HIBISCUS

BULLETIN OF ENVIRONMENTAL SCIENCE AND MANAGEMENT

Website: http://journal.hibiscuspublisher.com

In silico study to breaking mystery of bioluminescence protein structure of bacterial luciferase and firefly luciferase

Abdul Rachman, A.R[1]., Shaharudin, N.A[1]., Sabullah, M.K[2]., Shukor, M.Y.[1]

[1]Department of Biochemistry, Faculty of Biotechnology and Biomolecular Sciences, University Putra Malaysia, UPM 43400 Serdang, Selangor, Malaysia.
[2] School of Food Science and Nutrition, University Malaysia Sabah, Jalan UMS, 88400, Kota Kinabalu, Sabah, Malaysia.
Corresponding author; Shukor, M.Y., E-mail: yunus@biotech.upm.edu.my

History

Received: 23 April 2014
Received in revised form: 15 Mei 2014
Accepted: 4 July 2014
Available online: 25 July 2014

Keyword

solid waste generation;
solid waste management;
organic solid waste;
composting;
organic additive

Abbreviations

PDB;	Protein Data Bank
α/β;	alpha/beta
pI;	isoelectric point
Å;	Angstrom
kDa;	kilodalton
ATP;	Adenosine triphosphate

Abstract

Bioluminescent proteins such as luciferase or green fluorescent protein become an interesting protein to be studied particularly on its structure, mechanism of light emission in certain wavelength and also its contribution in several applications. This study was conducted to understand more on the luciferase enzymes from bacteria and firefly based on *in silico* studies of their protein sequences, protein surface and interactions with different substrates. Protein crystal structures of bacterial luciferase and firefly luciferase that been deposited on Protein Data Bank (PDB) were used as model proteins for this study. Computer analysis of luciferase proteins such as pI prediction, multiple sequence alignment and molecular docking with different luciferase substrates (Flavin mononucleotide, L-luciferin, D-luciferin, vargulin, coelectrazine, Dinoflagellate luciferin, Latia luciferin) by using iGEMDOCK docking software had been conducted. Final analysis showed that, the bacterial luciferase prefers flavin mononucleotide and latia luciferin as its substrates and for firefly luciferase, it prefers dinoflagellate luciferin, L-luciferin and D-luciferin as the substrates. Bacterial luciferase and firefly luciferase shared very low protein sequence homology (<30%) although they have similar catalyzing function. Prediction of pI by using ProtParam tool showed that bacterial luciferase and firefly luciferase has pI values of 4.95 and 6.42, respectively. In conclusion, bacterial and firefly luciferase can consider several different of luciferin in enzyme catalysis.

INTRODUCTION

Luminous organisms such as marine bacteria, insects, marine coelenterates and crustacean had been widely classified [1]. Bioluminescent properties of these organisms were used to distract predator, attract prey and mating partners. Luminescence occurred when the substrate (S) been converted into the excited electronic state product (P*) that can emit photon of visible light requiring oxygen in its enzymatic mechanism [2-3]. P* will be converted into product (P) by oxidation process [1]. This reaction was catalyzed by luciferases. General mechanism of reaction was shown below

$$S + O_2 \longrightarrow P^* \longrightarrow P + h\nu$$

Luciferase from bacteria and firefly were extensively studied to explain the mechanism on light emission at certain wavelength. The first bacterial luciferase protein structure had been solved with a resolution of 2.4 Å from *Vibrio harveyi* [4] and the first firefly luciferase protein structure had been solved by Conti and co-worker [5]. Bacterial luciferase consists of 2 subunits $(\alpha/\beta)_8$ barrel and each subunit can work independently though at 6 times lower magnitude of light compared with combined subunits [6]. But the heterodimer fails to fold properly if the 2 subunits are combined outside of their original system combination [7-8]. They can only be folded properly after refolding process using high concentration of urea [9].

Luciferase protein structure from firefly has been extensively studied from firefly *Photinus pyralis* species. This protein is found in the insect's abdomen called lantern [5]. This enzyme is 62 kDa in molecular weight [10] and requires oxygen molecule, ATP as well as luciferin to generate light in a two-step process [11]. Firefly luciferase was extensively used in molecular biology as a reporter gene to detect the presence of desired gene [12]. Luciferase was also used as a model to investigate the possible protein-anaesthetic interactions for sensitivity of certain protein to the general anaesthetic [13]. In this research, we reported on the characteristics of both proteins specifically on the molecular docking of the proteins with different substrate in order to elucidate the most compatible substrates the proteins prefer the most. The interaction between enzymes and substrates were determined based on its total binding energy and binding with critical amino acids in the protein structure. This study will lead researcher for further studies on how both luciferase in bacteria and firefly can bind with different substrates.

MATERIAL AND METHODS

Protein and substrate structure retriever

Both protein structures (PDB:1BRL and 1LCI) were retrieved from Protein Data Bank (http://www.rcsb.org/pdb/home/home.do) and substrate structures (L-luciferin, D-luciferin, Dinoflagellate luciferin, Coelectrazine and Latia luciferin) were retrieved from PubChem (http://pubchem.ncbi.nlm.nih.gov/search/search.cgi).

Protein sequence analysis

Multiple sequence alignment of bacterial and firefly luciferase protein sequence was done by using biology workbench software (http://workbench.sdsc.edu/) to identify the similarity between both protein sequences.

pI protein determination

pI protein determination of bacterial and firefly luciferase were determined by using ProtParam tools (http://web.expasy.org/protparam/).

Molecular docking

Interaction of both luciferase proteins and their substrates were done by molecular docking using iGEMDOCK molecular software. Five different substrates (L-luciferin, D-luciferin, Dinoflagellate luciferin, Vargulin and Coelectrazine) were used for substrate docking process.

Protein and amino acid viewer

Important amino acids and potential electrostatic in bacterial and firefly proteins were viewed by using Pymol viewing software.

RESULTS AND DISCUSSION

Based on the protein multiple sequence alignment (Figure 1), both luciferases shared very low homology (<30% homology). No conserved amino acid can be found in both enzymes because of the mechanism of catalyzing substrate is different although they exhibit similar function. This is probably because firefly luciferase has two binding sites but only one binding site for bacterial luciferase [14]. pI of both proteins were calculated by using ProtParam tool. 1BRL and 1LCI have pI values of 4.95 and 6.42, respectively (data not shown). pI value was define as the overall net charge of the protein surface and is determined based on the electrostatic potential of the protein surface. Figure 2 shows the electrostatic potential of both enzyme and the location of important amino acid for both enzymes. Based on this, important amino acids of 1BRL were surrounded by negatively charged amino acids (red colour) while for 1LCI, the important amino acids were surrounded by negatively and positively charged amino acids (red and blue colour). This means 1LCI is more stable at neutral pH compared to 1BRL that require a slight acidic conditions to stabilize its structure although previous reported have shown that firefly luciferase exhibited optimum pH at 7.8 [14] whereas bacterial luciferase showed optimum pH at 6.8 [15].

Molecular docking for both enzymes was carried out using iGEMDOCK software with different substrates (Flavin mononucleotide, D-luciferin, L-luciferin, Vargulin, Coelectrazine, Dinoflagellate luciferin, Latia luciferin). Docking is a result of interaction between amino acids of the enzymes with the substrates based on its total energy. Total energy was calculated from the important energy binding in hydrogen bond, Van Der Waals and electrostatic interaction. The preferred substrates for 1BRL based on total binding energy are latia luciferin > vargulin > flavin mononucleotide > coelectrazine > D-luciferin > L-luciferin > dinoflagellate luciferin. For 1LCI, the preferred substrates are latia luciferin > vargulin > coelectrazine > flavin mononucleotide > L-luciferin > D-luciferin > dinoflagellate luciferin (Table 1). Both enzymes preferred to bind with the substrate latia luciferin as shown by the highest total energy obtained compared to other substrates (Table 1). However, the binding energy of the substrates with important amino acids that play critical role in catalysis in both enzymes must also be considered in order to determine the suitable substrates. Several amino acids that play important role in catalysis are shown in table 2 and the positions of that amino acid are shown in figure 3.

Interaction of protein with its substrate depends on weak non covalent bonds such as hydrogen bond, Van der Waals, electrostatic interactions and hydrophobic interactions [16]. In the structure of 1BRL, the amino acid histidine at position 45 and glutamic acid at position 175 are critical amino acids in catalysis based on previous findings of Li and Hosseinkhani [17-18]. However, *in silico* analysis indicated that the 1BRL protein preferred latia luciferin as the first substrate based on total binding energy even though latia luciferin only showed hydrogen bond and Van der Waals interaction with glutamic acid at position 175. No sign of binding with histidine at position 45 can be observed (Table 3 (a)). So, latia luciferin cannot be considered as the most suitable substrate for 1BRL protein. In contrast, flavin mononucleotide showed binding energy with both critical amino acids in bacterial luciferase (1BRL)-glutamic acid 175 (Van der Waals interaction) and histidine 45 (Hydrogen bond and Van der Waals interaction). But, latia luciferin can become potential substrate for 1BRL because it exhibits binding energy with glutamic acid 175. Hosseinkhani and coworker (2005) concluded the important role of glutamic acid at position 175 for substrate binding affinity. The amino acid determines either 'slow' or 'fast' luminescence decay rate in luciferases catalysis. They reported on the role of glutamic acid at position 175 for aldehyde binding for flavin mononucleotide substrate. Change to other amino acids at position 175 will make aldehyde binding weaken and decrease its affinity towards flavin substrates [17]. Histidine at position 45 also plays important role in substrate binding. Mutation at histidine 45 at the alpha subunit can lead to a 300 fold reduction of luminescence activity of luciferase [18]. But then, there is still no further information on other important amino acids for substrate binding site rather than histidine 45 and glutamic acid 175 in 1BRL protein.

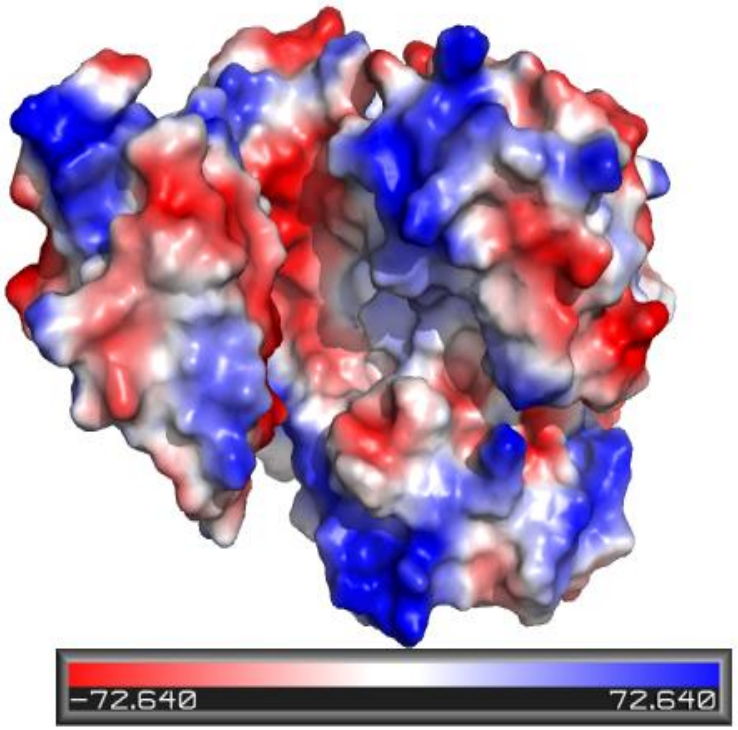
Important amino acid for firefly luciferase [1LCI] structure for substrate catalysis is shown in table 3(b). But then, only L-luciferin, D-luciferin and dinoflagellate luciferin exhibited binding with important amino acids between the position 218 and 348. There are several reports on the used of L-luciferin and D-luciferin as primary substrates of firefly luciferase from *Photinus pyralis* (1LCI) but to the best of our knowledge, there is no report on the binding of dinoflagellate luciferin with firefly luciferase. Branchini [19] had studied the influence of this region to the intensity of bioluminescence emission by mutating the amino acids one by one.

They conclude that, amino acid on this region also play same catalytic role on all luciferin substrate. Lysine at position 529 was claimed to be the important amino acid for substrate binding. However, from the docking result, no substrates showed binding energy with lysine 529. This means that no interaction *in silico* occurred between the substrates with lysine 529. Furthermore, the role of lysine 529 is still putative although previous research have suggested this amino acid play a role in bioluminescence intensity [20].

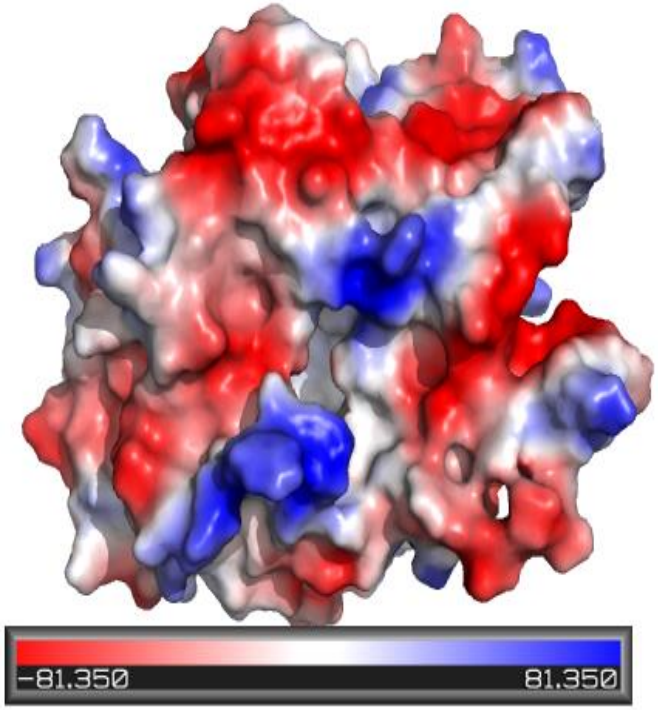
Figure 1. Protein multiple sequence alignment of firefly luciferase protein and bacterial luciferase. Upper sequence: Firefly luciferase, Below sequence: Bacterial luciferase. Red colour indicated conserved amino acids.

```
Firefly luciferase      MEDAKNIKKGPAPFYPLEDGTAGEQLHKAMKRYALVPGTIAFTDAHIEVN
Bacterial luciferase     ------MKFG----------------------------NFLLTYQPPE
                        :* *                        * ::  :

Firefly luciferase      ITYAEYFEMSVRLAEAMKRYGLNTNHRIVVCSENSLQFFMPVLGALFIGV
Bacterial luciferase     LSQTEVMKRLVNLGKASEGCGFDT----VWLLEHHFTEFG-LLGNPYVAA
                      :: :* ::  *.*.:* :  *::*    *   *: :  * :**  ::..

Firefly luciferase      AVAPANDIYNERELLNSMNISQPTVVFVSKKGLQKILNVQKKLPIIQKII
Bacterial luciferase     AHLLG--------ATETLNVGTAAIVLPTAHPVRQAEDVN----------
                      *   .        :::*:. .::*: : : :::  .:*:

Firefly luciferase      IMDSKTDYQGFQSMYTFVTSHLPPGFNEYDFVPESFDRDKTIALIMNSSG
Bacterial luciferase     LLDQMSKGR--------FRFGICRGLYDKDFRVFGTDMDNSRALMD----
                      ::*. :. :       .  :  *: : **   . * *:: **:

Firefly luciferase      STGLPKGVALPHRTACVRFSHARDPIFGNQIIPDTAILSVVPFHHGFGMF
Bacterial luciferase     ----------------CWYDLMKEGFNEGYIAADNEHIKFPKIQLNPSAY
                               :.  :: :  . * .*.  :..  :: . . :

Firefly luciferase      TTLGYLICGFRVVLMYRFEEELFLRSLQDYKIQSALLVPTLFSFFAKSTL
Bacterial luciferase     TQGGAPVY-------------VVAESASTTEWAAERGLPMILSWIIN-TH
                      *  *  :           :. .* .  : :  :* ::*:: : *

Firefly luciferase      IDKYDLSNLHEIASGGAPLSKEVGEAVAKRFHLPGIRQGYGLTETTSAIL
Bacterial luciferase     EKKAQLDLYNEVAT----------------------EHGYDVTKIDHCLS
                       .* :*.  :*:*:                  .:**.:*:   .:

Firefly luciferase      ITPEGDDKPGAVGKVVPFFEAKVVDLDTGKTLGVNQRGELCVRGPMIMSG
Bacterial luciferase     YITS-------------------VDHDS------NRAKDICRN--FLGHW
                      ..              ** *:      *:  ::* .  ::
Firefly luciferase      YVNNPEATNALIDKDGWLHSGDIAYWDEDEHFFIVDRLKSLIKYKGYQVA
Bacterial luciferase     YDSYVNATKIFDDSD--------------------------QTKGYDFN
                      * .  :**: : *.*                    : ***:.

Firefly luciferase      PAELESILLQHPNIFDAGVAG-LPDDDAGELPAAVVVLEHGKTMTEKEIV
Bacterial luciferase     KGQWRDFVLKGHKDTNRRIDYSYEINPVGTPEECIAIIQQDIDATG---I
                       .: ..::*:  : : :     : .*    .:.:::::.   *    :

Firefly luciferase      DYVASQVTTAKKLRGGVVFVDEVPKGLTGKLDARKIREILIKAKKGGKSK
Bacterial luciferase     DNICCGFEANGSEEEIIASMKLFQSDVMPYLKEKQ----------------
                      * :.. . :  . .  :. :. . ..:   *. ::

Firefly luciferase      L
Bacterial luciferase     -
```

Figure 2. Electrostatic potential of both luciferases on protein structure. Circle shown position of critical amino acid. (A): 1LCI protein structure, (B):1BRL protein structure



(A)



(B)

Table 1. Molecular docking result for both luciferase enzymes by using iGEMDOCK molecular docking software. (A): result docking of bacteria luciferase (1BRL), (B): result docking of firefly luciferase (1LCI).

Ligand	Total Energy	VDW	H Bond	Electrostatic
D-Luciferin	-90.6944	-67.5932	-21.2705	0
Flavin mononucleotide	-112.283	-79.0471	-31.2953	0
L-luciferin	-90.2973	-67.0706	-21.2166	0
Vargulin	-118.0	-82.1361	-32.5796	-3.28048
Coelectrazine	-110.8	-98.5648	-12.252	0
Dinoflagellate luciferin	-	-	-	-
Latia luciferin	-122.4	-114.391	-8.55838	0.567942

(A)

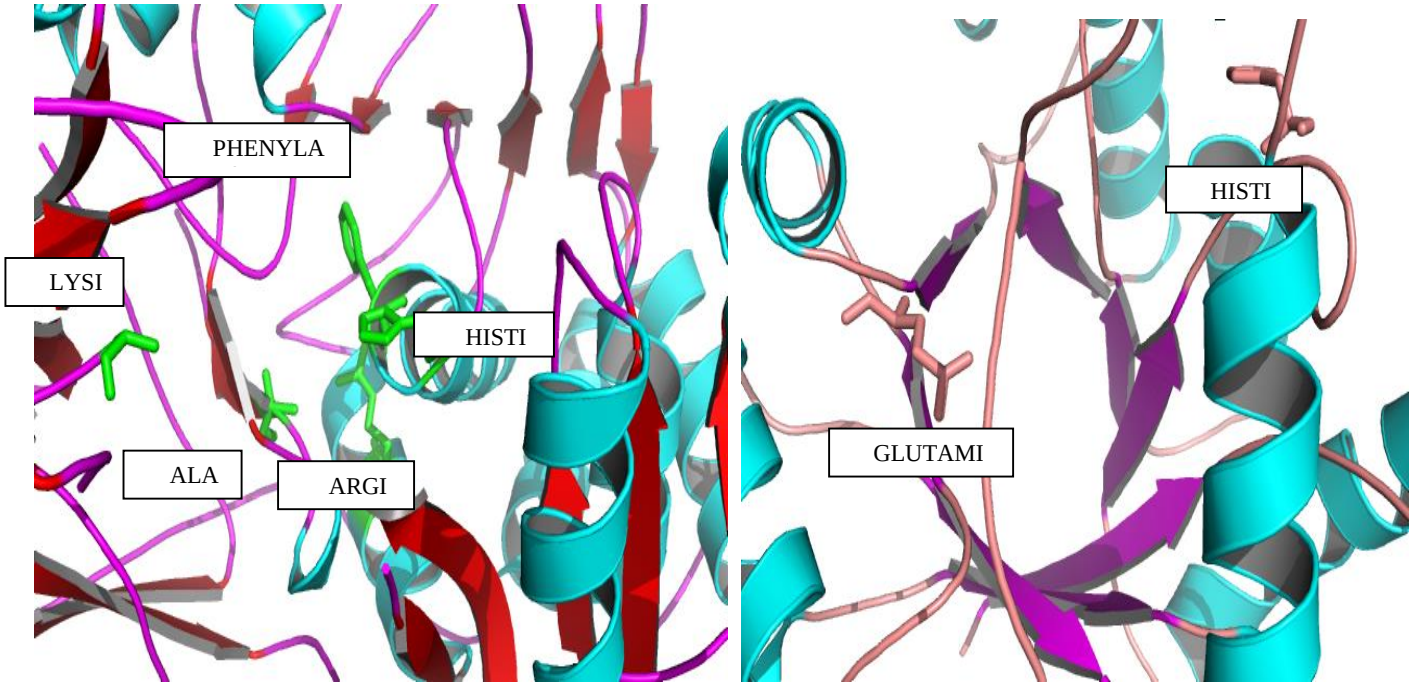
Ligand	Total Energy	VDW	H Bond	Electrostatic
D-Luciferin	-79.5618	-65.9847	-13.5772	0
Flavin mononucleotide	-107.007	-67.8653	-34.9556	-4.18656
L-luciferin	-90.6656	-69.009	-19.5516	2.10494
Vargulin	-114.998	-94.9923	-16.3514	-3.65437
Coelectrazine	-112.406	-96.8512	-15.5552	0
Dinoflagellate luciferin	-78.3983	-71.2386	-7.15967	0
Latia luciferin	-118.302	-88.2368	-37.4356	7.37079

(B)

Table 2. Summary of critical amino acids that responsible in enzyme- substrate binding for both luciferase enzymes.

Protein structure	**Amino acid**
1BRL	His45,Glu175
1LCI	Arg218,His245,Phe247,Ala348,Lys529

Figure 3. Location of critical amino acid in both luciferases on protein structures based on previous investigation. (A):1LCI protein structure, (B):1BRL protein structure.



(A) (B)

Table 3. Summary of amino acid interaction with substrates for both luciferase enzyme. (A): Bacteria luciferase (1BRL), (B): Firefly luciferase (1LCI). Bold word indicated critical amino acids on substrate binding.

Substrates	Hydrogen bond	Van Der Waals	Electrostatic
Flavin mononucleotide	Glu43,**His45**,Gly53, Asn54,Val77	**His**45,Gly53,Asn54, Tyr56,Thr80,Phe117, **Glu175**	-
D-luciferin	Asp223,Thr318, Asp321,Lys354	Val219,Thr220,His224, Thr318,Gly319,Ile320	-
L-luciferin	Asp223,Thr318, Asp321,Lys354	Val219,Thr220,His224, Thr318,Gly319,Ile320	-
Vargulin	Asp120,Asp122, Glu149,His150, Thr267	Gly118,Thr119,Asp120, Asp122,Asn123,Thr267, Lys268,Gly269	Glu149
Coelectrazine	Thr220,Asp223, Asp321	Thr220,Lys221,Ile222, Asp223,His224,Gly319, Ile320,Asp321	-
Dinoflagellate luciferin	-	-	-
Latia luciferin	**Glu175**	Arg107,Leu109,Val173, Ala174,**Glu175**,Ser176, Arg289	-

(A)

Substrates	Hydrogen bond	Van Der Waals	Electrostatic
Flavin mononucleotide	Arg77,Ile120,Ser121, Asn403	Glu122,Ile192,Pro233, Val402,Asn403	Arg77
D-luciferin	Gly316,Gly341, **Ala348**	**His245**,Gly246,Gly316,Ala317,Gly341,Leu342	-
L-luciferin	**His245**,Gly316, Gly341,**Thr343**	**His245**,Gly246,Gly316,Ala317,Gly341,Leu342	-
Vargulin	Asp234,Thr235	Tyr33,Gly38,Thr39,Ile40,Pro233, Asp234,Gly259, Phe260,Arg261	Asp234
Coelectrazine	Gly38,Asp234, Lys281,Gln283	Gly38,Thr39,Ile40, Pro233,Asp234,Gly246, Gly259,Phe260, Arg261	-
Dinoflagellate luciferin	**Thr343**	**His245**,Gly246,Gly316,Gly341,Leu342,**Thr343**	-
Latia luciferin	Thr21,Ser185,Phe186, Asp187,Arg188, Asp189	Glu184,Ser185,Phe186, Asp187,Arg188	-

(B)

CONCLUSION

From this study, we conclude that 1BRL and 1LCI enzymes can consider several luciferin substrates such as latia luciferin and dinoflagellate luciferin instead of flavin mononucleotide, L-luciferin and D-luciferin in catalysis. Bacterial luciferin (1BRL) also showed energy binding with the substrate latia luciferin- a specific substrate for latia luciferase. The pursuit of the *in vitro* study for the activity of both luciferase enzymes with different substrates should complement the *in silico* results in this work. This study will lead researcher to pursue the potential of bacterial and firefly luciferase binding towards different substrates and their specificity to different substrates.

ACKNOWLEDGEMENT

This project was supported by funds from The Research University Grant Scheme (RUGS), Project No: 05-01-09-0750RU/F1.

REFERENCES

[1] Hastings J.W. Biological diversity, chemical mechanism and evolutionary origins of bioluminescent system. *Journal Molecular Environment*, vol.19, pp.309-321, 1983.

[2] Harvey E.N. *Bioluminescence*. New York: Academic Press, pp.649, 1952

[3] Shimomura O. Bioluminenescence: Chemical principles and methods. *Singapore World Scientific*, pp.470, 2006

[4] Fisher A.J., Raushel F.M., Baldwin T.O., Rayment I. Three dimensional structure of bacterial luciferase from *Vibrio harveyi* at 2.4 Å resolution. *Biochemistry*, vol.34, pp.6581-6586, 1995.

[5] Conti E.,Franks N.P.,Brick P. Crystal structure of firefly luciferase throws light on a superfamily of adenylate-forming enzymes. *Structure*, vol.4 no 3, 1996.

[6] Waddle J., and Baldwin T.O. *Biochemistry Biophysical Resource Communication*, vol.178, pp.1188-1193, 1991.

[7] Sinclair J.F., Waddle J.J., Waddill E.F.,Baldwin T.O. Purified native subunits of bacterial luciferase are active in bioluminescence reaction but fail to assemble into the alpha beta structure. *Biochemstry*, vol.32, pp.5036-5044,1993.

[8] Waddle J.J., Johnston T.C., Baldwin T.O. Polypeptide folding and dimerization in bacterial luciferase occur by a concerted mechanism *in vivo*. *Biochemistry*, vol.26, pp.4917-4921,1987.

[9] Baldwin T.O., Ziegler M.M., Chaffotte A.F., Goldberg M.E. Contribution of folding steps involving the individual subunits of bacterial luciferase to the assembly of the active heterodimeric enzyme. *Journal Biology Chemistry*, vol. 268, pp. 10766-10772. 1993.

[10] De Wet J.R., Wood K.V., Helinski D.R and DeLuca M. Cloning of firefly luciferase cDNA and the expression of active luciferase in *E.coli* .*Protein Natural Academic Science* (USA), vol. 82, pp.7870-7873.1985.

[11] DeLuca M. Firefly luciferase. *Advance Enzymology*, vol.44, pp. 37-68.1976.

[12] Gould S.J., Subramani S. Firefly luciferase as a tool in molecular and cell biology. *Analytical Biochemistry*, vol. 175, pp. 5-13.1988.

[13] Frank N.P., Lieb W.R. Do general anaethetics act by competitive binding to the specific receptors?. *Nature*, vol. 310,pp. 599-601.1984.

[14] Steghens J.P., Min K.L., Bernengo J.C. Firefly luciferase has two nucleotide binding sites: effect of nucleoside monophosphate and CoA on the light-emission spectra. *Journal Biochemistry*, vol. 336 no 1, pp. 109–113.1998.

[15] Hastings, J., Weber, K., Friedland, J., Eberhard, A., Mitchell, G., and Gunsalus, A. Structurally distinct bacterial luciferases. *Biochemistry*, vol. 8 no. 12, pp. 4681-4689.1969.

[16] Albert B., Johnson A., Lewis J., The components of the cell: *Molecular Biology of the Cell*, 4 Edition. Garland Science, New York, 2002.

[17] Hosseinkhani S., Szittner R., Meighen E.A., Random mutagenesis of bacterial luciferase: critical role of Glu175 in the control of luminescence decay. *Journal Biochemistry*, vol. 385, pp. 575-580. 2005.

[18] Li.H., Ortego B.C., Maillard K.I., Wilson R.C., Tu S.C. Effects of mutations of the alpha His 45 residue of *Vibrio harveyi* luciferase on the yield and reactivity of flavin peroxide intermediate. *Biochemistry*, vol. 38, pp. 4409-4415.1999.

[19] Branchini B.R., Southworth T. L., Murtiashaw M. H., Boije H., Fleet S.E. A Mutagenesis Study of the putative luciferin binding site residues of firefly luciferase. *Biochemistry*, vol. 42, pp. 10429-10436. 2003.

[20] Branchini B.R., Magyar R.A., Muutiashaw M.H., Anderson S.M., Helgerson L.C., Zimmer M. Site-directed mutagenesis of firefly luciferase active site amino acid: A proposed model for bioluminescence color. *Biochemistry*, vol. 38, pp. 13223-13280.1999.