Cryo-EM structure of the fatty acid reductase LuxC–LuxE complex provides insights into bacterial bioluminescence

Qingwei Tian[1,2,‡], Jingting Wu[1,‡], Haifeng Xu[3,4,‡], Zhangli Hu[1,2], Yangao Huo[3,*], and Liyan Wang[1,*]

From the [1]Shenzhen Key Laboratory of Marine Bioresource and Eco-Environmental Science, College of Life Sciences and Oceanography, and [2]Key Laboratory of Optoelectronic Devices and Systems, College of Physics and Optoelectronic Engineering, Shenzhen University, Shenzhen, China; [3]National Laboratory of Macromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing, China; [4]College of Life Sciences, University of Chinese Academy of Sciences, Beijing, China

Edited by Ruma Banerjee

The discovery of reduced flavin mononucleotide and fatty aldehydes as essential factors of light emission facilitated study of bacterial luminescence. Although the molecular mechanisms underlying bacterial luminescence have been studied for more than 60 years, the structure of the bacterial fatty acid reductase complex remains unclear. Here, we report the cryo-EM structure of the *Photobacterium phosphoreum* fatty acid reductase complex LuxC–LuxE to a resolution of 2.79 Å. We show that the active site Lys238/Arg355 pair of LuxE is >30 Å from the active site Cys296 of LuxC, implying that catalysis relies on a large conformational change. Furthermore, mutagenesis and biochemical experiments support that the L-shaped cleft inside LuxC plays an important role in substrate binding and reaction. We obtained a series of mutants with significantly improved activity as measured by *in vitro* bioluminescence assays and demonstrated that the double mutant W111A/F483K displayed the highest activity (370% of the WT). Our results indicated that the activity of LuxC significantly affects the bacterial bioluminescence reaction. Finally, we expressed this mutated *lux* operon in *Escherichia coli* but observed that the *in vivo* concentrations of ATP and NADPH limited the enzyme activity; thus, we conclude that the luminous intensity mainly depends on the level of metabolic energy.

Bioluminescence is a natural phenomenon in which luciferase catalyzes the oxidation of luciferin, the corresponding substrate, to produce photons and emit visible light (1). Luminous organisms are widely distributed in bacteria, mushrooms, dinoflagellates, and various animals. There are more than 30 independent evolutionary bioluminescence systems in nature (2). Presently, 10 different luciferins have been isolated and their structures determined (2, 3), including bacterial luciferin, a reduced type of FMN (FMNH$_2$), which is abundant in all cells (4). The other luciferins are produced by organisms expressing corresponding luciferases, and the biosynthetic pathway of fungal luciferin 3-OH-hispidin was clarified (5).

Luminous bacteria are the most abundant and widely distributed light-emitting organisms that mediate bioluminescence through *lux* gene clusters (6). Although the *lux* operon organization exhibits a broader diversity, the canonical *lux* operons adopt the *CDAB(F)E(G)* gene order, and *luxA*, *B*, *C*, *D*, *E*, and *G* are considered the core genes (Fig. 1) (4, 6). Luciferases, LuxA and LuxB, belong to class C flavin monooxygenases and form heterodimers (7). The LuxAB complex (Protein Data Bank [PDB] code: 3FGC) catalyzes the monooxygenation reaction of long-chain fatty aldehydes (myristic aldehyde is proposed to be the natural substrate) with FMNH$_2$ and O$_2$ to produce the corresponding fatty acids, FMN, and H$_2$O and emit λ = 490 nm visible blue-green light (8, 9). FMNH$_2$ and fatty aldehyde must be continuously regenerated to continue the bioluminescence reaction. LuxG is an NAD(P)H-dependent flavin reductase, which is highly homologous to *Escherichia coli* Fre (PDB code: 1QFJ) and exhibits a similar function (10–12). Some *lux* gene clusters lacked *luxG*, indicating that other genes may perform corresponding functions.

LuxD is a thioesterase responsible for the cleavage of the myristoyl chain from the acyl carrier protein of the fatty acid synthase complex (PDB code: 1THT) (13–15). LuxE is an acyl-protein synthetase, which has <18% sequence identity with acyl-CoA synthetase but lacks the C-terminal subdomain. The substituent is a short C-terminal motif (approximately 20 amino acids) containing a conserved cysteine (Fig. S1). Although LuxE shares similar reaction mechanisms with acyl-CoA synthetase, coenzyme A is not involved in the reaction. After myristic acid reacts with ATP to form an acyl-AMP intermediate, the acyl chain is transferred to the sulfhydryl of cysteine to obtain acyl-LuxE (16, 17). LuxC is an acyl-CoA reductase, about 55 kDa, which has <18% sequence identity with the aldehyde dehydrogenase (ALDH) superfamily and catalyzes the opposite reaction. LuxC also has a conserved cysteine (Fig. S2), which is the covalent binding site of the substrate (18). Acyl-LuxE or acyl-CoA can be used as acyl donors to transfer acyl groups to cysteine to obtain acyl-LuxC. Subsequently, acyl-LuxC uses NADPH as a cosubstrate to catalyze the thioester bond broken *via* hydrogenation reduction to form the corresponding fatty aldehyde (4, 19). Biochemical studies have shown that LuxE is a monomer in

‡ These authors contributed equally to this work.
* For correspondence: Liyan Wang, lwang@szu.edu.cn; Yangao Huo, huoyangao@moon.ibp.ac.cn.
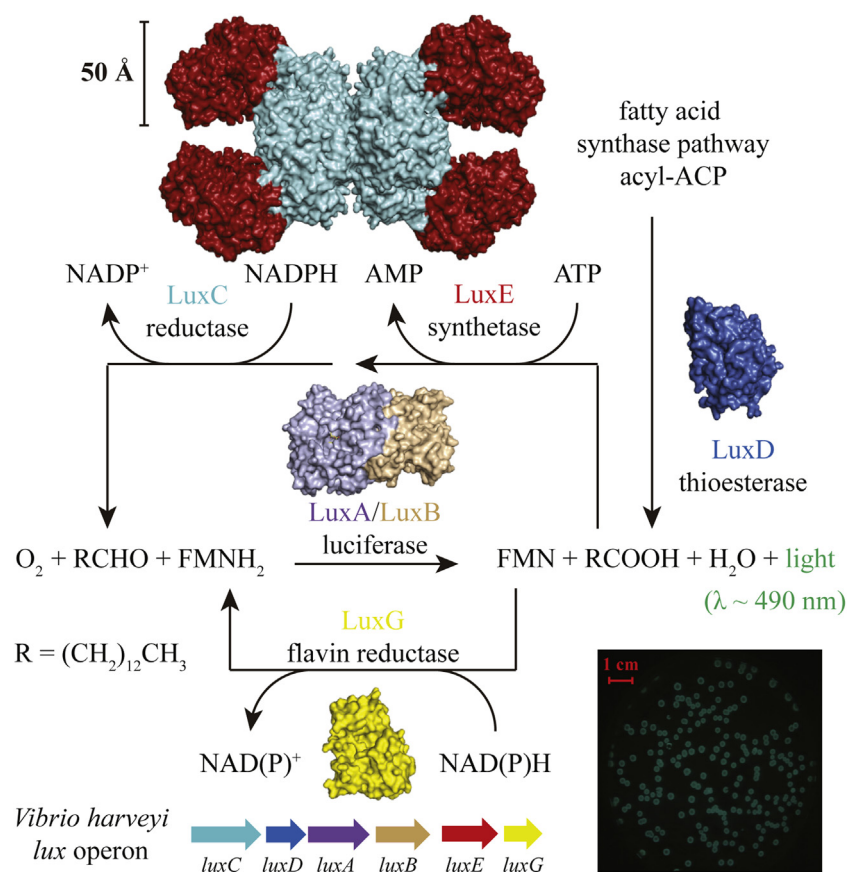
Figure 1. Molecular mechanism of bacterial bioluminescence. The heterodimeric luciferase LuxAB (shown in *light blue*/wheat, PDB code: 3FGC) performs the oxidation of myristic aldehyde to the myristic acid accompanied by light emission. The required $FMNH_2$ is provided by a NAD(P)H-dependent FMN reductase LuxG (structure of the closely related enzyme Fre is shown in *yellow*, PDB code: 1QFJ). The myristic aldehyde is synthesized through the fatty acid reductase LuxCE (structure of LuxE is predicted by the alpha-fold server and shown in *red*; structure of LuxC is resolved in this study and shown in *cyan*). The myristic acid is hydrolyzed from the fatty acid synthase by thioesterase LuxD (shown in *marine*, PDB code: 1THT). The photograph of luminous *Vibrio harveyi* and the *lux* gene order is shown at the bottom. PDB, Protein Data Bank.

solution rather than a dimer like acyl-CoA synthase. Although LuxC and LuxD have no direct interaction, they directly interact with LuxE, suggesting that they may form a LuxDEC fatty acid reductase complex (20).

Bioluminescent genes are important molecular tools because they are easily detected. For example, luciferases of fireflies and jellyfish are widely used in biological research as reporter genes (21). Furthermore, the luciferase in fireflies or jellyfish is expressed by one gene, but the substrate should be made available to the cell by adding it to the medium. However, as reporter genes, the bacterial *lux* gene cluster must simultaneously express five genes (luxCDABE), which is relatively difficult for genetic manipulation (22–25). However, its significant advantage is that it can emit light sustainably without adding a substrate, which is convenient for continuous detection. While bioluminescent bacteria are often used as biosensors in ecotoxicological studies, the *lux* operon has been used to produce autoluminescent organisms such as human cell lines and tobacco plants (23, 26, 27). Compared with fluorescence, bioluminescence imaging provides additional benefits. There is no background because of the lack of autofluorescence, no excitation light source and filters are required, and no phototoxicity or bleaching occurs (28). However, its main limitation is its poor brightness. This study

aimed to elucidate the molecular mechanism of fatty acid reductase complex by determining the cryo-EM structure of *Photobacterium phosphoreum* LuxC–LuxE and guide mutagenesis to enhance its activity.

Results

Reconstruction of the bacterial bioluminescence system in vitro

We cloned *the luxCDABEG* DNA segment from *Vibrio harveyi*, constructed two expression vectors, pACYCDuet1-LuxCDA and pRSFDuet1-LuxBEG, and cotransferred them into *E. coli* BL21(DE3) to obtain a luminous strain (Fig. 2A). The growth and luminous intensity curves were continuously detected at 30 °C (Fig. 2B). Furthermore, pET24a-VhLuxC and pET22b-VhLuxE-His with a C-terminal hexahistidine tag were constructed and cotransferred to BL21(DE3) to express *V. harveyi* LuxC–LuxE complex (termed VhLuxCE). However, we found that the VhLuxCE was easy to precipitate. Therefore, we tried homologous genes from other species and obtained the *P. phosphoreum lux*C and *lux*E genes through chemical synthesis. After similar construction and expression, the soluble PpLuxCE was obtained by nickel-nitrilotriacetic acid (Ni-NTA) affinity column chromatography and size-exclusion
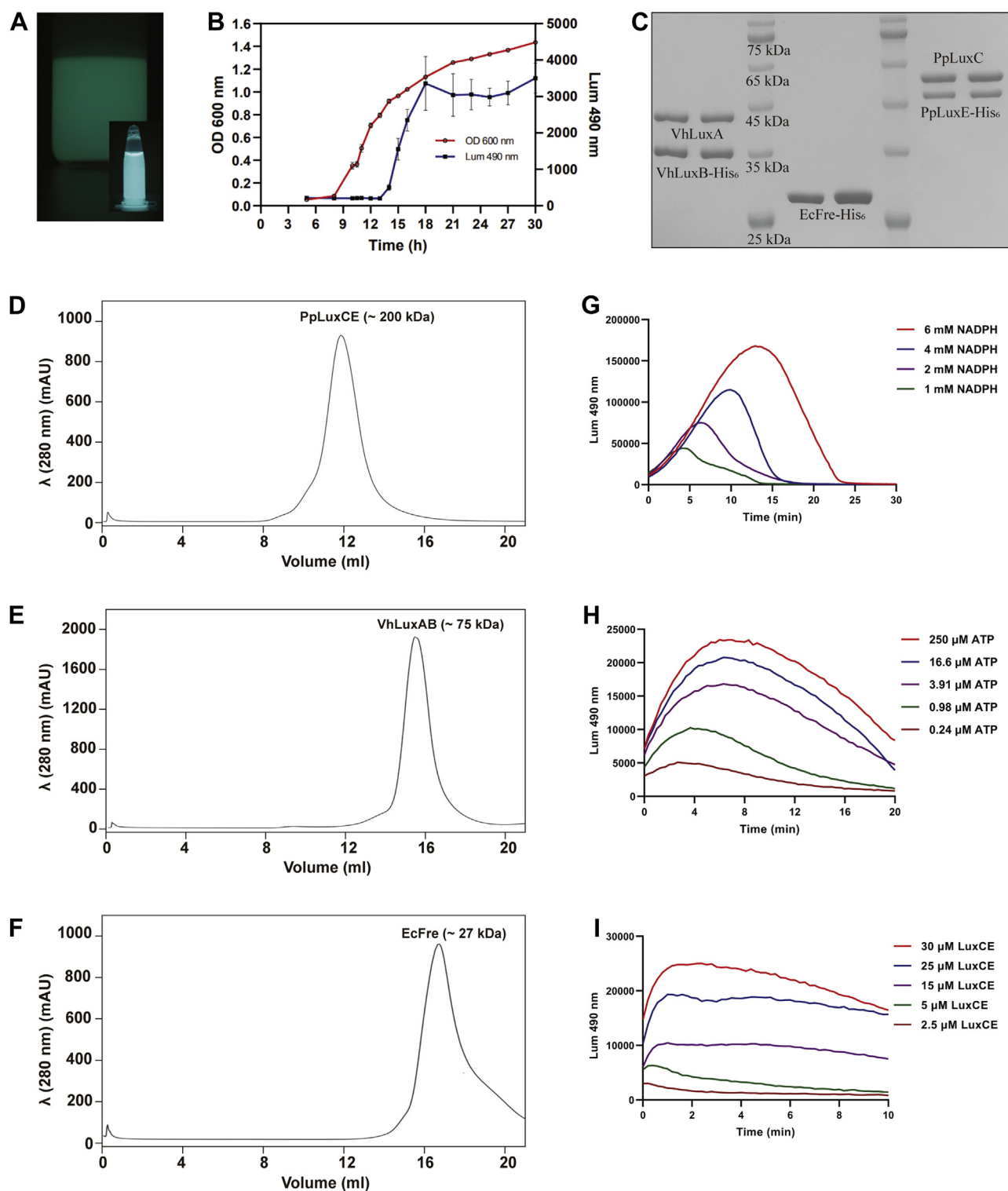
Figure 2. *In vitro* reconstruction of the bacterial bioluminescence. *A*, the photograph of luminous *E. coli* BL21(DE3) *via* recombinant expression of the *V. harveyi luxCDABEG*. The photograph of *in vitro* bioluminescence reaction is shown in the *inset*. *B*, the growth curve and bioluminescence intensity curve of luminous *E. coli* cultured at 30 °C. *C*, SDS-PAGE analysis of the purified recombinant proteins VhLuxAB, EcFre, and PpLuxCE. *D–F*, gel-filtration profiles of the recombinant PpLuxCE, VhLuxAB, and EcFre. *G–I*, the intensity curves of the *in vitro* bioluminescence reaction under different concentrations of NADPH, ATP, and LuxCE. Default concentrations: 1 mM NADPH, 1 mM ATP, 100 μM LuxCE.

column chromatography. SDS-PAGE showed that the ratio of LuxC to LuxE was about 1:1 (Fig. 2, *C* and *D*). The VhLuxAB complex and EcFre were expressed and purified to verify the activity of LuxCE *in vitro* (Fig. 2, *C*, *E* and *F*). When the components were mixed with myristic acid and FMN, a bioluminescence reaction occurred after adding ATP and NADPH (Fig. 2*A* insert). The reaction curves for different NADPH, ATP, and LuxCE concentrations were determined.

The luminous intensity was positively correlated with the concentrations of NADPH, ATP, or LuxCE (Fig. 2,*G*, *H* and *I*). The results indicated that the purified PpLuxCE complex exhibited good activity *in vitro*, and its relative enzyme activity could be detected by an *in vitro* bioluminescence system.

The overall structure of *P. phosphoreum* *LuxC–LuxE complex*

PpLuxCE samples purified in different conditions were screened by negative-staining EM. Protein purification in optimized conditions yielded homogenous particles that were suitable for analysis by cryo-EM. Representative 2D class average indicated an assembly of D2 symmetry, consistent with the previously proposed tetrameric organization of LuxC (4, 20). Furthermore, 229,033 selected particles yielded a 3D reconstruction with an overall resolution of 2.79 Å for the LuxC tetramer complex with a portion of LuxE (Table 1, Figs. 3*A* and S3–S5). In each LuxC protomer, 477 residues (12–488) could be resolved and were constructed from 18 α helices (α1–α18) and 14 β strands (βa–βn) (Fig. 3*C*). LuxC adopted the canonical aldehyde dehydrogenase fold (29), including an N-terminal cofactor-binding domain (residues 12–133 and 152–263), a catalytic domain (residues 264–453), and an oligomerization domain (residues 134–151 and 454–488) (Fig. 3*B*). The cofactor-binding domain consists of alternating α/β secondary structural elements that give rise to a five-stranded parallel β sheet flanked by a two-stranded antiparallel β sheet and 11 α helices. The catalytic domain consists of alternating α/β secondary structural elements that give rise to a five-stranded mixed parallel/antiparallel β sheet flanked by five α helices. The oligomerization domain consists of the remaining two α helices and a two-stranded antiparallel β sheet.

However, the density of LuxE was weak, implying that LuxE is dynamic. Therefore, only 118 of the 373 residues of LuxE (residues 33–91, 290–338, and 348–357) could be resolved. To form the interface between LuxC and LuxE, 927 Å² of the surface area was buried (Fig. 4*A*). Nonpolar residues comprise ~60% of the interface and are accompanied by six hydrogen bonds and one salt bridge (LuxC-D335 and LuxE-R354) (Fig. 4*B*). We predicted the complete structure of LuxE using an alpha-fold server and superimposed it on our cryo-EM structure (RMSD is 1.2 Å on 118 α carbons) (Fig. 4*C*). Thus, we simulated the model of the LuxC₄–LuxE₄ complex, which is in line with the 1:1 heterodimer determined by biochemical experiments (Fig. 4*D*) (20).

The catalytic domain of LuxC contained the key cysteine (Cys296), located at the entry of a deep cleft formed between the cofactor-binding domain and the catalytic domain (Fig. 3*B*). According to the NAD-binding funnel of ALDHs (discussed later), the broad groove outside the cleft is speculated to play an important role in cosubstrate binding (Fig. 3*F*). Autodock 4.2 software (https://autodock.scripps.edu/) was used to simulate the binding of NADPH (30). Among the molecular docking results, an ideal model showed that an NADPH binds in the NADPH-binding groove formed by the cofactor-binding domain, especially the C4 of nicotinamide moiety that was located 3.6 Å away from the sulfhydryl of Cys296 (Fig. 3, *D* and *E*). Although the predicted result is a reasonable model for reaction, the actual binding site can be clarified by resolving the structure of LuxC complex with NADPH. A distinctive structural feature of LuxC is the presence of a novel structural element (P460–P488), represented by the oligomerization domain. This short segment from a neighboring subunit extends over the cleft to form a long L-shaped hydrophobic cleft (Fig. 5*A*).

The active site of LuxE and the fatty acyl transferring

LuxE is evolutionarily related to the adenylate-forming enzyme family, which includes enzymes involved in the metabolism of fatty acids and aromatic compounds. The adenylate-forming enzyme family adopts a general architecture of two α/β domains with the active site at the interface of the N-terminal domain and the C-terminal subdomain, and catalysis relies on a large conformational change of the subdomain. Intriguingly, LuxE lacks this subdomain. We tried to solve its active site by comparing it to the structure of a well-studied phenylacetic-CoA ligase PaaK1 (PDB code: 2Y27) (16). The PaaK1/ATP costructure exhibits the preadenylation state, in which the catalytically essential Lys422/Arg326 pair orients the phosphate for nucleophilic attack. Superimposition of the predicted structure of LuxE implies a similar function for the Lys238/Arg355 pair in LuxE (Fig. 6). While Lys422 of PaaK1 is located at the subdomain, the conserved residue Lys238 of LuxE is located at a loop (²³⁴GGG WKTKQKQALN²⁴⁶) longer than the corresponding loop of

Table 1

Summary of cryo-EM data collection, processing, and structure refinement

| LuxC–LuxE (EMDB-33113, PDB 7XC6)	
Data collection and processing	
Magnification	29,000×
Voltage (kV)	300
Electron exposure (e−/Å²)	50
Defocus range (μm)	0.5–2.3
Pixel size (Å)	0.833
Symmetry imposed	C1
Initial particle images (no.)	1,012,247
Final particle images (no.)	229,033
Map resolution (Å)	2.79
FSC threshold	0.143
Map resolution range (Å)	2.7–6.0
Refinement	
Initial model used (PDB code)	No
Model resolution (Å)	2.79
FSC threshold	0.143
Map sharpening *B* factor (Å²)	–76.4
Model composition	
Nonhydrogen atoms	16,176
Protein residues	2026
Ligands	0
B factors (mean, Å²)	46.72
RMSD	
Bond lengths (Å)	0.008
Bond angles (°)	0.781
Validation	
MolProbity score	2.06
Clash score	7.82
Poor rotamers (%)	1.55
Ramachandran plot	
Favored (%)	92.99
Allowed (%)	7.01
Disallowed (%)	0.00

Abbreviations: FSC, Fourier shell correlation.

Figure 3. Cryo-EM structure of *P. phosphoreum* LuxC–LuxE complex. *A*, the overall structure of the LuxCE is shown as *cartoon*. Four subunits of LuxC tetramer are colored in *cyan*, *green-cyan*, *magenta*, and *lemon*; LuxE is colored in *orange*. *B*, LuxC–A is shown as *cartoon*, LuxC–B is shown as *surface*. Each protomer includes an N-terminal cofactor-binding domain (*cyan/green-cyan*), a catalytic domain (*pale cyan/pale green*), and an oligomerization domain (*pink/wheat*). The catalytic Cys296 (*red*) is shown as *ball* and *stick*. *C*, a schema of LuxC protomer with the strands and helices labeled and colored in *cyan* and *red*, respectively. *D*, molecular docking shows the NADPH (*yellow*) binding site of LuxC. *E*, the magnified region shows the details of the NADPH binding site. *F*, the surface of the cofactor-binding domain (*cyan*) and catalytic domain (*pale cyan*) of LuxC–A, and the oligomerization domain of LuxC-B (*wheat*).

PaaK1. In addition, the proposed active site Arg355 is stabilized by the salt bridge formed between LuxE-R354 and LuxC-D335. The reaction activities of mutants K238A, R354A, and R355A were measured using the aforementioned *in vitro* bioluminescence system, and their relative activities were characterized by the maximum luminous intensity. While K238A was completely inactivated, the reaction activity of R354A (34%) and R355A (10%) decreased significantly. These results support the theory that the Lys238/Arg355 pair of LuxE plays a critical role in the adenylation (myristic acid + ATP → myristoyl adenylate + pyrophosphate) and thioesterification (myristoyl adenylate + Cys364 → myristoyl-Cys364 + AMP) reactions.

However, the side chain of Arg355 of LuxE placed >30 Å away from the Cys296 of LuxC, indicating that the fatty acyl chain transferring between these two active sites relies on a conformational change of the C-terminal of LuxE (Fig. 4*E*). Although the conserved C-terminal motif (^{358}TRxxKGCxLS^{367}) containing Cys364 is not resolved in our structure, the structural prediction shows it as a flexible loop, which is consistent with the proposed functional role for the conformational change. Thus, we simulated two functional conformations and displayed them in Figure 4, *E* and *F*: (I) thioesterificaion state, LuxE-C364 engage into the active stie (Lys238/Arg355) of LuxE; (II) acyl-transferring state, the C-terminal of LuxE moves to the LuxE-binding groove formed by
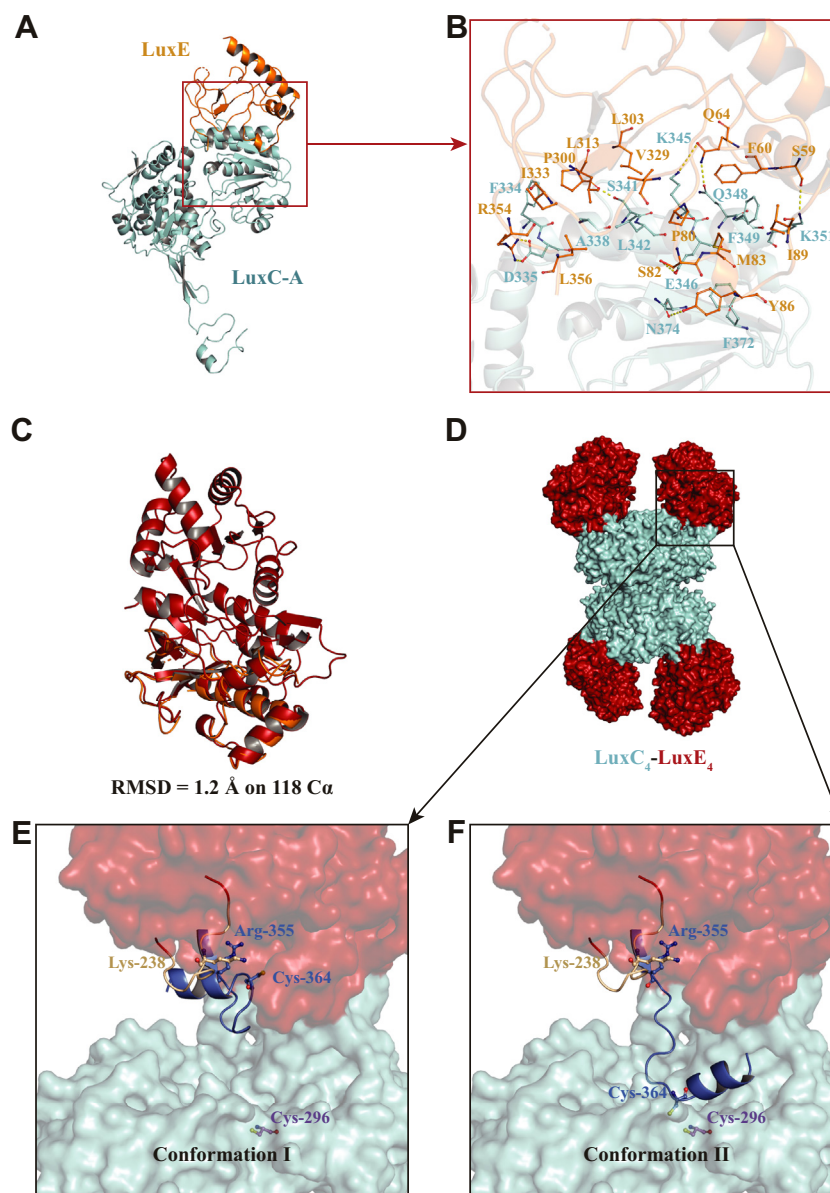
Figure 4. The interaction between LuxE and LuxC. *A*, LuxC–A (*cyan*) and LuxE (*orange*) are shown as *cartoon*. *B*, the magnified region shows the details of the interface between LuxE and LuxC–A. *C*, superimposition between the predicted structure (*red*) and cryo-EM structure (*orange*) of LuxE. *D*, overall structure of the simulated LuxC$_4$–LuxE$_4$ complex (*cyan/red*) is shown as surface. *E* and *F*, the magnified regions show the conformational change between thioesterificaion state (I) and fatty acyl transferring state (II). The residues 234 to 246 (*wheat*) and 356 to 373 (*marine*) of LuxE are shown as illustrations. Key residues are labeled and shown as *ball* and *stick*.
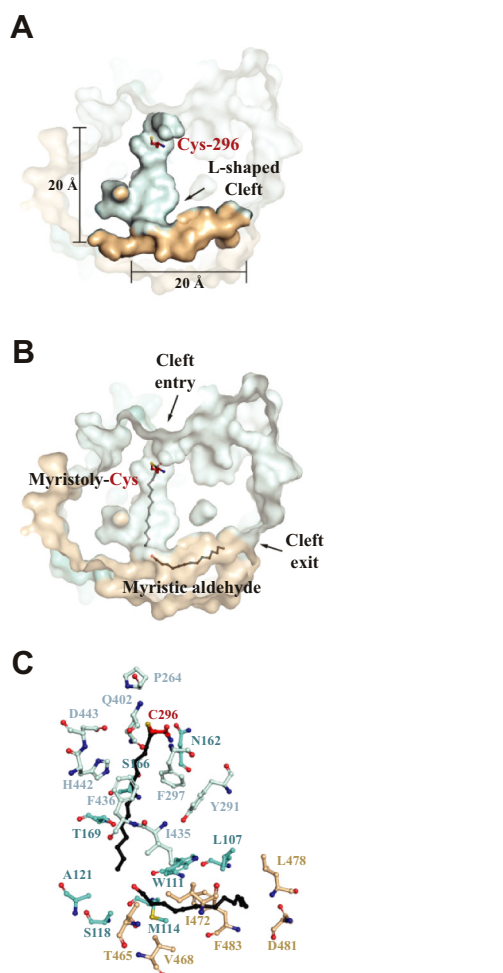
the catalytic domain of LuxC, and the LuxE-C364 engage into the active site (Cys296) of LuxC.

L-shaped cleft inside LuxC

The L-shaped cleft was formed by three domains, and the length of the cleft was more than 40 Å (Fig. 5A). It is long enough to accommodate a substrate myristoyl covalently connected to Cys296 and a product myristic aldehyde (Fig. 5B). The cleft was mainly composed of hydrophobic residues that match the structures of the substrate and the product. Most of these residues are highly conserved. Thus, we hypothesized that in the initial of the reaction, the acyl group of acyl-LuxE (or acyl-CoA) enters the cleft from the entry and

transfers to the Cys296 of LuxC. When NADPH is combined in the NADPH-binding groove, its nicotinamide moiety is close to acyl-Cys to cleave the thioester bond and generate fatty aldehydes. We hypothesized that the product would move to the bottom along the cleft and release from the exit. Changing the shape or polarity of the cleft *via* mutagenesis may affect the reaction process, thus affecting enzyme activity. Further study is needed to clarify this hypothesis, especially regarding the structures of the LuxC complex with substrate and/or product.
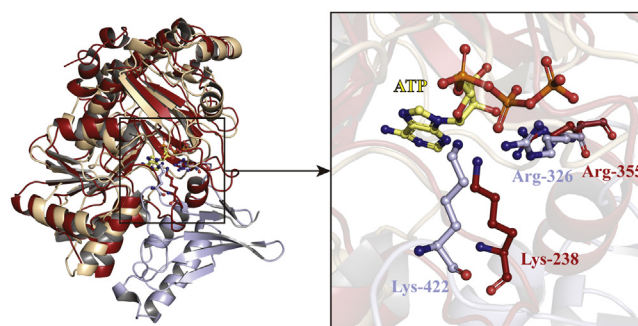
We selected some conserved residues constituting the cleft and constructed a series of mutants to test the hypothesis (Figs. 5C and S2). The reaction activities of different mutants were measured at the same protein concentration using the

Figure 5. Structure of the L-shaped cleft inside LuxC. *A*, cross-section through the L-shaped cleft formed by the cofactor-binding domain (*cyan*) and catalytic domain (*pale cyan*) of LuxC-A, and the oligomerization domain of LuxC-B (*wheat*). The length of the cleft is marked. The catalytic Cys296 (*red*) is located near the cleft entry. *B*, the substrate myristoyl is covalently connected to Cys296, and the product myristic aldehyde is placed at the bottom of the L-shaped cleft. *C*, a series of conserved residues constituting the L-shaped cleft selected for the mutagenesis study.

Figure 6. Structural comparison between predicted LuxE and phenylacetic-CoA ligase PaaK1 (PDB code: 2Y27). LuxE (*red*) is predicted by alpha-fold server. N-terminal domain and C-terminal subdomain of PaaK1 are colored in *wheat* and *light blue*, respectively. The magnified region shows the details of the active site. The critical residues and ATP (*yellow*) are shown as *ball* and *stick*. PDB, Protein Data Bank.

aforementioned *in vitro* bioluminescence system, and their relative activities were characterized by the maximum luminous intensity (Fig. 7*A*). We found that the enzyme activity of 10 mutants was almost stopped (<10%). C296A prevents the substrate from covalently binding to LuxC. S166F, F297A, Q402E, Q402L, F436E, H442R, and D443A were located near the cleft entry site. S166F and H442R increased the side-chain volume, potentially hindering substrate binding. The remaining mutants did not increase the volume of their side chains; however, they showed significantly reduced activities. In addition, L478W and F483W were located near the exit of the cleft, implying that they may block the exit and hinder product release.

In contrast, the activity of the 11 mutants was significantly higher (>120%) than that of the WT: L107F, L107W, W111A, W111Q, M114W, S118F, A121Q, T169F, Y291F, I472Y, and F483K. The residues Leu107, Trp111, Met114, Ile472, and Phe483 were far from the cleft entry and closer to the exit.

Changes in the side chains may promote the release of the product and accelerate the turnover of the enzyme. We selected W111A (220%), I472Y (240%), and F483K (250%) single mutations with the highest activity to form three double mutants (Fig. 8). W111A/F483K exhibited the highest activity, reaching 370% of the WT. Although a systematic kinetic analysis was not performed in this study, the results demonstrate that the activity of LuxC affected the bioluminescence system significantly, indicating that the activity of WT LuxC has great potential for enhancement.

We constructed identical mutants of VhLuxC based on pACYCDuet1-LuxCDA to improve the luminous intensity *in vivo* and cotransferred them with pRSFDuet1-LuxBEG into *E. coli* BL21(DE3). The strains were cultured and induced under the same conditions. The luminous intensity of each mutant relative to the WT was characterized by its maximum value (Fig. 7*B*). The luminous intensity of all mutants did not exceed the WT, and the data deviation between multiple repetitions was large. Moreover, the growth rate of bacteria and the time of maximum luminous intensity fluctuated significantly. Although bioluminescence requires continuous consumption of ATP and NADPH as energy sources, the excessive synthesis of fatty aldehydes may cause cell damage and inhibit cell growth. However, the cell density and metabolic energy limit the luminous intensity and complex effects disturb measurement accuracy. After repeated measurements, the enzyme activity of W111Q, S118F, S166F, C296A, Q402A, Q402L, H442R, T465F, L478W, and F483K was abolished. Comparisons with the inactivated mutants of PpLuxC, S166F, C296A, Q402L, H442R, and L478W, were consistent. However, the activities of the other PpLuxC mutants were significantly higher than that of the WT, while the corresponding VhLuxC mutants were completely inactivated. Thus, although the sequence identity between the two LuxC is up to 60%, the structural details of the same mutation may produce very different results.

Structural comparison between LuxC and ALDHs

The sequence identity of LuxC to ALDHs is rather low (<18%), and they catalyze the opposite reactions. LuxC is
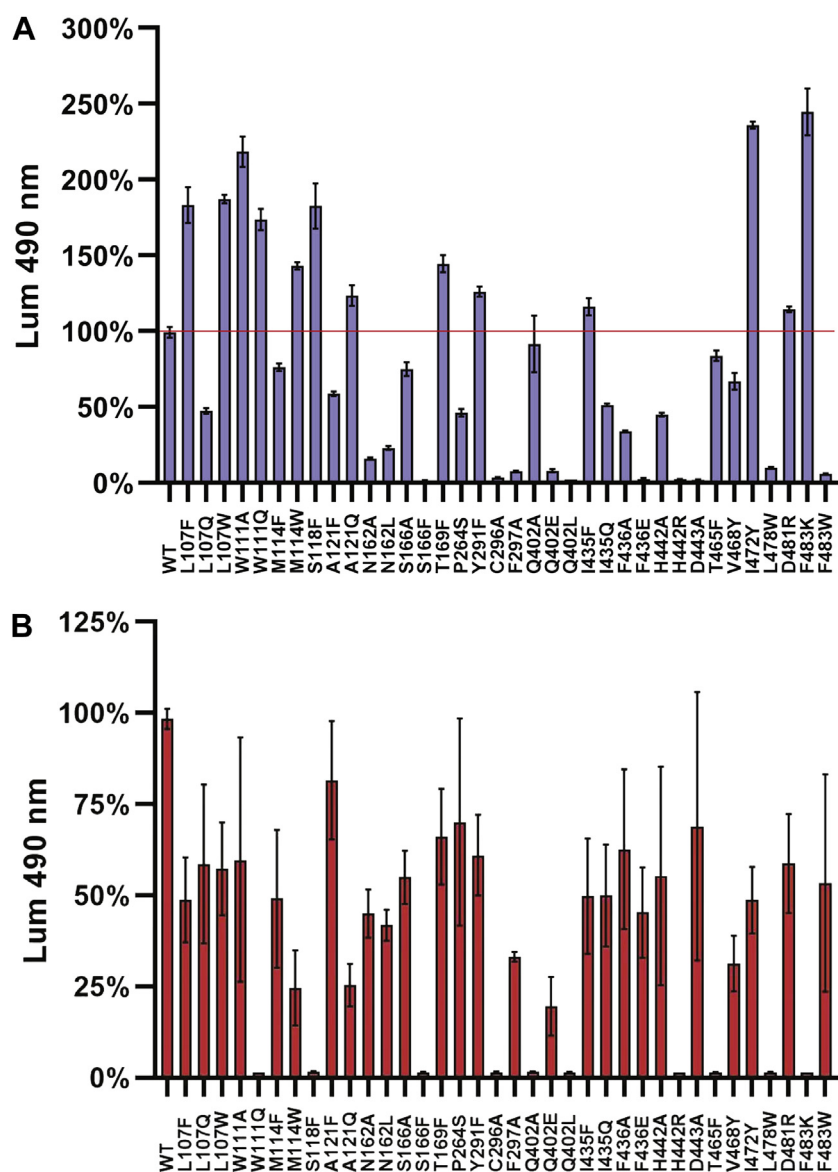
Figure 7. Activity assay of LuxC. *A,* the activities of WT and mutants PpLuxC are assayed *in vitro*. The *red line* indicates the maximum luminous intensity of *in vitro* bioluminescence reaction involving the WT PpLuxC, specified as 100%. *B,* the activities of WT and mutants VhLuxC are assayed *in vivo* (*E. coli* BL21(DE3)). The maximum luminous intensity of the strain containing WT VhLuxC is specified as 100%.
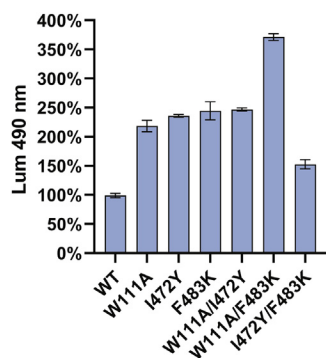
Figure 8. Activity assay of double mutation of PpLuxC. The activities of WT and mutants PpLuxC are assayed *in vitro*. The maximum luminous intensity of *in vitro* bioluminescence reaction involving the WT PpLuxC is specified as 100%.

structurally similar to the human cytosolic ALDH1A1 (PDB code: 4WB9) and the human membrane-associated long-chain fatty aldehyde dehydrogenase (FALDH) (PDB code: 4QGK) (29, 31). Superimposition showed that their cofactor-binding and catalytic domains share a similar architecture but their oligomerization domains differed (Fig. 9*A*). Comparisons of their internal clefts and active centers showed that their catalytic cysteines were in almost the same position (Fig. 7*B*). In ALDHs, cysteine divides the cleft into NAD-binding funnel and substrate-binding funnel (Fig. 9, *C* and *D*). NAD-binding sites are highly conserved in the superfamily of ALDHs (29), but the corresponding residues of LuxC are different, which is consistent with the LuxC usage of NADPH but not NAD as a cosubstrate. The reaction mechanism for ALDHs is based on the nucleophilic attack of the aldehyde substrate by cysteine.
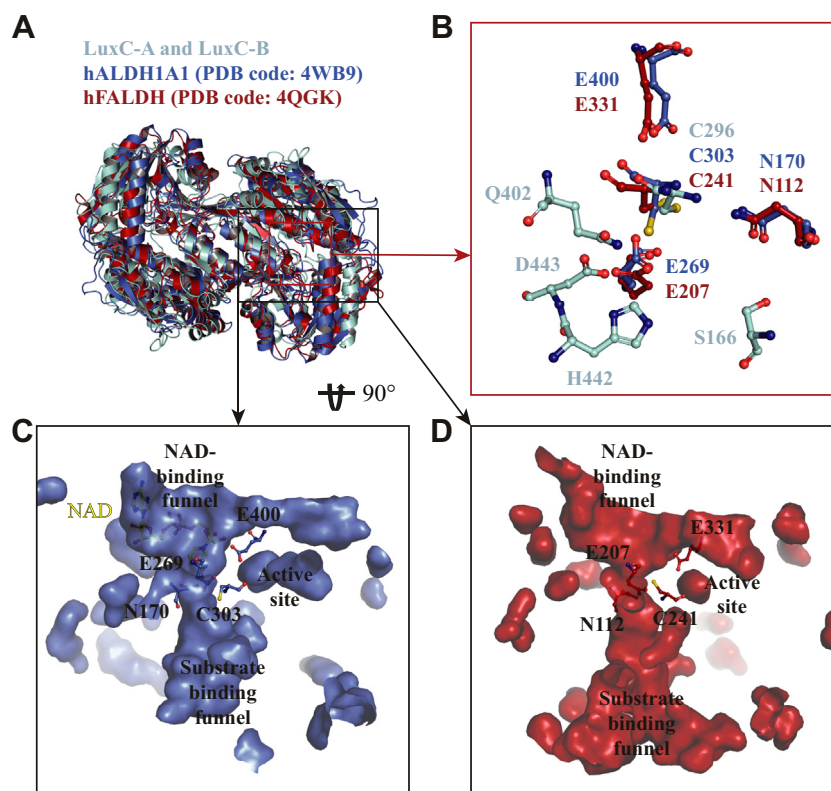
Figure 9. Structural comparison between LuxC and ALDHs. *A*, the overall architectures of LuxC (*cyan*), human ALDH1A1 (*blue*, PDB code: 4WB9), and human FALDH (*red*, PDB code: 4QGK) are similar. *B*, the magnified region shows the details of the active sites of ALDH1A1 (*blue*) and FALDH (*red*), and the polar residues around the Cys296 of LuxC (*cyan*). *C*, cross-section through the inside cleft of hALDH1A1 complex with NAD (*yellow*). *D*, cross-section through the inside cleft of hFALDH. ALDH, aldehyde dehydrogenase; PDB, Protein Data Bank.

First, cysteine is deprotonated and initiates a nucleophilic attack on the carbonyl carbon of the aldehyde and forms a thioester. Subsequently, a water molecule is deprotonated and triggers a nucleophilic attack of the hydroxide anion on the thioester to release the carboxylic acid product. Three conserved residues around cysteine (two glutamic acids and one asparagine) were involved in the catalytic process (29). While these key residues are lacking in LuxC, there are four polar residues (Ser166, Asn402, His442, and Asp443) around Cys296 (Fig. 9*B*). Mutagenesis analysis showed that the S166A, Q402A, and H442A mutants of PpLuxC and the S166A, H442A, and D443A mutants of VhLuxC exhibited normal activity. Therefore, the four polar side chains were not necessary for the reaction. These results revealed the main differences between the active centers of LuxC and ALDHs. They explained why acyl-LuxC does not catalyze the hydrolysis of thioester and generates fatty acid. Furthermore, the L-shaped cleft may promote the product aldehyde to leave the active center and prevent the reverse reaction.

Discussion

Although several studies have confirmed the interaction between LuxC and LuxE, the mechanism for the fatty acyl transferring between two critical cysteines remains a major challenge for mechanistic and structural biochemists since high-resolution data on LuxCE complex remain lacking. We

determined the cryo-EM structure of the fatty acid reductase LuxCE from luminous bacteria for the first time. Even though the structure of LuxE is fragmentary, we clearly observed the interface with LuxC and determined the coordinates of the proposed active site Arg355. As a flexible loop predicted by alpha-fold, the conserved C-terminal motif of LuxE may undergo a large conformational change to transport the covalently connected substrate. We verified that the L-shaped cleft inside LuxC is an important structural basis for substrate binding and reaction by combining mutagenesis and activity assays. The fatty acyl chain accesses the active site from the L-shaped cleft entry and transfers to the Cys296 of LuxC. The cosubstrate NADPH plays a direct role in the thioester bond cleavage and the production of fatty aldehyde. However, all polar residues around the Cys296 of LuxC are not directly involved in catalysis. Therefore, the mechanistic details for acyl transfer and the reduction reaction remain unclear.

Bacterial bioluminescence consumes a large amount of energy. LuxE and LuxC consume ATP and NADPH to convert fatty acids into fatty aldehydes, a rare biochemical reaction. ALDHs transfer the hydride from free aldehydes to NAD to produce NADH and transfer hydroxyl from a water molecule to thiohemiacetal to release the corresponding carboxylic acids. However, the bacterial luciferase catalyzes the oxidation of free aldehydes to fatty acids *via* the FMN-4a-OOH flavoprotein mechanism, which is different from ALDHs (32). In the process, $FMNH_2$ is converted into an electronically excited

FMN state, emitting a photon on return to the ground state, with visible light emitted. Most luminous organisms separate the synthesis of luciferin from the bioluminescence reaction, avoiding excessive energy consumption and improving brightness ([2](#)). Fatty acid reductase and flavin reductase must work with luciferase at the same time because fatty aldehyde and $FMNH_2$ cannot accumulate in the bacterial bioluminescence system. We suggest that luminous intensity depends on metabolic energy rather than enzyme activity *in vivo*. The reaction directly consumes ATP and NADPH, direct energy sources used in almost all anabolisms. Thus, the *lux* system can be used as a general reporting system to characterize the level of metabolic energy in synthetic biology, such as screening high-performance chassis cells, exploring new gene expression regulatory elements, and optimizing cell culture schemes. In addition, enhanced LuxC *via* structure-guided mutagenesis exhibits a significant impact on the bacterial bioluminescence *in vitro*, which expands the application for future studies.

Experimental procedures

Cloning and site-directed mutagenesis

The gene encoding LuxC (GenBank accession number: BAF40855.1) from the *P. phosphoreum* strain was chemically synthesized and inserted into the pET24a vector using the restriction enzymes *Nde*I and *Xho*I to obtain pET24a-LuxC. The gene encoding LuxE from the *P. phosphoreum* strain was chemically synthesized and inserted into the pET22b vector using the restriction enzymes *Nde*I and *Xho*I to obtain pET22b-LuxE-His. The vector allowed the expression of LuxE in *E. coli* to fuse with a C-terminal hexahistidine tag. DNA, including *lux*CDABEG, was obtained from *V. harveyi* by PCR. Fragments of *lux*CDA and *lux*BEG were cloned into pACYCDuet1 and pRSFDuet1 vectors to obtain pACYC-Duet1-LuxCDA and pRSFDuet1-LuxBEG, respectively. In addition, the *lux*AB fragment was subcloned into the pET28a vector with *EcoRI/XhoI* to obtain pET28a-LuxAB-His. The vector allowed the expression of LuxB to fuse with a C-terminal hexahistidine tag. The *fre* gene was cloned from *E. coli* K12 by PCR and inserted into pET28a with *EcoRI/XhoI* to obtain pET28a-Fre-His.

Site-directed mutagenesis of LuxC in the pET24a-LuxC and pACYCDuet1-LuxCDA vectors was performed using appropriate primers. Mutants L107F, L107Q, L107W, W111A, W111Q, M114F, M114W, S116A, S118F, A121F, A121Q, N162A, N162L, S166F, T169F, P264S, Y291F, C296A, F297A, Q402A, F436A, F436E, H442A, H442R, D443A, T465F, V468Y, I472Y, L478W, D481R, F483K, and F483W were generated by site-directed mutagenesis. The correct integration of all constructions was verified using sequence analysis.

Protein expression and purification

First, pET24a-LuxC (WT or mutant) and pET22b-LuxE-His plasmids were cotransformed into *E. coli* BL21(DE3) cells. Next, cultures were inoculated in 10 ml of LB medium containing 50 μg/ml ampicillin and 50 μg/ml kanamycin

containing the transformed BL21(DE3) cells incubated at 37 °C overnight. The next day, the cultured cells were transferred to 750 ml of LB medium. Cells were grown at 37 °C until the A_{600} nm reached 0.6 to 0.8. Protein expression was induced with 1 mM IPTG. The induced cells were grown at 16 °C for 16 to 18 h. Cells were harvested by centrifugation at 3500 rpm at 4 °C for 10 min and subsequently lysed in buffer A (20 mM imidazole, 300 mM NaCl, 50 mM Tris–HCl, pH 8.0). The lysed cells were sonicated with 4s/4s pulses on an ultrasonic processor and centrifuged at 10,000 rpm at 4 °C for 40 min. The clarified supernatant containing LuxC and LuxE-His was loaded onto a Ni-NTA affinity column pre-equilibrated with buffer A. The His-tag protein was eluted by increasing the imidazole concentration in buffer A from 20 to 300 mM. The protein LuxCE complex was concentrated using an ultrafiltration centrifuge tube (30 kDa, Millipore). The sample was subjected to gel filtration using a Superdex 200 increase column (GE Healthcare) with 20 mM Hepes, pH 7.5, 150 mM NaCl, and 5% glycerol as a buffer. Fractions containing the LuxCE complex were concentrated again in ultrafiltration centrifuge tubes (30 kDa, Millipore), flash-frozen in liquid nitrogen, and subsequently stored at –80 °C at a concentration of 45 mg/ml. The pET28a-LuxAB-His and pET24a-Fre-His were transformed into *E. coli* BL21(DE3) cells. LuxAB and Fre were purified as described for LuxCE. LuxAB was concentrated to 40 mg/ml, and Fre was concentrated to 30 mg/ml. After the protein purity was tested by SDS-PAGE, it was frozen in liquid nitrogen and stored at –80 °C.

VhLuxC activity assay *in vivo*

pACYCDuet1-LuxCDA (WT or mutant) and pRSFDuet1-LuxBEG plasmids were cotransformed into BL21(DE3) to obtain luminescent *E. coli*. The transformants were cultured overnight in 5 ml of LB medium containing 50 μg/ml chloramphenicol and 50 μg/ml kanamycin. The sample was inoculated into 50 ml LB medium and cultured to logarithm at 30 °C, and IPTG (0.5 mM) was added to continue the culture. Luminescence was detected using a FlexStation 3 microplate reader (Molecular Devices) at a detection wavelength of 490 nm. The samples were continuously sampled and detected several times to determine the maximum luminescence intensity of *E. coli*. Each group of experiments was repeated three times, and GraphPad Prism 8 (GraphPad) was used for statistical analysis.

PpLuxC activity assay *in vitro*

PpLuxC (WT or mutant) activity was determined by monitoring the luminescence at 490 nm. A typical experimental sample contained reaction mixtures A and B. Reaction mixture A was prepared as follows: 20 mM Hepes (pH 7.5), 150 mM NaCl, 5% glycerol, 1 mM $MgCl_2$, 2 mM DTT, 20 μM myristic acid, 20 μM FMN, 1 mM ATP, 10 μM LuxAB, and 2 μM Fre to facilitate the measurement. In each experiment, the reaction was initiated by adding reaction mixture B: 4 mM NADPH and 20 μM LuxCE (WT or mutant) to the reaction mixture A. FlexStation 3 microplate reader (Molecular

Devices) was used for all measurements in a 100 µl reaction system in a 96-well microplate. The microplate reader was used as the continuous detection mode for 30 min. Each group of experiments was repeated three times, and GraphPad Prism 8 was used for statistical analysis.

EM sample preparation and data collection

For negative-staining sample preparation, holey grids coated with continuous carbon film were glow discharged. After that, 4 µl of the protein sample was loaded. The grids were first blotted using filter paper and sequentially washed with water and 2% uranyl formate. The samples were observed using a 120 kV Tecnai Spirit (FEI) equipped with a Gatan Ultrascan 4000 camera at a magnification of ×62,000, corresponding to pixel size of 1.35 Å on the images. Defocus ranging from −1.0 to −2.5 µm and a total dose of ∼50 electrons per square Å were used. For cryo-EM sample preparation, Quantifoil holey carbon grid (R1.2/1.3, 300 mesh) were glow discharged with air for 40 s. Next, 4 µl of purified PpLuxCE was applied to glow-discharged grids. After 60 s, the grid was blotted with filter paper for 4.0 s (under 100% humidity and 4 °C) and plunged into liquid ethane cooled with liquid nitrogen using Vitrobot IV (FEI). A total of 8119 micrograph stacks were automatically collected with UCSFImage4 (https://cryoem.ucsf.edu/software/) on Titan Krios microscope (FEI) at 300 kV equipped with a Gatan K3 Summit direct electron detector, at a nominal magnification of ×29,000 with defocus values range from 1.5 µm to 2.0 µm. Data were collected at 0.833 Å per pixel with a dose rate of 8.2 counts per physical pixel per second. Images were recorded for 0.25 s exposure time in 32 subframes with a total dose of 50 electrons per square Å.

Cryo-EM data processing, model building, and refinement

Beam-induced drift was corrected using the MOTION-COR2 program (https://msg.ucsf.edu/software) (33). The contrast transfer function parameters of micrographs were estimated using the CTFFIND4 program (34). All particles were autopicked using the Relion3.0 program (https://hpc.nih.gov/apps/relion/index.html) and total of 1,012,247 particles were picked and extracted (pixel size: 0.833 Å, Box size: 256 pixels, mask size: 160 Å). Following 2D classification, 778,247 particles were used for 3D classification in Relion3.0 (35), and the particles were then classified into five classes. Apart from the junk class, the dominant class (40.0% particles) was re-extracted with a larger box size (pixel size: 0.833 Å, box size: 340 pixels, mask size: 220 Å). The re-extracted particles were further submitted to three rounds of 3D classification, and the main classes were further autorefined, yielding the best density map with a resolution of 2.79 Å based on gold-standard Fourier shell correlation = 0.143 cutoff criteria. Local resolution analysis was performed using the ResMap program. The atomic models were built in COOT (36). Briefly, the initial atomic model was obtained through the automated protein structure homology-modeling server SWISS-MODEL (https://swissmodel.expasy.org/). The predicted model was docked into the density map as a rigid body in Chimera and then manually adjusted using COOT according to the density. The final model was refined using real-space refinement implemented in PHENIX (http://www.phenix-online.org/) (37). All images used for the model were created using PyMOL (www.pymol.org/) and Chimera (38).

Data availability

The 3D cryo-EM density maps have been deposited into the Electron Microscopy Data Bank under accession numbers EMD-33113. The coordinates have been deposited into the Protein Data Bank (PDB) with accession numbers 7XC6.

References

1. Widder, E. A. (2010) Bioluminescence in the ocean: origins of biological, chemical, and ecological diversity. *Science* **328**, 704–708
2. Kaskova, Z. M., Tsarkova, A. S., and Yampolsky, I. V. (2016) 1001 lights: luciferins, luciferases, their mechanisms of action and applications in chemical analysis, biology and medicine. *Chem. Soc. Rev.* **45**, 6048–6077
3. Kotlobay, A. A., Dubinnyi, M. A., Purtov, K. V., Guglya, E. B., Rodionova, N. S., Petushkov, V. N., *et al.* (2019) Bioluminescence chemistry of fireworm odontosyllis. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 18911–18916
4. Brodl, E., Winkler, A., and Macheroux, P. (2018) Molecular mechanisms of bacterial bioluminescence. *Comput. Struct. Biotechnol. J.* **16**, 551–564
5. Kotlobay, A. A., Sarkisyan, K. S., Mokrushina, Y. A., Marcet-Houben, M., Serebrovskaya, E. O., Markina, N. M., *et al.* (2018) Genetically encodable bioluminescent system from fungi. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 12728–12732
6. Vannier, T., Hingamp, P., Turrel, F., Tanet, L., Lescot, M., and Timsit, Y. (2020) Diversity and evolution of bacterial bioluminescence genes in the global ocean. *NAR Genom. Bioinform.* **2**, lqaa018
7. van Berkel, W. J., Kamerbeek, N. M., and Fraaije, M. W. (2006) Flavoprotein monooxygenases, a diverse class of oxidative biocatalysts. *J. Biotechnol.* **124**, 670–689

8. Ulitzur, S., and Hastings, J. W. (1979) Evidence for tetradecanal as the natural aldehyde in bacterial bioluminescence. *Proc. Natl. Acad. Sci. U. S. A.* **76**, 265–267

9. Campbell, Z. T., Weichsel, A., Montfort, W. R., and Baldwin, T. O. (2009) Crystal structure of the bacterial luciferase/flavin complex provides insight into the function of the beta subunit. *Biochemistry* **48**, 6085–6094

10. Ingelman, M., Ramaswamy, S., Niviere, V., Fontecave, M., and Eklund, H. (1999) Crystal structure of NAD(P)H:flavin oxidoreductase from Escherichia coli. *Biochemistry* **38**, 7040–7049

11. Nijvipakul, S., Wongratana, J., Suadee, C., Entsch, B., Ballou, D. P., and Chaiyen, P. (2008) LuxG is a functioning flavin reductase for bacterial luminescence. *J. Bacteriol.* **190**, 1531–1538

12. Fieschi, F., Niviere, V., Frier, C., Decout, J. L., and Fontecave, M. (1995) The mechanism and substrate specificity of the NADPH:flavin oxidoreductase from Escherichia coli. *J. Biol. Chem.* **270**, 30392–30400

13. Ferri, S. R., and Meighen, E. A. (1991) A lux-specific myristoyl transferase in luminescent bacteria related to eukaryotic serine esterases. *J. Biol. Chem.* **266**, 12852–12857

14. Lawson, D. M., Derewenda, U., Serre, L., Ferri, S., Szittner, R., Wei, Y., *et al.* (1994) Structure of a myristoyl-ACP-specific thioesterase from vibrio harveyi. *Biochemistry* **33**, 9382–9388

15. Li, J., Szittner, R., and Meighen, E. A. (2000) Hyperactivity and interactions of a chimeric myristoryl-ACP thioesterase from the lux system of luminescent bacteria. *Biochim. Biophys. Acta* **1481**, 237–246

16. Law, A., and Boulanger, M. J. (2011) Defining a structural and kinetic rationale for paralogous copies of phenylacetate-CoA ligases from the cystic fibrosis pathogen Burkholderia cenocepacia J2315. *J. Biol. Chem.* **286**, 15577–15585

17. Lin, J. W., Chao, Y. F., and Weng, S. F. (1996) Nucleotide sequence and functional analysis of the luxE gene encoding acyl-protein synthetase of the lux operon from Photobacterium leiognathi. *Biochem. Biophys. Res. Commun.* **228**, 764–773

18. Lee, C. Y., and Meighen, E. A. (1997) Cysteine-286 as the site of acylation of the Lux-specific fatty acyl-CoA reductase. *Biochim. Biophys. Acta* **1338**, 215–222

19. Riendeau, D., Rodriguez, A., and Meighen, E. (1982) Resolution of the fatty acid reductase from Photobacterium phosphoreum into acyl protein synthetase and acyl-CoA reductase activities. Evidence for an enzyme complex. *J. Biol. Chem.* **257**, 6908–6915

20. Wall, L., and Meighen, E. A. (1986) Subunit structure of the fatty acid reductase complex from Photobacterium phosphoreum. *Biochemistry* **15**, 4315–4321

21. Syed, A. J., and Anderson, J. C. (2021) Applications of bioluminescence in biotechnology and beyond. *Chem. Soc. Rev.* **50**, 5668–5705

22. Kurvet, I., Ivask, A., Bondarenko, O., Sihtmae, M., and Kahru, A. (2011) LuxCDABE–transformed constitutively bioluminescent Escherichia coli for toxicity screening: comparison with naturally luminous vibrio fischeri. *Sensors (Basel)* **11**, 7865–7878

23. Krichevsky, A., Meyers, B., Vainstein, A., Maliga, P., and Citovsky, V. (2010) Autoluminescent plants. *PLoS One* **5**, e15461

24. Gupta, R., Patterson, S., Ripp, S., Simpson, M., and Sayler, G. (2003) Expression of the Photorhabdus luminescens *lux* genes (*luxA, B, C, D, and E*)in *Saccharomyces cerevisiae. FEMS Yeast Res.* **4**, 305–313

25. Kaku, T., Sugiura, K., Entani, T., Osabe, K., and Nagai, T. (2021) Enhanced brightness of bacterial luciferase by bioluminescence resonance energy transfer. *Sci. Rep.* **11**, 14994

26. Gregor, C., Pape, J. K., Gwosch, K. C., Gilat, T., Sahl, S. J., and Hell, S. W. (2019) Autonomous bioluminescence imaging of single mammalian cells with the bacterial bioluminescence system. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 26491–26496

27. Su, L., Jia, W., Hou, C., and Lei, Y. (2011) Microbial biosensors: a review. *Biosens. Bioelectron.* **26**, 1788–1799

28. Gregor, C., Gwosch, K. C., Sahl, S. J., and Hell, S. W. (2018) Strongly enhanced bacterial bioluminescence with the ilux operon for single-cell imaging. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 962–967

29. Keller, M. A., Zander, U., Fuchs, J. E., Kreutz, C., Watschinger, K., Mueller, T., *et al.* (2014) A gatekeeper helix determines the substrate specificity of Sjogren-Larsson Syndrome enzyme fatty aldehyde dehydrogenase. *Nat. Commun.* **5**, 4439

30. Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., *et al.* (2009) AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J. Comput. Chem.* **30**, 2785–2791

31. Morgan, C. A., and Hurley, T. D. (2015) Development of a high-throughput *in vitro* assay to identify selective inhibitors for human ALDH1A1. *Chem. Biol. Interact.* **234**, 29–37

32. Ellis, H. R. (2010) The FMN-dependent two-component monooxygenase systems. *Arch. Biochem. Biophys.* **497**, 1–12

33. Zheng, S. Q., Palovcak, E., Armache, J. P., Verba, K. A., Cheng, Y., and Agard, D. A. (2017) MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332

34. Mindell, J. A., and Grigorieff, N. (2003) Accurate determination of local defocus and specimen tilt in electron microscopy. *J. Struct. Biol.* **142**, 334–347

35. Scheres, S. H. (2012) Relion: implementation of a bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519–530

36. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132

37. Liebschner, D., Afonine, P. V., Baker, M. L., Bunkoczi, G., Chen, V. B., Croll, T. I., *et al.* (2019) Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in phenix. *Acta Crystallogr. D Struct. Biol.* **75**, 861–877

38. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., *et al.* (2004) UCSF Chimera–a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612