

Structure and Folding of Bacterial Luciferase

Thomas O. Baldwin, Miriam M. Ziegler, A. Clay Clark, James F. Sinclair, Frank M. Raushel
Center for Macromolecular Design and the Departments of Biochemistry and Biophysics and of
Chemistry
Texas A&M University
College Station, Texas, 77843-2128 U.S.A.

Ivan Rayment, Hazel M. Holden, Andrew J. Fisher, Thomas B. Thompson, James B. Thoden
Institute for Enzyme Research
Department of Biochemistry
University of Wisconsin
Madison, Wisconsin, 53705 U.S.A.

Introduction

Bacterial luciferase is a heterodimeric (ab) enzyme which employs reduced flavin mononucleotide as a substrate (see reference 1 and other references contained therein). The enzyme binds FMNH₂, which reacts with molecular oxygen and an aliphatic aldehyde to yield the carboxylic acid, oxidized FMN, and blue-green light that is emitted from an excited state flavin intermediate on the surface of the enzyme. Based on biochemical and genetic investigations, it has been reported that the heterodimeric enzyme has a single active center that resides primarily, and probably exclusively, on the a subunit. As a consequence of these and other features, bacterial luciferase is not only an excellent model system with which to investigate the flavin-mediated monooxygenation reaction, it is a robust system with which to investigate protein subunit folding and oligomer assembly. The individual subunits support an extremely low level bioluminescence reaction (2-4), but the high quantum yield reaction requires catalysis by the heterodimer (5). Formation of the active enzyme is accompanied by an increase in light intensity from the refolding reaction mixture; light intensity from the reaction is proportional to the number of active, heterodimeric enzyme molecules. The folding and assembly of the enzyme have been studied in some detail, and recently, the structures of the heterodimeric enzyme (6-8) and of a kinetically-trapped species, the b₂ homodimer (9), have been determined. The discussion which follows will describe the current state of our understanding of the folding and assembly of the enzyme, and the structures of the two major products of the folding reactions.

Results and Discussion

It has been over 3 decades since Anfinsen and colleagues demonstrated that the information that dictates the structure of the native form of an enzyme is resident in the amino acid sequence of the protein (10,11). Their conclusions, and much data gathered over the period since, have led to almost universal acceptance of the thermodynamic hypothesis, which, simply stated, maintains that the native form of an enzyme resides at a global minimum of free energy. Indeed, the thermodynamic hypothesis is virtually axiomatic in the field of protein folding today. Nonetheless, there is a growing body of evidence that suggests that the thermodynamic hypothesis might not explain the folding and assembly of large, multidomain or multisubunit proteins. Rather, it appears for some proteins that the distribution of products of the protein folding reaction is determined by the relative rate constants with which they form, not by the relative stability of the products (12-17). For example, in the folding of bacterial luciferase, it appears that the yield of the ab heterodimer relative to the other potential products is determined by the relative rates with which the products form, rather than by their relative stabilities (3,13,14,18-21); this particular characteristic constitutes the definition of kinetic control.

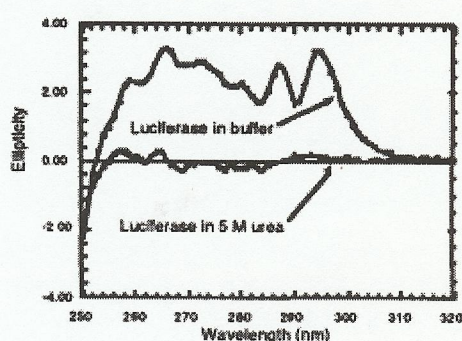


Fig. 1

Circular dichroism spectrum of luciferase in 50 mM phosphate buffer, pH 7.0 at 18°C with and without 5 M urea.

Subunit folding and enzyme assembly. The unfolding and refolding of luciferase is reversible without the involvement of chaperones. Treatment of luciferase with 5 M urea at pH 7, 18°C, leads to rapid unfolding as monitored by circular dichroism and by fluorescence spectroscopy (18). The near UV circular dichroism spectra of the protein in buffer with and without 5 M urea, shown in Figure 1, clearly demonstrate unfolding of the enzyme under these conditions. Extensive studies of the refolding reaction have demonstrated that the heterodimeric enzyme unfolds and refolds reversibly in urea-containing buffers following the same transition, whether unfolding or refolding (20). These same experiments have demonstrated the existence of an inactive, dimeric intermediate that is well-populated at ca. 2.2 M urea and 25 µg/ml total protein. The individual unfolded subunits undergo multiple isomerizations during refolding from urea, ultimately achieving forms that are competent to interact to form the inactive heterodimeric structure, which must itself then undergo subsequent isomerization to assume the active conformation. The kinetic

mechanism describing these and other features is shown in Scheme 1. The b subunit, allowed to fold without the a subunit either *in vivo* or *in vitro*, assumes a b_2 homodimeric structure, while the a subunit folds into a native-like monomer that remains competent to interact with refolding b subunit. The b_2 homodimer remains fully native in 5 M urea (Fig. 2), whereas the heterodimer is rapidly and apparently fully unfolded under the same conditions (Fig. 1). When the subunits are refolding together, the preferred pathway is formation of the biologically active heterodimer, since formation of the homodimer is much slower. The second order rate constant for formation of the homodimer has been determined in 50 mM phosphate, pH 7.0 and 18° to be ca. $150 \text{ M}^{-1}\cdot\text{s}^{-1}$ (21), while for the heterodimer, the heterodimerization rate constant under similar conditions has been determined to be ca. $2400 \text{ M}^{-1}\cdot\text{s}^{-1}$ (22).

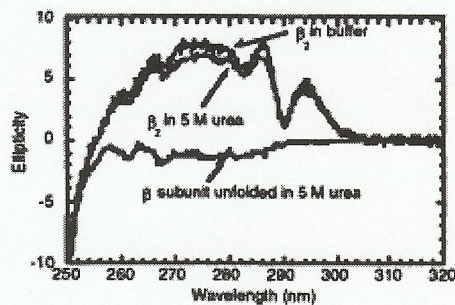
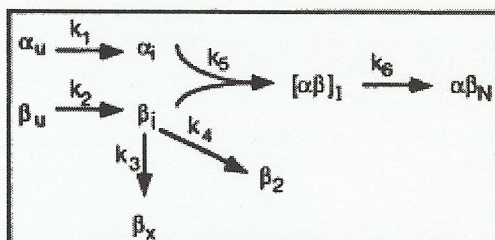


Fig. 2
Near ultraviolet circular dichroism spectra of β_2 in 50 mM phosphate, pH 7.0 at 18° with and without 5 M urea. Also shown is the circular dichroism spectrum of unfolded b subunit isolated by ion exchange chromatography of luciferase in 5 M urea. The protein concentration was the same for all three samples.



Scheme 1
Proposed kinetic mechanism for the refolding of luciferase subunits following dilution from 5 M urea into buffer (22). The individual subunits exhibit multiple kinetic phases which were combined to yield the single first order steps indicated in the Scheme. The kinetically preferred reaction is combination of the heterodimerization-competent species to yield the inactive heterodimeric species, $[ab]_I$, which isomerizes to yield the active enzyme. The b has two other kinetic options; when the concentration of the a subunit is very low and the concentration of b is comparatively high, the second order reaction yielding the b_2 homodimer becomes the kinetically preferred pathway. When the concentration of both subunits is low,

the b subunit will preferentially fold via
the first order pathway to yield the
folding-incompetent species, b_x .

The heterodimer and the homodimer have essentially the same thermodynamic stability. These and other observations regarding the folding and assembly of the luciferase subunits suggest that the basic tenet of protein folding, that the native

structure of a protein resides at a global minimum of free energy, may not be entirely correct. It appears that the formation of the native luciferase, or alternatively, formation of the b_2 homodimer, is kinetically determined, not thermodynamically determined. To advance our understanding of the structural basis of the remarkable kinetic stability of both the heterodimer and the b_2 homodimer, as well as to contribute to our understanding of the bioluminescence reaction, we have determined the structures of the heterodimer and the homodimer to high resolution by x-ray diffraction methods. The following discussion presents our current understanding of the structure and the implications of the structure in the context of the folding and assembly of the ab and b_2 structures.

Structure of the Luciferase Heterodimer

The subunits of bacterial luciferase assume the $(b/a)_8$ barrel structure (6,8) that appears to be one of the most common of protein folding motifs (23). The a and b subunits have been shown to be homologous, so it is not surprising that they assume similar structures (24-26).

A ribbon diagram of the enzyme, viewed down the local two-fold axis between the subunits, is shown in Figure 3. The rms deviation of one subunit from the other in the heterodimer is given in Table 1. It is obvious that the structures of the two subunits are very similar.

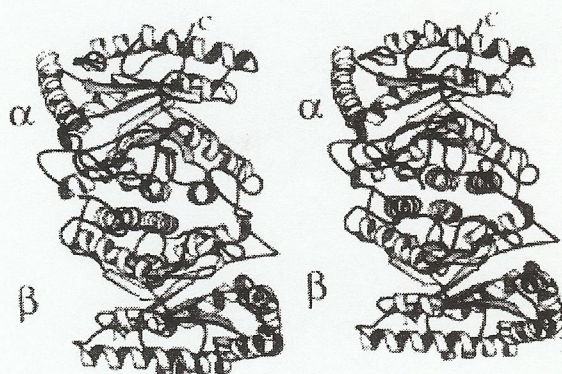


Fig. 3
Stereoview of the luciferase ab heterodimer down the pseudo two-fold symmetry axis. The N- and C-termini of both subunits are indicated. Note the location of the disordered loop of the a subunit between residues Asp 271 and Asn 289.

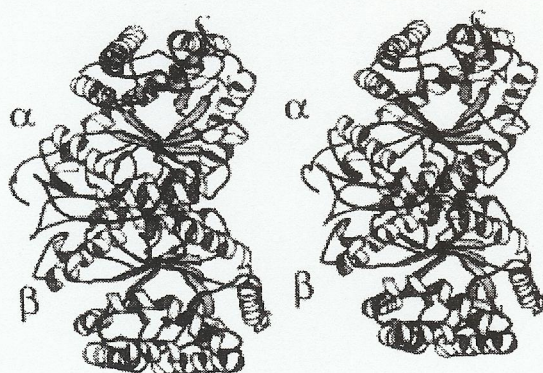


Fig. 4
Stereoview of the luciferase ab heterodimer perpendicular to the local two-fold axis between the two subunits. The view presented is a rotation of the view in Figure 3 to the left 90° .

Table 1. Comparisons of the rms differences between structurally-equivalent atoms in the subunits of the ab heterodimer and the b₂ homodimer.

	a subunit	b subunit	Subunit 2
	ab heterodimer	ab heterodimer	b ₂ homodimer
Subunit 1, b ₂ homodimer	6.2 Å ^a (6.0 Å) ^b	0.7 Å (1.2 Å)	0.63 Å (1.1 Å)
Subunit 2, b ₂ homodimer	6.2 Å (6.0 Å)	0.6 Å (1.1 Å)	

^aComparisons of a-carbon coordinates only.

^bComparisons of all atoms.

The a subunit of the heterodimer is thought to possess most if not all of the residues that comprise the active center of luciferase (1). While the structure that has been determined did not have flavin bound, it is possible to draw some conclusions based on the locations of residues thought to be involved in substrate binding and/or the chemical mechanism of the bioluminescence reaction. It is well known that the enzyme is exquisitely protease sensitive, and that flavin binding greatly reduces the accessibility of the protease labile region to protease attack. The protease sensitive region that is the site of the first, inactivating cleavage by proteases, is near residues 280-290 of the a subunit (25, 27-29). This region of the protein is disordered in the crystal structure, even at 1.5 Å resolution, suggesting that residues in this region are highly mobile, in accord with predictions based on limited proteolysis (30-32). It is highly suggestive that immediately "beneath" the region that must be occupied by the disordered loop resides a narrow opening into a large internal cavity that has been proposed to be the active center (6-8). Several of the mutant luciferases described by Cline and Hastings (33) as altered kinetics mutants have been cloned and the locations of the lesions determined (34,35). All of the residues that have been identified from these studies have amino acid side chains that contact this internal cavity. In addition, several mutants having dramatically altered activities in the bioluminescence reaction, constructed by Tu and his colleagues (36), also contact this internal cavity.

It has been proposed that the disordered loop binds to the surface of the luciferase following binding of the reduced flavin substrate (1,30-32). Such movement and binding could cover the opening to the internal cavity, thereby blocking access of the internal cavity to solvent water, which would likely be very effective both in causing nonluminescent breakdown of the C4a peroxydihydroflavin or the hemiacetal intermediates (37), and in quenching the excited state of the flavin, should it be formed. Much chemical modification work has focused on the reactive thiol of the a subunit; this residue resides at the opening to the internal cavity. Chemical modification studies have shown that the thiol must reside in or near a large hydrophobic pocket (38-40);

the structure confirms these conclusions (6-8). Clearly, chemical modification of this site would reduce access to the internal cavity, thereby hindering substrate binding. Furthermore, if substrate binding causes the disordered loop to become more ordered and to bind over the entrance to the internal cavity, the disordered loop could reduce access of chemical modification reagents to the thiol, thereby effectively protecting the thiol from chemical modification. It is well known that substrate binding protects the thiol from modification, and it is also well known that the thiol remains protected long after the end of a single catalytic cycle, suggesting that the altered conformation associated with binding of reduced flavin is slow to relax to the original form (32).

Another aspect of the protein that is of particular interest is the subunit interface. The two subunits are homologous, and the region having the greatest similarity in sequence is at the subunit interface (7). There is a local two-fold axis of symmetry between the two subunits that runs through the interface region, and between two parallel helices that are packed very closely together in the structure (6-8). The a subunit has 355 residues, while the b subunit has 324; the difference, 31 residues, resides in a long helix in which extends across the C-terminal end of the barrel structure to a stretch of coil which extends over the interface before becoming disordered in the structure. The disordered series of amino acid residues then reaches 25.4 Å across the barrel from residue Asp 271 to Asn 289, and the ordered structure then continues through folding patterns that are homologous to those of the b subunit. The loop that extends across the interface is located near the local two-fold axis and above a cleft that communicates with a small pocket in the middle of the interface region that is lined with charged/polar side chains (7,8). It has been proposed that the active center of luciferase resides at the subunit interface (41,42), and it appeared that if the active center were to reside at the interface, this pocket could be the active center. However, close inspection demonstrates that this region is highly unlikely to be the flavin or aldehyde binding site. First, the channel and pocket are too small to admit the isoalloxazine of the flavin. Second, the polar nature of the channel and pocket would not seem to be a good location for binding of the aliphatic aldehyde substrate. Third, there is now a very long list of enzymes that fold into the (b/a)₈ motif, and for all of these enzymes, the active centers consist of residues that are in the loops at the C-terminal ends of the b strands in the middle of the barrels. Fourth, the only lesion at the interface that alters the activity is the His 45 position described by Tu and his colleagues (36), and this residue also contacts the internal cavity accessed by the opening near the reactive thiol. Based on these many considerations, we believe that the active center of luciferase resides in the internal cavity, not at the subunit interface.

Structure of the b₂ Homodimer

Based on investigations of the unfolding and refolding of the b₂ homodimer, we had fully expected to find some form of domain swapping in its structure, thereby explaining to some extent the exceedingly slow association rate constant, and truly remarkable kinetic stability of the homodimer thus formed. However, the structure of the homodimer, refined to 1.95 Å and shown in Figure 5,

demonstrates no such excursion of one subunit into the other. Rather, the structure of the b subunit in the heterodimer is virtually identical to both of the subunits in the homodimer. The rms difference of the two subunits of the homodimer from each other and from the b subunit of the heterodimer is given in Table 1. It is apparent from these data that there is no greater difference between the structures of the two subunits in the homodimer than there is between either of the subunits in the homodimer and the b subunit of the heterodimer.

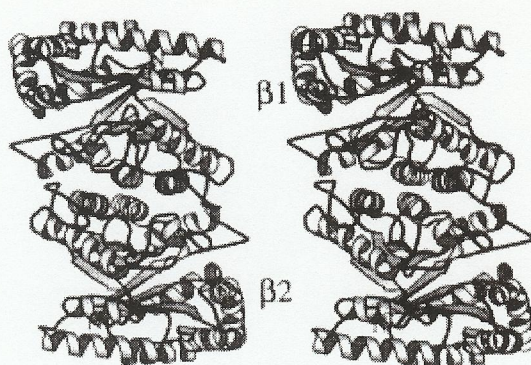


Fig. 5

Stereoview of the b₂ homodimer. The view is down the two-fold symmetry axis, as is the view of the heterodimer shown in Figure 3. The structure shown is of the kinetically stable species that does not unfold, even after prolonged incubation in 5 M urea at 18° (21).

The studies presented here were conducted both to support ongoing efforts to understand the luminescent flavin monooxygenase reaction that is catalyzed by bacterial luciferase, and to begin to develop an appreciation of a fundamental feature of protein structure and stability, the structural basis of slow conformational changes in proteins. The b₂ homodimer forms very slowly, and dissociates in buffer exceedingly slowly. We have estimated the half-time for the dissociation reaction at 18° in 50 mM phosphate, pH 7.0 to be on the order of 1×10^6 years (21). In 5 M urea, the homodimeric structure is stable indefinitely; at higher urea concentrations, the subunits do dissociate and unfold in an experimentally accessible time. This behavior is very different from that of the ab heterodimer, which unfolds very quickly indeed in 5 M urea. The two structures are very similar, and the interface regions are the most highly conserved between the two structures. Yet it is clear that there is an enormous kinetic barrier to be overcome upon unfolding of the homodimer relative to that of the heterodimer. It should be stressed here that the stability of the homodimer is a kinetic stability, *not* a thermodynamic stability. The thermodynamic stabilities of the homodimer and of the heterodimer are similar, but the kinetic stability of the homodimer is much greater than that of the heterodimer.

The association and dissociation reactions of the ab heterodimer and of the b₂ homodimer are excellent processes by which to investigate the underlying structural basis of slow conformational changes in proteins. It is clear from the results of our studies thus far that high resolution knowledge of the final structures of protein folding reactions does not immediately define the pathway by which the structures form, or the

pathways by which they unfold. Nor does such high resolution structural information explain how a structure can continue to exist indefinitely when the thermodynamic driving force strongly favors the unfolded species. To more fully understand and appreciate these critical details of protein structure and stability, we must develop a better understanding of the structures of the individual protomers as they are in the process of associating to form the biologically active species.

Acknowledgements

This research was supported in part by grants from the NIH (AR35186 and DK47814 to I.R. and H.M.H., respectively), NSF (BIR-9317398 shared instrumentation to I.R.), the Office of Naval Research (N00014-96-1-0087 and N00014-93-1-1345 to T.O.B.), and the Robert A. Welch Foundation (A-865 to T.O.B.).

References

1. Baldwin, T., Ziegler, M. (1992) in *Chemistry and Biochemistry of Flavoenzymes*, vol. III (F. Müller, ed.) CRC Press, Boca Raton, FL, p. 467.
2. Waddle, J., Baldwin, T. 1991. *Biochem. Biophys. Res. Commun.* **178**, 1188.
3. Sinclair, J., Waddle, J., Waddill, E., Baldwin, T. 1993. *Biochemistry* **32**, 5036.
4. Choi, H., Tang, C.-K., Tu, S.-C. 1995. *J. Biol. Chem.* **270**, 16813.
5. Friedland, J., Hastings, J.W. 1967. *Proc. Natl. Acad. Sci. U.S.A.* **58**, 2336.
6. Fisher, A., Raushel, F., Baldwin, T., Rayment, I. 1995. *Biochemistry* **34**, 6581.
7. Baldwin, T., Christopher, J., Raushel, F., Sinclair, J., Ziegler, M., Fisher, A., Rayment, I. 1995. *Curr. Opin. Struct. Biol.* **5**, 798.
8. Fisher, A., Thompson, T., Thoden, J., Baldwin, T., Rayment, I. 1996. *J. Biol. Chem.* **271**, 21956.
9. Thoden, J., Holden, H., Fisher, A., Sinclair, J., Wesenberg, G., Baldwin, T., Rayment, I. 1996. (submitted for publication).
10. Anfinsen, C. 1973. *Science* **181**, 223.

11. Epstein, C., Goldberger, R., Anfinsen, C. 1963. Cold Spring Harbor Symposium on Quantitative Biology **28**, 439.
12. Gething, M., McCammon, K., Sambrook, J. 1986. Cell **46**, 939.
13. Waddle, J., Johnston, T., Baldwin, T. 1987. Biochemistry **26**, 4917.
14. Sugihara, J., Baldwin, T. 1988. Biochemistry **27**, 2872.
15. Baker, D., Sohl, J., Agard, D. 1992. Nature **356**, 263.
16. Mottonen, J., Strand, A., Symersky, J., Sweet, R., Danley, D., Geoghegan, K., Gerard, R., Goldsmith, E. 1992. Nature **355**, 270.
17. Baker, D., Agard, D. 1994. Biochemistry **33**, 7505.
18. Ziegler, M., Goldberg, M., Chaffotte, A., Baldwin, T. 1993. J. Biol. Chem. **268**, 10760.
19. Baldwin, T., Ziegler, M., Chaffotte, A., Goldberg, M. 1993. J. Biol. Chem. **268**, 10766.
20. Clark, A., Sinclair, J., Baldwin, T. 1993. J. Biol. Chem. **268**, 10773.
21. Sinclair, J., Ziegler, M., Baldwin, T. 1994. Nature Struct. Biol. **1**, 320.
22. Clark, A., Raso, S., Sinclair, J., Ziegler, M., Chaffotte, A., Baldwin, T. 1996. (submitted for publication).
23. Farber, G., Petsko, G. 1990. Trends Biochem. Sci. **15**, 228.
24. Baldwin, T., Ziegler, M., Powers, D. 1979. Proc. Natl. Acad. Sci. **76**, 4887.
25. Cohn, D., Mileham, A., Simon, M., Nealson, K., Rausch, S., Bonam, D., Baldwin, T. 1985. J. Biol. Chem. **260**, 6139.
26. Johnston, T., Thompson, R., Baldwin, T. 1986. J. Biol. Chem. **261**, 4805.
27. Holzman, T., Baldwin, T. 1980. Proc. Natl. Acad. Sci. USA **77**, 6363.
28. Holzman, T., Riley, P., Baldwin, T. 1980. Arch. Biochem. Biophys. **205**, 554.

29. Rausch, S., Dougherty, J., Jr., Baldwin, T. (1982) in *Flavins and Flavoproteins* (V. Massey and C. Williams, eds.) Elsevier, Amsterdam, p. 375.
30. Baldwin, T., Riley, P. (1980) in *Flavins and Flavoproteins* (K. Yagi and T. Yamano, eds.) Japan Scientific Societies Press, Tokyo; University Park Press, Baltimore, p. 139.
31. Dougherty, J., Jr., Rausch, S., Baldwin, T. (1982) in *Flavins and Flavoproteins* (V. Massey and C. Williams, eds.) Elsevier, Amsterdam, p. 379.
32. AbouKhair, N., Ziegler, M., Baldwin, T. 1985. *Biochemistry* **24**, 3942.
33. Cline, T., Hastings, J. 1972. *Biochemistry* **11**, 3359.
34. Chen, L., Baldwin, T. 1989. *Biochemistry* **28**, 2684.
35. Chlumsky, L., Chen, L., Clark, C., Abu-Soud, H., Ziegler, M., Raushel, F., Baldwin, T. (1991) in *Flavins and Flavoproteins 1990* (B. Curti, S. Ronchi and G. Zanetti, eds.) Walter de Gruyter, Berlin, p. 261.
36. Xin, X., Xi, L., Tu, S.-C. 1991. *Biochemistry* **30**, 11255.
37. Eberhard, A., Hastings, J. 1972. *Biochem. Biophys. Res. Commun.* **47**, 348.
38. Nicoli, M., Meighen, E., Hastings, J. 1974. *J. Biol. Chem.* **249**, 2385.
39. Nicoli, M., Hastings, J. 1974. *J. Biol. Chem.* **249**, 2393.
40. Merritt, M., Baldwin, T. 1980. *Arch. Biochem. Biophys.* **202**, 499.
41. Fried, A., Tu, S.-C. 1984. *J. Biol. Chem.* **259**, 10754.
42. Paquette, O., Fried, A., Tu, S.-C. 1988. *Arch. Biochem. Biophys.* **264**, 392.