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 Raymont‡

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 monoxygenase 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 1.5-Å 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 heterodimer, and the individual subunits fold into a single domain (β/α)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 enzyme catalyzes the reaction of FMN, the aliphatic carboxylic acid and blue-green light. All luciferases on 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.

Living organisms that radiate light have been captivating people throughout the ages. Bioluminescent organisms such as fireflies, glowworms, mushrooms, fish, or bacteria represent a diverse range of species, which are widely dispersed in nature (1, 2). The enzymes that catalyze the bioluminescence reactions are named luciferases, and in most cases, their substrates are designated luciferins. These enzymes comprise a large evolutionarily diverse group, and the chemistry they catalyze is quite varied. Indeed, the only common factors of these enzymes is the requirement of O2, which was first established by Robert Boyle (3) more than 3 centuries ago. Today, it is known that all luciferase reactions are oxidative processes that convert a substrate to an electronically excited intermediate. Light emission occurs when the excited-state intermediate reverts back to the ground state resulting in the final product.

Luminous bacteria are the most abundant and widely distributed of all bioluminescent organisms and are found in marine, freshwater, and terrestrial environments. Bacterial luciferase has been studied extensively and is the best understood of all luciferases. The luciferase of luminous bacteria is a flavin monoxygenase. Bacterial luciferase is an uncommon flavoprotein in that it employs reduced flavin as a substrate rather than a tightly bound cofactor. The enzyme catalyzes the reaction of FMNH2, O2, and a long-chain aliphatic aldehyde to yield FMN, the aliphatic carboxylic acid and blue-green light. All bacterial luciferases studied so far appear to be homologous, and all catalyze the same reaction:

\[
\text{FMNH}_2 + \text{O}_2 + \text{RCHO} \rightarrow \text{FMN} + \text{RCOOH} + \text{H}_2\text{O} + \text{hv}(\lambda_{\text{max}} = 490 \text{ nm})
\]

The reaction proceeds through a series of intermediates leading to the formation of a C4a hydroxyflavin (for review see Ref. 4). Light emission apparently occurs from this hydroxyflavin, which dehydrates to yield FMN.

Bacterial luciferase is a heterodimeric enzyme of 77 kDa, composed of α and β subunits with molecular masses of 40 and 37 kDa, and in the case of Vibrio harveyi, 355 and 324 residues, respectively. The two polypeptides, encoded on adjacent genes, luxA and luxB in the lux operon, display sequence homology and appear to have arisen by gene duplication (4). There is a single active center in the luciferase heterodimer that resides on the α subunit (5) and binds one reduced flavin molecule (6, 7). The role of the β subunit is not clear at this time but is essential for a high quantum yield reaction (8). Amino acid sequence alignment between the two subunits reveals that there share 32% sequence identity. The α subunit contains 29 additional amino acid residues inserted between residues 258 and 259 of the β subunit (9, 10). This region of the α subunit is known to be sensitive to proteolytic digestion in the absence of substrates (11, 12). A single proteolytic cleavage in the region of residues 274-291 in the α subunit inactivates the enzyme (11, 13). The protease labile region of the α subunit appears to move during the catalytic cycle and becomes protected from protease treatment. Binding of FMNH2 to the α subunit reduces vulnerability to proteolytic inactivation (11, 14, 15). The β subunit is insensitive to proteases, and the quaternary structure of the α-β complex is not altered by proteolytic cleavage (13).

Last year, we reported the crystal structure of bacterial
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 forms 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 peptide bond between residues Ala74 and peptide bond between residues Ala74 and 18.2. The resulting map was of excellent quality, and manual adjustments of the model and adding 302 water molecules with the conjugate direction algorithm implemented in TNT (25), lowering the R-factor to 30.3%. An initial electron density map was computed at 2.4-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 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

### 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 microbatch method at 4 °C in 17% methyl ether polyethylene glycol (ME-PEG)1 (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 preincubation experiments. Crystals utilized for seeding was induced by introduction of micro- or macroseeds obtained 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 peptide bond between residues Ala74 and Ala18. The resulting map was of excellent quality, and manual adjustments of the model and adding 302 water molecules with the conjugate direction algorithm implemented in TNT (25), lowering the R-factor to 30.3%. An initial electron density map was computed at 2.4-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.

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 $\times$ 45 $\times$ 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 $\beta7$ and $\alpha7$.

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 ($\alpha7a$ and $\alpha7b$) emerge from the $\beta7$-$\alpha7$
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 III 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\(\delta-1 \) 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 $\alpha2$ and $\alpha3$ to form the four-helix bundle. Helix $\alpha2$ lies very close to the pseudo 2-fold axis resulting in an unusually close packing of the $\alpha2$ helices from each subunit. At one point, the main chain atoms from one helix reside within 3.2 Å 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 overall shape of bio luminescence activity (35). Another intriguing hydrogen bond occurs between residues His45 and Glu88. These two residues are conserved among all bacterial luciferase subunits (34). Twenty-two intersubunit hydrogen bonds help tether the two subunits together (Table II). An overall shape of bio luminescence activity (35).

<table>
<thead>
<tr>
<th>Residue</th>
<th>Atom</th>
<th>Residue</th>
<th>Atom</th>
<th>Bond distance (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gln117</td>
<td>Oe-1</td>
<td>His61</td>
<td>N</td>
<td>2.92</td>
</tr>
<tr>
<td>Thr118</td>
<td>Oγ-1</td>
<td>Gln95</td>
<td>Oe-1</td>
<td>3.12</td>
</tr>
<tr>
<td>His15</td>
<td>Ne-1</td>
<td>Glu88</td>
<td>Oe-1</td>
<td>2.57</td>
</tr>
<tr>
<td>His24</td>
<td>Ne-1</td>
<td>Glu88</td>
<td>Oe-2</td>
<td>3.08</td>
</tr>
<tr>
<td>Asn174</td>
<td>Nε-2</td>
<td>Glu88</td>
<td>Oe-2</td>
<td>2.97</td>
</tr>
<tr>
<td>Thr40</td>
<td>O</td>
<td>Arg85</td>
<td>Nγ-2</td>
<td>3.02</td>
</tr>
<tr>
<td>Thr40</td>
<td>Oγ-1</td>
<td>Arg85</td>
<td>Nγ-2</td>
<td>3.14</td>
</tr>
<tr>
<td>Arg85</td>
<td>Nγ-2</td>
<td>Thr80</td>
<td>O</td>
<td>2.92</td>
</tr>
<tr>
<td>Arg85</td>
<td>Nγ-2</td>
<td>Thr80</td>
<td>Oγ-1</td>
<td>3.00</td>
</tr>
<tr>
<td>Arg85</td>
<td>Nγ-2</td>
<td>Thr80</td>
<td>Oγ-1</td>
<td>2.67</td>
</tr>
<tr>
<td>Glu48</td>
<td>Oε-1</td>
<td>His45</td>
<td>Nε-1</td>
<td>3.10</td>
</tr>
<tr>
<td>Glu48</td>
<td>Oε-2</td>
<td>His45</td>
<td>Nε-1</td>
<td>2.59</td>
</tr>
<tr>
<td>Asp89</td>
<td>Oδ-1</td>
<td>Thr57</td>
<td>Oγ-1</td>
<td>2.82</td>
</tr>
<tr>
<td>Glu56</td>
<td>Nε-2</td>
<td>Asp85</td>
<td>Oε-2</td>
<td>2.98</td>
</tr>
<tr>
<td>Arg115</td>
<td>Nγ-1</td>
<td>Ser152</td>
<td>O</td>
<td>2.71</td>
</tr>
<tr>
<td>Val116</td>
<td>O</td>
<td>His82</td>
<td>N-2</td>
<td>3.05</td>
</tr>
<tr>
<td>Gin157</td>
<td>Nε-2</td>
<td>Asp48</td>
<td>Oδ-1</td>
<td>2.81</td>
</tr>
<tr>
<td>Asn159</td>
<td>Nε-2</td>
<td>Ser47</td>
<td>O</td>
<td>3.22</td>
</tr>
<tr>
<td>Asn159</td>
<td>Nε-2</td>
<td>Ser47</td>
<td>O</td>
<td>2.91</td>
</tr>
<tr>
<td>Ser161</td>
<td>N</td>
<td>Ser17</td>
<td>Oγ</td>
<td>3.35</td>
</tr>
<tr>
<td>Ser161</td>
<td>Oγ</td>
<td>Ser17</td>
<td>Oγ</td>
<td>3.05</td>
</tr>
<tr>
<td>Ser161</td>
<td>Oγ</td>
<td>Asp48</td>
<td>Oδ-1</td>
<td>3.05</td>
</tr>
</tbody>
</table>

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 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 fit 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 (β-loops) barrels exhibit their active sites (40). Ala74 and Ala75 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 Cys\textsuperscript{106} on the other. Mutation of His\textsuperscript{44} to Ala or Asp results in inactivation of the enzyme (35). Cys\textsuperscript{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\textsuperscript{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\textsuperscript{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\textsuperscript{88} from the \(\beta\) subunit hydrogen bonds to His\textsuperscript{395}, 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\textsuperscript{185}...
forms the back of the cavity. Arg95 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 Oe-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\(^{237}\) 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\(^{237}\) 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\(^{237}\), and the distance between Oe-2 of Glu\(^{340}\) 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. Oε-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 hydrogen bonds to water 3335. The amine group of Lysα222 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 αβ 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. Hisβ82 from the α subunit hydrogen bonds to a glycol hydroxyl oxygen, which in turn hydrogen bonds to the carbonyl oxygen of Pheβ226 in the β subunit. In all five cases, one or both ethylene glycol hydroxyl oxygens hydrogen bond 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 α-carbons (Fig. 11a). The structures of the β-barrels are very similar with only a 0.61 Å r.m.s. deviation in the superposition of the barrel’s 39 α-carbons. Most of the differences in the α-β superposition occur in the exterior α-helices, which are slightly displaced relative to their pseudo 2-fold-related subunit. The largest displacement appears in the short helix α4a near the C-terminal end of the β-barrel. Helix α4a in the α 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 α2 and α3 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 α and β 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 β-barrel surrounded by seven α-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 (βα)8 fold. The structure of NFP superimposes surprisingly well with the individual luciferase α and β subunits, 2.48 and 1.55 Å r.m.s. deviation, respectively (Fig. 11b). The NFP β-barrel is mostly parallel (strands β3 and β4 form an antiparallel hairpin) and contains a considerable gap between strands β2 and β3. The seven NFP β-strands align well structurally with seven of the eight luciferase strands. In the superposition, strand β3 of luciferase (α and β subunits) resides in the gap between strands β2 and β3 of NFP that would complete an eight-stranded β-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 β2 and β3 (50). It is interesting to point out that the
strand missing in the NFP structure is the same strand that terminates with a non-prolyl cis peptide bond in the luciferase α subunit. Strand β3 of NFP structurally aligns with β4 of luciferase in the superposition but runs in the reverse direction.

The seven helices of NFP align with helices α1, α5, α6, α7a, α7b, α7, and α8 of the luciferase α and β 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 α and β subunits reveals a gap in the NFP sequence corresponding to helices α2, α3, β-strand β3, 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 β-strands (49, 50) and not helices as observed in luciferase. Additionally, the β7-α7 loop in the luciferase β subunit, which contains helices α7a and α7b and the short strand β7a that augments β7, 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 β 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 β-barrel between the surface helices. These binding sites probably do not reveal the flavin active site in luciferase, because helices α4 and α8 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 α subunit’s β-barrel, which is the location of the active site in enzymes that exhibit the (βα)₈ motif (40). The phosphate moiety of the flavin molecule that binds near the N terminus of helix α5 in NFP is 3.2 Å away from the phosphate binding site seen in the luciferase structure previously determined in ammonium sulfate (16). This region was hypothesized to bind the phosphate moiety of reduced flavin in luciferase (16). By anchoring the phosphate moiety of FMNH₂ at this site, the flavin can be modeled extending across the C-terminal portion of the α subunit β-barrel. This positions the isoxazoline ring next to Trp₁⁹⁴ and Trp₂⁵⁰ (16), which have been implicated to interact with the flavin ring as measured by fluorescence spectroscopy and circular dichroism spectroscopy. A phosphate ion was not observed in the ME-PEG structure because phosphate was not included in the crystallization conditions.

The regions of high temperature factors in the α subunit correspond to loops that have been proposed to bind flavin (16). The peaks in the α subunit temperature factor plot (Fig. 1a) around residues 109, 121, and 175 all map to loops in the vicinity of the phosphate binding site that was observed in the structure of luciferase derived from crystals grown in ammonium sulfate. The high thermal parameters in the phosphate-free structure suggest flexibility in these areas, which would become stationary upon binding reduced flavin.

There appears to be no significant differences between the structure of luciferase solved in ME-PEG and the original structure solved in ammonium sulfate. Superposition of 630 backbone atoms results in an r.m.s. deviation of 0.59 Å between the two structures. This value falls in the range observed for other proteins whose structures have been determined from different space groups (52, 53). Furthermore, in addition to the differences described above, more of the α subunit’s α₇a-α₇b loop is disordered in the ME-PEG structure, which contains 13 fewer ordered residues.

Active Site Pocket—The active sites of all (βα)₈ barrel enzymes reside at the C-terminal end of the β-barrel (40). In most cases, residues in the loops that connect the β-strand to the subsequent α-helix fabricate the active site. Many flavoenzymes employ the TIM barrel motif to bind flavin (40). Glycolate oxidase (54), flavocytochrome b₂ (55), trimethylamine dehydrogenase (56), and old yellow enzyme (57) are all (βα)₈ barrels that tightly bind flavin mononucleotide as a coenzyme. In these enzymes the phosphate moiety of FMN binds between the β₇-α₇ loop and the NH₂ terminus of an additional small helix in the β₈-α₈ loop. Similar interactions are also observed in other TIM barrels that secure phosphate components in their substrates (58). Bacterial luciferase does not contain a small helix in the β₈-α₈ loop. Its absence could explain why luciferase utilizes FMNH₂ as a substrate and not as a prothetic group as in other (βα)₈ flavoenzymes.

The structure of luciferase reveals a large deep pocket entering the C-terminal end of the α subunit’s β-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 FMNH₂, O₂, and a long-chain aldehyde. Furthermore, the pocket is expected to exclude access water from the C₄₈ hydroperoxylflavin 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 α 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 α and β 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 α and β subunits are combined (64, 65). It has recently been demonstrated that purified luciferase β subunit forms a very stable β₂ 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-Pro¹⁰⁹, which is found in both 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-

---

3 T. O. Baldwin et al., unpublished results.
Energy calculations predict that non-prolyl peptidase with the calculated values (70, 71). The presence of a 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 help in data acquisition. We are grateful to Vicki Green for supplying the enzyme used in this study.

REFERENCES