The 1.5-Å Resolution Crystal Structure of Bacterial Luciferase in Low Salt Conditions*

Andrew J. Fisher†, Thomas B. Thompson‡, James B. Thoden‡, Thomas O. Baldwin‡, and Ivan Rayment‡

From the †Institute for Enzyme Research and Department of Biochemistry, University of Wisconsin, Madison, Wisconsin 53705 and the ‡Department of Biochemistry and Biophysics, Texas A & M University, College Station, Texas 77843

Bacterial luciferase is a flavin monooxygenase that catalyzes the oxidation of a long-chain aldehyde and releases energy in the form of visible light. A new crystal form of luciferase cloned from Vibrio harveyi has been grown under low-salt concentrations, which diffract x-rays beyond 1.5Å resolution. The x-ray structure of bacterial luciferase has been refined to a conventional R-factor of 18.2% for all recorded synchrotron data between 30.0 and 1.50-Å resolution. Bacterial luciferase is an α/β heptamer, and all the individual subunits fold into a single domain (βα)8 barrel. The high resolution structure reveals a non-prolyl cis peptide bond that forms between Ala74 and Ala75 in the α subunit near the putative active site. This cis peptide bond may have functional significance for creating a cavity at the active site. Bacterial luciferase employs reduced flavin as a substrate rather than a cofactor. The structure presented was determined in the absence of substrates. A comparison of the structural similarities between luciferase and a nonfluorescent flavoprotein, which is expressed in the lux operon of one genus of bioluminescent bacteria, suggests that the two proteins originated from a common ancestor. However, the flavin binding sites of the nonfluorescent protein are likely not representative of the flavin binding site on luciferase. The structure presented here will furnish a detailed molecular model for all bacterial luciferases.

*This research was supported in part by National Institutes of Health Grant AR35186 (to I. R.) and Fellowship AR08304 (to A. J. F.), the Robert A. Welch Foundation Grant A-865 (to T. O. B.), and the NIH Health Grant AR35186 (to I. R.) and Fellowship AR08304 (to A. J. F.).

From the Office of Naval Research Grants N00014-93-1-0901, N00014-96-1-00, N00014-96-1-87, and N00014-93-J-1345 (to T. O. B.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (file accession 1LUC, tracking number T-8732) have been deposited in the Protein Data Bank, Brookhaven National Laboratory, Upton, NY.

†Present address: Dept. of Chemistry, University of California, Davis, CA 95616.
‡To whom correspondence should be addressed: Institute for Enzyme Research, 1710 University Ave., Madison, WI 53705. Tel.: 608-262-0529; Fax: 608-265-2904; E-mail: ivan@enzyme.wisc.edu.
luciferase from V. harveyi at 2.4-Å resolution (16). That structure was determined by multiple isomorphous replacement from crystals grown in 1.4 M ammonium sulfate, 0.2 M phosphate. Each subunit folds into a single domain (β/α)8 barrel motif. Dimerization is mediated through a parallel four-helix bundle centered on a pseudo 2-fold axis that relates the structurally similar subunits. Recently, Conti and co-workers (17) ascertained the crystal structure of firefly luciferase. Firefly luciferase is a 62-kDa monomer that folds into a structure different from bacterial luciferase. This was anticipated since the protein sequence and the chemistry catalyzed by these two luciferases are considerably different.

We report here a new crystal form of bacterial luciferase grown in low-salt conditions that diffract x-rays to significantly higher resolution than the previous crystals grown in high-salt concentrations. The structure was determined at 1.50-Å resolution from a single crystal of bacterial luciferase grown in methyl ether polyethylene glycol, which was frozen at −160 °C. This higher resolution structure has revealed many new features of luciferase including the solvent structure and the observation of a non-prolyl cis peptide bond between residues Ala92 and Ala98 of the α subunit, which may have functional significance. The higher resolution structure categorically defined the geometry of all the residues of the 77-kDa bioluminescent enzyme, with the exception of the protease-sensitive loop, and provides a molecular framework for all bacterial luciferases.

**EXPERIMENTAL PROCEDURES**

Crystallization and Data Collection—Luciferase, cloned from V. harveyi, was expressed in Escherichia coli and purified as described earlier (18). A new crystal form of bacterial luciferase was grown under conditions different than previously reported (16, 19). Crystals were grown by micro batch method at 4 °C in 17% methyl ether polyethylene glycol (ME-PEG) (M, 5000), 250 mM MgCl2, buffered at pH 6.5 with 100 mM MES. The final protein concentration was 7.5 mg/ml. Crystallization was induced by introduction of micro- or macroseeds obtained from preliminary hanging drop experiments. Crystals utilized for seeding were prepared by soaking in 14% ME-PEG for 5 min to dissolve any additional nucleation points. The new crystal form of bacterial luciferase grown in methyl ether polyethylene glycol grew to the size of 0.3 x 0.7 x 1.5 mm in 7 days and diffracts x-rays to better than 1.5-Å resolution at a synchrotron radiation source. A crystal large enough for data collection was transferred directly into a cryoprotectant consisting of 20% ethylene glycol, 22% ME-PEG, 300 mM MgCl2, 50 mM MES, pH 6.5.

After transferring to the cryoprotectant, the crystal was immediately mounted in a loop (20) constructed of 20-μm thick surgical suture and was frozen in the stream of nitrogen (−160 °C) directly on the rotation camera at the Stanford synchrotron radiation laboratory, beam line 7-1 (λ = 1.08 Å). Strong diffraction maxima were observed beyond Bragg spacings of 1.5 Å. Data were collected by oscillation photography. Two scans of a single frozen crystal were employed for data acquisition strategy. The first scan consisted of 1° oscillations for long doses to increase the intensity of the high resolution reflections but resulted in overloading the low resolution data. The exposure time depended on beam energy and varied during the ring fill to equalize the number of photons per exposure. Diffraction data were collected on a MAR image plate system with a crystal to detector distance of 160 mm. After the high resolution data were collected, the crystal to detector distance was increased to 210 mm, and 2.5° oscillation photographs were taken for lower photon counts to record the low resolution data. Diffraction intensities were measured and scaled together with the programs DENZO and SCALEPACK (21, 22). Partial reflections recorded on adjacent images were added together to approximate full reflections. The crystals belong to the monoclinic space group C2 with unit cell parameters: a = 150.5 Å, b = 59.0 Å, c = 76.5 Å, β = 93.86°. There is one αβ heterodimer in each asymmetric unit (V_m = 2.20 Å³, solvent content ~44%). A 99% complete data set to 1.5-Å resolution was collected using a MAR image plate camera from a single crystal that was frozen to −160 °C. The overall Rmerge is 4.1% for all data to 1.5-Å resolution. Table I gives the data collection statistics.

**Structure Determination and Refinement**—The structure of luciferase grown in ME-PEG was solved by the molecular replacement method (23). The 2.4-Å resolution ammonium sulfate structure was used as a search model in the molecular replacement program AMORE (24). Data between 10.0- and 4.0-Å resolution were used in the rotation search that resulted in a peak of 14.6 σ, the highest false peak was 7.3 σ. The rotated model was applied in a translation search resulting in a single large peak of 47.6 σ and an R-factor of 37.9%. The luciferase structure was then refined against all recorded data to 2.0-Å resolution by the conjugate direction algorithm implemented in TNT (25), lowering the R-factor to 30.3%. An initial electron density map was computed at 2.0-Å resolution employing SIGMAA coefficients to suppress model bias (26). The resulting map was of excellent quality, and manual adjustments were made with the program O (27). The ensuing model was then subjected to another round of TNT refinement, which lowered the R-factor to 24.1% for all data to 2.0-Å resolution. Subsequent refinement against all recorded data to 1.50-Å resolution resulted in an R-factor of 26.0%, which was reduced to 21.1% with minor manual adjustments of the model and adding 302 water molecules with the programs PEKPIK in TNT (25) and WATPEAK in the CCP4 program suite (28). The final model of the α subunit consists of residues 1–261 and 291–355. The 29 residues (262–290) for which there is no electron density corresponds to the protease-sensitive loop that is also disordered in the ammonium sulfate structure. SDS gel analysis of luciferase

---

1 The abbreviations used are: ME-PEG, methyl ether polyethylene glycol; MES, 2-(N-morpholino)ethanesulfonic acid; TIM, triose-phosphate isomerase; NFP, nonfluorescent flavoprotein.
crystals indicates that both subunits are intact (data not shown). All but the last four amino acids of the \( \beta \) subunit were traced in the electron density map. The final model has an \( R \)-factor of 18.2\% for all recorded data (\( uF_u \)) to 1.50 Å resolution where the root mean square (r.m.s.) deviation from ideal bond lengths, angles, and planes are 0.015 Å, 2.40°, and 0.007 Å, respectively. Table I presents the final refinement statistics including number of atoms and average \( B \)-values.

A plot of the average main-chain temperature factors and correlation coefficient are shown in Fig. 1, a and b. The correlation coefficient is a measure of how well the atoms fit the electron density as calculated by the molecular graphics program \( O \) (27). The mean main-chain temperature factor and correlation coefficient for the \( \alpha \) subunit are 16.9 Å\(^2\) and 0.948, respectively, and for the \( \beta \) subunit 16.6 Å\(^2\) and 0.945, respectively.

**RESULTS AND DISCUSSION**

Structure Description—A Ramachandran plot (29) of the main-chain conformation angles indicates that 99\% of the non-glycine residues lie in the allowed regions as defined by the program PROCHECK (30). The average coordinate error in the final model, as estimated from a Luzzati plot (31), is between 0.125 and 0.15 Å. Figs. 7a and 9 show a region of representative electron density at 1.50 Å resolution computed with the coefficients of \( 2F_o - |F_c| \) and phases calculated from the final model.

The \( \alpha \)-\( \beta \) heterodimer has a parallelepiped shape with dimensions roughly 75 × 45 × 40 Å (Fig. 2). As expected from the sequence similarity, the \( \alpha \) and \( \beta \) subunits display similar tertiary structures. Both subunits contain a single \((\beta/\alpha)_8\) barrel that was first observed in the crystal structure of triose-phosphate isomerase (TIM) (32). The \( \alpha \) and \( \beta \) subunits have identical topologies (Fig. 3), with the most outstanding loop of the \((\beta/\alpha)_8\) motif existing between \( \beta 7 \) and \( \alpha 7 \).

Hydrophobic residues pack in the \( \beta \)-barrel inner core of both subunits. However, NH\(_2\)-terminal residues of the \( \beta \)-strands are hydrophilic and exposed to solvent. Part of the \( \beta \)-barrel’s C-terminal end is hydrophobic and shielded from solvent by two \( \alpha \)-helices. The \( \alpha \)-helices (\( \alpha 7a \) and \( \alpha 7b \)) emerge from the \( \beta 7-\alpha 7 \)
loop. This feature is observed in both \( \alpha \) and \( \beta \) subunits. In the \( \beta \) subunit, helix \( \alpha 7a \) extends along the top of the barrel, followed by a tight turn then helix \( \alpha 7b \), which runs antiparallel to helix \( \alpha 7a \). In the \( \alpha \) subunit, helix \( \alpha 7a \) stretches toward the subunit interface. The loop that connects helices \( \alpha 7a \) to \( \alpha 7b \) is disordered in the electron density map. Residues 262–290 of the \( \alpha \) subunit are not seen in the electron density map. The disordered loop in the \( \alpha \) subunit corresponds to the 29-residue insert when compared with the \( \beta \) subunit (residues 258–286) and is the loop that is readily cleaved by proteases in the absence of substrates (11, 12, 14). In the \( \alpha \) subunit, helix \( \alpha 7b \) is short, consisting of 5 residues, although its true length may be obscured by the flexibility in the preceding loop. After helix \( \alpha 7b \), both subunits contain a 3-residue \( \beta \)-strand (\( \beta 7a \)) that runs parallel to and augments \( \beta 7 \), which extends past the other \( \beta \)-strands of the \( \beta \)-barrel.

The only other deviations from the \((\beta/\alpha)_8\) topology is a small helix (\( \alpha 4a \)) that is positioned at the C-terminal end of the \( \beta \)-barrel of each subunit near the subunit interface. There is also a hairpin loop structure in both subunits that runs along the periphery of the subunit interface and embraces the parallel four-helix bundle at the dimer interface. This hairpin loop contains internal main-chain hydrogen bonds, but the main-chain torsion angles are inconsistent with \( \beta \) structure required to designate the loop an antiparallel \( \beta \)-hairpin. Pro\(^{154} \), conserved in both subunits, disrupts the possible \( \beta \)-strand (Fig. 4, a and b). Furthermore, Pro\(^{146} \) disrupts the other strand in the \( \beta \) subunit opposite Pro\(^{154} \) (Fig. 4b). The reverse turn at the apex of the hairpin loop structure closely resembles a \( \beta \) type I\( \beta \) turn, but the carbonyl oxygen of residue \( i \) does not hydrogen bond with the main-chain amide nitrogen of residue \( i + 3 \). This structure is observed in the hairpin loops of both subunits. Also in both subunits, the residue at position \( i \) of the turn is Asn\(^{148} \), which favors \( \beta \) reverse turns because the O\( i \)-1 atom hydrogen bonds to the main-chain amide nitrogen of residue \( i + 2 \) as is observed in both luciferase subunits. The hairpin loops in both subunits terminate with Pro\(^{150} \) whose peptide bond adopts the cis configuration in both subunits. Pro\(^{150} \) is conserved among all luciferase \( \alpha \) and \( \beta \) subunits suggesting the importance of a cis peptide bond conformation at this position.

Dimerization is mediated through a parallel four-helix bun-
dle, which is centered on a pseudo 2-fold axis that relates the \( \alpha \) and \( \beta \) subunits (Fig. 2). Each subunit contributes helices \( \alpha_2 \) and \( \alpha_3 \) to form the four-helix bundle. Helix \( \alpha_2 \) lies very close to the pseudo 2-fold axis resulting in an unusually close packing of the \( \alpha_2 \) helices from each subunit. At one point, the main chain atoms from one helix reside within 3.2 \( \AA \) from the main chain atoms in the pseudo 2-fold-related helix in the other subunit. In this region, glycines and alanines shape the surface of the helix allowing for the close contact. In particular, Gly\(^{64}\) is totally conserved in all luciferase \( \alpha \) and \( \beta \) subunits permitting this intimate contact.

There are a considerable number of intersubunit interactions...
arising from the dimer interface. Most of these contacts occur in the parallel four helix bundle. The majority of intersubunit contacts established in the four helix bundle are van der Waals interactions. 2150 Å² of accessible surface area is buried upon dimer formation based on a search probe radius of 1.4 Å (33). This value falls in the expected range given the size of the luciferase subunits (34). Twenty-two intersubunit hydrogen bonds help tether the two subunits together (Table II). An interesting hydrogen bond occurs between residues His45 and Glu88. These two residues are conserved among the α and β subunits creating similar intersubunit hydrogen bonds related by pseudo 2-fold symmetry (His45-Glu88). Both of these residues are conserved among all luciferase subunits, but is 3.7 and 4.4 Å away from the carbonyl oxygen of Thr80 and Thr80, respectively.

Non-prolyl Cis Peptide Bond—In the α subunit, β-strand 3 terminates with a bulge that protrudes into the core of the β-barrel. This bulge contains a cis peptide bond between residues Ala74 and Ala75. Fig. 7a illustrates the conformation of β3 with the bulge and the cis peptide displayed with the electron density map. The 1.5-Å resolution electron density map unequivocally demonstrates the cis conformation of the peptide bond between residues Ala74 and Ala75 in the α subunit (Fig. 7a). In the 2.4-Å resolution ammonium sulfate structure, the bulge did not fill the density extremely well, but the map was not high enough resolution to confidently build a cis peptide bond. A similar bulge terminates β3 in the β subunit, but the density clearly indicates a trans peptide bond between Leu74 and Asn75. Fig. 7b illustrates the similarity of the bulge and overall shape of β3 between the two subunits.

Non-prolyl cis peptide bonds are rare (37, 38) but have been observed in a few other crystal structures, and almost all play significant roles in positioning crucial residues to carry out ligand binding and/or catalysis (39). In luciferase, the cis peptide bond occurs in a bulge at the end of β3 positioning it at the C-terminal end of the barrel, where all (βα)8 barrels exhibit their active sites (40). Alα74 and Alα75 form the bottom floor at the entrance of a small cavity projecting off the larger and deeper pocket in the center of the β-barrel of the α subunit (Fig. 8). The walls of this smaller cavity include His45 on one side and Thr45 on the other subunit. The C-terminal region of helixα3 establishes hydrophobic contacts with the N-terminal region of helix α1 in the other subunit. Most of the hydrophobic interactions are conserved among the subunits resulting in pseudo symmetric interactions.

A great deal of sequence conservation exists between the two luciferase subunits. Many of the conserved residues are also preserved in the luciferases from other bioluminescent bacteria. Thirty residues are totally conserved among all bacterial luciferase α and β subunits whose sequences are known presently (8, 36). A majority of the conserved residues dwell near the luciferase α-β dimer interface (Fig. 6). This demonstrates that the pseudo 2-fold axis, which relates the α and β subunits, is also evident at the level of the primary structure. Furthermore, conservation of the 2-fold symmetry at the interface suggests its significance for dimerization and enzyme function. This is confirmed by the mutational and structural studies described above. Similar intersubunit interactions are also observed between these residues in the crystal structure of the LuxB homodimer.2

With the luciferase structure in hand, the beta sandwich fold is seen to be an evolutionarily conserved motif. The structure of the luciferase dimer provides a basis for understanding the evolution of the beta barrel from a beta sandwich.

Table II

<table>
<thead>
<tr>
<th>Residue</th>
<th>Atom</th>
<th>Bond distance Å</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gln117</td>
<td>Oe-1</td>
<td></td>
</tr>
<tr>
<td>Thr18</td>
<td>Oγ-1</td>
<td></td>
</tr>
<tr>
<td>His17</td>
<td>Nε-1</td>
<td></td>
</tr>
<tr>
<td>His28</td>
<td>Nε-1</td>
<td></td>
</tr>
<tr>
<td>Asn34</td>
<td>Nα-2</td>
<td></td>
</tr>
<tr>
<td>Asn44</td>
<td>Oe-1</td>
<td></td>
</tr>
<tr>
<td>Thr40</td>
<td>O</td>
<td></td>
</tr>
<tr>
<td>Thr40</td>
<td>Oγ-1</td>
<td></td>
</tr>
<tr>
<td>Arg45</td>
<td>Nγ-2</td>
<td></td>
</tr>
<tr>
<td>Arg45</td>
<td>Nγ-2</td>
<td></td>
</tr>
<tr>
<td>Glu48</td>
<td>Oe-1</td>
<td></td>
</tr>
<tr>
<td>Glu48</td>
<td>Oe-2</td>
<td></td>
</tr>
<tr>
<td>Asp50</td>
<td>Oδ-1</td>
<td></td>
</tr>
<tr>
<td>Gin56</td>
<td>Nε-2</td>
<td></td>
</tr>
<tr>
<td>Arg75</td>
<td>Nγ-1</td>
<td></td>
</tr>
<tr>
<td>Val116</td>
<td>O</td>
<td></td>
</tr>
<tr>
<td>Gin157</td>
<td>Nε-2</td>
<td></td>
</tr>
<tr>
<td>Asn159</td>
<td>Nε-2</td>
<td></td>
</tr>
<tr>
<td>Asn159</td>
<td>Nε-2</td>
<td></td>
</tr>
<tr>
<td>Ser161</td>
<td>Oγ</td>
<td></td>
</tr>
<tr>
<td>Ser161</td>
<td>Oγ</td>
<td></td>
</tr>
<tr>
<td>Ser161</td>
<td>Oγ</td>
<td></td>
</tr>
</tbody>
</table>

and Cys$^{106}$ on the other. Mutation of His$^{44}$ to Ala or Asp results in inactivation of the enzyme (35). Cys$^{106}$ of the $\alpha$ subunit is a highly reactive thiol whose chemical modification resulted in inactivation of the enzyme (41, 42). However, site-directed mutagenesis experiments have clearly demonstrated that the reactive thiol is not involved in the bioluminescence reaction (18). Binding of either FMN or FMNH$_2$ in the presence of O$_2$ to luciferase protects the reactive thiol from modification (41), and modification of the Cys$^{106}$ thiol substantially decreases the affinity of the protein for FMNH$_2$ (43). However, modification of the reactive thiol has little effect on the binding of FMN (44). These observations suggest that there is not a direct interaction of the flavin with the thiol that affects protection but rather a conformational change resulting from flavin binding.

These data demonstrate the importance of the small cavity projecting off the central large pocket and could justify the reason for the cis peptide bond, because a trans conformation would decrease the size of the opening. Residue 75 of the $\alpha$ subunit is either an alanine or glycine residue in all luciferases. Proline, which is more energetically favorable in cis peptides, would introduce a larger side chain and reduce the size of the opening. Additionally, the main-chain dihedral angles for Ala$^{75}$ ($\phi = -153.1, \psi = 164.5$) are unfavorable for proline residues, which prefer to reside around $\phi \approx -60$.

As seen in Fig. 8, two residues from the $\beta$ subunit also play a role in the small cavity. Glu$^{88}$ from the $\beta$ subunit hydrogen bonds to His$^{37}$, which forms part of the cavity sidewall. This intersubunit interaction, as pointed out above, is conserved in all luciferase subunits. In addition, the guanido group of Arg$^{96}$

---

**Fig. 5. Luciferase surface.** Stereo view surface rendering of bacterial luciferase generated with the program GRASP (73). The surface formed from $\alpha$ subunit residues is colored white and the $\beta$ subunit is colored red. The deep pocket at the subunit interface is clearly visible as well as part of the active site pocket in the $\alpha$ subunit.

**Fig. 6. Conserved residues.** Stereo $\alpha$-carbon trace mapping the location of 30 residues that are conserved in all luciferase $\alpha$ and $\beta$ subunits. The luciferase backbone is drawn in blue and red lines to designate the $\alpha$ and $\beta$ subunits, respectively, and $\alpha$-carbon of each conserved residue is drawn as a gray ball. The locations of the three Mg$^{2+}$ ions are also identified by green balls. Mg$^{2+}$ ions labeled 2001 and 2002 are involved in crystal contacts while Mg$^{2+}$ 2003 is likely the result of nonspecific binding.
forms the back of the cavity. Arg85 is also mentioned above for its conserved interactions between subunits. These two residues might suggest a possible role for the β subunit during the bioluminescent reaction if this cavity, which extends off the larger pocket at the C-terminal end of the barrel, is part of the active site. The cavity in the β-subunit is more confined because of a trans peptide bond between positions 74 and 75, and larger residues line the cavity's entrance; Asn replaces Ala at position 75 and Tyr substitutes for Leu at position 42.

Magnesium Binding and Crystal Packing—Crystallization of luciferase in methyl ether polyethylene glycol requires the presence of magnesium. Omission of magnesium results in no crystal growth. Removal of magnesium from crystals, by addition of EDTA, results in cracking. During the first manual rebuilding, it became evident why magnesium was required for crystallization. A total of three magnesium ions were observed in the crystal structure (Fig. 6). Two magnesium ions are involved in crystal contacts between symmetry related dimers and the third ion binds to the α subunit but does not have any functional or structural capacities.

One magnesium ion (Mg\(^{2+}\) 2002) involved in crystal packing is coordinated by Oe-1 of Glu\(^{237}\) and Oδ-1 of Asp\(^{346}\) in a symmetry related subunit (Fig. 9) (prime Greek letters represent crystallographic symmetry related subunits). Four ordered water molecules complete the octahedral geometry. The Mg\(^{2+}\)-oxygen coordination distances range from 2.03 to 2.33 Å. Oe-1 of Glu\(^{129}\) coordinates the other magnesium ion involved in crystal packing (Mg\(^{2+}\) 2001), and the five remaining Mg\(^{2+}\) ligands are ordered water molecules. The carboxylate group of Glu\(^{130}\) from a symmetry related subunit hydrogen bonds to two of the five Mg\(^{2+}\) water ligands. Water molecule 3101 is 2.75 Å away from Oe-1 of Glu\(^{130}\), and the distance between Oe-2 of Glu\(^{130}\) and water 3177 is 2.76 Å.

The third magnesium ion seen in the crystal structure (Mg\(^{2+}\) 2003) is not involved in any crystal contacts but binds to the periphery of the α subunit. No protein atoms directly ligate the Mg\(^{2+}\). Six ordered water molecules coordinate the ion with octahedral geometry. This magnesium ion binds near the N-
terminal opening of the β-barrel and interacts with residues in the loops preceding β strands β7 and β8 of the α subunit (Fig. 6). Six protein atoms hydrogen bond to five of the water ligands. The Oδ-2 atom of Aspα223 and the Oδ-2 of Aspα321 hydrogen bond to the same water ligand (3418), 2.75 and 2.70 Å, respectively. Oδ-1 of Aspα321 hydrogen bonds to water 3335. The main-chain carbonyl oxygen of Lysα221 is 2.91 Å away from water ligand 3352, and the carbonyl oxygen of Ileα222 is 3.18 Å away from water ligand 3380. The amine group of Lysα2 hydrogen bonds to Mg2⁺ ligating waters 3418 and 3486. The binding site of this magnesium ion was unexpected but probably does not have any functional significance since no protein atoms directly ligate the Mg2⁺. Furthermore, the B-factors of the third Mg2⁺ and its coordinated waters are approximately 15 Å² higher than for the other two Mg2⁺ sites involved in crystal contacts, suggesting this binding might be nonspecific and a result of the high MgCl2 concentration used in crystallization.

In addition to the magnesium ions binding to the luciferase structure, five well ordered ethylene glycol molecules are apparent in the solvent structure. Ethylene glycol was used as a cryo-protectant to preserve the crystal during freezing at –160 °C. All five ethylene glycols bind at the protein surface, three of them at the α-β subunit interface. One of the ethylene glycol molecules binds in a small cavity that is formed between helices α1 and α2 of the α subunit. Another ethylene glycol molecule mediates an intersubunit contact. His82 from the α subunit hydrogen bonds to a glycol hydroxyl oxygen, which in turn hydrogen bonds to the carbonyl oxygen of Pheα226 in the β subunit. In all five cases, one or both ethylene glycol hydroxyl oxygens hydrogen bonds to the protein.

Structural Similarities—There is extensive structural conservation between the α and β subunits confirming their common origin (4). The topology of the α and β subunits is identical, and the secondary structural elements align exactly with the sequence (Fig. 10). The two luciferase subunits superimpose
with a root mean square deviation of 1.99 Å for 300 equivalent \(\alpha\)-carbons (Fig. 11a). The structures of the \(\beta\)-barrels are very similar with only a 0.61 Å r.m.s. deviation in the superposition of the barrel’s 39 \(\alpha\)-carbons. Most of the differences in the \(\alpha\)-\(\beta\) superposition occur in the exterior \(\alpha\)-helices, which are slightly displaced relative to their pseudo 2-fold-related subunit. The largest displacement appears in the short helix \(\alpha_4a\) near the C-terminal end of the \(\beta\)-barrel. Helix \(\alpha_4a\) in the \(\alpha\) subunit shifts approximately 3 Å along the helix axis away from the barrel’s center permitting a larger opening to the active site. The regions involved with dimerization, helices \(\alpha_2\) and \(\alpha_3\) and the hairpin loop structure, are exceptionally similar in structure.

Structural similarities were also observed between the luciferase subunits and the nonfluorescent flavoprotein (NFP) (45, 46) from Photobacterium leiognathi. Bioluminescent bacteria belonging to the genus Photobacterium contain an additional gene located between luxB and luxE in the lux operon. This gene now known as luxF, was originally designated luxG (47), and independently as luxN (48). The luxF gene encodes a 24-kDa nonfluorescent flavoprotein whose function is unknown at present but binds two molecules of an unusual flavin mononucleotide adduct (45, 46, 49). Myristic acid is covalently linked to C-6 of the isoalloxazine ring of the flavin mononucleotide. Interestingly, both myristic acid and FMN are end products of the luciferase bioluminescence reaction. However, the connection between the nonfluorescent flavoprotein and bioluminescence remains unclear. NFP displays sequence similarity to both luciferase subunits and is 22.4 and 33.3% identical in amino acid sequence to the luciferase \(\alpha\) and \(\beta\) subunits, respectively (Fig. 10).

The crystal structure of NFP has been recently determined (49) and refined to high resolution (50). The crystal structure revealed that NFP forms a homodimer, and each monomer folds into a novel seven-stranded \(\beta\)-barrel surrounded by seven \(\alpha\)-helices. Given the NFP structure and sequence alignment of luxF to luxA and luxB, Moore and James (51) correctly predicted the structure of the luciferase monomer to have a \((\beta\alpha)_8\) fold. The structure of NFP superimposes surprisingly well with the individual luciferase \(\alpha\) and \(\beta\) subunits, 2.48 and 1.55 Å r.m.s. deviation, respectively (Fig. 11b). The NFP \(\beta\)-barrel is mostly parallel (strands \(\beta_3\) and \(\beta_4\) form an antiparallel hairpin) and contains a considerable gap between strands \(\beta_2\) and \(\beta_3\). The seven NFP \(\beta\)-strands align well structurally with seven of the eight luciferase strands. In the superposition, strand \(\beta_3\) of luciferase (\(\alpha\) and \(\beta\) subunits) resides in the gap between strands \(\beta_2\) and \(\beta_3\) of NFP that would complete an eight-stranded \(\beta\)-barrel (Fig. 11b). In the NFP structure, this gap is filled with ordered water molecules that fasten the two ends of the barrel together with a hydrogen bonding network to strands \(\beta_2\) and \(\beta_3\) (50). It is interesting to point out that the

Fig. 10. Sequence alignment. Protein sequence alignment of luciferase \(\alpha\) and \(\beta\) subunits from V. harveyi (LuxA and LuxB) and nonfluorescent protein from P. leiognathi (LuxF). Alignment was based on sequence and secondary structural elements from the crystal structures. The luciferase \(\alpha\) and \(\beta\) subunits share 31.9% sequence identity, whereas NFP is 22.4 and 33.3% identical to the luciferase \(\alpha\) and \(\beta\) subunits, respectively. The secondary structural elements, as observed in the crystal structures, are displayed above the sequence and by colored boxes. The asterisks label residues that are conserved in all three proteins.
strand missing in the NFP structure is the same strand that terminates with a non-prolyl cis peptide bond in the luciferase $\alpha$ subunit. Strand $\beta$3 of NFP structurally aligns with $\beta$4 of luciferase in the superposition but runs in the reverse direction.

The seven helices of NFP align with helices $\alpha$1, $\alpha$5, $\alpha$6, $\alpha$7a, $\alpha$7b, $\alpha$7, and $\alpha$8 of the luciferase $\alpha$ and $\beta$ subunits (Figs. 10 and 11b). NFP does not contain residues or secondary structural elements corresponding to the helices and the hairpin loop involved in the luciferase dimerization. Sequence alignment of NFP to luciferase $\alpha$ and $\beta$ subunits reveals a gap in the NFP sequence corresponding to helices $\alpha$2, $\alpha$3, $\beta$-strand $\beta$3, and the hairpin loop (Fig. 10). However, homo-dimerization of NFP still occurs along the same relative region of the molecule, but the intersubunit interactions occur between $\beta$-strands (49, 50) and not helices as observed in luciferase. Additionally, the $\beta$7-$\alpha$7 loop in the luciferase $\beta$ subunit, which contains helices $\alpha$7a and $\alpha$7b and the short strand $\beta$7a that augments $\beta$7, is also seen in NFP and superimposes with an r.m.s. deviation of 0.79 Å.

This evidence suggests that LuxF may have arisen from gene duplication of LuxB (luciferase $\beta$ subunit) and subsequently lost its ability to associate with the LuxA gene product by deletion of the residues involved in dimerization. Yet LuxF still maintained (or developed) its ability to form homodimers. The function of LuxF, which is found in only one genus of bioluminescent bacteria, is unknown, but is not required for bioluminescence (36).

The locations of the two unique flavin adducts that bind to NFP are shown in Fig. 11b. Both flavin cofactors bind on the side of the $\beta$-barrel between the surface helices. These binding sites probably do not reveal the flavin active site in luciferase, because helices $\alpha$4 and $\alpha$8 in luciferase extend over the equivalent flavin binding sites and would occlude FMN binding. Moreover, both sites are distant from the C-terminal end of the
The structure of luciferase reveals a large deep pocket entering the C-terminal end of the \( \alpha \)-subunit's \( \beta \)-barrel (Fig. 12). Projecting off this large central pocket is a smaller accessible cavity formed by the non-prolyl cis peptide bond highlighted above (Fig. 8). These pockets are sufficiently large enough to accommodate FMNH\(_2\), O\(_2\), and a long-chain aldehyde. Furthermore, the pocket is expected to exclude access water from the C4a hydroperoxyflavin intermediate and the excited flavin that is formed following the decay of the tetrahedral intermediate (8). The disordered loop is likely to achieve this task by blocking the entrance to the pocket after substrate binding, thus protecting itself from proteolysis (11, 14, 15).

The current 1.5-Å resolution structure contains a few ordered water molecules in the pocket. Even though the structure of luciferase was determined in the absence of substrates, we feel confident that the active site resides within this large internal cavity of the \( \alpha \)-subunit. It should be noted that every amino acid implicated as an active center residue, either by mutagenesis or chemical modification, contacts this internal cavity. Unfortunately, attempts to soak in both oxidized and reduced flavin with and without additional substrates into the crystal have proven unsuccessful.

Folding and Assembly of Luciferase—Protein unfolding, refolding, and assembly of bacterial luciferase has been extensively studied (59–62). It has been demonstrated that separate \( \alpha \) and \( \beta \) subunits, purified from recombinant E. coli independently bearing the luxA or luxB genes, carry out a bioluminescent reaction, but at a quantum efficiency 6 orders of magnitude below that of the heterodimer (63). Moreover, the active dimer fails to assemble when the purified folded \( \alpha \) and \( \beta \) subunits are combined (64, 65). It has recently been demonstrated that purified luciferase \( \beta \) subunit forms a very stable \( \beta_2 \) homodimer that is trapped in a heterodimerization-incompetent complex and is unable to form functional heterodimers due to kinetic partitioning of the folding pathway (62).

Functional dimers can assemble upon renaturation of the unfolded individual subunits (60, 62). Equilibrium unfolding studies of the luciferase heterodimer have shown that the enzyme unfolds through a well-populated non-native intermediate (59, 61). Conversion from the non-native heterodimeric intermediate to a functional enzyme is independent of protein concentration.

Some of the protein folding and assembly observations might be explained in part by the presence of the protein's two prolyl and one non-prolyl cis peptide bonds. It has been substantiated that cis/trans isomerization of the peptide bond preceding proline residues can limit the rate at which a protein can fold into its native conformation (66–68). The cis-Pro\(^{690} \) residue in luciferase subunits, is located at the end of the hairpin loop structure that forms extensive intersubunit contacts. The trans isomer of this peptide bond would cause minor perturbations in the loop that would affect dimerization contacts. In both subunits the residue preceding the cis-proline is an aspar-
Energy calculations predict that non-prolyl cis peptide bonds should destabilize a folded protein by 10–20 kJ/mol (37). Protein thermal stability measurements yield results that compare with the calculated values (70, 71). The presence of a non-prolyl cis peptide bond in the α subunit alone could explain why the heterodimer is apparently less stable than the β2 homodimer (59–62).

Acknowledgments—We thank Drs. M. M. Benning and L. C. Pedersen for helping in data acquisition. We are grateful to Vicki Green for supplying the enzyme used in this study.

REFERENCES

