luxA gene from *Enhygromyxa salina* encodes a functional homodimeric luciferase

Anna Yudenko[1,#], Sergey V. Bazhenov[1,#], Vladimir A. Aleksenko[1,#], Ivan M. Goncharov[1], Oleg Semenov[1], Alina Remeeva[1], Vera V. Nazarenko[1], Elizaveta Kuznetsova[1], Vadim V. Fomin[1], Maria N. Konopleva[1], Rahaf Al Ebrahim[1], Nikolai N. Sluchanko[2], Yury Ryzhykau[1,3], Yury S. Semenov[1], Alexander Kuklin[1,3], Ilya V. Manukhov[1], Ivan Gushchin[1,*]

[1] Research Center for Molecular Mechanisms of Aging and Age-Related Diseases, Moscow Institute of Physics and Technology, 141700 Dolgoprudny, Russia

[2] A.N. Bach Institute of Biochemistry, Federal Research Center of Biotechnology of the Russian Academy of Sciences, 119071, Moscow, Russia

[3] Frank Laboratory of Neutron Physics, Joint Institute for Nuclear Research, 141980, Dubna, Russia

[#] These authors have contributed equally to this work

* Corresponding author. E-mail: ivan.gushchin@phystech.edu

Abstract

Several clades of luminescent bacteria are known currently. They all contain similar *lux* operons, which include the genes *luxA* and *luxB* encoding a heterodimeric luciferase. The aldehyde oxygenation reaction is catalyzed by the subunit LuxA, while LuxB is inactive. Recently, genomic analysis identified a subset of bacterial species with rearranged *lux* operons lacking *luxB*. Here, we show that *Escherichia coli* transformed with a synthetic *luxA* gene from the reduced *luxACDE* operon of *Enhygromyxa salina* is luminescent upon addition of aldehydes. Overall, *Es*LuxA is less bright compared to luciferases from *Aliivibrio fischeri* (*Af*LuxAB) and *Photorhabdus luminescens* (*Pl*LuxAB), and most active with medium-chain C4-C9 aldehydes. Crystal structure of *Es*LuxA determined at the resolution of 2.71 Å reveals a classical monooxygenase fold, and the protein preferentially forms a dimer in solution. The mobile loop residues 264-293, which form a β-hairpin or a coil in *Vibrio harveyi* LuxA, form α-helices in *Es*LuxA. Phylogenetic analysis shows *Es*LuxA and related proteins may be bacterial protoluciferases that arose prior to duplication of the *luxA* gene and its speciation to *luxA* and *luxB* in the previously described luminescent bacteria. Our work paves the way for discovery of new luciferases that have an advantage of being encoded by a single gene.

Keywords

Bacterial luciferase, luminescence, phylogeny, small-angle X-ray scattering, X-ray crystallography

1. Introduction

Luminescence is widespread in the living world, with diverse organisms using it for visual communication to scare off predators, attract prey, or in courtship behavior [1]. More than 30 types of bioluminescence systems, usually consisting of a protein called luciferase and its substrate, luciferin, are currently known [2], with new ones still being discovered [3] and being used for many different applications, such as engineering biosensors and reporters [4] or developing autoluminescent plants [5].

One of the best studied bioluminescence systems is found in bacteria: *Vibrio harveyi, Aliivibrio fischeri, Photobacterium mandapamensis, Photobacterium leiognathi, Photorhabdus luminescens* and others [6,7]. In the ecology of modern marine microbiota, bacterial bioluminescence is believed to be used for detoxification of reactive oxygen species [8], photoreactivation of DNA repair, providing cells with additional protection from the lethal effects of short-wave solar UV radiation [9], in symbiotic relationships with eukaryotic organisms [10] or for attracting them [11]. The active protein complex is a heterodimer LuxAB of the α (LuxA) and β (LuxB) subunits. Bound to a reduced flavin mononucleotide ($FMNH_2$), LuxAB utilizes molecular oxygen to convert a long chain aldehyde to a fatty acid and emit light. Alongside the *luxA* and *luxB* genes, corresponding operons also contain the genes encoding a NADPH-dependent acyl protein reductase (*luxC*), an acyl-transferase (*luxD*), and an acyl-protein synthetase (*luxE*), usually in the order *luxCDABE* [6,7]. LuxC and LuxE form a complex [12]. Additionally, *luxF* and *luxG* genes, encoding a probable myristylated flavin scavenger [13] and a flavin reductase, respectively, may also be present [6,7].

Overall, bacterial luciferases have been extensively studied. Crystallographic structure of *Vibrio harveyi* LuxAB (*Vh*LuxAB) in complex with $FMNH_2$ has been determined, revealing the general architecture of the heterodimer as well as the binding site of the flavin [14]. The structure also revealed presence of a long flexible loop covering the active site; the loop were shown later to be important for luminescence [15,16]. The reaction was shown to occur in the α subunit [14] starting with peroxide formation [17]. When expressed separately, the subunits display activity at the level of 10^{-5}–10^{-6} of that of the heterodimer [18], and *V. harveyi* LuxA remains monomeric, whereas LuxB forms a homodimer [19]. Crystallographic studies of the latter revealed similar structure to that in the heterodimer [20,21]. Interestingly, *Photobacterium* LuxF also forms a homodimer with a similar monooxygenase fold [22–24]. LuxA and LuxB can be expressed in eukaryotic cells [25] and are active when fused [26].

Recently, a bioinformatic study revealed presence of luciferase operons in hundreds of genomes, and identified *luxA*-like genes in species that were not reported previously to display luminescence [27]. These *luxA*-like genes are located in putative *luxCEDA*, *luxCAED* and *luxAxxCE* operons that miss *luxB* [27]. Here, we describe a study of the product of the peculiar *luxA* gene from *Enhygromyxa salina*. We show that it is a functional homodimeric luciferase and describe the implications.

2. Materials and Methods

2.1. Bioinformatics

Seed alignment for the luciferase-like monooxygenase domain (Pfam family PF00296) was downloaded from Pfam [28] on April 12th, 2020. This alignment was used to build an initial PSSM matrix as an input for PSI-BLAST [29] to perform a search against the non-redundant set of sequences in the NCBI database (all non-redundant GenBank coding sequence translations, PDB, SwissProt, PIR and PRF excluding environmental samples from whole genome sequencing projects), also on April 12th, 2020. The cutoff E-value was chosen at 0.001. Resulted set of sequences was clustered using UCLUST [30] at the 40% identity level and centroids were used as a dataset for further phylogenetic analysis. Additionally, a set of known bacterial luciferase sequences was used to perform PSI-BLAST searches against the same database. Top 1000 results from this search were added into the dataset for further phylogenetic analysis. Finally, the resulting dataset was combined with the sequences representative of PF00296 family members with available experimentally determined structures. Multiple sequence alignment of the collected sequences was calculated using MAFFT v. 7.402 [31] with FFT-NS-2 algorithm [32]. Columns containing more than 50% gaps were removed using trimAl [33]. The phylogenetic tree was constructed using FastTree2 [34]. LG amino acid substitution model [35] was used, with 20 rate categories of CAT approximation, 44/22 rounds of ME/ML nearest-neighbor interchanges with turned off heuristics, 4 rounds of subtree-prune-regraft moves and the maximum length of a move of 10. Illustrations of the phylogenetic trees were prepared using FigTree v. 1.4.4 [36].

2.2. Cloning

Bacterial strains and plasmids used in the study are listed in Table S1. The gene encoding the putative LuxA protein from *E. salina* (GenBank ID WP_106392368) optimized for *E. coli* expression and harboring C-terminal 6×His-tag was synthesized *de novo*; the nucleotide sequence is presented in the supplementary materials. The gene was cloned into the pET11 expression vector (Novagen, Merck, Germany) using BamHI and NdeI restriction sites. The pABX plasmid encoding the *luxAB* genes from *P. luminescens* under control of P_{lac} was prepared by linearizing the pUC19 vector by KpnI restriction endonuclease, amplifying the *luxAB* fragment from the pXen7 plasmid [37], and ligating the resulting fragments with the use of Gibson Assembly [38].

2.3. Functional characterization

To characterize the luciferase activity, cell cultures were grown overnight on agarized (1.5% w/v) autoinduction ZYP-5025 medium [39] at 22 °C. Immediately prior to luminescence measurements, the overnight cultures were resuspended in liquid LB to reach the final optical density (OD) at 600 nm of 0.05 for *E. coli* K12 MG1655 pABX and K12 MG1655 pF2; and to OD of 2 for *E. coli* BL21(DE3) pSluxA, because of its very low luminescence. *E. coli* BL21(DE3) transformed with pET11 harboring a gene encoding a flavin-binding protein CagFbFP [40] was used as a negative control. The effect of aldehyde chain length on enzyme activity in investigated cells was obtained by subjecting the cell suspensions to different aldehydes at final concentration of 10 mM and then measuring the maximum light intensity using Biotox-7BM (BioPhysTech, Russia). Michaelis constants were determined by subjecting the cell suspensions to different aldehyde concentrations and fitting the luminescence curves using nonlinear least squares regression as implemented by Angel Herráez (https://biomodel.uah.es/en/metab/enzimas/MM-regresion.htm) based on the algorithm by John C. Pezzullo (https://statpages.info/nonlin.html).

2.4. Protein expression and purification

The protein was expressed in *Escherichia coli* strain C41 (DE3). Cells were cultured in shaking baffled flasks in LB medium containing 150 mg/l ampicillin. Protein expression was induced with 1 mM IPTG and continued for 22 hours at 20 °C. Harvested cells were resuspended in phosphate-buffered saline with addition of 1 mM PMSF and disrupted in M-110P Lab Homogenizer (Microfluidics, USA) at 25000 psi. The lysate was clarified by removal of cell membrane fraction by ultracentrifugation at 40000 g for 1 h at 10º C. The supernatant was incubated with Ni-nitrilotriacetic acid (Ni-NTA) resin (Qiagen, Germany) on a rocker for 1 hour at 4º C. The Ni-NTA resin and supernatant were loaded on a gravity flow column and washed with buffer containing 300 mM NaCl and 50 mM Tris-HCl, pH 8.0, supplemented with 20 mM imidazole. The protein was eluted in a buffer containing 50-200 mM imidazole, 300 mM NaCl and 50 mM Tris-HCl, pH 8.0. The eluate was concentrated and subjected to size-exclusion chromatography (SEC) on Superdex® 200 Increase 10/300 GL column (GE Healthcare Life Sciences, USA) in a buffer containing 150 mM NaCl and 25 mM Tris-HCl, pH 8.0.

2.5. Crystallization

Protein-containing fractions from SEC were pooled and, to prevent formation of disulfide bridges between surface-exposed cysteines, incubated with 15 mM iodoacetamide, which caused precipitation. Consequently, the protein was resuspended in a buffer containing 150 mM NaCl, 25 mM Tris-HCl and 15% glycerol, pH 8.0 and concentrated to 11.6 mg/ml for crystallization. Sitting drop vapor diffusion crystallization trials were set up using an NT8 robotic system (Formulatrix, USA) and stored at 20 °C. The drops contained 100 nL concentrated protein solution and 100 nL reservoir solution. The best crystals were obtained using the following solution as the precipitant: PEG3350 25%, lithium sulfate 0.2 M, BIS-TRIS 0.1 M pH 5.5. Cubic crystals appeared and reached the final size of ~100 μm within a month. The crystals were harvested using micromounts and then flash-cooled and stored in liquid nitrogen.

2.6. Diffraction data collection and processing

The diffraction data were collected at 100 K at the European Synchrotron Radiation Facility (ESRF) beamline ID23-1 equipped with an EIGER2 S 16M detector (Dectris, Switzerland). Diffraction images were processed using XDS [41]. AIMLESS [42] was used to merge and scale the data. The data collection and processing statistics are reported in Table 1.

2.7. Structure determination and refinement

The structure was solved using molecular replacement with MOLREP [43] and AlphaFold [44] model of *Es*LuxA as a search model. The resulting model was refined manually using Coot [45] and REFMAC5 [46].

2.8. Multi-angle light scattering (MALS)

SEC coupled with MALS (SEC-MALS) was done using a Superdex® 75 10/300 GL column (GE Healthcare Life Sciences, USA) and a combination of a UV-Vis Prostar 335 detector (Varian, Australia) and a miniDAWN detector (Wyatt Technology, USA) connected in this order. The *Es*LuxA sample (1 mg/ml in 120 μl) was loaded on the column equilibrated with filtered (0.1 μm) and degassed 25 mM Tris-HCl buffer, pH 8.0, containing 150 mM NaCl at 0.8 ml/min flow rate. Data were processed in ASTRA 8.0 (Wyatt Technology, USA) using dn/dc equal to 0.185 and protein extinction coefficient ε(0.1%) at 280 nm of 1.239.

2.9. Small-angle X-ray scattering (SAXS)

The SAXS experiments were carried out on the Rigaku instrument at the Moscow Institute of Physics and Technology (Dolgoprudny, Russia) as described previously [47,48]. SAXS profiles were obtained for the peak fraction after gel-filtration with protein concentration of 3.5 mg/ml, in a buffer containing 150 mM NaCl and 25 mM Tris-HCl, pH 8.0. For information on *Es*LuxA oligomerization polydispersity, we prepared models of the protein in monomeric, dimeric and tetrameric (dimer of dimers) states based on the crystal structure. The missing unstructured parts of the models were completed with fragments generated using AlphaFold [44]. Form-factors were calculated using CRYSOL [49]. The experimental data were approximated by superposition of form-factors using OLIGOMER [50] program from ATSAS software suite [51]. Details of SAXS measurements and data treatment are presented in Table S2. Guinier approximation, Kratky plot and pair distribution function are shown in Fig. S1.

3. Results

3.1. Phylogenetic analysis of luxA-related genes

We began our study with reanalysis of the available genomic data. We obtained the sequences of the monooxygenase fold proteins from InterPro and constructed a phylogenetic tree (Fig. 1). In accordance with the findings by Vannier and colleagues [27], we observed that *luxA* and *luxB* genes form two well-defined clusters, separate from other monooxygenase family members. *luxF* is highly dissimilar from *luxA* and *luxB* and is unlikely to be a direct descendant of a luciferase gene (Fig. 1). Several *luxA*-like genes are observed near the branching point of *luxA* and *luxB* clusters as parts of diverse operons also including *luxC*-like, *luxD*-like and *luxE*-like genes (Fig. 1). Compared to the findings of Vannier and colleagues [27], the genes from *Leptospira fletcheri, Leptospira fluminis, Leptospira perolatii, Nocardia brasiliensis* and *Nocardia uniformis* appear to be new (Table S3).

3.2. Luminescence properties of Enhygromyxa salina LuxA

Willing to test the properties of the proteins encoded by *luxA*-like genes experimentally, we synthesized a gene encoding putative *E. salina* luciferase LuxA (*Es*LuxA, GenBank ID WP_106392368, sequence identity to *Vh*LuxA is 41%) with a C-terminal histidine tag and cloned it into the pET11 plasmid. *E. coli* cells, transformed with the plasmid, displayed moderate luminescence upon addition of various aldehydes to the media (Fig. 2). Compared to canonical heterodimeric luciferases from *A. fischeri* (*Af*LuxAB) and *P. luminescens* (*Pl*LuxAB), which are most active with longer-chain aldehydes, *Es*LuxA was most active with medium-chain C4-C9 aldehydes. Normalized for the aldehyde concentration and cell density, the activity of *Es*LuxA was around 10^{-5} of that of *Af*LuxAB and *Pl*LuxAB.

Next, we tested the dependence of luminescence on the aldehyde concentration *in vivo*. Apparent Michaelis constants (K_M) determined for *Es*LuxA-expressing cells are 9±3 mM both for hexanal and nonanal, whereas for *Pl*LuxAB-expressing cells K_M is much lower for nonanal (25±2 μM) than for hexanal (7±2 mM).

3.3. Structural characterization of Enhygromyxa salina LuxA

To gain insight into the *Es*LuxA structural properties, we overexpressed the protein in *E. coli* and purified it using metal-affinity and size-exclusion chromatography. The purified protein was colorless and, presumably, in the apo-form. We determined the crystal structure of *Es*LuxA at the resolution of 2.71 Å. The P1 unit cell contains two *Es*LuxA homodimers.

*Es*LuxA has an $(\alpha/\beta)_8$ barrel fold characteristic for LuxA/LuxB and other monooxygenase family members [14] (Fig. 3). The structures of all *Es*LuxA protomers within the crystal are essentially identical with the exception of the amino acids 264-293, corresponding to the mobile loop of the α-subunit of classical heterodimeric bacterial luciferases (residues 259-283 in *Vh*LuxA, Fig. 3A and B). Presence of this elongated partially unstructured loop confirms annotation of the *E. salina luxA* gene product as LuxA-like, rather than LuxB-like. While the loop is disordered in protomers B and D, some of its residues form α-helices in protomers A and C (Fig. 3B). No putative cofactors or substrates are observed in the electron densities; a single sulfate group is observed in the *Es*LuxA structure at the place of the $FMNH_2$'s phosphate in the *Vh*LuxAB structure.

To determine the oligomeric state of *Es*LuxA in solution, we conducted SEC-MALS and small-angle X-ray scattering (SAXS) experiments. While the MW of His-tagged *Es*LuxA monomer is expected to be 42.2 kDa, SEC-MALS revealed that the major fraction contained the species with MW of 88.3 kDa, and the minor fraction contained the species with MW of 169.5 kDa (Fig. 3C), corresponding to *Es*LuxA dimers and tetramers, respectively. In accordance with this, the SAXS profile could be reasonably fitted with the theoretical scattering of 82% to 18% mixture of *Es*LuxA dimers and tetramers, respectively (Fig. 3D), which is in accordance with the MALS data.

4. Discussion

Recently, Vannier and colleagues identified *luxA*-like genes in species that were not reported to display luminescence [27]. Phylogenetic analysis shows that previously known *luxA* and *luxB* probably arose from duplication of some protoluciferase gene and underwent asymmetric divergence [52]: all *luxB* have approximately the same, relatively large, distance to the branching point, whereas some of the *luxA* genes are relatively close to it (Fig. 1). *luxA*-like genes are similarly closest to the *luxA*/*luxB* branching point. Given that *Es*LuxA and related proteins all have an elongated 'mobile loop' covering the active site, characteristic for α, but not β subunits of heterodimeric luciferases [27], it is reasonable to denote them *luxA* and not *luxB*. At the same time, it is important to keep in mind that the main function of these genes and proteins might be not luminescence *per se*, and thus they might represent protoluciferases rather than actual luciferases.

Here, we showed that the *luxA*-like gene from *E. salina* indeed encodes a functional homodimeric luciferase, *Es*LuxA, which, in contrast with *Vh*LuxA, remains monomeric [19]. Crystallographic structure of *Es*LuxA is similar to those of *Vh*LuxAB heterodimer and *Vh*LuxB homodimer. MALS and SAXS analyses show that in solution *Es*LuxA preferentially forms dimers, with a minor addition of tetramers.

E. salina is a slightly halophilic myxobacterium isolated from coastal areas of Japan [53] that has not been described as being luminescent. Accordingly, we found that *Es*LuxA is dim; its level of luminescence is several orders of magnitude lower compared to classical heterodimeric luciferases (Fig. 2). Coexpression of *E. salina luxA* with *P. luminescens luxB* or *A. logei luxB* did not lead to an increase in luminescence (data not shown). Previously reported attempts to hybridize other bacterial luciferases were also not successful [54]. Evidently, the subunit interaction interface is important for proper functioning of bacterial luciferases [55] and evolved for preferential binding of the native partner.

The optimal aldehyde tail length for *Es*LuxA is around 6 carbon atoms (Fig. 2); in contrast, heterodimeric luciferases *Af*LuxAB and *Pl*LuxAB are more active with aldehydes with chain lengths of 8-10 carbons (Fig. 2 and refs. [54,56]). Due to low activity of *Es*LuxA, it was difficult to assess the properties of the purified enzyme. Previously, the activity and K_M values of the *A. fisheri* luciferase obtained *in vivo* and *in vitro* were found to be almost the same for aldehydes shorter than octanal, while long chain aldehydes resulted in higher K_M values *in vivo* [57]. We found that for *Pl*LuxAB-expressing cells the K_M constant is ~300 times lower for nonanal (25±2 μM) than for hexanal (7±2 mM). Interestingly, for *Es*LuxA-expressing cells, there is no difference between hexanal and nonanal, as the K_M constants for both are 9±3 mM.

Altogether, our results show that *Es*LuxA is dim and rather nonselective, and thus is an unusual luciferase. Possible explanations for these findings are a) accidental loss of the *luxB* gene in the *lux*-operon of *E. salina*; b) requirement of native *luxC*, *luxD* and *luxE* genes for optimal functioning, and/or preference for an unorthodox aldehyde as the preferred substrate; c) that the main function of *Es*LuxA is not to emit light, but rather to be a terminal oxidase, providing an alternate respiratory pathway for cells grown under conditions where the cytochrome system is functionally impaired [58].

To summarize, our work paves the way for discovery of new unusual bacterial luciferases and for description of new luminescent microorganisms. Homodimeric luciferase encoded by a single gene is also promising for development of new tools for molecular biology:

bacterial luciferases are used ubiquitously to construct whole cell biosensors or molecular reporters [59–61], yet their heterodimeric nature complicates heterologous expression and engineering efforts [25,26,62].

Supplementary material

Supplementary text, supplementary tables 1-4 and supplementary figure 1 are available.

Declaration of competing interest

The authors declare no conflicts of interest related to this work.

Acknowledgements

Atomic coordinates and structure factors for the reported crystal structure have been deposited in the Protein Data Bank [63] under the accession code 8CBB. The diffraction experiments were performed on the beamline ID23-1 at the European Synchrotron Radiation Facility (ESRF), Grenoble, France [64]. We are grateful to Sergey Bukhdruker for conducting preliminary diffraction tests and to the ID23-1 beamline staff for providing assistance in using the beamline. Small angle scattering data have been deposited in the SASBDB [65] under the accession code SASDQA7. Bioinformatic analyses and characterization of *Es*LuxA was supported by the Russian Science Foundation, grant number 21-64-00018. Crystallographic data collection was supported by the Ministry of Science and Higher Education of the Russian Federation grant No. 075-15-2021-1354. Assessment of protein engineering prospects of bacterial luciferases was supported by the Ministry of Science and Higher Education of the Russian Federation grant No. 075-03-2023-106 (project FSMG-2021-0002).

References

[1] A. Fleiss, K.S. Sarkisyan, A brief review of bioluminescent systems (2019), Curr Genet. 65 (2019) 877–882. https://doi.org/10.1007/s00294-019-00951-5.

[2] Z.M. Kaskova, A.S. Tsarkova, I.V. Yampolsky, 1001 lights: luciferins, luciferases, their mechanisms of action and applications in chemical analysis, biology and medicine, Chem. Soc. Rev. 45 (2016) 6048–6077. https://doi.org/10.1039/C6CS00296J.

[3] A.A. Kotlobay, K.S. Sarkisyan, Y.A. Mokrushina, M. Marcet-Houben, E.O. Serebrovskaya, N.M. Markina, L. Gonzalez Somermeyer, A.Y. Gorokhovatsky, A. Vvedensky, K.V. Purtov, V.N. Petushkov, N.S. Rodionova, T.V. Chepurnyh, L.I. Fakhranurova, E.B. Guglya, R. Ziganshin, A.S. Tsarkova, Z.M. Kaskova, V. Shender, M. Abakumov, T.O. Abakumova, I.S. Povolotskaya, F.M. Eroshkin, A.G. Zaraisky, A.S. Mishin, S.V. Dolgov, T.Y. Mitiouchkina, E.P. Kopantzev, H.E. Waldenmaier, A.G. Oliveira, Y. Oba, E. Barsova, E.A. Bogdanova, T. Gabaldón, C.V. Stevani, S. Lukyanov, I.V. Smirnov, J.I. Gitelson, F.A. Kondrashov, I.V. Yampolsky, Genetically encodable bioluminescent system from fungi, Proceedings of the National Academy of Sciences. 115 (2018) 12728–12732. https://doi.org/10.1073/pnas.1803615115.

[4] X. Zhou, S. Mehta, J. Zhang, Genetically Encodable Fluorescent and Bioluminescent Biosensors Light Up Signaling Networks, Trends Biochem Sci. 45 (2020) 889–905. https://doi.org/10.1016/j.tibs.2020.06.001.

[5] T. Mitiouchkina, A.S. Mishin, L.G. Somermeyer, N.M. Markina, T.V. Chepurnyh, E.B. Guglya, T.A. Karataeva, K.A. Palkina, E.S. Shakhova, L.I. Fakhranurova, S.V. Chekova, A.S. Tsarkova, Y.V. Golubev, V.V. Negrebetsky, S.A. Dolgushin, P.V. Shalaev, D. Shlykov, O.A. Melnik, V.O. Shipunova, S.M. Deyev, A.I. Bubyrev, A.S. Pushin, V.V. Choob, S.V. Dolgov, F.A. Kondrashov, I.V. Yampolsky, K.S. Sarkisyan, Plants with genetically encoded autoluminescence, Nat Biotechnol. 38 (2020) 944–946. https://doi.org/10.1038/s41587-020-0500-9.

[6] E. Brodl, A. Winkler, P. Macheroux, Molecular Mechanisms of Bacterial Bioluminescence, Comput Struct Biotechnol J. 16 (2018) 551–564. https://doi.org/10.1016/j.csbj.2018.11.003.

[7] P. Dunlap, Biochemistry and genetics of bacterial bioluminescence, Adv Biochem Eng Biotechnol. 144 (2014) 37–64. https://doi.org/10.1007/978-3-662-43385-0_2.

[8] G.S. Timmins, S.K. Jackson, H.M. Swartz, The evolution of bioluminescent oxygen consumption as an ancient oxygen detoxification mechanism, J Mol Evol. 52 (2001) 321–332. https://doi.org/10.1007/s002390010162.

[9] O.E. Melkina, V.Yu. Kotova, M.N. Konopleva, I.V. Manukhov, K.S. Pustovoit, G.B. Zavilgelsky, Photoreactivation of UV-exposed Escherichia coli K12 AB1886 uvrA6 via luminescence of Photobacterium leiognathi luciferase, Mol Biol. 49 (2015) 928–932. https://doi.org/10.1134/S0026893315060175.

[10] K.L. Visick, E.G. Ruby, Vibrio fischeri and its host: it takes two to tango, Curr Opin Microbiol. 9 (2006) 632–638. https://doi.org/10.1016/j.mib.2006.10.001.

[11] M. Zarubin, S. Belkin, M. Ionescu, A. Genin, Bacterial bioluminescence as a lure for marine zooplankton and fish, Proc Natl Acad Sci U S A. 109 (2012) 853–857. https://doi.org/10.1073/pnas.1116683109.

[12] Q. Tian, J. Wu, H. Xu, Z. Hu, Y. Huo, L. Wang, Cryo-EM structure of the fatty acid reductase LuxC–LuxE complex provides insights into bacterial bioluminescence, Journal of Biological Chemistry. 298 (2022). https://doi.org/10.1016/j.jbc.2022.102006.

[13] C.R. Tabib, E. Brodl, P. Macheroux, Evidence for the generation of myristylated FMN by bacterial luciferase, Molecular Microbiology. 104 (2017) 1027–1036. https://doi.org/10.1111/mmi.13676.

[14] Z.T. Campbell, A. Weichsel, W.R. Montfort, T.O. Baldwin, Crystal structure of the bacterial luciferase/flavin complex provides insight into the function of the beta subunit, Biochemistry. 48 (2009) 6085–6094. https://doi.org/10.1021/bi900003t.

[15] J.M. Sparks, T.O. Baldwin, Functional implications of the unstructured loop in the (beta/alpha)(8) barrel structure of the bacterial luciferase alpha subunit, Biochemistry. 40 (2001) 15436–15443. https://doi.org/10.1021/bi0111855.

[16] Z.T. Campbell, T.O. Baldwin, Two lysine residues in the bacterial luciferase mobile loop stabilize reaction intermediates, J Biol Chem. 284 (2009) 32827–32834. https://doi.org/10.1074/jbc.M109.031716.

[17] N. Lawan, R. Tinikul, P. Surawatanawong, A.J. Mulholland, P. Chaiyen, QM/MM Molecular Modeling Reveals Mechanism Insights into Flavin Peroxide Formation in Bacterial Luciferase, J. Chem. Inf. Model. 62 (2022) 399–411. https://doi.org/10.1021/acs.jcim.1c01187.

[18] J.F. Sinclair, J.J. Waddle, E.F. Waddill, T.O. Baldwin, Purified native subunits of bacterial luciferase are active in the bioluminescence reaction but fail to assemble into the .alpha..beta. structure, Biochemistry. 32 (1993) 5036–5044. https://doi.org/10.1021/bi00070a010.

[19] H. Choi, C.-K. Tang, S.-C. Tu, Catalytically Active Forms of the Individual Subunits of Vibrio harveyi Luciferase and Their Kinetic and Binding Properties (∗), Journal of Biological Chemistry. 270 (1995) 16813–16819. https://doi.org/10.1074/jbc.270.28.16813.

[20] J.J. Tanner, M.D. Miller, K.S. Wilson, S.-C. Tu, K.L. Krause, Structure of Bacterial Luciferase β2 Homodimer: Implications for Flavin Binding, Biochemistry. 36 (1997) 665–672. https://doi.org/10.1021/bi962511x.

[21] J.B. Thoden, H.M. Holden, A.J. Fisher, J.F. Sinclair, G. Wesenberg, T.O. Baldwin, I. Rayment, Structure of the β2 homodimer of bacterial luciferase from vibrio harveyi: X-ray analysis of a kinetic protein folding trap, Protein Science. 6 (1997) 13–23. https://doi.org/10.1002/pro.5560060103.

[22] S.A. Moore, M.N. James, Structural refinement of the non-fluorescent flavoprotein from Photobacterium leiognathi at 1.60 A resolution, J Mol Biol. 249 (1995) 195–214. https://doi.org/10.1006/jmbi.1995.0289.

[23] A. Kita, S. Kasai, M. Miyata, K. Miki, Structure of flavoprotein FP390 from a luminescent bacterium Photobacterium phosphoreum refined at 2.7 Å resolution, Acta Cryst D. 52 (1996) 77–86. https://doi.org/10.1107/S0907444995009796.

[24] T. Bergner, C.R. Tabib, A. Winkler, S. Stipsits, H. Kayer, J. Lee, J.P. Malthouse, S. Mayhew, F. Müller, K. Gruber, P. Macheroux, Structural and biochemical properties of LuxF from Photobacterium leiognathi, Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics. 1854 (2015) 1466–1475. https://doi.org/10.1016/j.bbapap.2015.07.008.

[25] T. Xu, S. Ripp, G.S. Sayler, D.M. Close, Expression of a Humanized Viral 2A-Mediated lux Operon Efficiently Generates Autonomous Bioluminescence in Human Cells, PLOS ONE. 9 (2014) e96347. https://doi.org/10.1371/journal.pone.0096347.

[26] B. Cui, L. Zhang, Y. Song, J. Wei, C. Li, T. Wang, Y. Wang, T. Zhao, X. Shen, Engineering an Enhanced, Thermostable, Monomeric Bacterial Luciferase Gene As a Reporter in Plant Protoplasts, PLOS ONE. 9 (2014) e107885. https://doi.org/10.1371/journal.pone.0107885.

[27] T. Vannier, P. Hingamp, F. Turrel, L. Tanet, M. Lescot, Y. Timsit, Diversity and evolution of bacterial bioluminescence genes in the global ocean, NAR Genomics and Bioinformatics. 2 (2020). https://doi.org/10.1093/nargab/lqaa018.

[28] S. El-Gebali, J. Mistry, A. Bateman, S.R. Eddy, A. Luciani, S.C. Potter, M. Qureshi, L.J. Richardson, G.A. Salazar, A. Smart, E.L.L. Sonnhammer, L. Hirsh, L. Paladin, D. Piovesan, S.C.E. Tosatto, R.D. Finn, The Pfam protein families database in 2019, Nucleic Acids Res. 47 (2019) D427–D432. https://doi.org/10.1093/nar/gky995.

[29] S.F. Altschul, T.L. Madden, A.A. Schäffer, J. Zhang, Z. Zhang, W. Miller, D.J. Lipman, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, Nucleic Acids Res. 25 (1997) 3389–3402. https://doi.org/10.1093/nar/25.17.3389.

[30] R.C. Edgar, Search and clustering orders of magnitude faster than BLAST, Bioinformatics. 26 (2010) 2460–2461. https://doi.org/10.1093/bioinformatics/btq461.

[31] K. Katoh, D.M. Standley, MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability, Mol Biol Evol. 30 (2013) 772–780. https://doi.org/10.1093/molbev/mst010.

[32] K. Katoh, K. Misawa, K. Kuma, T. Miyata, MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform, Nucleic Acids Res. 30 (2002) 3059–3066. https://doi.org/10.1093/nar/gkf436.

[33] S. Capella-Gutiérrez, J.M. Silla-Martínez, T. Gabaldón, trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses, Bioinformatics. 25 (2009) 1972–1973. https://doi.org/10.1093/bioinformatics/btp348.

[34] M.N. Price, P.S. Dehal, A.P. Arkin, FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments, PLOS ONE. 5 (2010) e9490. https://doi.org/10.1371/journal.pone.0009490.

[35] S.Q. Le, O. Gascuel, An Improved General Amino Acid Replacement Matrix, Molecular Biology and Evolution. 25 (2008) 1307–1320. https://doi.org/10.1093/molbev/msn067.

[36] A. Rambaut, FigTree v1.4, 2012.

[37] G.B. Zavilgelsky, A.P. Zarubina, I.V. Manukhov, Sequencing and Comparative Analysis of the lux Operon of Photorhabdus luminescens Strain Zm1: ERIC Elements as Putative Recombination Hot Spots, Molecular Biology. 36 (2002) 637–647. https://doi.org/10.1023/A:1020663128043.

[38] D.G. Gibson, L. Young, R.-Y. Chuang, J.C. Venter, C.A. Hutchison, H.O. Smith, Enzymatic assembly of DNA molecules up to several hundred kilobases, Nat Methods. 6 (2009) 343–345. https://doi.org/10.1038/nmeth.1318.

[39] F.W. Studier, Protein production by auto-induction in high-density shaking cultures, Protein Expression and Purification. 41 (2005) 207–234. https://doi.org/10.1016/j.pep.2005.01.016.

[40] V.V. Nazarenko, A. Remeeva, A. Yudenko, K. Kovalev, A. Dubenko, I.M. Goncharov, P. Kuzmichev, A.V. Rogachev, P. Buslaev, V. Borshchevskiy, A. Mishin, G.V. Dhoke, U. Schwaneberg, M.D. Davari, K.-E. Jaeger, U. Krauss, V. Gordeliy, I. Gushchin, A thermostable flavin-based fluorescent protein from Chloroflexus aggregans: a framework for ultra-high resolution structural studies, Photochem. Photobiol. Sci. 18 (2019) 1793–1805. https://doi.org/10.1039/c9pp00067d.

[41] W. Kabsch, XDS, Acta Crystallographica Section D Biological Crystallography. 66 (2010) 125–132. https://doi.org/10.1107/S0907444909047337.

[42] P. Evans, Scaling and assessment of data quality, Acta Crystallographica Section D Biological Crystallography. 62 (2005) 72–82. https://doi.org/10.1107/S0907444905036693.

[43] A. Vagin, A. Teplyakov, Molecular replacement with MOLREP, Acta Crystallographica Section D Biological Crystallography. 66 (2009) 22–25. https://doi.org/10.1107/S0907444909042589.

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

[45] P. Emsley, B. Lohkamp, W.G. Scott, K. Cowtan, Features and development of Coot, Acta Cryst D. 66 (2010) 486–501. https://doi.org/10.1107/S0907444910007493.

[46] G.N. Murshudov, P. Skubák, A.A. Lebedev, N.S. Pannu, R.A. Steiner, R.A. Nicholls, M.D. Winn, F. Long, A.A. Vagin, REFMAC5 for the refinement of macromolecular crystal structures, Acta Crystallographica Section D Biological Crystallography. 67 (2011) 355–367. https://doi.org/10.1107/S0907444911001314.

[47] T.N. Murugova, O.I. Ivankov, Y.L. Ryzhykau, D.V. Soloviov, K.V. Kovalev, D.V. Skachkova, A. Round, C. Baeken, A.V. Ishchenko, O.A. Volkov, A.V. Rogachev, A.V. Vlasov, A.I. Kuklin, V.I. Gordeliy, Mechanisms of membrane protein crystallization in 'bicelles,' Sci Rep. 12 (2022) 11109. https://doi.org/10.1038/s41598-022-13945-0.

[48] G.V. Tsoraev, E.A. Protasova, E.A. Klimanova, Y.L. Ryzhykau, A.I. Kuklin, Y.S. Semenov, B. Ge, W. Li, S. Qin, T. Friedrich, N.N. Sluchanko, E.G. Maksimov, Anti-Stokes fluorescence excitation reveals conformational mobility of the C-phycocyanin chromophores, Structural Dynamics. 9 (2022) 054701. https://doi.org/10.1063/4.0000164.

[49] D. Svergun, C. Barberato, M.H.J. Koch, CRYSOL – a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates, J Appl Cryst. 28 (1995) 768–773. https://doi.org/10.1107/S0021889895007047.

[50] P.V. Konarev, V.V. Volkov, A.V. Sokolova, M.H.J. Koch, D.I. Svergun, PRIMUS: a Windows PC-based system for small-angle scattering data analysis, J Appl Cryst. 36 (2003) 1277–1282. https://doi.org/10.1107/S0021889803012779.

[51] D. Franke, M.V. Petoukhov, P.V. Konarev, A. Panjkovich, A. Tuukkanen, H.D.T. Mertens, A.G. Kikhney, N.R. Hajizadeh, J.M. Franklin, C.M. Jeffries, D.I. Svergun, ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions, J Appl Cryst. 50 (2017) 1212–1225. https://doi.org/10.1107/S1600576717007786.

[52] S. Mallik, D.S. Tawfik, E.D. Levy, How gene duplication diversifies the landscape of protein oligomeric state and function, Current Opinion in Genetics & Development. 76 (2022) 101966. https://doi.org/10.1016/j.gde.2022.101966.

[53] T. Iizuka, Y. Jojima, R. Fudou, M. Tokura, A. Hiraishi, S. Yamanaka, Enhygromyxa salina gen. nov., sp. nov., a Slightly Halophilic Myxobacterium Isolated from the Coastal Areas of Japan, Systematic and Applied Microbiology. 26 (2003) 189–196. https://doi.org/10.1078/072320203322346038.

[54] J.W. Hastings, K. Weber, J. Friedland, A. Eberhard, G. Mitchell, A. Gunsalus, Structurally distinct bacterial luciferases, Biochemistry. 8 (1969) 4681–4689. https://doi.org/10.1021/bi00840a004.

[55] J.K. Inlow, T.O. Baldwin, Mutational Analysis of the Subunit Interface of Vibrio harveyi Bacterial Luciferase, Biochemistry. 41 (2002) 3906–3915. https://doi.org/10.1021/bi012113g.

[56] P. Colepicolo, K.W. Cho, G.O. Poinar, J.W. Hastings, Growth and luminescence of the bacterium Xenorhabdus luminescens from a human wound, Applied and Environmental Microbiology. 55 (1989) 2601–2606. https://doi.org/10.1128/aem.55.10.2601-2606.1989.

[57] P. Rogers, W.D. McElroy, Enzymic determination of aldehyde permeability in luminous bacteria. I. Effect of chain length on light emission and penetration, Archives of Biochemistry and Biophysics. 75 (1958) 87–105. https://doi.org/10.1016/0003-9861(58)90400-4.

[58] J.W. Hastings, Biological diversity, chemical mechanisms, and the evolutionary origins of bioluminescent systems, J Mol Evol. 19 (1983) 309–321. https://doi.org/10.1007/BF02101634.

[59] R. Tinikul, P. Chunthaboon, J. Phonbuppha, T. Paladkong, Bacterial luciferase: Molecular mechanisms and applications, Enzymes. 47 (2020) 427–455. https://doi.org/10.1016/bs.enz.2020.06.001.

[60] S. Bazhenov, U. Novoyatlova, E. Scheglova, V. Fomin, S. Khrulnova, O. Melkina, V. Chistyakov, I. Manukhov, Influence of the luxR Regulatory Gene Dosage and Expression Level on the Sensitivity of the Whole-Cell Biosensor to Acyl-Homoserine Lactone, Biosensors. 11 (2021) 166. https://doi.org/10.3390/bios11060166.

[61] A.G. Kessenikh, U.S. Novoyatlova, S.V. Bazhenov, E.A. Stepanova, S.A. Khrulnova, E.Y. Gnuchikh, V.Y. Kotova, A.A. Kudryavtseva, M.V. Bermeshev, I.V. Manukhov, Constructing of Bacillus subtilis-Based Lux-Biosensors with the Use of Stress-Inducible Promoters, International Journal of Molecular Sciences. 22 (2021) 9571. https://doi.org/10.3390/ijms22179571.

[62] C. Gregor, K.C. Gwosch, S.J. Sahl, S.W. Hell, Strongly enhanced bacterial bioluminescence with the ilux operon for single-cell imaging, Proceedings of the National Academy of Sciences. 115 (2018) 962–967. https://doi.org/10.1073/pnas.1715946115.

[63] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, The Protein Data Bank, Nucleic Acids Res. 28 (2000) 235–242.

[64] A. Gabdulkhakov, I. Gushchin, I. Kapranov, K. Boyko, M. Shevtsov, V. Borshchevskiy, V. Samygina, Russian BAG for Xtallography and BioSAXS, (2021). https://doi.org/10.15151/ESRF-ES-573998198.

[65] A.G. Kikhney, C.R. Borges, D.S. Molodenskiy, C.M. Jeffries, D.I. Svergun, SASBDB: Towards an automatically curated and validated repository for biological scattering data, Protein Science. 29 (2020) 66–75. https://doi.org/10.1002/pro.3731.

Table 1. Crystallographic data collection and refinement statistics.

Data collection	
Space group	P1
Cell dimensions	-
a, *b*, *c* (Å)	51.8, 73.4, 94.0
α, β, γ (°)	88.7, 77.9, 87.1
Wavelength (Å)	0.9763
Resolution (Å)	91.89–2.71 (2.84–2.71) *
R_{merge} (%)	11.0 (31.0) *
R_{pim} (%)	11.0 (31.0) *
<I/σI>	4.0 (2.0) *
$CC_{1/2}$ (%)	98.1 (89.4) *
Completeness (%)	97.0 (96.2) *
Multiplicity	1.8 (1.8) *
Unique reflections	35,543 (4,681) *
Refinement	
Resolution (Å)	91.89–2.71
No. reflections	73,799
R_{work}/R_{free} (%)	19.1/26.9
No. atoms	-
Protein	10897
SO_4 ions	20
Water	165
Average *B* factors (Å^2)	-
Protein	41.3
SO_4 ions	66.1
Water	26.4
R.m.s. deviations	-
Protein bond lengths (Å)	0.001
Protein bond angles (°)	0.33
Ramachandran analysis	-
Favored (%)	97.1
Outliers (%)	0.3

* The data for the highest resolution shell is shown in parenthesis

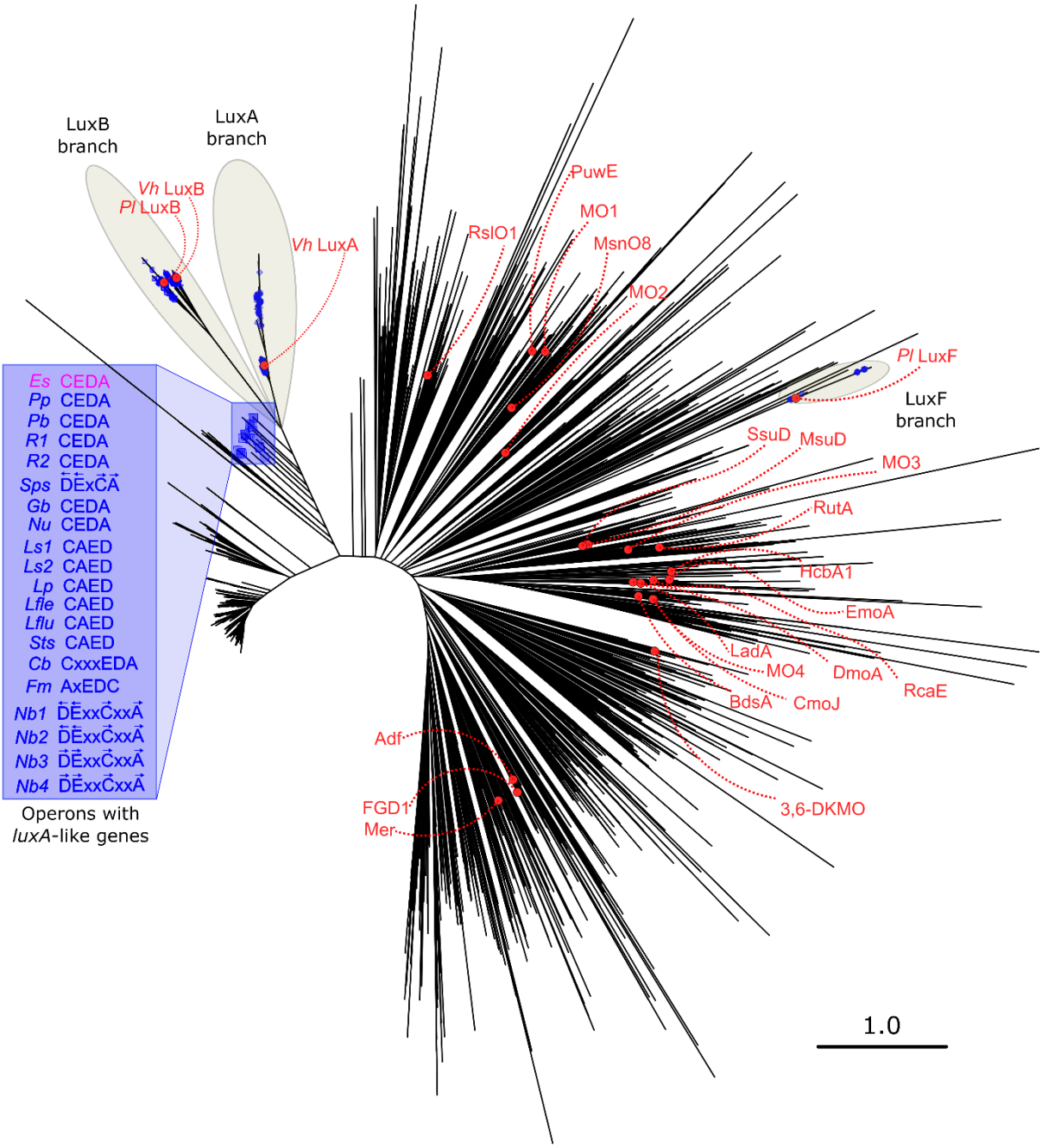

Fig. 1. Phylogenetic tree of *luxA*, *luxB* and *luxA*-like genes. The genes with *luxC*/*luxD*/*luxE* in vicinity are marked in blue. The inset shows the gene order in the putative luciferase operons; details are presented in Table S3. Proteins with available experimentally determined structures are labeled in red; details are presented in Table S4.

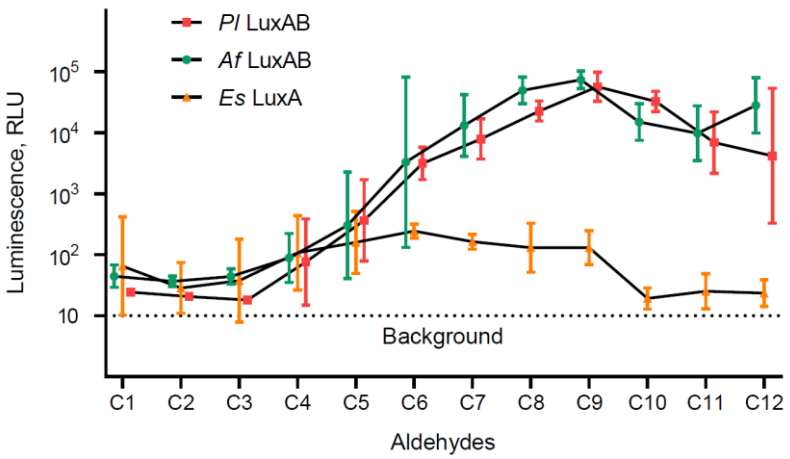

Fig. 2. Dependence of luminescence on the aldehyde aliphatic chain length for *Pl*LuxAB (*E. coli* MG1655 pABX), *Af*LuxAB (*E. coli* MG1655 pF2), and *Es*LuxA (*E. coli* BL21 (DE3) pSluxA). Aldehydes were used at the final concentration of 1 mM. Cells with *Es*LuxA are assayed at 40× optical density of cells with *Pl*LuxAB or *Af*LuxAB because of their low luminescence. Data points and error bars correspond to means and standard deviations of logarithms of luminescence values determined in three independent experiments. RLU is a relative luminescence unit.

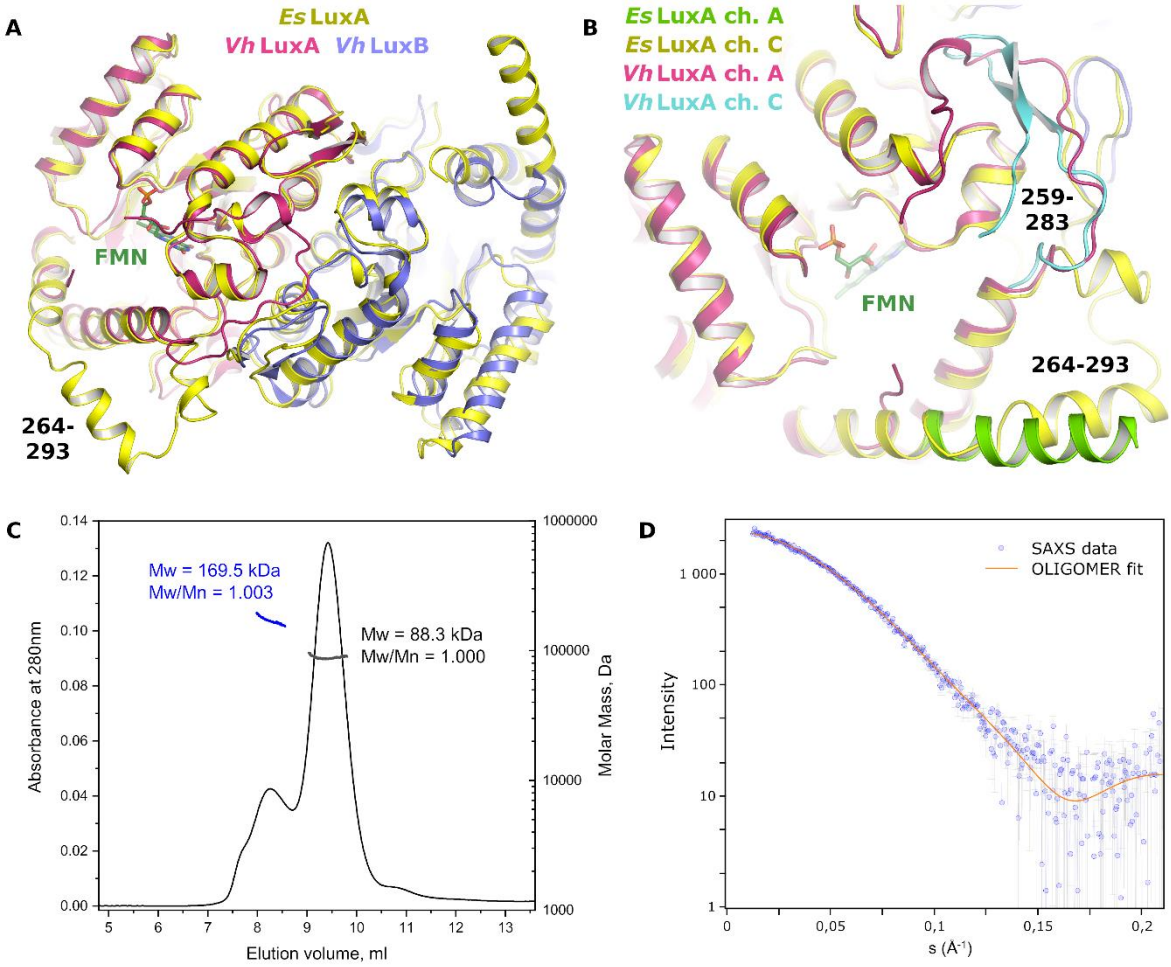

Fig. 3. Structural properties of *Es*LuxA in crystal and in solution. (A) Overlay of the crystallographic structure of homodimeric *Es*LuxA with heterodimeric *Vh*LuxAB (PDB ID 3FGC, [14]). (B) Overlay of the active site of *Es*LuxA with that of *Vh*LuxAB. For both complexes, different molecules in the crystallographic unit cells (chains A and C) show different arrangements of the elongated loop (α-helix and coil in *Es*LuxA, β-hairpin and coil or just coil in *Vh*LuxA). (C) SEC-MALS chromatogram. Two major fractions correspond to *Es*LuxA dimers and tetramers. (D) SAXS data (blue) and theoretical scattering curve (orange) corresponding to a mixture of 82% dimers and 18% tetramers of *Es*LuxA.