Kinetic Mechanism of Luciferase Subunit Folding and Assembly†

A. Clay Clark,[‡,§,||] Stephen W. Raso,[‡,⊥] James F. Sinclair,[‡,§,#] Miriam M. Ziegler,[‡,§] Alain F. Chaffotte,[‡,§,○] and Thomas O. Baldwin*[,‡,§,⊥]

Center for Macromolecular Design of the Institute of Biosciences and Technology and Departments of Biochemistry & Biophysics and of Chemistry, Texas A&M University, College Station, Texas 77843

Received October 2, 1996; Revised Manuscript Received November 25, 1996[⊗]

ABSTRACT: The kinetic mechanism *in vitro* of the folding and assembly of the heterodimeric flavin monooxygenase bacterial luciferase has been defined by a unique set of rate constants which describe both the productive refolding pathway and competing off-pathway reactions in 50 mM phosphate, pH 7.0 at 18 °C. The individual α and β subunits fold independently to form heterodimerization-competent species, α_i and β_i. The α_i and β_i species can interact to form an inactive heterodimeric intermediate, $[\alpha\beta]_I$, which isomerizes to form the active $\alpha\beta$ structure; the structure of the enzyme has been determined to 1.5 Å resolution [Fisher, A. J., Thompson, T. B., Thoden, J. B., Baldwin, T. O., & Rayment, I. (1996) *J. Biol. Chem. 271*, 21956−21968]. In the absence of α_i, β_i can form a kinetically trapped homodimer, β_2, with a second-order rate constant of about 180 M^{-1} s^{-1} [Sinclair, J. F., Ziegler, M. M., & Baldwin, T. O. (1994) *Nat. Struct. Biol. 1*, 320−326]; the structure of β_2 has recently been reported [Thoden, J. B., Holden, H. M., Fisher, A. J., Sinclair, J. F., Wesenberg, G., Baldwin, T. O., & Rayment, I. (1997) *Protein Sci. 6*, 13−23]. The β_i species, or some other form that precedes β_i on the refolding pathway, can also undergo a first-order conversion into a form (designated β_x) that cannot associate with α_i to form the native enzyme. The rate constant for this process, assigned here, accounts well for the previously observed dependence of final yield on concentration of refolding species [Ziegler, M. M., Goldberg, M. E., Chaffotte, A. F., & Baldwin, T. O. (1993) *J. Biol. Chem. 268*, 10760−10765]. In simulations of the refolding reaction, all processes associated with the refolding of the individual subunits were combined into single first-order rate constants for each subunit which were consistent with the rate constants determined from stopped-flow circular dichroism studies. The first-order rate constant for the folding of the α subunit, estimated from the concentration-independent lag preceding the appearance of active enzyme, and the second-order rate constant for assembly of α_i and β_i into the heterodimer, estimated from the concentration-dependent rate of appearance of active enzyme, were consistent with the rates of first- and second-order processes monitored by changes in fluorescence of an extrinsic probe [the product of modification with *N*-(4-anilino-1-naphthyl)maleimide] on the α subunit during refolding. The rate constant for the isomerization of $[\alpha\beta]_I$ to form the active heterodimer was estimated from the kinetic data of a secondary dilution experiment and from fluorescence measurements of protein diluted 20-fold from 2.1 M urea-containing buffer. The rate constants reported here for the kinetic mechanism of refolding permitted simulation of the time courses and yields for activity recovery during the refolding of luciferase from about 1 to 25 μg/mL which are in excellent agreement with our previously reported data.

The pioneering work of Anfinsen and colleagues led to the "thermodynamic hypothesis" of protein folding which states that "the three-dimensional structure of a native protein in its normal physiological milieu...is the one in which the Gibbs free energy of the whole system is lowest" (Anfinsen, 1973). This concept remains today axiomatic in the field of protein conformation and folding. Several models have been proposed to explain the folding properties and stabilities of globular proteins, all relying to a greater or lesser extent on the thermodynamic hypothesis (Srinivasan & Rose, 1995; Jaenicke, 1991; Creighton, 1990; Dill, 1985).

Because many proteins do not fold by a simple two-state mechanism, but rather form multiple transient intermediates during refolding, the possibility cannot be neglected that these intermediates may partition into nonproductive, off-pathway structures (Goldberg, 1985; Tandon & Horowitz, 1986; Waddle *et al.*, 1987; Sugihara & Baldwin, 1988; Goldberg *et al.*, 1991; Baker & Agard, 1994; Ranson *et al.*, 1995). Folding studies with α-lytic protease (Baker *et al.*, 1992), the serpin plasminogen activator inhibitor-1 (PAI-1) (Carrell *et al.*, 1991), and bacterial luciferase (Sinclair *et al.*, 1994) have shown that formation of the biologically active state of these proteins depends on the activation barrier between

† This work was supported by grants from the National Science Foundation (DMB 87-16262), the Office of Naval Research (N00014-93-1-0991 and N00014-93-1-1345), the Robert A. Welch Foundation (A865), and the Texas Agricultural Experiment Station.

* Address correspondence to this author at the Department of Biochemistry and Biophysics, Texas A&M University, College Station, TX 77843-2128. Telephone: (409) 845-1782. FAX: (409) 845-4946. E-mail: BALDWIN@bioch.tamu.edu.

‡ Center for Macromolecular Design of the Institute of Biosciences and Technology.

§ Department of Biochemistry & Biophysics.

|| Present address: Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, Box 8231, 660 S. Euclid Ave., St. Louis, MO 63110.

⊥ Department of Chemistry.

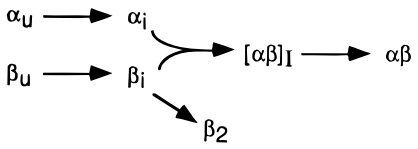
Present address: Department of Biological Chemistry, Johns Hopkins School of Medicine, 725 N. Wolfe St., Baltimore, MD 21205.

○ Present address: Institut Pasteur, 28 rue du Dr. Roux, 75724 Paris (CEDEX 15), France.

⊗ Abstract published in *Advance ACS Abstracts,* January 15, 1997.

S0006-2960(96)02477-4 CCC: $14.00

Scheme 1



the native state and other kinetically stable structures. Earlier work with bacterial luciferase showed that formation of the active heterodimer required that the two subunits fold within the same cell at the same time (Waddle *et al.*, 1987). In kinetically controlled processes, the relative yield of products formed is dependent on the relative rates of formation of those products rather than on the differences in free energies of those products. These examples show that a complete understanding of a protein folding pathway for most larger proteins may be attained only with the inclusion of detailed information on rates of formation and conversion of intermediate structures that occur on the pathway to the formation of the native conformation.

The folding pathway of bacterial luciferase has been shown to involve several intermediates, including both on-pathway and off-pathway species (Ziegler *et al.*, 1993; Baldwin *et al.*, 1993; Clark *et al.*, 1993; Sinclair *et al.*, 1994). This pathway has been summarized in a recent review of the luciferase structure by Baldwin *et al*. (1995). The unfolded subunits, α_u and β_u, undergo isomerizations to form heterodimerization-competent species, α_i and β_i, as shown in Scheme 1. These species can interact to form an inactive heterodimeric intermediate, $[\alpha\beta]_I$, which isomerizes to form the active $\alpha\beta$ structure. An inactive heterodimeric intermediate was also shown to be present under equilibrium conditions (Clark *et al.*, 1993). In the absence of α_i, β_i can form a kinetically trapped homodimer, β_2, whose dissociation rate is so slow that it is not in equilibrium with β_i and thus cannot associate with α_i to form native enzyme (Sinclair *et al.*, 1994). The structure of this β_2 species at 1.95 Å resolution has been determined (Thoden *et al*., 1997).

We describe here the kinetic mechanism *in vitro* of folding and assembly of bacterial luciferase in 50 mM phosphate, pH 7.0 at 18 °C. The rate constants for formation of secondary structure upon dilution of unfolded α and β subunits from urea-containing buffer were determined by monitoring the circular dichroism at 222 nm; both subunits undergo multiple transitions during refolding. The rate of heterodimerization to form $[\alpha\beta]_I$ was estimated by simulations of previously reported data on the rate of recovery of activity during refolding (Ziegler *et al.*, 1993) and agrees well with the apparent second-order rate constant determined by measuring the effect of the β subunit concentration on refolding of α subunit labeled with the fluorescent probe *N*-(4-anilino-1-naphthyl)maleimide (ANM).[1] Simulations of these fluorescence data as well as those for recovery of activity during refolding reported by Ziegler *et al.* (1993) indicate that the heterodimerization-competent form of the β subunit, β_i, or some species that precedes it on the refolding pathway, undergoes a first-order transition to form a kinetically stable species, which we have called β_x, that is not in equilibrium with β_i. These data show that formation of heterodimer is the kinetically preferred process when the α and β subunits are both present in a (re)folding mixture. When the concentration of α subunit is low, such that the heterodimerization reaction becomes rate-limiting, the kinetically preferred folding pathway of the β subunit yields β_2. When the β subunit concentration is very low, the competing first-order formation of β_x results in reduced yields of both $\alpha\beta$ and β_2.

MATERIALS AND METHODS

Enzyme and Chemicals. *Vibrio harveyi* luciferase was overexpressed in *Escherichia coli* and purified as described previously (Baldwin *et al.*, 1989). Our overexpression protocol results in accumulations of luciferase to levels over 50% of the soluble protein in *E. coli*. Urea (Ultra Pure) was purchased from Schwarz/Mann Biotech or Gibco BRL, ANM and β-mercaptoethanol were obtained from Sigma, and dithiothreitol and Tween-20 were from Boehringer Mannheim Biochemicals. "Phosphate buffer" refers to NaH_2PO_4/K_2HPO_4, pH 7.0. All other chemicals were reagent grade.

Stopped-Flow Circular Dichroism Measurements of the Refolding of Luciferase Subunits. The α and β subunits of luciferase for stopped-flow CD studies were separated using ion exchange chromatography (DEAE Sephadex A-50) of luciferase samples in 5 M urea, as described previously (Baldwin *et al.*, 1993), and were diluted 1:50 into 50 mM phosphate buffer, pH 7.0, to initiate the refolding reaction. Data were collected on a CD6 spectrodichrograph (Jobin-Yvon Instruments, Longjumeau, France) equipped with a Biologic stopped-flow attachment essentially as previously described (Chaffotte *et al.*, 1992) except that samples were maintained at 18 °C.

Separation of Luciferase α *and* β *Subunits by Ion Exchange Chromatography on Q Sepharose.* An ion exchange column of Q Sepharose (2.6 × 10 cm; Pharmacia) was equilibrated at 4 °C with 20 mM phosphate buffer containing 5 M urea, 1 mM EDTA, and 1 mM dithiothreitol, pH 7.0. Approximately 300 mg of heterodimer in 5 M urea was applied to the column, and the column was washed with 75 mL of 20 mM phosphate buffer containing 5 M urea. The subunits were eluted with a linear gradient (total volume of 425 mL) from 20 mM phosphate to 120 mM phosphate (5 M urea, 1 mM EDTA, and 1 mM dithiothreitol, pH 7.0) with a flow rate of 2 mL/min. Protein in the eluate was monitored by absorbance at 280 nm, and fractions with the highest A_{280} were analyzed by polyacrylamide gel electrophoresis in sodium dodecyl sulfate (Laemmli, 1970) to determine which fractions to pool. Base line resolution was obtained between elution positions of the subunits. The isolated subunits were concentrated using an Amicon concentrator and stored at −20 °C. The isolated β subunit was used in the refolding experiments described below.

Carbamoylation of proteins may occur as a result of incubation in high concentrations of urea for prolonged periods of time [cf. Pace (1986)]. Such modification may affect the behavior of a refolding protein. Appropriate control experiments were performed to assure that no such modification was observed during the refolding of luciferase. First, Clark *et al*. (1993) showed that the native fluorescence and CD signals returned completely after 18−24 h of incubation in 6 M urea. Second, Ziegler *et al*. (1993) demonstrated that the rate of activity return upon 50-fold

[1] Abbreviations: ANM, *N*-(4-anilino-1-naphthyl)maleimide; CD, circular dichroism; DTT, dithiothreitol; em, emission; exc, excitation; Tween-20, poly(oxyethylene)$_{20}$ sorbitan monolaurate.

dilution out of 5 M urea was independent of the time of incubation, whether it be several minutes, 3−4 h, or about 24 h.

Labeling of α Subunit with N-(4-Anilino-1-naphthyl)-maleimide and Refolding of the Labeled Enzyme. The cysteinyl residue at position 106 of the α subunit (Nicoli & Hastings, 1974; Nicoli *et al.*, 1974; Cohn *et al.*, 1985) was labeled with the fluorescent probe ANM (Kanaoka *et al.*, 1973) by mixing heterodimeric luciferase (50 μM) in 50 mM phosphate, pH 7.0, at room temperature with the reagent. Modification of this residue results in the loss of enzymatic activity. The labeling reaction was quenched with excess β-mercaptoethanol when the residual activity was approximately 12%. Labeled protein was separated from excess reagents by Sephadex G25 column chromatography; from the absorbance at 280 and 355 nm and the extinction coefficients of 1.13 $(mg/mL)^{-1}$ cm^{-1} at 280 nm (protein; Sinclair *et al.*, 1993) and 13 180 M^{-1} cm^{-1} at 355 nm (ANM; Kanaoka *et al.*, 1973), the molar ratio of probe to protein was estimated to be about 0.84. The fluorescence of the ANM-labeled α subunit was monitored during refolding from urea-containing buffer using an SLM 8000C spectrofluorometer. An excitation wavelength of 355 nm and an emission wavelength of 428 nm were used.

Determination of the Rate Constant for Isomerization of the Heterodimeric Intermediate to Native Luciferase. Luciferase (38 μg/mL) was incubated in 50 mM phosphate buffer containing 2.1 M urea, pH 7.0 at 18 °C for 16 h. One hundred microliters of the solution was mixed rapidly with 1.90 mL of phosphate buffer, stirring in a 1 × 1 cm fluorescence cuvette at 18 °C. The final luciferase concentration was 1.9 μg/mL. The decrease in tryptophanyl fluorescence at 330 nm was monitored using an SLM 8000C spectrofluorometer. The samples were maintained at 18 °C, and an excitation wavelength of 295 nm was used.

Simulation of the Time Course of Activity Recovery for Refolding Luciferase. The data analyzed here, described previously by Ziegler *et al.* (1993), were simulated using the program Hopkinsim [a version of Kinsim (Barshop *et al.*, 1983) developed for use with a Macintosh computer], the kinetic model shown in Scheme 2 of the text, and the set of rate constants shown in Table 2. The rate constants were determined from the experimental results of Ziegler *et al.* (1993), Baldwin *et al.* (1993), and Sinclair *et al.* (1994) and by agreement of the simulated time courses with the experimental data presented here, as described under Results.

RESULTS

Rate of Formation of Secondary Structure during Refolding of Luciferase α and β Subunits. The luciferase subunits used in these experiments were separated by ion-exchange chromatography in 5 M urea-containing buffers; the refolding of each unfolded subunit was monitored by circular dichroism at 222 nm following rapid 50-fold dilution into phosphate buffer. For both subunits, the formation of secondary structure, as monitored by circular dichroism at 222 nm, was a multistep process (Figure 1). Approximately 63% of the final amplitude for each sample was obtained within the dead-time of mixing, ca. 4 ms. The initial burst phase was followed by two kinetic phases. The observed rate constants for the transitions detected by the far-UV CD signal are given in Table 1.

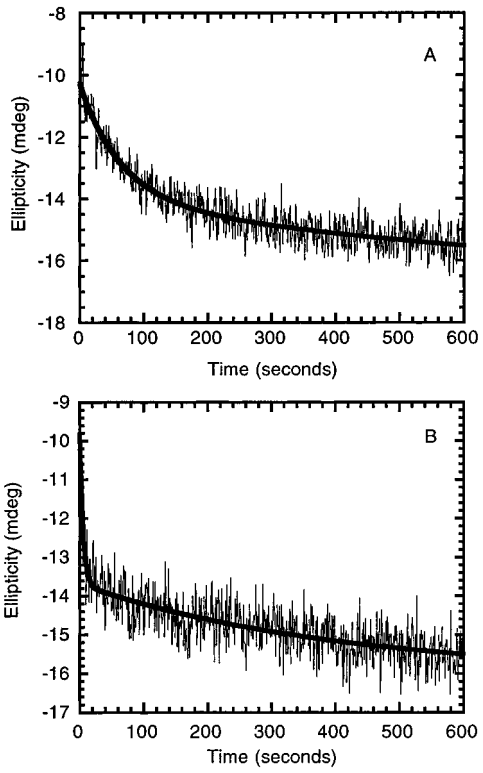


FIGURE 1: Stopped-flow circular dichroism measurements of the refolding of luciferase subunits. The α and β subunits of luciferase were separated as described in under Materials and Methods and rapidly diluted 50-fold from 5 M urea-containing buffer into 50 mM phosphate buffer (pH 7.0) to initiate the refolding reaction. The temperature was held constant at 18 °C, and the final protein concentration was 50 μg/mL. The initial signal (in 5 M urea) was 0 mdeg, and the final signal approached −15.8 mdeg. The data are an average of 50 refolding experiments. (A) Refolding of luciferase α subunit. (B) Refolding of luciferase β subunit. The thick solid lines are fits of the data to a sequential reaction mechanism (initial burst followed by two exponential phases). The rate constants for the two exponential phases obtained from these fits are given in Table 1.

Table 1: Rate Constants and Amplitudes of Circular Dichroism Detected Transitions

	α subunit		β subunit	
phase	rate constant (s^{-1})	amplitude	rate constant (s^{-1})	amplitude
burst	>250	0.66	>250	0.62
intermediate	0.0238	0.16	0.209	0.23
slow	0.0045	0.19	0.0027	0.15

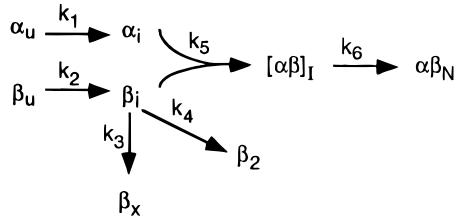
Simulations of Luciferase Folding and Assembly. Earlier studies of the kinetics of refolding of luciferase based on the recovery of active enzyme ($\alpha\beta$) (Ziegler *et al.*, 1993; Baldwin *et al.*, 1993), and refolding of the β subunit alone (Sinclair *et al.*, 1993, 1994), demonstrated several features which must be accounted for in any kinetic scheme:

(1) A concentration-independent lag of about 3 min prior to appearance of active enzyme (Ziegler *et al.*, 1993), shown to be due to folding of the individual subunits. The overall folding is somewhat slower for the α subunit than for the β subunit (Baldwin *et al.*, 1993).

(2) A concentration-dependent rate of formation of active enzyme following the lag, presumably reflecting the heterodimerization step.

(3) A saturation of the concentration dependence of the rate of formation of active enzyme at concentrations above

Scheme 2



about 25 μg/mL, suggesting that at high concentrations a first-order process following heterodimerization becomes rate-limiting.

(4) A continuation of the rapid rate of activity recovery characteristic of the high protein concentration for about 2 min following a secondary 10-fold dilution of the refolding mixture, eventually slowing to a rate characteristic of the lower protein concentration, a result which was interpreted as evidence for an isomerization step between inactive heterodimer and active heterodimer (Ziegler *et al.*, 1993).

(5) The loss of ability of β subunit folded in the absence of α to interact subsequently with α (Baldwin *et al.*, 1993; Sinclair *et al.*, 1993), shown to be due at least in part to formation of a kinetically trapped β_2 species (Sinclair *et al.*, 1994).

(6) A markedly reduced yield of active heterodimer at low protein concentrations (Ziegler *et al.*, 1993), reflecting an additional off-pathway process which appears to be first order. Since the α subunit refolded in the absence of β retains its ability to associate with β and subsequently to form active enzyme, this first-order off-pathway process must involve the β subunit.

The sixth condition required modification of Scheme 1 by addition of a first-order conversion of β_i to a species (β_x) incompetent to associate with α, i.e., not in equilibrium with β_i. A kinetic mechanism accounting for all of these observations is shown in Scheme 2.

The following constraints were used to obtain the best agreement of the simulated time course with the experimental data of Ziegler *et al.* (1993). Initial estimates for the rate constants k_1 and k_2 for the first-order folding steps of the individual α and β subunits, respectively, were derived from the slow phases of the data presented in Figure 1, and were subsequently adjusted to accommodate the duration of the lag shown in Figure 2. It is clear that multiple rate constants for the isomerizations that occur during the transformations $\alpha_u \rightarrow \alpha_i$ and $\beta_u \rightarrow \beta_i$ are contained within k_1 and k_2. However, for simplicity, we have reduced the numerous rate constants for folding of each subunit to the single first order rate constants k_1 and k_2; the best-fit rate constants of 0.003 and 0.006 s^{-1} for the folding of α and β, respectively, are of the same order as the slowest observed transitions in the recovery of the circular dichroism signal (Figure 1 and Table 1).

The rate constant for the isomerization step $[\alpha\beta]_I \rightarrow \alpha\beta_N$, k_6, was estimated from the secondary dilution experimental data of Ziegler *et al.* (1993), and from fluorescence changes of protein diluted 20-fold from 2.1 M urea, both discussed below. The second-order rate constant for formation of the β_2 homodimer, k_4, was determined previously (Sinclair *et al.*, 1994) to be 180 M^{-1} s^{-1}. The rate constant for the proposed first-order transition of $\beta_i \rightarrow \beta_x$, k_3, was adjusted to account for the concentration-dependent yield of active

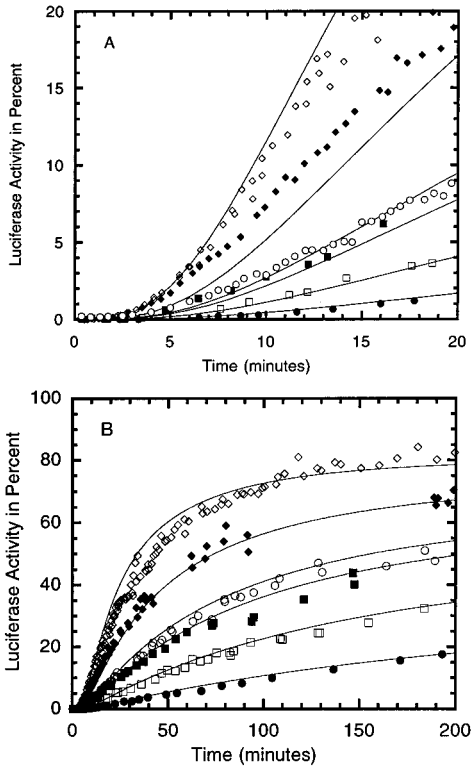


FIGURE 2: Simulation of refolding of luciferase as monitored by recovery of activity following 50-fold dilution of protein from 5 M urea into 50 mM phosphate buffer, pH 7.0, 18 °C. The data, reported in Ziegler *et al.* (1993), are shown in panels A and B with different time axes. Each data point represents a single enzyme activity determination in which an aliquot of the refolding reaction mixture was diluted into assay buffer and the luciferase activity determined by rapid injection of $FMNH_2$. Refolding reactions typically were monitored for a period of 20−24 h. The solid lines are calculated time courses based on the kinetic model shown in Scheme 2 and the rate constants given in Table 2. Each curve was simulated with the same set of rate constants, but the protein concentrations were varied as in the original experiment. The protein concentrations used were the following (Ziegler *et al.*, 1993): (●) 0.8 μg/mL; (□) 2 μg/mL; (■) 4 μg/mL; (○) 5 μg/mL; (◆) 10 μg/mL; (◇) 25 μg/mL.

heterodimer at low concentrations; at high protein concentrations, the yield is also compromised, apparently due to aggregation of some species on the refolding pathway (Ziegler *et al.*, 1993), but no attempt has been made to simulate the aggregation phenomenon. The second-order heterodimerization rate constant, k_5, was estimated from the effect of protein concentration on the rate of formation of active enzyme so as to obtain the best agreement between the simulated time courses and the experimental data. Finally, all of the rate constants were systematically varied both independently and in combinations until a unique set of rate constants was obtained that gave the best correlation between the experimental data and the simulated time courses for all of the protein concentrations. These rate constants are given in Table 2, and the results of the simulations are shown as the solid lines in Figure 2.

There is good agreement between the simulated time courses of refolding and the experimental data from approximately 1 to 25 μg/mL luciferase (Figure 2). The simulations show a protein concentration-independent lag of approximately 3 min followed by a protein concentration-dependent increase in the rate of formation of active protein. The rates of the first-order folding processes of the individual subunits prior to heterodimerization reflect the slower folding

Table 2: Rate Constants of Luciferase Refolding[a]

	rate constants					
experiment	k_1 (s^{-1})	k_2 (s^{-1})	k_3 (s^{-1})	k_4 ($M^{-1}\ s^{-1}$)	k_5 ($M^{-1}\ s^{-1}$)	k_6 (s^{-1})
activity	0.003	0.006	0.00004	180	2400	0.008
ANM	0.002	0.006	0.00004	180	2400	0.00027

[a] Simulations of the data in Figures 2 through 5 were done using the folding model shown in Scheme 2 as described in the text. Activity data for wild-type luciferase are from Ziegler *et al.* (1993). ANM refers to the refolding of ANM-labeled α subunit in the presence of various concentrations of unlabeled β subunit (Figure 3).

of the α subunit (Tables 1 and 2), and are in approximate agreement with the rates of formation of secondary structure for both subunits (Figure 1 and Table 1). It is known that at concentrations of 25 μg/mL and above, the yield of active enzyme is compromised, apparently due to aggregation (Ziegler *et al.*, 1993). In this higher concentration range, the calculated activities were higher than the experimental data, reflecting the aggregation process which was not included in the simulations. Measurements of light scattering have shown that aggregation occurs during refolding in this concentration range, while at 10 μg/mL and below, aggregation is minimal (data not shown). Little is known about this aggregation beyond the fact that it occurs, so it was not included in the simulations presented here.

The rate constants given in Table 2 were also used to simulate the time courses of several other refolding experiments, as described below, to test the kinetic mechanism of Scheme 2 (and the rate constants themselves) against (a) the time course of the change in signal of a probe other than activity recovery (an extrinsic fluorescent probe), (b) data suggesting a first-order off-pathway process (reduced yield of active enzyme at low protein concentrations; Ziegler *et al.*, 1993), and (c) data suggesting a final isomerization of an inactive heterodimer $[\alpha\beta]_I$ to the active heterodimer (a secondary dilution experiment; Ziegler *et al.*, 1993).

Effect of β Subunit Concentration on Refolding of α Subunit Labeled with the Fluorescent Probe ANM. The α subunit of luciferase was labeled specifically at the highly reactive thiol at position α106 (Baldwin & Ziegler, 1992; Baldwin *et al.*, 1995) using the fluorescent probe *N*-(4-anilino-1-naphthyl)maleimide (ANM). The fluorescence quantum yield of this probe is highly sensitive to solvent polarity (Kanaoka *et al.*, 1973), so it provides an excellent signal for monitoring changes in environment during refolding. As shown in Figure 3, upon 50-fold dilution of labeled luciferase from urea-containing buffer, the fluorescence of the extrinsic probe increased rapidly in an initial burst phase that encompassed approximately 75% of the maximum fluorescence signal, similar to the burst phase for formation of secondary structure in the α subunit as measured by circular dichroism at 222 nm (Figure 1A). The initial burst phase was followed by a slower fluorescence increase with a rate independent of the β subunit concentration. Following these first-order processes, the probe fluorescence emission decreased at a rate that was dependent on the concentration of the β subunit, suggesting that this third phase reflected the (second-order) association of the α and β subunits. The results also suggest that the environment of the probe in the α subunit is more nonpolar than the environment of the probe in the heterodimer.

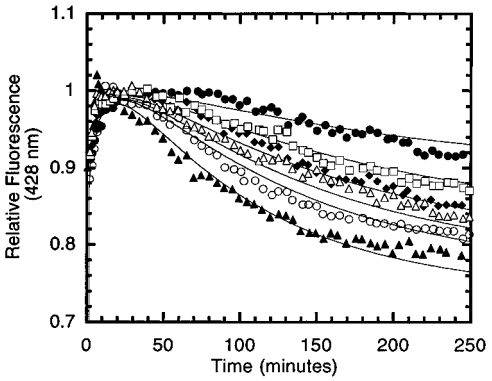


FIGURE 3: Effect of β subunit concentration on the refolding of ANM-labeled α subunit. The α subunit was labeled in the heterodimer as described under Materials and Methods; the heterodimer was unfolded in 5 M urea-containing buffer (50 mM phosphate, pH 7.0, 1 mM dithiothreitol, and 0.005% Tween-20). The unfolded ANM-labeled heterodimer was mixed with various concentrations of unfolded β subunit (in 5 M urea) that had been separated from the α subunit by ion-exchange chromatography (see Materials and Methods). Refolding was initiated by a rapid 50-fold dilution into buffer (50 mM phosphate, pH 7.0, 1 mM dithiothreitol, and 0.005% Tween-20) at 18 °C and monitored by the fluorescence of the probe at 428 nm (excitation wavelength 355 nm). The concentration of ANM-labeled α subunit was constant at 1 μg/mL. The final concentrations of β subunit were the following: (●) 1 μg/mL (sample with denatured heterodimer alone); (□) 2 μg/mL; (◆) 3 μg/mL; (△) 4 μg/mL; (○) 5 μg/mL; (▲) 10 μg/mL. All data were normalized between the starting signal and the maximum signal to give a range between 0 and 1. The amplitudes for α_u, α_i, $\alpha\beta_i$, and $\alpha\beta_N$ are 0, 0.722, 0.804, and 0.604, respectively. Solid lines represent simulations of the data using the folding model shown in Scheme 2 and the rate constants from Table 2, except that the faster of the two measurable rate constants obtained from the CD data (Figure 1, Table 1) was included in order to simulate the burst phase.

The rate constants given in Table 2, derived from the kinetics of recovery of luciferase activity during refolding, were used to simulate the kinetics of the change in fluorescence of the ANM-labeled luciferase during refolding in the presence of different concentrations of β subunit, and adjusted to obtain the best agreement between the experimental data and the calculated progress curves. The resulting simulations are in excellent agreement with the experimental data (Figure 3); the rate constants used to calculate the simulated curves are given in Table 2. In the simulations we assume that both the burst phase and the slower first-order increase in fluorescence report folding events that occur within the α subunit prior to heterodimerization. The burst phase accounted for ca. 85% of the maximum amplitude and had a rate constant $>0.06\ s^{-1}$. To better simulate the slower first-order increase in fluorescence, a rate constant of 0.002 s^{-1} was used, rather than the rate constant of 0.003 s^{-1} that was used to simulate the lag in appearance of activity (Figure 2 and Table 2). The slowest observed rate constant for formation of the circular dichroism signal at 222 nm was 0.0027 s^{-1} (Table 1). The first-order transitions were followed by a quenching of fluorescence emission at a rate dependent on the concentration of the β subunit. The rate of heterodimerization, k_5, obtained from simulations of these data was in exact agreement with that obtained from simulations of the return of activity for the β subunit concentration range 1−10 μg/mL (Figure 2 and Table 2). It is not clear why simulations for the folding of the labeled protein (Figure 3 and Table 2) indicate a much slower rate for the final isomerization, k_6, than the activity measurements

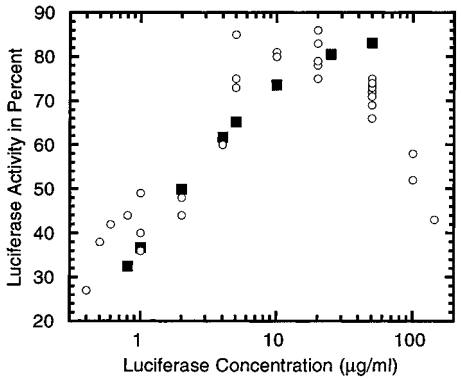


FIGURE 4: Comparison of the final yields of active luciferase determined experimentally with those from the simulated folding reaction. Luciferase was refolded from urea-containing buffers as described (Ziegler *et al.*, 1993) at the final concentrations shown. Simulations were performed as described under Materials and Methods using the folding model shown in Scheme 2 and the rate constants given in Table 2. To determine the final yield of active luciferase, the folding reactions were simulated for 24 h (■). The experimentally determined yields are from Ziegler *et al.* (1993) (○).

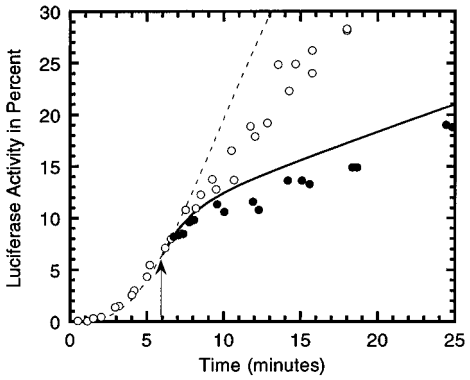


FIGURE 5: Comparison of experimentally determined and simulated secondary dilution reactions of refolding luciferase. The experimental results are from Ziegler *et al.* (1993). The concentration of each species shown in Scheme 2 at the time of the secondary 10-fold dilution (6 min after initial 50-fold dilution out of 5 M urea) was determined from simulations and divided by 10. Simulation of the secondary dilution reaction was performed using those concentrations of species. (○) original sample of refolding luciferase, 50 μg/mL (Ziegler *et al.*, 1993); (●) secondarily diluted sample of luciferase, 5 μg/mL (Ziegler *et al.*, 1993). The dashed line represents a simulation of the time course of refolding of the original sample of luciferase as shown in Figure 2, and the solid line represents a simulation of the time course of refolding of the secondarily diluted sample of luciferase. The arrow indicates the time of the secondary dilution.

(Figure 2 and Table 2). However, this isomerization rate constant was determined independently by other methods (see below) to be 0.0085 s^{-1}, in good agreement with the activity data. The ANM label may be sensitive to slower conformational changes which occur after native activity is obtained. This possibility is explored further under Discussion.

It has been shown that the final yield as well as the rate of formation of active protein is compromised at low protein concentrations (Ziegler *et al.*, 1993). The data for the return of activity upon dilution from urea-containing buffer, shown in Figure 2, were simulated for a 24 h period, and the calculated final yields of active protein were compared to the experimentally determined values as shown in Figure 4. The calculated final yields of active luciferase are in excellent agreement with the experimentally determined yields over a protein concentration range of 0.8−25 μg/mL. At protein concentrations above 25 μg/mL, the simulated final yield of active protein continued to increase with concentration, whereas the observed percent yields decreased. As noted above, the lower final yield at high protein concentrations is due to aggregation of an intermediate species on the folding pathway.

The rate of isomerization of $[\alpha\beta]_I \rightarrow \alpha\beta_N$, k_6, was determined in two ways. First, the rate constant was estimated from a secondary dilution experiment (Ziegler *et al.*, 1993): luciferase was diluted 50-fold from 5 M urea to 0.1 M urea and a protein concentration in the refolding mixture of 50 μg/mL, giving the maximum rate of formation of active enzyme. After 6 min of refolding, the sample was diluted another 10-fold so that the final protein concentration was 5 μg/mL, and the rate of formation of new heterodimer should have been limited by the protein concentration (see Figure 2). Following the secondary dilution, the rate of formation of active protein continued for approximately 1−2 min at the rate obtained for the *higher* protein concentration, as if no secondary dilution had occurred, and then the rate decreased to that obtained for the lower protein concentration. This behavior was interpreted as evidence for the existence of an inactive heterodimeric intermediate present at the time of the secondary dilution, which undergoes isomerization to the active structure ($[\alpha\beta]_I \rightarrow \alpha\beta_N$ in Scheme 2). Above 25 μg/mL protein, the rate of isomerization of the inactive heterodimeric intermediate to the native state becomes limiting. As shown in Figure 5, simulation of the time course of the secondary dilution experiment agrees well with the experimental data of Ziegler *et al.* (1993). The differences between the simulated and experimental data probably reflect error in the concentrations of the species calculated by simulations at the time of the secondary dilution, due to the aggregation occurring in the 50 μg/mL sample following dilution.

To obtain an independent measure of the isomerization rate constant, luciferase was incubated in buffer containing 2.1 M urea, a concentration of urea which has been shown to maximally populate the $[\alpha\beta]_I$ species (Clark *et al.*, 1993). The protein was diluted into buffer, and the change in fluorescence was monitored as a function of time (Figure 6). This transition represents the first-order isomerization of the heterodimeric intermediate to the native conformation, the rate of which is independent of protein concentration. The data were fit to a single exponential process which yielded a rate constant of 0.0085 s^{-1}. The change in tryptophanyl fluorescence upon similar treatment of the ANM-labeled luciferase showed similar kinetics to those of the unmodified enzyme (data not shown), demonstrating that the slower rate observed for this transition is not due to physical interference by the probe.

DISCUSSION

The kinetic scheme and the associated rate constants reported here were developed by a combination of fitting of experimental data to exponential rate expressions and numerical simulation of processes that were not directly approachable by experiment. The only process for which the rate constant could not be directly measured and for which we have no independent verification is the isomer-

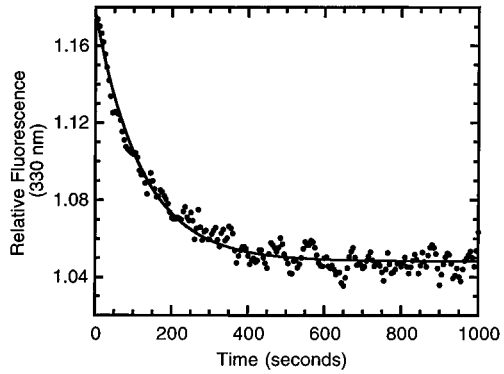


FIGURE 6: Rate of isomerization from heterodimeric intermediate $[\alpha\beta]_I$ to native luciferase. Luciferase (38 μg/mL) was incubated at 18 °C in 50 mM phosphate, 1 mM DTT, pH 7.0, containing 2.1 M urea for 18 h, conditions under which the heterodimeric intermediate is maximally populated. The protein was diluted 20-fold out of urea, and the change in tryptophanyl fluorescence was monitored at 330 nm. The excitation wavelength was 295 nm, and the final protein concentration was 1.9 μg/mL. The data were fit to a single exponential, which yielded a first-order rate constant of 8.5×10^{-3} s^{-1}.

ization of $\beta_i \rightarrow \beta_x$. The other process that could not be measured directly was the second-order heterodimerization process, k_5. However, this process was confirmed by determination of the effect of various concentrations of β subunit on the rate of folding of the ANM-labeled α subunit (Figure 3). In the simulations, the duration of the lag phase (Figure 2) was sensitive to the values of k_1 and k_2, the concentration dependence of the rate of activity return was sensitive primarily to the heterodimerization rate constant, k_5, the final yield was determined by the off-pathway processes k_3 and k_4, and the isomerization rate constant k_6 determined the concentration range over which the rate of activity return became concentration-independent, as well as the behavior of the secondary dilution experiment depicted in Figure 5. Direct measurement of k_4 (Sinclair *et al*., 1994) and k_6 (Figure 6) afforded constraints on the numerical simulations and greatly enhanced our confidence in the validity of the set of rate constants. To further evaluate the rate constants, we have (Clark, 1994) varied each rate constant independently and in combination with other rate constants by at least an order of magnitude above and below the best fit value and evaluated the quality of the resulting simulations. This approach was intended to explore whether or not the set of rate constants presented in Table 2 comprises a global best fit to the data; the results of the evaluation of the data convince us that we do indeed have a global best fit and not a local minimum in the simulation process. This conclusion is further supported by the fact that using these rate constants and varying only the protein concentration, we can simulate all available kinetic folding data for studies with bacterial luciferase in 50 mM phosphate, pH 7.0 and 18 °C, over the concentration range 1−25 μg/mL.

The data presented here demonstrate that the second-order rate constant for heterodimerization, k_5, is over 10-fold greater than that for the formation of the homodimeric form of the β subunit, k_4. Since the rate of any bimolecular reaction is determined by the concentrations of the interacting species, at any given concentration of β subunit, the fractional partition between formation of $\alpha\beta$ luciferase and the β_2 homodimer would be determined by the concentration of the α subunit. At low concentrations of α, the slow homodimerization reaction would predominate. At low concentrations of both α and β, the very slow conversion of $\beta_i \rightarrow \beta_x$ would be expected to compromise the overall yield of both $\alpha\beta$ and β_2. A mixture of free α subunit and the β_2 homodimer at the appropriate concentrations would, based solely on thermodynamic considerations, tend to form $\alpha\beta$, but this process cannot occur since the rate of dissociation of β_2 is very slow (1.6×10^{-14} s^{-1}; half-time of about 1 million years; Sinclair *et al.*, 1994). Clearly, these data confirm our hypothesis that the process of folding and assembly of luciferase is kinetically controlled (Sugihara & Baldwin, 1988). The partition of folding polypeptide into $\alpha\beta$ and β_2 + α is based on kinetic, not thermodynamic, parameters. Failure of the system to achieve partitioning in proportion to thermodynamic parameters results from the enormous activation barriers that are apparent from the slow rates of interconversion of the various species.

The structure of the β_2 homodimer has been determined at 1.95 Å and recently reported (Thoden *et al*., 1997). Comparison with the structure of the $\alpha\beta$ heterodimer (Fisher *et al*., 1996) demonstrates that the structure of the β subunit in the heterodimer is the same as the structures of the two β subunits in the homodimer; the rms deviation of α carbon coordinates of the two β subunits in the homodimer was 0.63 Å, while the rms deviations of the α carbon coordinates of the β subunit of the heterodimer and each of the two subunits of the homodimer were 0.7 and 0.6 Å, respectively. The homology of the α and β subunits is apparent from the conserved three-dimensional fold (Fisher *et al*., 1995). The most conserved region of the protein is at the subunit interface of the heterodimer. It is indeed of great interest that the two species, $\alpha\beta$ and β_2, are so similar in three-dimensional structure and the subunit interfaces are very highly conserved, yet the association and dissociation reactions are kinetically so very different.

The discrepancy in the rate constant for the final isomerization to native luciferase is itself very interesting. Simulations of activity recovery, including the secondary dilution experiment, and the measurement of the tryptophan fluorescence during refolding indicate that the rate constant is on the order of 0.008 s^{-1}. However, simulations of the refolding monitored by the fluorescence of ANM covalently linked to the α subunit indicated a much slower rate constant of 0.00027 s^{-1}. All other rate constants determined for this mechanism (Scheme 2) from both the ANM and activity data agree very well.

A recent report describes very slow conformational rearrangements detected by phosphorescence, which occur after native biological activity is obtained from refolding alkaline phosphatase (Subramaniam *et al*., 1995). Earlier work demonstrated that native phosphoglycerate kinase exists as slowly interconverting conformations which lead to more stable forms (Rothstein, 1985; Yuh & Gafni, 1987; Zúñiga & Gafni, 1988). Although the ANM probe may perturb the system to some extent, it seems possible that the probe may report some slow structural rearrangements which occur in the active luciferase. Alternatively, the covalent modification of Cys106 of the α subunit may alter the folding kinetics of the subunit, actually causing the slow conformational change. At this time, we cannot distinguish between these two possibilities. However, it should be stressed that for the purpose of the studies reported here, the distinction is irrelevant. We have used the modified protein to test the

value of the second-order rate constant for formation of the heterodimeric intermediate extracted from the kinetics of activity return during refolding. The rate constant of 2400 $M^{-1}\cdot s^{-1}$ was obtained from both the simulation of the activity return data and the ANM fluorescence data. Whether folding beyond heterodimer formation is altered as a consequence of the chemical modification or the ANM probe is reporting processes not apparent from the intrinsic fluorescence measurements is irrelevant to the question that was being investigated; the distinction must await further experimentation.

The ANM fluorescence data were also well described by a mechanism in which a slow isomerization occurs *after* formation of active enzyme (not shown). The rate constants reported in Table 2 were used in these simulations, where 0.008 s^{-1} was used for the isomerization to native protein and a rate constant of 0.00027 s^{-1} was used to reflect the subsequent slow conformational rearrangement. Since we have no evidence for the slower isomerization in the unmodified enzyme, we have not included this process in Scheme 2.

As discussed above, the off-pathway species β_x was initially incorporated into the kinetic mechanism to account for the concentration dependence of the yield of active enzyme. Recently, we have obtained independent evidence of the existence of this species. First, when β subunit is refolded at higher temperatures (25−35 °C), the homodimer forms at lower yield, and in the presence of excess α, the yield of heterodimeric luciferase is reduced; both processes appear to be compromised by formation of β_x (Sinclair and Baldwin, unpublished results). Second, a point mutation (βD313P) that was selected for study based on its inability to form active luciferase at higher temperatures does not form the β_2 species even at 18 °C, but forms an alternative structure that resembles β_x (A. C. Clark, E. F. Waddill, A. F. Chaffotte, and T. O. Baldwin, submitted for publication).

The experiments described here provide quantitative evidence that the assembly of active heterodimer is the kinetically preferred reaction pathway under conditions of stoichiometric synthesis of both subunits at high levels. As a consequence of the cell density-dependent control mechanism which regulates luciferase synthesis, production of the subunits occurs at a high level for a brief time during development of the culture (Reeve & Baldwin, 1982). Both before induction and following the burst of synthesis, the rates of luciferase synthesis are presumably quite low, such that the concentrations of the heterodimerization-competent species would be expected to be very low. Under these conditions, one would expect the yield of active enzyme to be compromised by the first-order folding of $\beta_i \rightarrow \beta_x$. This process would prevent formation of active enzyme except under conditions that lead to high-level synthesis of both subunits. While this hypothesis explaining the physiological significance of the $\beta_i \rightarrow \beta_x$ transition is plausible, no conditions have been reported under which the β subunit is produced without coincident synthesis of the α subunit (Tu *et al.*, 1977). Therefore, we have no understanding of the physiological significance of the β_2 homodimer, if any.

It is clear that the thermodynamic hypothesis described by Anfinsen (1973) is not sufficient to explain the folding of bacterial luciferase. The work discussed here for the folding of luciferase, as well as that for α-lytic protease (Baker *et al.*, 1992) and PAI-1 (Carrell *et al.*, 1991), demonstrates that a complete understanding of a protein folding pathway must include detailed information on the rates of interconversion and structural features of intermediate species which may occur only transiently. Models for the prediction of tertiary structure based on the primary sequence of amino acids must include the kinetic partitioning that may occur as a result of formation of these structures. Only then will algorithms that correctly predict the structure of small globular proteins also correctly predict the structure of a large multidomain or multisubunit protein, such as bacterial luciferase.

ACKNOWLEDGMENT

We thank Dr. Daniel Lohse for expert assistance with FPLC and Jon Christopher for informative discussions.

REFERENCES

Anfinsen, C. B. (1973) *Science 181*, 223−230.
Baker, D., & Agard, D. A. (1994) *Biochemistry 33*, 7505−7509.
Baker, D., Sohl, J. L., & Agard, D. A. (1992) *Nature 356*, 263−265.
Baldwin, T. O., & Ziegler, M. M. (1992) in *Chemistry and Biochemistry of Flavoenzymes* (Müller, F., Ed.) Vol. III, pp 467−530, CRC Press, Boca Raton, FL.
Baldwin, T. O., Chen, L. H., Chlumsky, L. J., Devine, J. H., & Ziegler, M. M. (1989) *J. Biolumin. Chemilumin. 4*, 40−48.
Baldwin, T. O., Ziegler, M. M., Chaffotte, A. F., & Goldberg, M. E. (1993) *J. Biol. Chem. 268*, 10766−10772.
Baldwin, T. O., Christopher, J. A., Raushel, F. M., Sinclair, J. F., Ziegler, M. M., Fisher, A. J., & Rayment, I. (1995) *Curr. Opin. Struct. Biol. 5*, 798−809.
Barshop, B. A., Wrem, R., & Frieden, C. (1983) *Anal. Biochem. 130*, 134−145.
Carrell, R. W., Evans, D. L., & Stein, P. E. (1991) *Nature 353*, 576−578.
Chaffotte, A. F., Guillou, Y., & Goldberg, M. E. (1992) *Biochemistry 31*, 9694−9702.
Clark, A. C., Sinclair, J. F., & Baldwin, T. O. (1993) *J. Biol. Chem. 268*, 10773−10779.
Cohn, D. H., Mileham, A. J., Simon, M. I., Nealson, K. H., Rausch, S. K., Bonam, D., & Baldwin, T. O. (1985) *J. Biol. Chem. 260*, 6139−6146.
Clark, A. C. (1994) Ph.D Thesis, Texas A&M University, College Station, TX.
Creighton, T. E. (1990) *Biochem. J. 270*, 1−16.
Dill, K. A. (1985) *Biochemistry 24*, 1501−1509.
Fisher, A. J., Raushel, F. M., Baldwin, T. O., & Rayment, I. (1995) *Biochemistry 34*, 6581−6586.
Fisher, A. J., Thompson, T. B., Thoden, J. B., Baldwin, T. O., & Rayment, I. (1996) *J. Biol. Chem. 271*, 21956−21968.
Goldberg, M. E. (1985) *Trends Biochem. Sci. 10*, 388−391.
Goldberg, M. E., Rudolph, R., & Jaenicke, R. (1991) *Biochemistry 30*, 2790−2497.
Hastings, J. W., Baldwin, T. O., & Nicoli, M. Z. (1978) *Methods Enzymol. 57*, 135−152.
Jaenicke, R. (1991) *Biochemistry 30*, 3147−3169.
Kanaoka, Y., Machida, M., Machida., M., & Sekine, T. (1973) *Biochim. Biophys. Acta 317*, 563−568.
Laemmli, U. K. (1970) *Nature 227*, 680−685.
Nicoli, M. Z., & Hastings, J. W. (1974) *J. Biol. Chem. 249*, 2393−2396.
Nicoli, M. Z., Meighen, E. A., & Hastings, J. W. (1974) *J. Biol. Chem. 249*, 2385−2392.
Pace, C. N. (1986) *Methods Enzymol. 131*, 266−280.
Ranson, N. A., Dunster, N. J., Burston, S. G., & Clarke, A. R. (1995) *J. Mol. Biol. 250*, 581−586.
Reeve, C. A., & Baldwin, T. O. (1982) *J. Biol Chem. 257*, 1037−1043.

Rothstein, M. (1985) *Age-Related Changes in Enzymes Levels and Enzyme Properties*, in *Review of Biological Research in Aging 2*, pp 421−433, Alan R. Liss Inc., New York.
Sinclair, J. F., Waddle, J. J., Waddill, E. F., & Baldwin, T. O. (1993) *Biochemistry 32*, 5036−5044.
Sinclair, J. F., Ziegler, M. M., & Baldwin, T. O. (1994) *Nat. Struct. Biol. 1*, 320−326.
Srinivasan, R., & Rose, G. D. (1995) *Proteins: Struct., Funct., Genet. 22*, 81−99.
Subramaniam, V., Bergenhem, N. C. H., Gafni, A., & Steel, D. G. (1995) *Biochemistry 34*, 1133−1136.
Sugihara, J., & Baldwin, T. O. (1988) *Biochemistry 27*, 2872−2880.
Tandon, S., & Horowitz, P. (1986) *J. Biol. Chem. 261*, 15615−15681.
Thoden, J. B., Holden, H. M., Fisher, A. J., Sinclair, J. F., Wesenberg, G., Baldwin, T. O., & Rayment, I. (1997) *Protein Sci. 6*, 13−23.
Tu, S.-C. (1978) *Methods Enzymol. 57*, 171−174.
Tu, S.-C., Makemson, J. C., Becvar, J. E., & Hastings, J. W. (1977) *J. Biol. Chem. 252*, 803−805.
Waddle, J. J., Johnston, T. C., & Baldwin, T. O. (1987) *Biochemistry 26*, 4917−4921.
Yuh, K.-C. M., & Gafni, A. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 7458−7462.
Ziegler, M. M., Goldberg, M. E., Chaffotte, A. F., & Baldwin, T. O. (1993) *J. Biol. Chem. 268*, 10760−10765.
Zúñiga, A., & Gafni, A. (1988) *Biochim. Biophys. Acta 955*, 50−57.

BI962477M