CORE
Metadata, citation and similar papers at core.ac.uk
Provided by Siberian Federal University Digital Repository

Functional divergence between evolutionary related LuxG and Fre oxidoreductases of luminous bacteria

Anna A. Deeva[1,*], Evgenia A. Zykova[2,], Elena V. Nemtseva[1,4] and Valentina A. Kratasyuk[1,4]

[1]Laboratory of Bioluminescent Biotechnologies, Siberian Federal University, Krasnoyarsk, Russian Federation,

[2]Mechanism of Cell Genome Functioning Laboratory, Institute of Cell Biophysics RAS, Moscow Region, Pushchino, Russian Federation,

[3]Department of applied research informatization, State Institute of Information Technologies and Telecommunications (SIITandT "Informika"), Moscow, Russian Federation,

[4]Laboratory of Photobiology, Institute of Biophysics SB RAS, Federal Research Center "Krasnoyarsk Science Center SB RAS", Krasnoyarsk, Russian Federation

Short title: A functional divergence between LuxG and Fre oxidoreductases

Five to ten key words: NAD(P)H:flavin-oxidoreductase, Fre, LuxG, gene duplication, functional divergence, bacterial bioluminescence

*Contact: Address: Laboratory of Bioluminescent Biotechnologies, Siberian Federal University, Svobodny prosp. 79, Krasnoyarsk, 660041, Russian Federation, E-mail: adeeva@sfu-kras.ru

Abstract

In luminous bacteria NAD(P)H:flavin-oxidoreductases LuxG and Fre there are homologous enzymes that could provide a luciferase with reduced flavin. While Fre functions as a housekeeping enzyme, LuxG appears to be a source of reduced flavin for bioluminescence as it is transcribed together with luciferase. This study is aimed at providing the basic conception of Fre and LuxG evolution and revealing the peculiarities of the active site structure resulted from a functional variation within the oxidoreductase family. A phylogenetic analysis has demonstrated that Fre and LuxG oxidoreductases have evolved separately after the gene duplication event, and consequently, they have acquired changes in the conservation of functionally related sites. Namely, different evolutionary rates have been observed at the site responsible for specificity to flavin substrate (Arg 46). Also Tyr 72 forming a part of a mobile loop involved into FAD binding has been found to be conserved among Fre in contrast to LuxG oxidoreductases. The conservation of different amino acid types in NAD(P)H binding site has been defined for Fre (arginine) and LuxG (proline) oxidoreductases.

Introduction

In bacteria a bioluminescent reaction is regulated by a set of proteins encoded in *lux*-operon.[1] Two genes, *luxA* and *luxB*, code for a bacterial luciferase that is the most studied enzyme playing a major role in the bioluminescent reaction. The rest of a multienzyme bioluminescent system involved in continuous light emission is poorly investigated and the role of each enzyme is still under consideration.[2,3]

The gene product of *luxG* that flanks *luxAB* is supposed to function as NAD(P)H:flavin oxidoreductase. Its function has been hypothesized on the basis of LuxG homology to the flavin oxidoreductase found in *Escherichia coli* (Fre).[4,5] Fre is the essential enzyme involved into different metabolic pathways in bacteria, including the reduction of ferric iron integrated into active sites of various enzymes.[6] Precisely it catalyzes the reduction of soluble flavins (FMN, FAD or riboflavin) associated with NADH or NADPH oxidation.[7] Hence, LuxG has been proposed as a source of reduced flavin, $FMNH_2$, for bacterial luciferase.[8] Moreover, LuxG has also turned out to reduce the ferric irons *in vitro*[9] and *vice versa* Fre has proved to provide $FMNH_2$ to the bioluminescent reaction catalyzed by a bacterial luciferase in recombinant *E. coli*.[10]

In many luminous species (i.e. *Aliivibrio fischeri*, *Photorhabdus luminescens* and others) not only LuxG, but also Fre-like oxidoreductases have been found.[5] Probably, they are not involved into the regulation of bioluminescence *in vivo* except for in *Photorhabdus* species which lack *luxG* gene and apparently compensate oxidoreductase activity by Fre.

LuxG crystal structure has not been identified yet, however the folding patterns of this subfamily could be analyzed using the known spatial structure of Fre oxidoreductase. The latter has turned out to be structurally similar to the members of the ferredoxin:$NADP^+$ reductase (FNR) family.[11] At the same time, amino acid sequence similarity between Fre and FNR is very low.[12] An exception is the identical functional residues involved into the binding of flavins and NAD(P), and these regions are uniformly ordered along the sequence.[11]

LuxG and Fre oxidoreductases were discovered when only a few luminous species were described. Since then a numerous sequences of *lux* and *fre* genes have become available, but all the experimental investigations of functional properties have been carried out using only either LuxG from *P. leiognathi* or Fre from *E. coli.* It has been revealed that LuxG functions as a homodimer and its unique feature is the half-sites reactivity which is impossible for Fre that exists exclusively as a monomer.[13] Additionally the difference in substrate specificity has been shown: for Fre FAD is a preferred substrate and an inhibitor while for LuxG it is not the case.[14] Also the possibility for Fre or LuxG to form a complex with bacterial luciferase has been studied and functioning of LuxG as a source of reduced flavins for the ferric iron reduction process has been revealed.[3,9,10]

Regardless the incomplete knowledge about the characteristics of proteins involved into bacterial bioluminescence, it is widely applied as an enzyme-based bioassay and whole-cell biosensor for ecotoxicology, foods quality control, etc.[15,16]Research of LuxG and Fre evolution and specific structural changes occurred to LuxG upon functioning as a part of *lux*-operon could shed light on the mechanisms of bacterial bioluminescence. This study is aimed at providing a basis conception of Fre and LuxG evolution and revealing the peculiarities of an active site structure resulted from a functional variation within the oxidoreductase family. The phylogenetic relationships between Fre and LuxG oxidoreductases have been analyzed and critical changes of functionally important sites have been identified.

Methods

Sequence data collection and multiple sequence alignment

Amino acid sequences of LuxG oxidoreductases from *Vibrio harveyi* (accession No ABX76849.1) and *Photobacterium leiognathi* (accession No KJF87632.1) have been used as an input to BLAST to identify homologous sequences available in NCBI.[17,18] We have used the protein-protein BLAST algorithm with an e-value of e-10 and the BLOSUM62 matrix. Each query result has contained 1,000 amino acid sequences. The dataset has been manually modified

in a following way: {1} two search results have been merged, that has led to a set of 1,072 unique sequences, {2} 13 partially sequenced proteins have been excluded from the dataset, {3} one sequence has been chosen as a representative from each group of sequences with a high identity (over 95%), so the resulting dataset has comprised 431 amino acid sequences (Table SI). In order to root the tree we have used an outgroup consisting of five amino acid sequences of enzymes belonging to FNR superfamily. The resulting sequences have been aligned using MAFFT software with default parameters implemented in the Jalview program package.[19,20]

Reconstruction of phylogenetic tree

The alignment has been used as the input to the ProtTest 3 program, which has identified LG+I+G as the best-fit amino acid replacement model for the evolution of LuxG and Fre oxidoreductases.[21] This model has been implemented for the reconstruction of a phylogenetic tree with PhyML software and support has been calculated with 1,000 bootstrap replicates.[22] Bayesian analysis has been conducted with MrBayes v3.2.6 program, computational runs have been performed for 3,000,000 generations sampled every 50 generations, and 25% of sampled trees have been used for building up a consensus tree.[23] The tree has been displayed using the interactive Tree of Life (iTOL) website.[24]

Analysis of conserved amino acids and functional divergence between Fre and LuxG

The amino acid sequence of *E. coli* Fre, 14 sequences of LuxG and 19 sequences of Fre oxidoreductases from confirmed luminous species with available sequenced genomes have been retrieved from the final data set. The sequences have been realigned and phylogenetic analysis has been performed according to the procedure described above. The analysis of sites conserved among collected 34 oxidoreductase sequences and within the Fre and LuxG subfamilies and different between them has been performed.

Additionally Diverge 3.0 has been used to identify the functional divergence that may have occurred after oxidoreductases split from a common ancestor.[25] We have determined the positions that refer to type I or type II functional divergence. Type I means a considerable rate

difference at a given site between Fre and LuxG oxidoreductases, it shows that functional constraints are different for two subgroups of the enzyme family. The coefficient of type I functional divergence, θ_I, has been estimated for two clusters of amino acid sequences. Then θ_I coefficient and its standard error have been used for statistical evaluation of each position of the alignment based on Z-score test. Sites that could be related to this type of functional divergence have posterior probability score under a cut-off value (0.75). Type II means a radical change in amino acid property of the alignment position, which is highly conserved within the subfamily. Similarly, coefficient θ_{II} has been estimated and used to predict the alignment positions responsible for type II functional divergence (with a posterior probability cut-off value $\geq$0.75).

Structure preparation and refinement

To explore the potential function of the identified sites a spatial structure of *E. coli* Fre oxidoreductase (PDB ID: 1QFJ) and a modelled structure of *V. harveyi* LuxG have been analyzed. The latter one has been obtained using the Swiss-MODEL server with the structure of *E. coli* Fre applied as a template.[26] The missing loop in *E. coli* Fre structure has been reconstructed with MODELLER.[27] After LuxG model preparation and the reconstruction of Fre mobile loop we have performed 20 ns molecular dynamics simulation for both structures at 298 K and a constant pressure of 1 atm, using a model of explicit water SPC/E and CHARMM36 force field in GROMACS (version 5.1.4).[28] Initially both structures have been put into a box with periodic boundary conditions and ions have been added to neutralize the system. Also before the production runs a minimization followed by NVT and NPT equilibration has been done in order to remove any bad contacts between the protein and the solvent, to heat the system and reach the pressure of 1 atm.

After the production runs the structural validation of Fre and a modelled LuxG oxidoreductase have been performed by creating a Ramachandran plot using the PROCHECK software.[29] The coordinates of FAD and 2'-phospho-5'-AMP have been derived from the structure of *Spinacia oleracea* oxidoreductase (PDB ID: 1FNC) using structural alignment with *E. coli* Fre. Then

LuxG and Fre (with substrates) structures have been superimposed with VMD and plugin MultiSeq for the analysis of ligand binding patterns.[30,31,32]

Results and discussion

Evolutionary relationships of LuxG and Fre

The BLAST searches with amino acid sequences of *V. harveyi* LuxG and *P. leiognathy* LuxG have identified 367 amino acid sequences of Fre from non-luminescent species, 31 sequences of Fre from luminous bacteria and 33 sequences of LuxG (Table SI). As a result, the tree reconstructed from 431 homologous sequences clearly displays three clades containing {1} LuxG proteins, {2} Fre oxidoreductases of luminous bacteria together with a few non-luminescent species and {3} Fre of non-luminescent bacteria and luminous *Photorhabdus* species (Figure 1 and more detailed Figure S1). Whereas the BLAST searches have been performed using two LuxG sequences as the query, a large representation of Fre oxidoreductases from luminescent and non-luminescent species has been observed in the dataset. A bias of Fre oxidoreductases should not affect the tree topology as e-values for the sequences have been belowthe selected threshold (Table SI). LuxG oxidoreductases have formed a single monophyletic clade {1}, while all Fre have fallen outside it. It has included some proteins of luminous bacteria that have been identified as general NAD(P)H:FMN-oxidoreductases according to the NCBI annotation, but the analysis of the associated genome, especially their neighboring genes has allowed assigning them to LuxG. Moreover, a few NAD(P)H:FMN-oxidoreductases from non-luminous species have occurred in the LuxG clade, precisely from *Photobacterium piscicola*, *Vibrio lentus* and *Vibrio* sp. BCB494. The latter one is a natural dark mutant having a deletion in the *luxC* gene.[33] Some isolates of *P. piscicola* containing a *luxA* gene have been described, therefore this sequence could be also designated as a LuxG oxidoreductase.[34] *V. lentus* is phenotypically close to luminous *V. splendidus,* probably the detailed study of *V. lentus* genome could reveal functionally redundant *lux* genes.[35]

LuxG clade {1} contains two large groups composed of (i) *Aliivibrio*, *Photobacterium* and *Shewanella* genus, and (ii) *Vibrio* genus and *Candidatus photodesmus katoptron*, plus minor third branch formed by LuxG from closely related *V. albensis* (*V. cholerae*) and *Vibrio sp.* RC586.[36] The division of luminous species majority within the clade {1} is consistent with two previously described groups of luciferases (*luxAB* gene product): “fast” and “slow” ones.[37] It supports the idea of *luxAB* and *luxG* genes co-evolution, opposite to Fre genes of luminous bacteria that do not follow the same partition tendency (Figure S2). The segregation of LuxG involving about 30 amino acid residues will be analyzed in detail elsewhere, here we indicate only the intriguing assumption of possible difference in functional properties of LuxG from two groups of luminous bacteria. The only LuxG oxidoreductase that has by now been purified and investigated is the one from *P. leiognathi*.[3,8,13]

Fre-like proteins found by BLAST are mainly represented by *Vibrionaceae* and *Enterobacteriaceae* families, with a few species from *Shewanellaceae*, *Idiomarinaceae*, *Chromatiaciae*, *Pseudoalteromonadaceae* and *Yersiniaceae* families (Figure S1). The first Fre clade {2} contains the majority of luminous species and consists of NAD(P)H:FMN-oxidoreductases only, while the second Fre clade {3} includes also aquacobalamin reductases, 2-polyprenylphenol hydroxylase-like oxidoreductases, and CDP-6-deoxy-delta-3,4-glucoseen reductases. These enzymes could function as a flavin oxidoreductase as well. Fre of luminous *Photorhabdus* species belong to the clade {3} that lacks oxidoreductases of other bioluminescent bacteria. Similar phylogenetic relationships have been observed for various housekeeping genes of luminous strains that have acquired *lux* genes horizontally.[38]

Conservation and coevolution of amino acids in LuxG and Fre sequences

The obtained phylogenetic tree indicates that LuxG and Fre oxidoreductase subfamilies are the result of a gene duplication. In this case both natural selection and neutral evolutionary processes could be the mechanisms that have caused the difference in their evolution. To examine the conservation patterns in Fre and LuxG oxidoreductases, the group of amino acid sequences from

confirmed luminous species with available sequenced genomes have been selected.[1] The extracted dataset consists of 14 primary sequences of LuxG, 19 sequences of Fre oxidoreductases and one *E. coli* Fre.

To study the potential function of the amino acids in Fre and LuxG oxidoreductases, missing structural motifs of *E. coli* Fre has been reconstructed, and *V. harveyi* LuxG 3D model has been constructed by homology modeling using Swiss-MODEL server with *E. coli* Fre structure as a template.[26] The alignment of the query sequence and the template has been calculated by the server (Figure S3) and relatively high identity and similarity (39.04 and 54.27%, respectively) have been obtained. Both structures have been refined using MD to correct possible errors that could occur in variable loops, the relative orientations of secondary structure elements, etc. (Figure S4). Structural validation of the model has been performed using PROCHECK software before and after MD simulation.[29] The analysis has showed that before MD refinement 87.3 % of LuxG residues have been found in most favored regions, while after the simulation there have been 90.7 % residues in most favored regions meaning a good quality of the obtained model (Figure S5). Both structures (PMDB ID: PM0081491 and PM0081492) have been used for further analysis of conserved positions and possible protein-ligand interactions.

We have first determined the conserved sites using a multiple sequence alignment of 34 amino acid sequences of LuxG and Fre oxidoreductases using *E. coli* Fre as the reference sequence to display the conservation of the residues. 47 alignment positions have been identified as highly conserved among all oxidoreductases (Figure 2), a number of them belong to the active center residues in flavin and nicotinamide binding sites.

Moreover, all Fre proteins have been found to share additional 22 highly conserved residues, which are distinct from the corresponding amino acids of LuxG subfamily (Table SII). Three of them have been identified as responsible for type I functional divergence (Lys45, Arg46 and Phe 203) and one – for type II functional divergence (Arg 202) (Table I). Moreover, four amino acid residues conserved among LuxG oxidoreductases have been identified as related to type I

functional divergence and one – to type II one. The determined specific residues can play a significant role in the division of oxidoreductases into Fre and LuxG subfamily and the mechanisms of their functioning.

We have explored possible functions of sites that have showed a specific conservation using superimposed structures of LuxG and Fre obtained with VMD and plugin MultiSeq.[31,32] FAD and 2'-phospho-5'-AMP coordinates have been derived from the structure of *Spinacia oleracea* oxidoreductase (PDB ID: 1FNC) using a structural alignment with *E. coli* Fre.[30]

It is known that LuxG enzymes are able to reduce FMN, FAD and riboflavin with comparable efficiency, while for Fre oxidoreductases FAD is a preferred substrate.[14] The difference in affinity to flavins could be partly attributed to the absence of the Arg46 in the structure of LuxG (Figure 3, A). This residue forms a conserved Arg46-Pro47-Phe48-Ser49 segment characteristic to all Fre oxidoreductases as well as to the members of FNR family, but not to LuxG oxidoreductases (Figure S7).[4,12]According to the alignment and the analysis of functionally diverged sites, Arg46 is characterized by different conservation between LuxG and Fre oxidoreductases indicating type I functional divergence (Table I and Table SII). In other words this site is characterized by significantly different evolutionary rates for Fre and LuxG meaning reduced functional constraints for LuxG oxidoreductases. This can lead to differences in flavin phosphate or pyrophosphate binding and consequently in the enzyme activity.

The second segment where LuxG sequences have less conservative residues compared to Fre is a short mobile loop (residues 67 – 72 in the *E. coli* Fre structure). Particularly, all Fre oxidoreductases have conserved Asn70 and Tyr72, while LuxG ones display no conservation of corresponding residues (Figure S8). Generally, for other members of FNR family, the similar but a longer loop is known to provide the tight binding of an adenosine diphosphate moiety of FAD that acts as a cofactor.[11] The shortened loops of Fre and LuxG could also perform the same function but with less efficiency because they use FAD as a substrate. Thus, the lack of

conservation in the loop sequence obtained for LuxG enzymes well corresponds to their less specificity to FAD in comparison with Fre-like reductases.

Both oxidoreductases use NADH as well as NADPH as donors of electrons. Previously it was shown that the nicotinamide ribose phosphate part of NAD(P)H plays a major role in binding to Fre, while the impact of adenosine phosphate part is negligible.[39] The functional divergence analysis revealed that individual amino acid site, namely Arg202, is related to a type II functional divergence (Table I and Table SII). This type of functional divergence refers to radical change of physical and chemical properties of amino acid residue. The position is conserved in both subfamilies: arginine in Fre and proline in LuxG (Figure 3, B). A functional divergence may have occurred after a gene duplication event and further purifying selection acted to maintain the related, but distinct, functions of the position in LuxG and Fre sub-families. In the active site of Fre oxidoreductases arginine residue could interact with the phosphate of adenosine or nicotinamide part of NAD(P)H. No arginine or a polar residue that could serve the same function has been found near proline in the active site of LuxG. Therefore the structural analysis indicates that Fre and LuxG oxidoreductases can apply different NAD(P)H binding patterns.

Conclusion

The presence of a Fre oxidoreductase homolog in the *lux*-operon of luminous bacteria was discovered over two decades ago, later a series of experiments demonstrated the role of LuxG in a coupled reaction with luciferase.[3,8] Here we have performed a comparative phylogenetic analysis to determine evolutionary relationship between Fre and LuxG including protein structural context of active site amino acid differences that have not been studied before.

The phylogenetic tree obtained from 431 homologous sequences indicated that Fre and LuxG evolved independently from a single ancestor. Probably the gene duplication led to the split of these oxidoreductases into two subfamilies. The evolution of LuxG as a part of *lux*-operon has resulted in substantial variation in amino acid composition of segments that are highly conserved among Fre oxidoreductases including those of substrate binding site. It seems that LuxG has lost

the specificity to FAD, while Fre has conserved this feature. This LuxG evolution pathway can be attributed to the bacterial luciferase specificity to reduced FMN, but not FAD. The alternative explanation can be based on the unknown role of LuxG in bacterial bioluminescence different from providing reduced flavin for the luciferase. A further comparative study of Fre and LuxG functioning *in vivo* and *in vitro* would shed light on probable additional functions of the latter oxidoreductase.

Acknowledgements

The study has been partially funded by the Russian Foundation for Basic Research, Government of Krasnoyarsk Territory, Krasnoyarsk Regional Fund of Science (project «Study of intermolecular interactions of toxicants and luminous bacteria enzymes using molecular modeling and QSAR» № 18-44-243009) and by the Ministry of Education and Science of the Russian Federation (projects № 6.7734.2017 and № 01201351504).

Conflict of Interest: none declared.

References

1. Dunlap PV, Urbanczyk H. Luminous Bacteria. In: Rosenberg E, DeLong EF, Lory S, Stackebrandt E, Thompson F, editors. The Prokaryotes. Springer, Berlin, Heidelberg; 2013. pp. 495-528.

2. Tabib CR, Brodl E, Macheroux P. Evidence for the generation of myristylated FMN by bacterial luciferase. Molecular Microbiology 2017;104(6):1027-1036.

3. Tinikul R, Pitsawong W, Sucharitakul J, Nijvipakul S, Ballou DP, Chaiyen P. The transfer of reduced flavin mononucleotide from LuxG oxidoreductase to luciferase occurs via free diffusion. Biochemistry 2013;52(39):6834-6843.

4. Swartzman E, Miyamoto C, Graham A, Meighen E. Delineation of the transcriptional boundaries of the lux operon of *Vibrio harveyi* demonstrates the presence of two new lux genes. J Biol Chem 1990;265(6):3513-3517.

5. Zenno S, Saigo K. Identification of the genes encoding NAD(P)H-flavin oxidoreductases that are similar in sequence to *Escherichia coli* Fre in four species of luminous bacteria: *Photorhabdus luminescens, Vibrio fischeri, Vibrio harveyi,* and *Vibrio orientalis*. Journal of bacteriology 1994;176(12):3544-3551.

6. Coves J, Fontecave M. Reduction and mobilization of iron by a NAD(P)H:flavin oxidoreductase from *Escherichia coli*. FEBS J 1993;211(3):635-641.

7. Fieschi F, Nivière V, Frier C, Décout JL, Fontecave M. The mechanism and substrate specificity of the NADPH: flavin oxidoreductase from *Escherichia coli*. J Biol Chem 1995;270(51):30392-30400.

8. Nijvipakul S, Wongratana J, Suadee C, Entsch B, Ballou DP, Chaiyen P. LuxG is a functioning flavin reductase for bacterial luminescence. Journal of Bacteriology 2008;190(5): 1531-1538.

9. Lee EH, Nam KS, Lee SK, Oh E, Lee CY. Ferric iron reductase activity of LuxG from *Photobacterium leiognathi*. Korean Journal of Microbiology 2016;52:495-499.

10. Campbell ZT, Baldwin TO. Fre is the major flavin reductase supporting bioluminescence from *Vibrio harveyi* luciferase in *Escherichia coli*. J Biol Chem 2009;284(13):8322-8328.

11. Ingelman M, Ramaswamy S, Nivière V, Fontecave M, Eklund H. Crystal structure of NAD(P)H:flavin oxidoreductase from *Escherichia coli*. Biochemistry 1999;38(22):7040-7049.

12. Nivière V, Fieschi F, Décout JL, Fontecave M. Is the NAD (P) H: flavin oxidoreductase from *Escherichia coli* a member of the ferredoxin-NADP+ reductase family? Evidence for the catalytic role of serine 49 residue. J Biol Chem 1996;271(28):16656-16661.

13. Nijvipakul S, Ballou DP, Chaiyen P. Reduction kinetics of a flavin oxidoreductase LuxG from Photobacterium leiognathi (TH1): half-sites reactivity. Biochemistry 2010;49(43):9241-9248.

14. Louie TM, Yang H, Karnchanaphanurach P, Xie XS, and Xun L. FAD is a preferred substrate and an inhibitor of *Escherichia coli* general NAD(P)H:flavin oxidoreductase. J Biol Chem 2002;277(42):39450-39455.

15. Esimbekova EN, Asanova AA, Deeva AA, Kratasyuk VA. Inhibition effect of food preservatives on endoproteinases. Food Chemistry 2017;235:294-297.

16. Gregor C, Gwosch KC, Sahl SJ, Hell SW Strongly enhanced bacterial bioluminescence with the *ilux* operon for single-cell imaging. Proc Nat Acad *Sci* 2018:201715946.

17. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol 1990;215(3):403-410.

18. Wheeler DL, Barrett T, Benson DA, Bryant SH, Canese K, Chetvernin V, ... Feolo M. Database resources of the national center for biotechnology information. Nucleic Acids Res 2007;36(suppl_1):D13-D21.

19. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 2013;30(4):772-780.

20. Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. Jalview Version 2 — a multiple sequence alignment editor and analysis workbench. Bioinformatics 2009;25(9):1189-1191.

21. Darriba D, Taboada GL, Doallo R, Posada D. ProtTest 3: fast selection of best-fit models of protein evolution. Bioinformatics 2011;27(8):1164-1165.

22. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol 2010;59(3):307-321.

23. Ronquist F, Huelsenbeck JP. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 2003;19(12):1572-1574.

24. Letunic I, Bork P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res 2016;44(W1):W242-W245.

25. Gu X, Zou Y, Su Z, Huang W, Zhou Z, Arendsee Z, Zeng Y. An update of DIVERGE software for functional divergence analysis of protein family. Mol Biol Evol 2013;30(7): 1713-1719.

26. Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, Kiefer F, Gallo T, Bertoni CM, Bordoli L, Schwede T. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res* 2014;42(W1):W252-W258.

27. Webb B, Sali A. Protein structure modeling with MODELLER. *Functional Genomics*. Humana Press, New York, NY; 2017. pp. 39-54.

28. Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, Lindahl E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX 2015;1:19-25.

29. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr 1993;26(2):283-291.

30. Bruns CM, Karplus AP. Refined crystal structure of spinach ferredoxin reductase at 1.7 Å resolution: oxidized, reduced and 2′-phospho-5′-AMP bound states. J Mol Biol 1995;247(1):125-145.

31. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. J Mol Graph 1996;14(1):33-38.

32. Roberts E, Eargle J, Wright D, Luthey-Schulten Z. MultiSeq: unifying sequence and structure data for evolutionary analysis. BMC bioinformatics 2006;7(1):382.

33. O'Grady EA, Wimpee CF. Mutations in the *lux* operon of natural dark mutants in the genus *Vibrio*. Appl Environ Microbiol 2008;74(1):61-66.

34. Figge MJ, Cleenwerck I, van Uijen A, De Vos P, Huys G, Robertson L. *Photobacterium piscicola* sp. nov., isolated from marine fish and spoiled packed cod. Syst Appl Microbiol 2014;37(5):329-335.

35. Macián MC, Ludwig W, Aznar R, Grimont PA, Schleifer KH, Garay E, Pujalte MJ. *Vibrio lentus* sp. nov., isolated from Mediterranean oysters. Int J Syst Evol Microbiol (2001);51(4): 1449-1456.

36. Haley BJ, Grim CJ, Hasan NA, Choi SY, Chun J, Brettin TS, Bruce DC, Challacombe JF, Detter JC, Huq A, Colwell RR. Comparative genomic analysis reveals evidence of two novel *Vibrio* species closely related to *V. cholerae*. BMC Microbiol 2010;10(1):154.

37. Deeva AA, Temlyakova EA, Sorokin AA, Nemtseva EV, Kratasyuk VA. Structural distinctions of fast and slow bacterial luciferases revealed by phylogenetic analysis. Bioinformatics 2016;32(20):3053-3057.

38. Forst S, Nealson K. Molecular biology of the symbiotic-pathogenic bacteria *Xenorhabdus* spp. and *Photorhabdus* spp. Microbiol Rev 1996;60(1):21.

39. Nivière V, Fieschi F, Décout JL, & Fontecave M. The NAD (P) H: Flavin Oxidoreductase from *Escherichia coli*. Evidence for a new mode of binding for reduced pyridine nucleotides. J Biol Chem 1999;274(26):18252-18260.

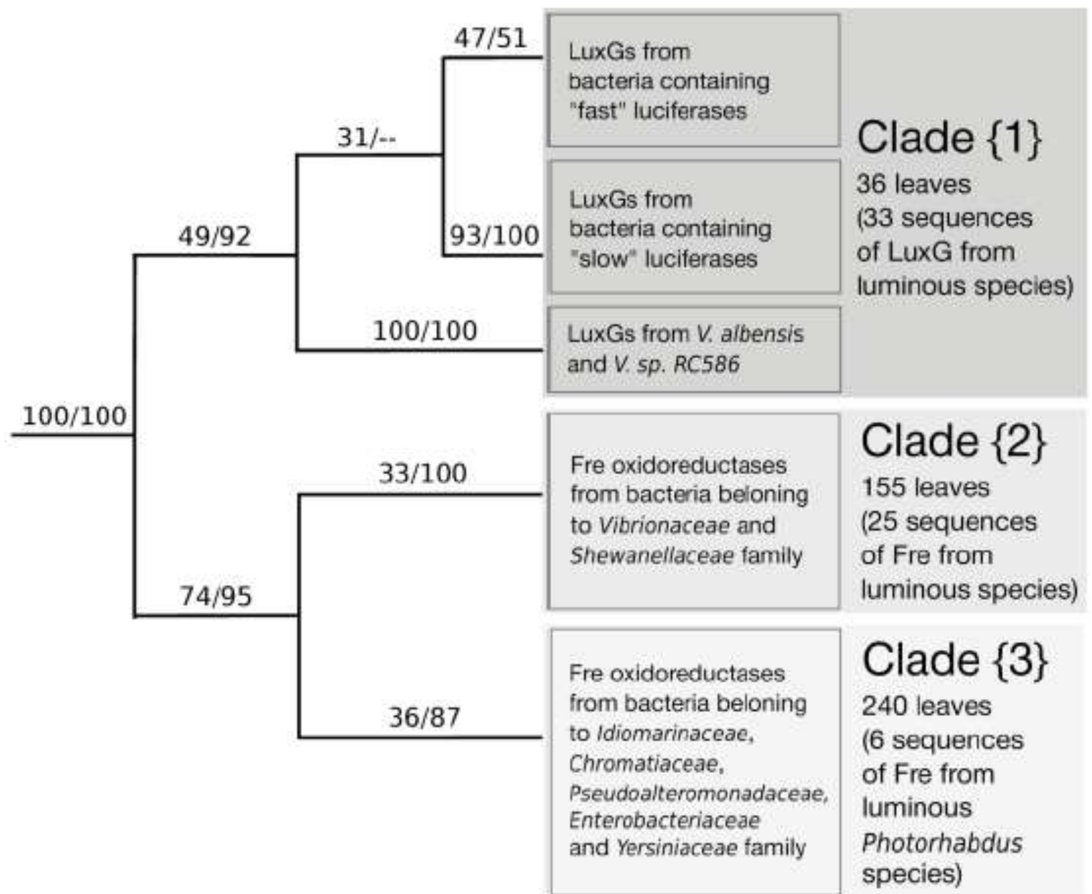

Figure 1 – A simplified representation of the phylogenetic relationship between Fre and LuxG oxidoreductases. The percentage of replicate trees in which the associated oxidoreductases clustered together in the bootstrap test (1000 replicates) is shown next to the branches together with Bayesian information criterion: bootstrap value/posterior probability

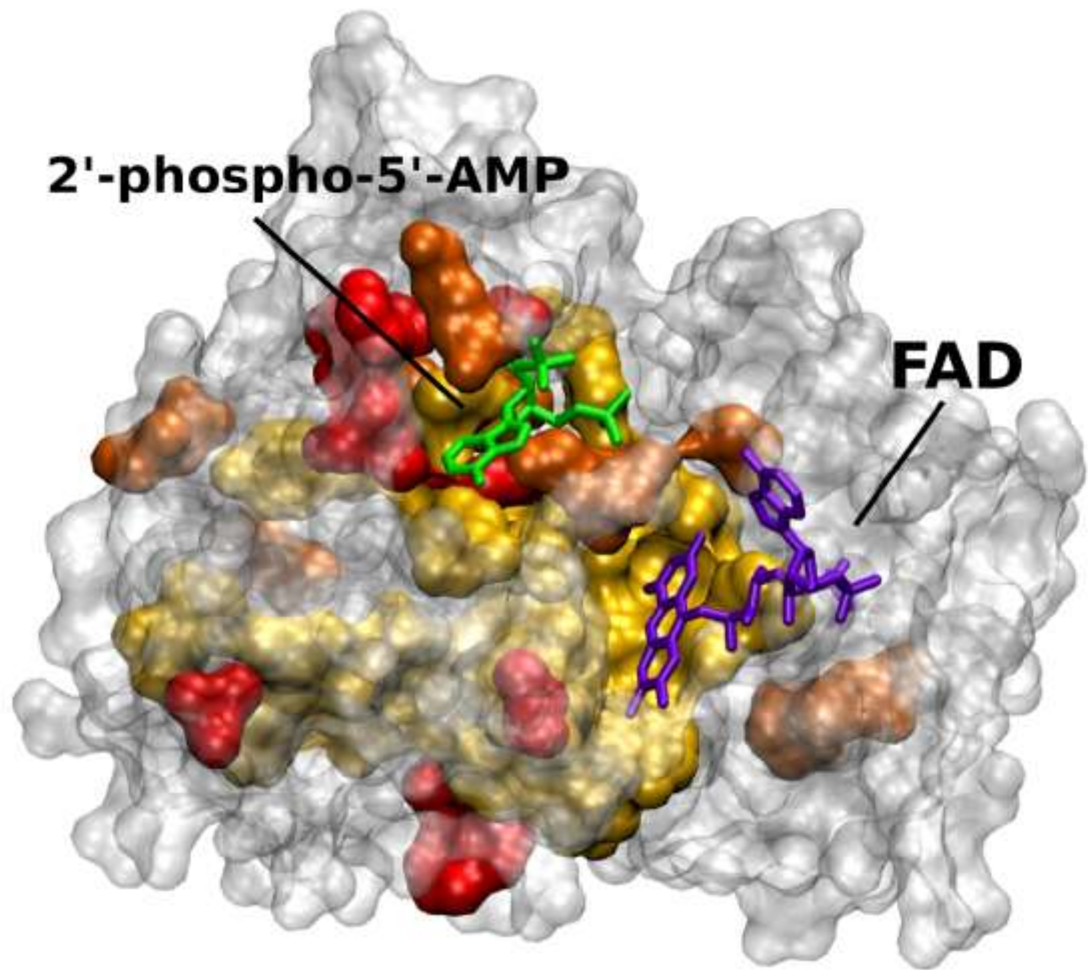

Figure 2 – Arrangement of conservative amino acid residues on the structure of E. coli Fre oxidoreductase. Identical residues among 14 LuxG and 19 Fre oxidoreductases from luminous bacteria and E. coli Fre oxidoreductase are represented by a gold surface, identical with one exception are colored orange, identical with two exceptions – red, the rest of the structure is shown by a transparent gray surface. These residues have been determined directly from the multiple sequence alignment.

The location of FAD (purple) and 2'-phospho-5'-AMP (green) in flavin- and NAD(P)-binding sites has been identified using structural alignment with *Spinacia oleracea* oxidoreductase (PDB ID: 1FNC), both molecules are shown in stick representation

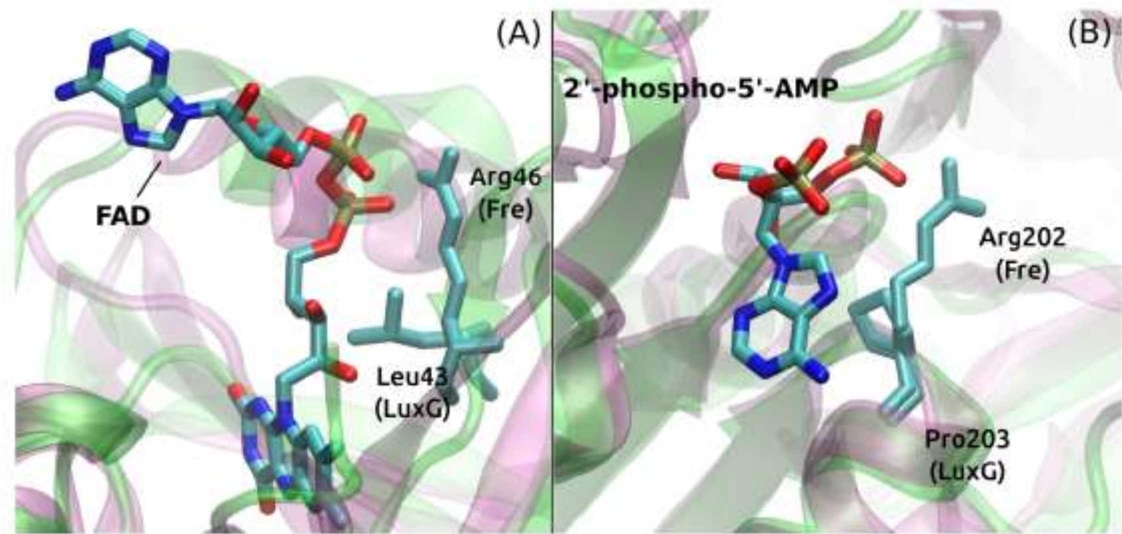

Figure 3 – Superposition of the active sites of *E. coli* Fre and *V. harveyi* LuxG. Ligands configuration has been retrieved from *Spinacia oleracea* oxidoreductase structure. Functionally important amino acids of Fre and LuxG found in the corresponding positions of the alignment are shown in stick representation. (A) depicts possible differences in phosphate binding of flavins; (B) shows adenine binding patterns distinctive for LuxG and Fre proteins