino acids have scored greater or equal to 0.2 in the 3D/1D profile. The quality of the 3D structure homology of the Bac_ luciferase protein from Verify3D is as shown below:

91.01% of the residues attained an averaged 3D-1D score >= 0.2 Pass: At least 80% of the amino acids must attained scored >= 0.2 in the 3D/1D profile.

3.5 Structural Comparison

3.5.1 Size

There are some differences between the modelBac_luciferase protein from *V. longisporum*and the template 5W48 Riboflavin protein Rcae from *E. coli* in term of their size. The Bac_luciferase protein and Riboflavin protein even though both are homodimers protein that composed of two identical subunits, but the model possess a longer amino acids chain when compared with that of template. The Bac_luciferase protein contains 956 number of amino acids while the Riboflavin protein RcaE contain 461 respectively. The number of residues in each protein was determine by the number of amino acids. Therefore, the model protein has 956 residues, while the template protein possesses 461 residues. This clear difference in term of both protein size can be seeing on the tertiary structures of both proteins (Fig. 19a, b).

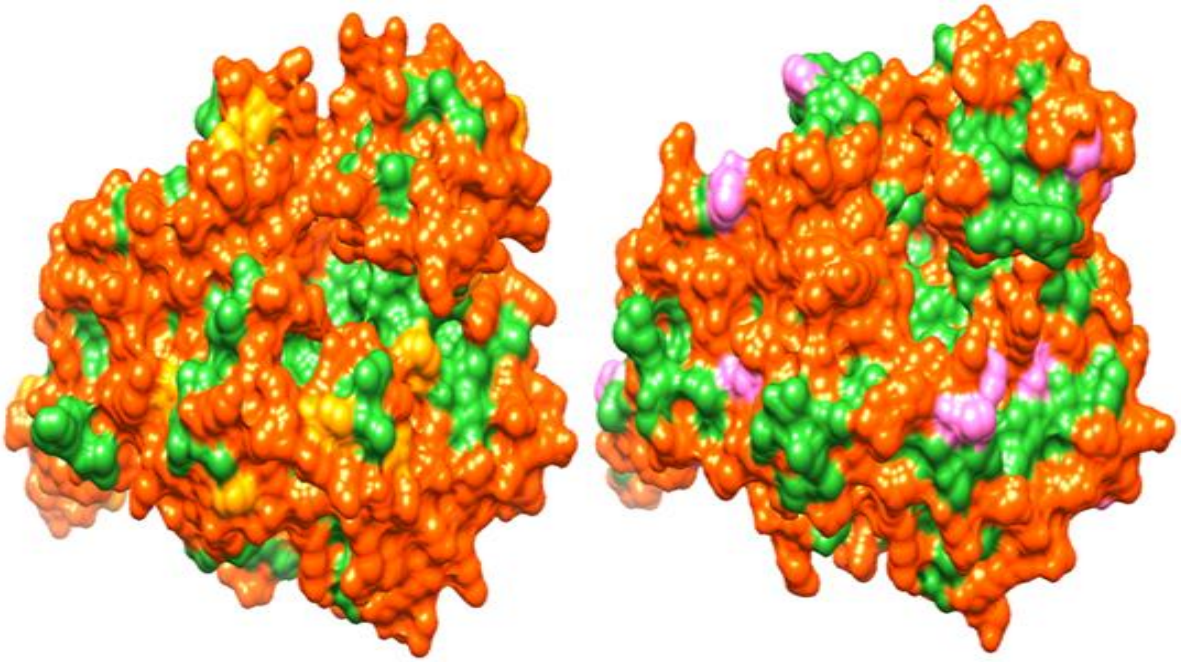
3.5.2 Elevated thermostability

Thermostability of Luciferase enzyme among different species of bioluminescent organisms had been established. The living Earth is comprised of a wide variety of diverse environments which necessitate a different ways of strategies for living organisms to be alive. Among the key factors is temperature that has direct effects to the biochemical adaptation by organisms to their environment for survival. Organisms living in extreme temperatures (hyperthermophilic organisms) are of particular importance due to the proteins they contain which can be isolated from them and can remain in a stable state and function in such environments. The proteins if isolated are of significant importance for industrial processes and engineering proteins from organisms that are surviving in moderate temperatures (mesophilic organisms). With these unique features of biochemistry they possess, they can provide an exceptional opportunity to researchers in understanding the relationships between structural features and biological functions [29].

The whole biotechnological exploitation of enzymes is still hampered by way of their low interest, low balance and excessive value in which luciferase as an enzyme is not in exception. Temperature-dependent catalytic properties of enzymes are a key to green and price-powerful translation to industrial programs. Organisms adapted to temperature extremes are a rich source of enzymes with wide ranging thermal homes which, if isolated, characterized and their structure–function–balance dating elucidated, should underpin a selection of technology. Enzymes from thermally-adapted organisms along with psychrophiles (low-temperature) and thermophiles (excessive-temperature) are a full-size herbal resource this is already below scrutiny for his or her biotechnological capacity, but psychrophilic and thermophilic enzymes display a pastime–stability exchange-off that necessitates using various genetic and chemical modifications to further improve their properties to suit various commercial packages. This evaluation describes in fact the properties and biotechnological applications of each in thermal-adapted and thermophilic enzymes. furthermore, the overview severely examines methods to improve their importance for biotechnology, concluding by using providing an incorporated approach concerning thermally-adapted, genetically and magnetically changed enzymes to make biocatalysis greater efficient and value-effective [30].

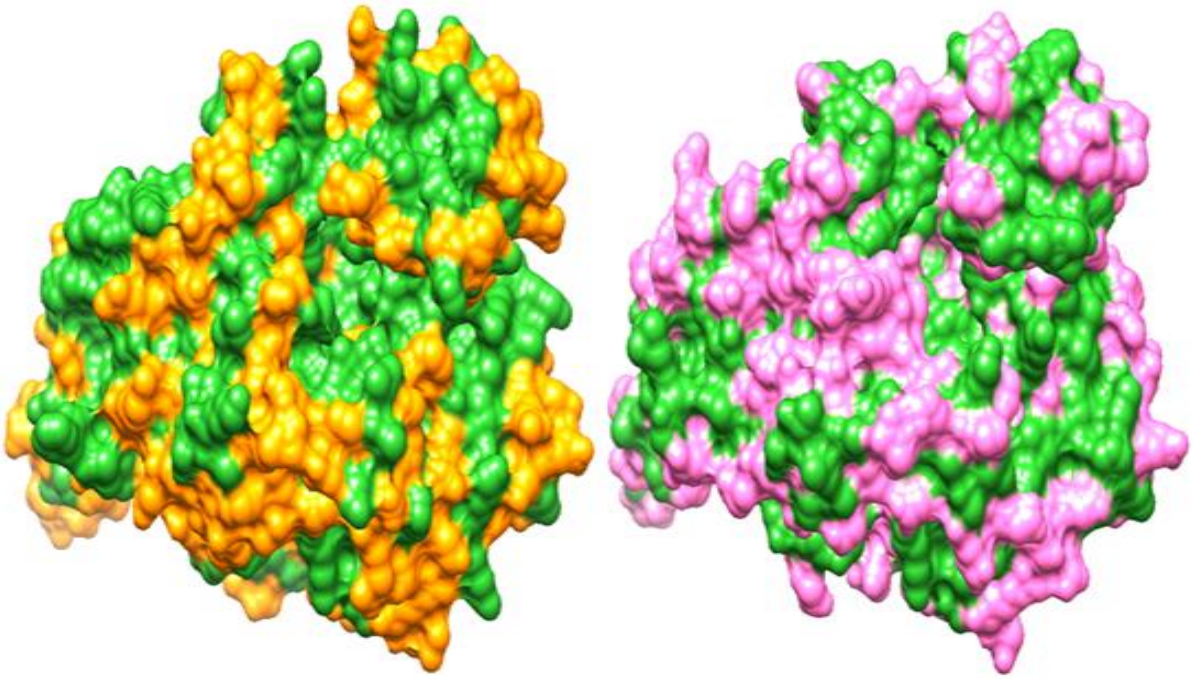
The amino acids composition commonly identified in thermophilic proteins, include an increase in residue hydrophobicity, a decrease in uncharged polar residues, a growth in charged residues, significantly high in aromatic residues [31]. In the model protein hydrophobic amino acid residues is (43.8%) and polar amino acid residues is (42.5%), respectively, while in template protein there are hydrophobic (43.9%) and polar (42.5 %) respectively (see Table 5). It has been widely acknowledged that the aliphatic

amino acids might make a contribution to the hydrophobic interplay that possess essential energy for maintaining conformational stability in internal part of protein at higher temperature [31]. Hydrophobic amino acid residues constituted between 6-8% and 12-14% in thermophilic proteins [28]. (see Table 5). (see Figs. 11, 12, and 13).



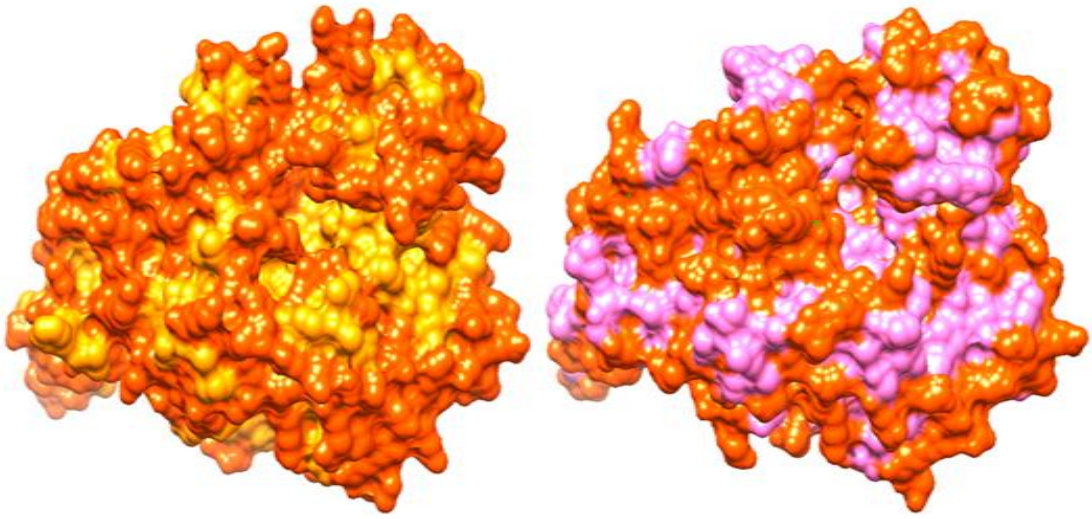
a. Model **b. Template**

Fig. 9. Shows hydrophobic and polar amino acid residues in both model and template proteins



a. Model **b. Template**

Fig. 10. Shows hydrophobic amino acid residues in both model and template proteins



a. Model **b. Template**

Fig. 11. Shows polar amino acid residues in both model and template proteins

Table 9. Shows amino acid binding sites of both model and template

Corresponding amino acids of Bac_luciferase domain-containing protein Model	Binding site amino acids of 5w48 Riboflavin Rcae Template
Asp 543	Asp 64
Thr 589	Thr 110

3.5.3 Binding sites

The template Riboflavin protein 5w48 possesses two binding sites which obtained from Uniprot. Both the binding site amino acids for the template protein and corresponding amino acids in the model protein Bac_luciferase are shown in the Table 8. On the other hand, the template used in modelling the Bac_luciferase protein Riboflavin 5w48 https://www.rcsb.org/structure/5w48. The binding site amino acids for 5w48 were obtained from Uniprot. Both the binding site amino acids for the template protein and corresponding amino acids in the model protein Bac_luciferase are shown in Table 6.

Fig. 12 shows that superimposed image of the luciferase protein model Bac_luciferase with that of template 5W48. The result shows the similarity of the model protein with template 5W48 from SWISS Model database is 33.48%. As finding shows that, there are no previous research showing the location of the residues that forms the binding site, the template 5W48 used as a yardstick to guide the possible location of similar amino acids that would form the binding site and to further find if there are any catalytic activities that are taking place in that location to qualify it to be an active site on the model protein. Two FMN binding sites has been identified from the template in Uniprot database that contain a total number of two important amino acid residues in the site for the template 5W48. The binding sites were confirm to be were the catalytic activities of Riboflavin enzyme RcaE are occurring, as mention RcaE catalyzes the formation of Lumichrome (fluorescent), [32]. Considering this as a reference point, the two amino acids were aligned with the model protein as highlighted in Fig. 12a. The amino acids in the template structure were highlighted using green colour, while in the model structure using red colour as shown in Figs. 12b and 12c. The result of 3D superimposed structure revealed that both binding sites exist at related location within the conformational structure of luciferase as shown that the proximity of the 5W48 residues responsible for forming the binding site are located near the modelled residues which also are in conformity with the alignment from the sequence, even though sequence number varied (Asp D64, D548 and Thr T110, T589 of template and model respectively as indicated in Table 9), the binding sites of both proteins are located at a similar position from each other along the protein sequence.

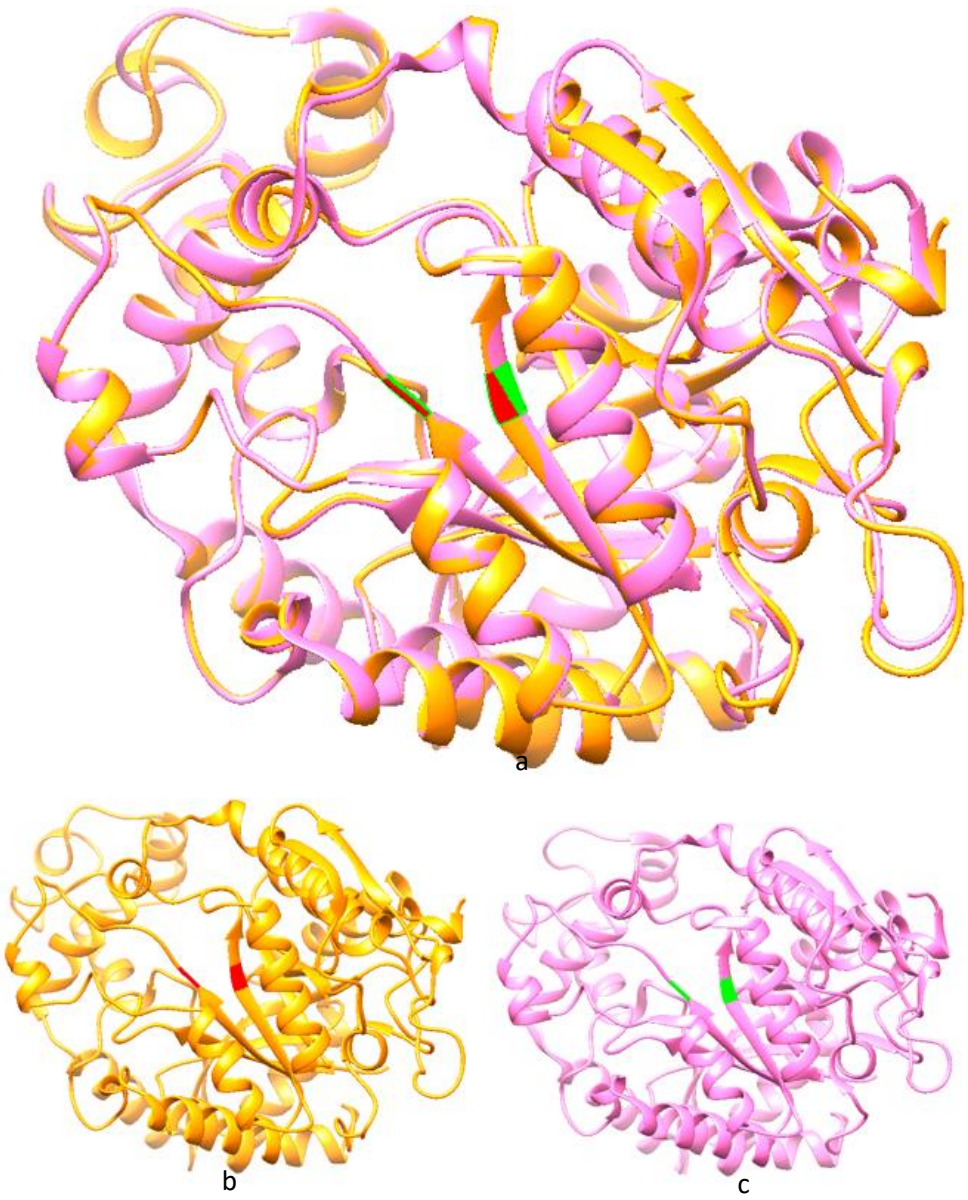


Fig. 12. (a) Showing superimposed image of model Bac_luciferase and template 5w48 where the binding site amino acids are highlighted in red and green accordingly. (b) Showing the binding site amino acids of the model Bac_luciferase. (c) Showing the binding site amino acids of template 5w48

Fig. 12 shows the binding sites amino acids for both model (orange) and the template (hot pink) in the ribbon 3D image is at the same location when comparing the surface view of the 3D image of the protein even though its position is at different amino acid residue (see Table 9).

Fig. 14 shows surface images of the two proteins the model Luciferase protein from *Verticillium longisporum* (4.16a), and the template 5W48 protein from Riboflavin Rcae of *E. coli* (4.16b) with all binding sites highlighted. The two proteins are classified as oxidoreductase, meaning they have the same basic function. The binding sites in the two proteins are similar and are in the same location.

The distance between the two binding sites and the distance between the binding sites and ligand has been measured from the model protein and compared it with the template protein. The distance between the two binding sites from the two proteins are 4.917Å resolution in the model protein and 4.809Å respectively. The distance

between ligand and threonine in the model protein is 7.898Å and that of the template is 7.795Å respectively. The distance between aspartic acid in the model protein is 7.075Å, while that of the template is 7.128Å (see Figs. 15 and 16).

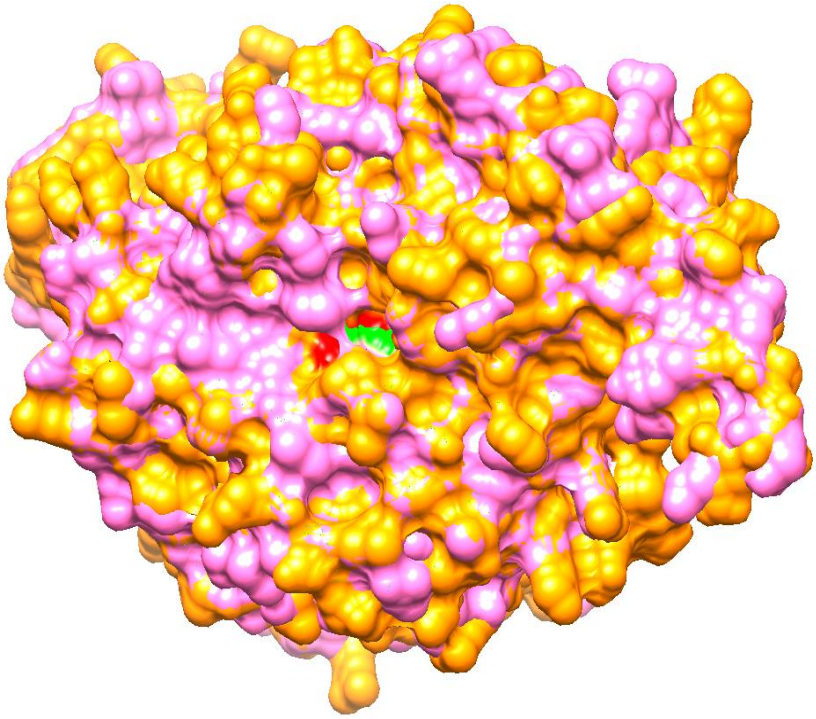
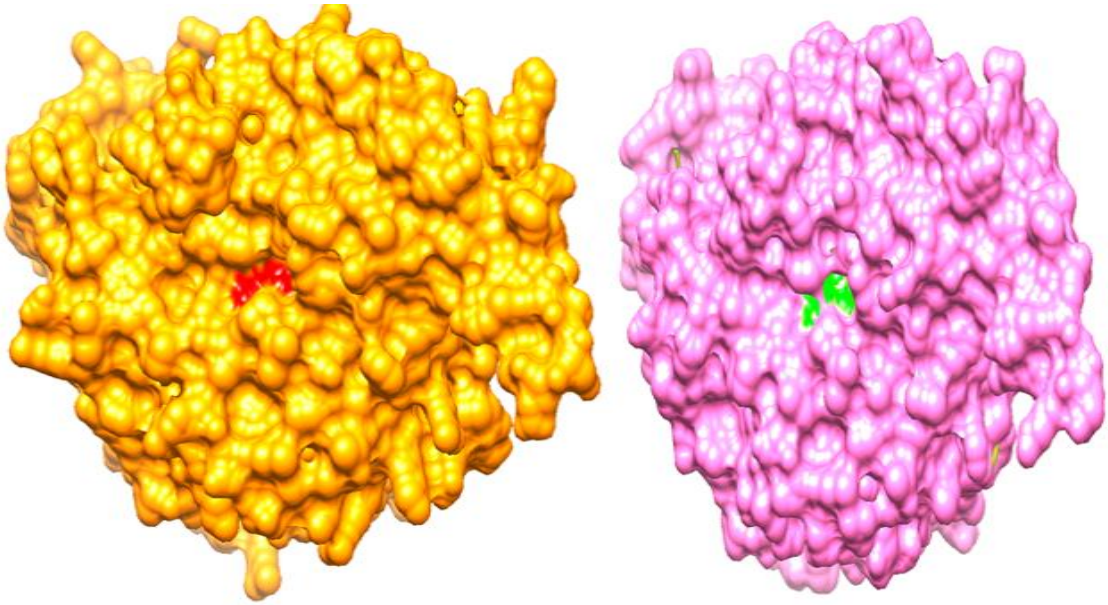


Fig. 13. Surface image of the binding sites of model Bac_luciferase protein from *V. longisporum* and template 5W48 Riboflavin lyaseRcaEprotein



a b

Fig. 14. (a) Surface image of model Luciferase from *V. longisporum* where the binding site amino acids are highlighted in red. (b) Surface image of the template Riboflavin RcaE 5w48 in which the binding site amino acids are highlighted in green

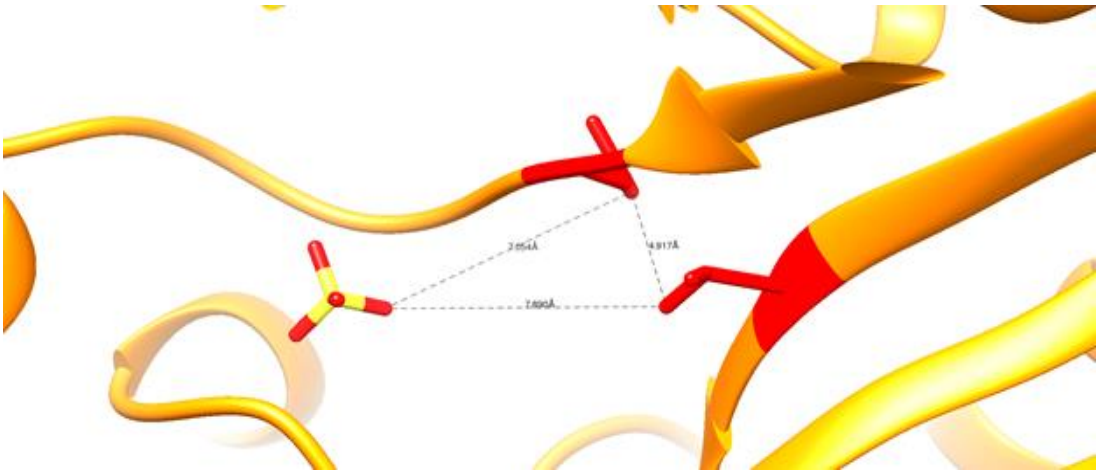


Fig. 15. Model Luciferase from *V. longisporum*

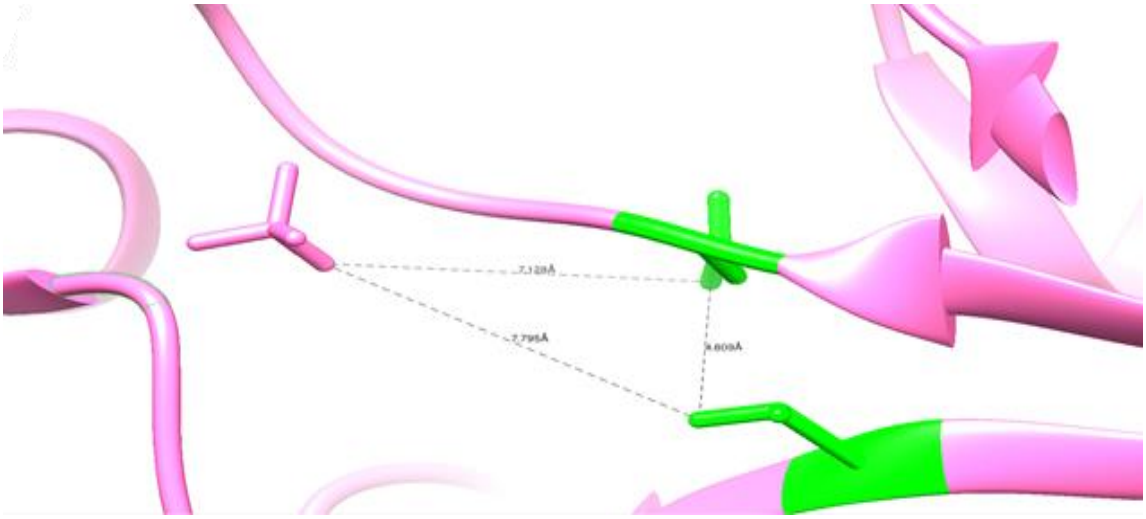


Fig. 16. Template RiboflavinRcaE 5w48

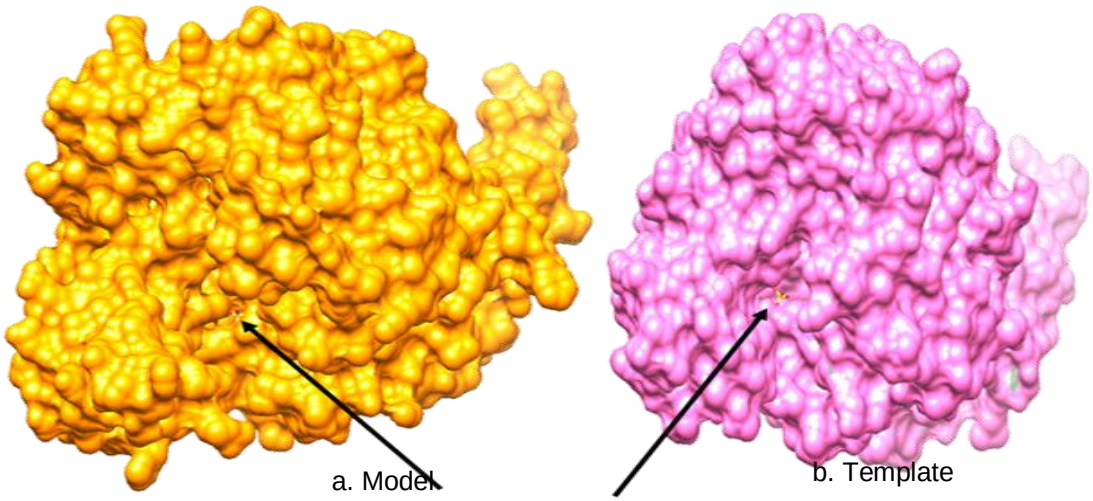


Fig. 17. The model and template proteins surface shows the position of ligand on similar position in both proteins

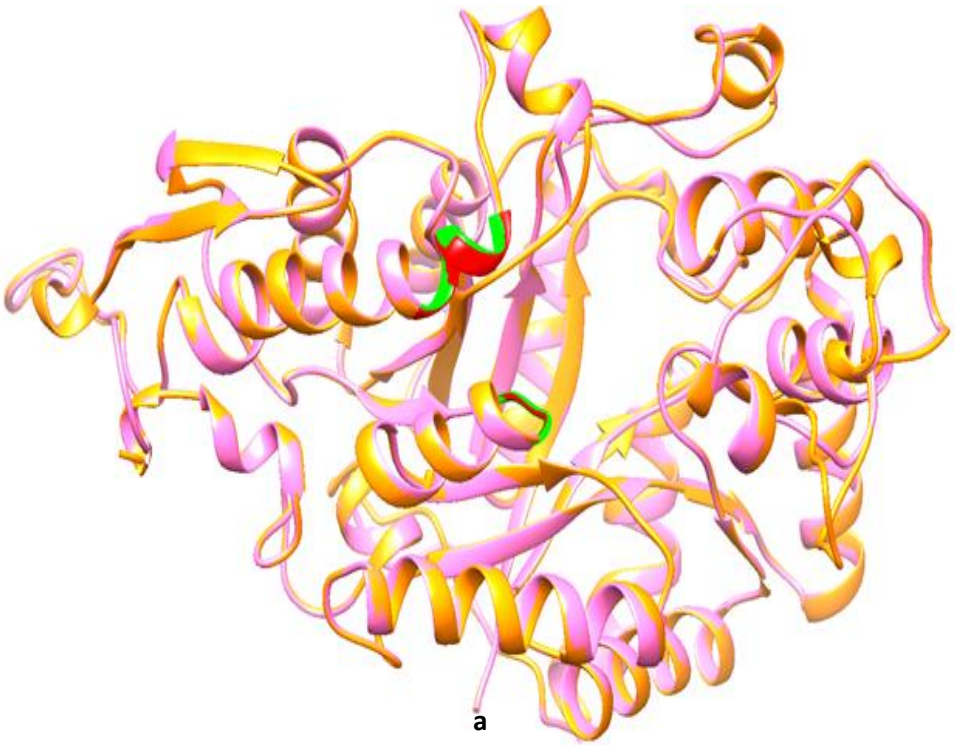
The position of ligand inside the two pocket proteins of the model Bac_luciferase from *V. longisporum* and template Riboflavin RcaE from *Herbiconiux* sp are in conformity with each other.

3.5.4 Nucleotide binding sites

The template protein 5w48 Riboflavin lyaseRcaE shows little differences in amino acid residues constituents from the model luciferase protein. The first binding from the template comprised HDDRY amino acids residues at 160, 161, 162, 163, and 164, while the model protein comprised of HDERY amino acids residues at 639,640,641, 642 and 643, which shows the substitution of D at 162 with E at 641 which belong to same R-group, negatively charged (acidic) amino acids. The second nucleotide binding from Template protein comprised of AGSS at 232, 233,234 and 235, while that of Model involved AGTS at 712,713,714 and 715, which shows substitution of Sat 234 with T at 714, hence belong to the same R-group polar amino acids as indicated in Table 10 and in Figs. 18, 20 and 21 respectively.

Table 10. Shows Nucleotide binding of the amino acid of both model and template

Nucleotide binding sites of Bac_luciferase domain-containing protein Model	Binding site amino acids of 5w48 Riboflavin Rcae Template
160 161 162 163 164 H D E R Y	639 640 641 642 643 H D D R Y
232 233 234 235 A G T S	712 713 714 715 A G S S



a

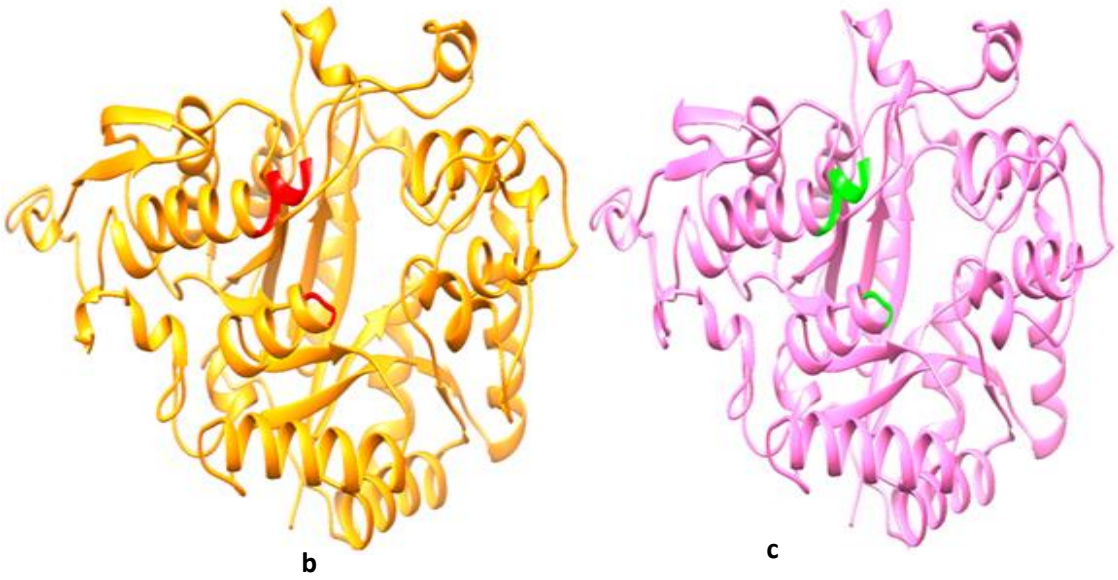


Fig. 18. Showing the Nucleotide binding in both the Model and Template proteins

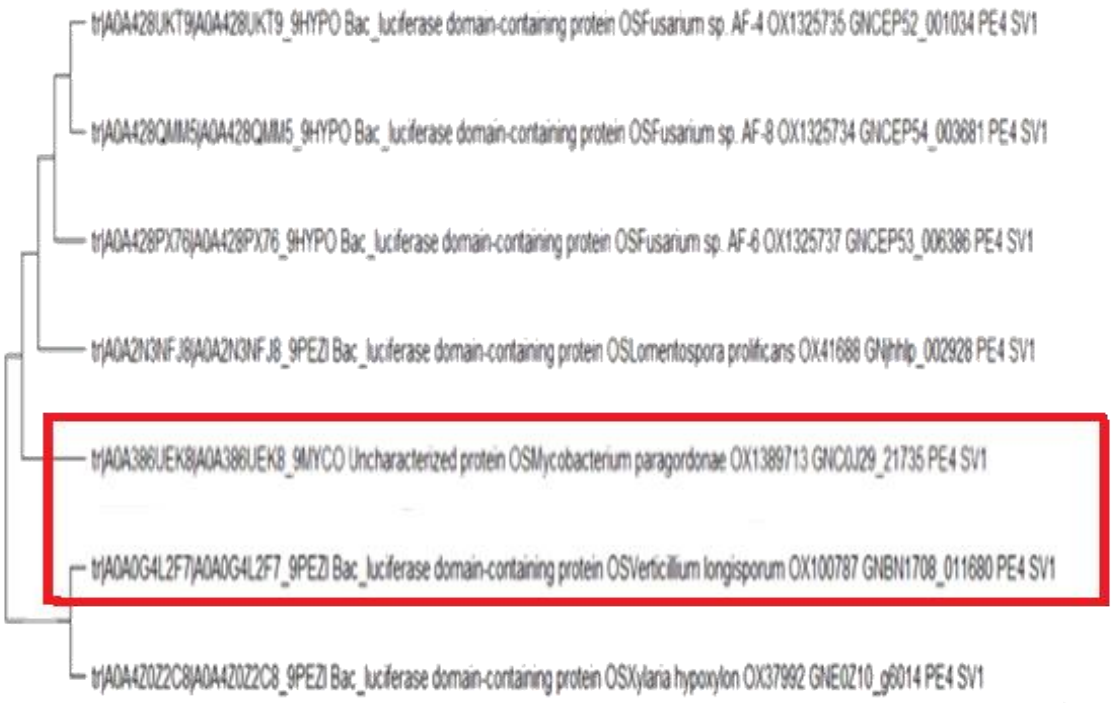


Fig. 19. Summary of Phylogenetic tree result using maximum likelihood tree neighbour-joining method from Mega X

3.6 Phylogenetic Study

The evolutionary relationship among the Bac_luciferase enzymes was investigated by phylogenetic analysis. A search for complete sequences was performed using the NCBI BlastP service. The alignment results were used to construct the phylogenetic tree using neighbor-joining method implemented in the MEGA X program [18]. Based on the phylogenetic tree in Fig. 19, it clearly shows how closely related *Herbiconiux* species with

Verticillium longisporum among all the six species in phylogenetic tree, it is the closest (see Fig. 19).

3.7 Molecular Function (MF) Evaluation

Gene ontology of A0A0G4L2F7 Bac_luciferase domain-containing protein a luciferase from *Verticillium longisporum* which obtain from http://www.uniprot.org/ and http://toolkit.tubeingen.mpg.de shows that the Oxidoreductase activity, which is acting on paired donors, incorporate or reduce molecular oxygen. The molecular function result to catalysing oxidation-reduction (redox) reaction due to this hydrogen or electrons transferred from each of two donors, and molecular oxygen is reduce or incorporate into a donor.

Fig. 20. Shows the alignmentsof the two positions of nucleotide binding Uniprot underline with red colour

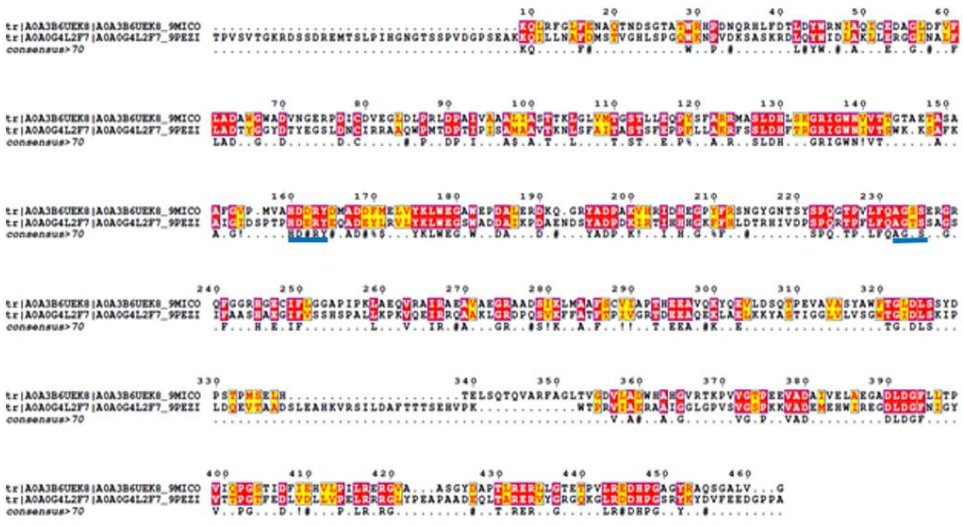
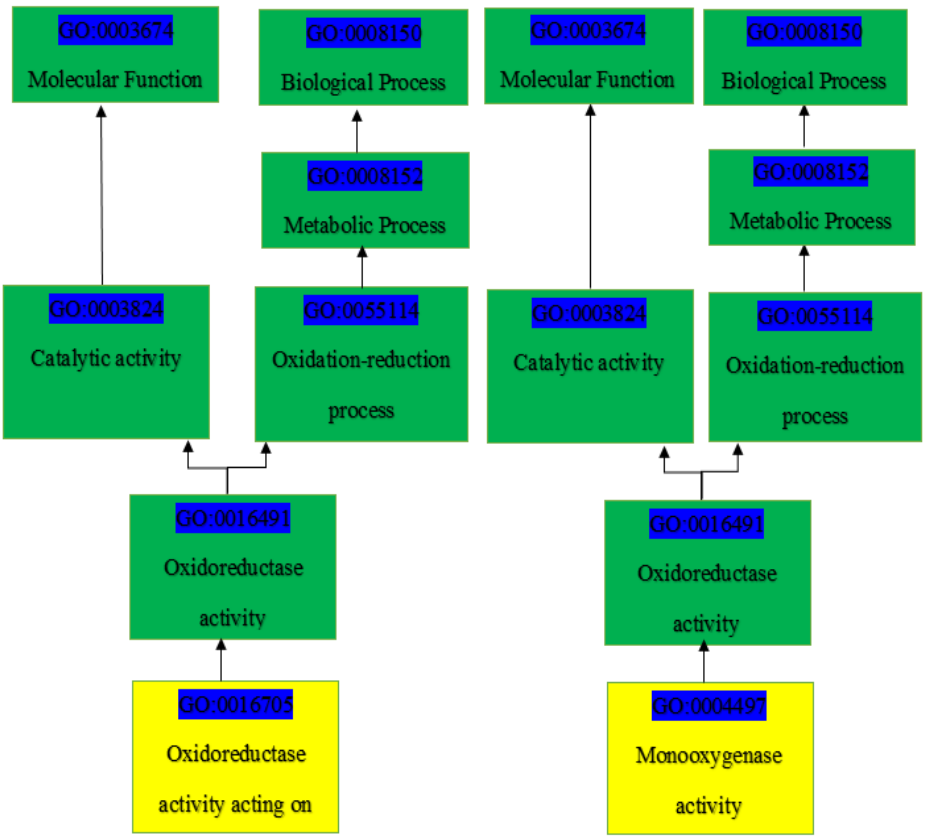


Fig. 21. Shows the alignments of the two positions of nucleotide binding from MultAlin underline with blue colour



(a) (b)

Fig. 22. Shows Gene ontology (GO) of the two proteins

4. CONCLUSION

In this study, a novel luciferase from *V. longisporum* was characterized and modeled. The modelled 3D structure was of high quality with a PROCHECK score of 99.5%, ERRAT2 value of 91.01%, and Verify3D score of 91.01%, which suggest that the protein conformational structure is acceptable. The luciferase enzyme was shown to have similar characteristics with existing luciferase representatives from various fungal and bacterial species. The distance between the ligand and the binding site of template protein and that of the luciferase model were also similar. Only minor differences that could be linked to evolutionary trends or enzyme specificity were detected. The study provided an avenue for the search of luciferase from fungal sources as against the well-established bacterial and insect based luciferase.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Kargar F, Mortazavi M, Savardashtaki A, Hosseinkhani S. Genomic and protein structure analysis of the luciferase from the Iranian bioluminescent beetle, Luciola sp. International Journal of Biological Macromolecules. 2019;124:689–698. Available:https://doi.org/10.1016/j.ijbiomac.2018.11.264.
2. Baldwin TO. Firefly luciferase : the structure is known , but the mystery remains. 1996;223–228.
3. Inouye S. Firefly luciferase: An adenylate-forming enzyme for multicatalytic functions. Cellular and Molecular Life Sciences. 2010;67(3):387–404. Available:https://doi.org/10.1007/s00018-009-0170-8.
4. Vongsangnak W, Chumnanpuen P, Sriboonlert A. Transcriptome analysis reveals candidate genes involved in luciferin metabolism in Luciola aquatilis (Coleoptera: Lampyridae). PeerJ. 2016;4:2534.

Available:https://doi.org/10.7717/peerj.2534.
5. Nakatsu T, Ichiyama S, Hiratake J, Saldanha A, Kobashi N, Sakata K, Kato H. Structural basis for the spectral difference in luciferase bioluminescence. Nature. 2006;440(7082):372–376. Available:https://doi.org/10.1038/nature04542.
6. 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):6493–6495. Available:https://doi.org/10.1021/bi300934s.
7. Smirnova DV, Ugarova NN. Invited Review Fire fly Luciferase-based Fusion Proteins and their Applications. 2017;436–447. Available:https://doi.org/10.1111/php.12656.
8. Dragulescu-Andrasi A, Chan CT, De A, Massoud TF, Gambhir SS. Bioluminescence resonance energy transfer (BRET) imaging of protein-protein interactions within deep tissues of living subjects. Proceedings of the National Academy of Sciences, 2011;108(29):12060–12065. Available:https://doi.org/10.1073/pnas.1100923108.
9. Jeske L, Placzek S, Schomburg I, Chang A, Schomburg D. BRENDA in 2019: A European ELIXIR core data resource. Nucleic Acids Research. 2019;47(D1):D542–D549. Available:https://doi.org/10.1093/nar/gky1048.
10. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A. Protein Identification and Analysis Tools on the ExPASy Server. The Proteomics Protocols Handbook. 2009;571–607. Available:https://doi.org/10.1385/1-59259-890-0:571.
11. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Higgins DG. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Molecular Systems Biology. 2011;7(539). Available:https://doi.org/10.1038/msb.2011.75.
12. Buchan DWA, Jones DT. The PSIPRED Protein Analysis Workbench: 20 years on. Nucleic Acids Research. 2019;47(W1):W402–W407. Available:https://doi.org/10.1093/nar/gkz297.
13. Laskowski RA, Jabłońska J, Pravda L, Vařeková RS, Thornton JM. PDBsum: Structural summaries of PDB entries. Protein Science, 2018;27(1):129–134. Available:https://doi.org/10.1002/pro.3289.
14. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Schwede T. SWISS-MODEL: Homology modelling of protein structures and complexes. Nucleic Acids Research. 2018;46(W1):W296–W303. Available:https://doi.org/10.1093/nar/gky427
15. Moreno-Ulloa A, Mendez-Luna D, Beltran-Partida E, Castillo C, Guevara G, Ramirez-Sanchez I, Villarreal F. The effects of (-)-epicatechin on endothelial cells involve the G protein-coupled estrogen receptor (GPER). Pharmacological Research. 2015;100:309–320. Available:https://doi.org/10.1016/j.phrs.2015.08.014
16. Paxman JJ, Heras B. Bioinformatics tools and resources for analyzing protein structures. Methods in Molecular Biology. 2017;1549:209–220. Available:https://doi.org/10.1007/978-1-4939-6740-7_16.
17. Lee J. Perspectives on Bioluminescence Mechanisms. Photochemistry and Photobiology. 2017;93(2):389–404. Available:https://doi.org/10.1111/php.12650.
18. Eljounaidi K, Lee SK, Bae H. Bacterial endophytes as potential biocontrol agents of vascular wilt diseases – Review and future prospects. Biological Control. 2016;103:62–68. Available:https://doi.org/10.1016/j.biocontrol.2016.07.013.
19. Poux S, Arighi CN, Magrane M, Bateman A, Wei CH, Lu Z. UniProt Consortium on expert curation and scalability: UniProtKB/Swiss-Prot as a case study. Bioinformatics (Oxford, England). 2017;33(21):3454–3460. Available:https://doi.org/10.1093/bioinformatics/btx439.
20. Burley SK, Berman HM, Bhikadiya C, Bi C, Chen L, Di Costanzo L, Zardecki C. RCSB Protein Data Bank: Biological macromolecular structures enabling research and education in fundamental

biology, biomedicine, biotechnology and energy. Nucleic Acids Research. 2019;47(D1):D464–D474. https://doi.org/10.1093/nar/gky1004.
21. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Bourne PE. The Protein Data Bank Helen. Nucleic Acids Research. 2000;28(1):235–242. Available:https://doi.org/10.1093/nar/28.1.235.
22. Boratyn GM, Camacho C, Cooper PS, Coulouris G, Fong A, Ma N, Zaretskaya I. BLAST: a more efficient report with usability improvements. Nucleic Acids Research. 2013;41(Web Server issue), 29–33. Available:https://doi.org/10.1093/nar/gkt282.
23. Kaskova ZM, Dörr FA, Petushkov VN, Purtov KV, Tsarkova AS, Rodionova NS, Yampolsky IV. Mechanism and color modulation of fungal bioluminescence. 2017;1–9.
24. Ghaemi B M, Bassami MR, Hashemi Tabar GR, Saberi MR, Haghparast AR, Dehghani H. A bioinformatic approach to check the spatial epitope structure of an immunogenic protein coded by DNA vaccine plasmids. Journal of Theoretical Biology. 2015;380:315–320. Available:https://doi.org/10.1016/j.jtbi.2015.04.030.
25. Dennis E. Desjardin AGO, Charles VS. Fungi bioluminescence revisited. Received 3rd September 2007, Accepted 3rd January 2008 First Published as an Advance Article on the Web 24th January 2008;7(2):170–182. Available:https://doi.org/DOI: 10.1039/b713328f
26. Söding J, Biegert A, Lupas AN. The HHpred interactive server for protein homology detection and structure prediction. Nucleic Acids Research. 2005;33(2):244–248. Available:https://doi.org/10.1093/nar/gki408.
27. 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 Reports, 2011;39(5):578–585. Available:https://doi.org/10.5483/bmbrep.2006.39.5.578.
28. Panja AS, Bandopadhyay B, Maiti S. Protein thermostability is owing to their preferences to non-polar smaller volume amino acids, variations in residual physico-chemical properties and more salt-bridges. PLoS ONE. 2015;10(7):1–21. Available:https://doi.org/10.1371/journal.pone.0131495.
29. Depotter JRL, Rodriguez-Moreno L, Thomma BPHJ, Wood TA. The emerging British Verticillium longisporum population consists of aggressive Brassica pathogens. Phytopathology. 2017;107(11):1399–1405. Available:https://doi.org/10.1094/PHYTO-05-17-0184-R.
30. Sheraz MA, Kazi SH, Ahmed S, Anwar Z, Ahmad I. Photo, thermal and chemical degradation of riboflavin. Beilstein Journal of Organic Chemistry. 2014;10:1999–2012. Available:https://doi.org/10.3762/bjoc.10.208.
31. Zhou XX, Wang YB, Pan YJ, Li, WF. Differences in amino acids composition and coupling patterns between mesophilic and thermophilic proteins. Amino Acids. 2008;34(1):25–33. Available:https://doi.org/10.1007/s00726-007-0589-x
32. Hui Xu, Yindrila C, Benjamin P, Angad PM, Dhananjay B, Hans-Peter H. Identification of the First Riboflavin Catabolic Gene ClusterIsolated from Microbacterium maritypicum G10. The Journal of Biological Chemistry. 2016;291(45):23506–23515.

Peer-review history:
The peer review history for this paper can be accessed here:
http://www.sdiarticle4.com/review-history/60696