Thermodynamic Analysis of Unfolding and Dissociation in Lactose Repressor Protein†

Jennifer K. Barry and Kathleen S. Matthews*

Department of Biochemistry and Cell Biology, Rice University, Houston, Texas 77251-1892

Received January 12, 1999; Revised Manuscript Received March 16, 1999

ABSTRACT: Lactose repressor protein, regulator of *lac* enzyme expression in *Escherichia coli*, maintains its structure and function at extremely low protein concentrations ($<10^{-12}$ M). To examine the unfolding and dissociation of this tetrameric protein, structural transitions in the presence of varying concentrations of urea were monitored by fluorescence and circular dichroism spectroscopy, analytical ultracentrifugation, and functional activities. The spectroscopic data demonstrated a single cooperative transition with no evidence of folded dimeric or monomeric species of this protein. These spectroscopic transitions were reversible provided a long incubation step was employed in the refolding reaction at ∼3 M urea. The refolded repressor protein possessed the same functional and structural properties as wild-type repressor protein. The absence of concentration dependence expected for tetramer dissociation to unfolded monomer ($M_4 \leftrightarrow 4U$) in the spectral transitions indicates that the disruption of the monomer−monomer interface and monomer unfolding are a concerted reaction ($M_4 \leftrightarrow U_4$) that may occur prior to the dissociation of the dimer−dimer interface. Thus, we propose that the unfolded monomers remain associated at the C-terminus by the 4-helical coiled-coil structure that forms the dimer−dimer interface and that this intermediate is the end point detected in the spectral transitions. Efforts to confirm the existence of this species by ultracentrifugation were inhibited by the aggregation of this intermediate. Based upon these observations, the wild-type fluorescence and CD data were fit to a model, $M_4 \leftrightarrow U_4$, which resulted in an overall $\Delta G°$ for unfolding of 40 kcal/mol. Using a mutant protein, K84L, in which the monomer−monomer interface is stabilized, sedimentation equilibrium results demonstrated that the dimer−dimer interface of *lac* repressor could persist at higher levels of urea than the monomer−monomer interface. The tetramer−dimer transition monitored using this mutant repressor yields a $\Delta G°$ of 20.4 kcal/mol. Using this free energy value for the dissociation process of $U_4 \leftrightarrow 4U$, an overall free energy change of ∼60 kcal/mol was calculated for dissociation of all interfaces and unfolding of the tetrameric *lac* repressor, reflecting the exceptional stability of this protein.

Lactose repressor protein regulates the transcription of the *lac* operon genes that are responsible for transport and metabolism of lactose in *E. coli* (*1*). *lac* repressor is a well-characterized tetrameric protein comprised of identical monomers with a molecular mass of ∼37 500 Da (*2*−*4*). The monomers are separated structurally into three domains: an N-terminal domain (amino acids ∼1−60) that binds operator DNA (*4*−*8*); a core domain (amino acids ∼60−340) that binds inducer sugars and forms the dimer interface (*8*−*16*); and a C-terminal domain (amino acids ∼340−360) that forms the tetramer interface (*8, 16*−*19*). Crystallographic studies determined that the two interfaces that must associate to produce the tetrameric protein are physically separated from each other in different regions of the monomer (*8, 16*). The monomer−monomer interface is formed as a result of numerous interactions between the core domains of two monomers (*8, 16*). In contrast, the dimer−dimer interface is formed by a short C-terminal sequence of ∼20 amino acids that contains a leucine heptad repeat sequence; this region generates an antiparallel 4-helical coiled-coil that serves to assemble 2 dimers into a tetramer (*8, 16*−*19*). The extent of surface area buried in these two structurally and physically distinct interfaces is similar based on the crystallographic structures (*8, 16*). The oligomeric structure of the *lac* repressor can be altered by mutation to generate both dimeric and monomeric derivatives (*11, 14, 17*−*20*). Dimeric proteins result from mutation or deletion of the C-terminal assembly motif, while point mutations within the monomer−monomer interface result in monomeric proteins (*11, 14, 17*−*20*).

The association of the repressor into a tetramer is important for its ability to diminish transcription of the genes for lactose metabolism. Tetrameric *lac* repressor possesses two DNA binding sites that allow the protein to interact with two separate operator sites and thereby form looped DNA structures (*21*−*27*). DNA looping is important for maximum repression, as dimeric repressors that bind only a single operator site or the presence of only a single operator in the DNA sequence diminishes repression in vivo (*28*). The contribution of DNA looping to repression depends on the

† Support for this project was provided by grants to K.S.M. from NIH (GM 22441) and the Robert A. Welch Foundation (C-576). J.K.B. was supported in part by National Institutes of Health Biotechnology Training Grant 5T32-GM-08362. Spectroscopic facilities utilized were provided by the Keck Center for Computational Biology and the Lucille P. Markey Charitable Trust.

* To whom correspondence should be addressed. Phone: 713-527-4871. Fax: 713-737-6149. E-mail: ksm@bioc.rice.edu.

10.1021/bi9900727 CCC: $18.00
Published on Web 04/29/1999

strength of these loops, which derives from both DNA−protein interaction and protein−protein interaction.

Attempts to determine the stability of the *lac* repressor protein have employed pressure dissociation, DNA binding methods, dilution of the protein, and urea denaturation of dimeric mutants (*29*−*35*). While dissociation of the wild-type protein was observed by pressure dissociation in the nanomolar range (*29, 30*), results from other experiments indicate that *lac* repressor does not dissociate in accessible concentration ranges ($<10^{-12}$ M) (*31, 32, 35*). For example, no evidence for the presence of *lac* repressor dimers was detected in the thermodynamic analysis of DNA binding over a wide range of conditions (*35*). These experiments set a lower limit for the stability of *lac* repressor protein subunit interfaces at femtomolar concentrations.

Protein unfolding studies using chemical denaturants have provided information about the energy involved in the folding of a large number of monomeric systems (*36*−*38*). This approach has also been used for examining folding and subunit interactions in dimeric and tetrameric proteins, including dimeric *lac* repressor mutants (*34, 39*−*42*). In the present study, we have measured the free energy of wild-type *lac* repressor dissociation and unfolding using urea denaturation coupled with multiple structural assessments. By employing a mutant *lac* repressor protein with a strengthened monomer−monomer interface (K84L), a pathway for wild-type repressor unfolding and dissociation was established. Based on these measurements, we propose that *lac* repressor denaturation occurs through an intermediate that consists of unfolded monomers still associated at the C-terminus.

MATERIALS AND METHODS

Protein Purification. Protein purification followed procedures that have been described previously (*43, 44*) with the following modifications. Wild-type and mutant proteins were expressed in BL-26 cells (Novagen) that were cured of the episome carrying the I^q promoter and the I gene (D. Wycuff, personal communication). The ammonium sulfate pellet was resuspended in 0.08 M potassium phosphate (pH 7.5) and dialyzed against the same buffer. This solution was centrifuged to remove any precipitate and then loaded onto a phosphocellulose column equilibrated in the same buffer. For tetrameric proteins, the column was washed with the same buffer, followed by elution with 0.12 M potassium phosphate buffer, and the *lac* repressor protein was eluted with a gradient from 0.12 to 0.3 M potassium phosphate. Dimeric *lac* repressor mutants, −11 aa and K84L/−11 aa, were eluted from the phosphocellulose column with 0.12 M potassium phosphate buffer after extensive washing with 0.08 M potassium phosphate. Protein concentrations for mutant repressors were determined by absorbance at 280 nm using the extinction coefficient for wild-type *lac* repressor.

Spectroscopic Studies of Protein Unfolding by Urea. Unfolding of tetrameric *lac* repressor in varying concentrations of urea was carried out in 0.1 M K_2SO_4, 0.01 M Tris-HCl (pH 7.4) as described previously for dimeric mutants (*34*). Ultrapure urea (Fluka) was prepared fresh daily and filtered before use. The concentration of the urea stock was determined by refractive index. In a denaturation experiment, 16 samples of a fixed amount of *lac* repressor (0.5−4 μM monomer) were mixed with the urea stock solution to final urea concentrations between 0 and 6 M. Samples were incubated at room temperature for 2 h before fluorescence and CD[1] signals were determined.

Measurements of the intrinsic tryptophan fluorescence of the repressor proteins were performed using SLM 8000 or AB2 spectrofluorometers. An excitation wavelength of 285 nm was used, and the fluorescence emission spectra were collected between 300 and 380 nm. The intensity of the fluorescence signal at 340 nm was used to monitor protein unfolding. CD spectra were measured using an Aviv Model 60 DS spectropolarimeter. The spectra were collected over a range of 210−300 nm in an 0.2 cm path length cuvette, and the intensity of the signal at 222 nm was used to monitor protein unfolding.

Renaturation of Lactose Repressor Protein from Urea. Wild-type *lac* repressor cannot be refolded from 6 M urea without precipitation using the procedure developed for the dimeric *lac* repressor mutants (*34*). An additional step involving a longer incubation at an intermediate concentration of urea was found to be necessary. Wild-type *lac* repressor was denatured for 2 h in urea concentrations >5 M at 28.5 μM monomer concentration at room temperature. Renaturation was initiated by diluting the sample to 2.6 M urea at 4 μM monomer concentration, and equilibrating for 2 h. This material was then further diluted to 0.5 μM monomer at the final urea concentration desired, and allowed to incubate for 2 h at room temperature before the signal was recorded. The lowest concentration of urea that could be reached using this renaturation scheme was 0.7 M urea. The renatured protein samples which had their oligomeric state determined in the ultracentrifuge were refolded by diluting the denatured protein to 4 μM monomer concentration at a urea concentration of 2.5 M urea. After a 2 h incubation, the samples were dialyzed into 0.1 M K_2SO_4, 0.01 M Tris-HCl (pH 7.4) buffer overnight.

Operator Binding Assay. For measurement of operator binding activity, protein was denatured in 6 M urea for 2 h and then renatured for 2 h at 2.5 M urea. This sample was diluted to 1 M urea and exhaustively dialyzed in 0.1 M K_2SO_4, 0.01 M Tris-HCl (pH 7.4) to remove the remaining urea. The operator assays were performed as described previously (*43*−*45*) using buffer containing 0.15 M KCl, 0.01 M Tris (pH 7.4), 0.1 mM EDTA, 5% DMSO, 0.1 mM DTT, and 100 μg/mL bovine serum albumin.

Inducer Binding Assay. An ammonium sulfate precipitation assay was used to monitor the binding of inducer by the denatured repressor proteins at different urea concentrations. This assay followed the protocol of Bourgeois (*46*) with the following modifications. The assay buffer was 0.1 M K_2SO_4, 0.01 M Tris-HCl (pH 7.4) as described for the denaturation experiments. The protein concentration was constant at 5 μM monomer, and the repressor was preincubated in urea concentrations between 0 and 6 M urea overnight with 5×10^{-7} M [^{14}C]-IPTG. Overnight incubation with inducer was necessary because of the very low association rate constant for K84L with this ligand (*47*). Each sample had 15 mg/mL lysozyme added before precipitation with 70% saturated

[1] Abbreviations: CD, circular dichroism; DMSO, dimethyl sulfoxide; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; IPTG, isopropyl-β,D-thiogalactoside.

ammonium sulfate. After these samples were centrifuged, the pellets were resuspended in buffer and the mixtures precipitated with 5% TCA. After centrifugation, the amount of radiolabel in the supernatant was determined by scintillation counting.

Sedimentation Equilibrium. The molecular mass at varying urea concentrations was determined for wild-type, −11 aa, K84L, and K84L/−11 aa by sedimentation equilibrium using a Beckman XL-A ultracentrifuge. The sedimentation equilibrium experiments were conducted in buffer containing 0.1 M K_2SO_4, 0.01 M Tris-HCl (pH 7.4). Initially, the proteins were characterized in buffer containing no urea at multiple concentrations (7×10^{-7} to 1×10^{-5} M monomer) and multiple speeds (10 000−16 000 rpm). Samples were loaded into a six-channel Epon charcoal-filled equilibrium cell and scanned at 280 nm or at 230 nm for the samples at 7×10^{-7} M monomer. For the urea experiments, the protein concentration was constant at 4 μM monomer, and multiple speeds were employed for each urea concentration. Equilibrium was presumed to be reached when no difference was detected between scans 4 h apart. The average time for a run consisting of three different speeds was 40 h.

When the protein is at equilibrium in the ultracentrifuge, the concentration gradient is a function of molecular mass. The concentration gradient can be analyzed to determine molecular mass using the equation:

$$c_r = c_{r0} \exp\left[\frac{\omega^2}{2RT} M(1 - \bar{v}\rho)(r^2 - r_0^2)\right] \quad (1)$$

where c_r = concentration at radius position r, c_{r0} = concentration of the monomer at the reference radius r_0, ω = angular velocity, R = gas constant [8.314×10^{-7} erg/(mol·K)], T = temperature in degrees kelvin, M = apparent molecular mass, $\bar{v}$ = partial specific volume of the solute, and ρ = density of the solvent. Data were fit to this equation using the Beckman Optima XL-A Data Analysis package. Density was calculated based upon the method of Kawahara and Tanford (*48*). The partial specific volume was calculated as 0.7411 for wild-type *lac* repressor, 0.7413 for K84L, 0.7420 for K84L/−11 aa, and 0.7425 for −11 aa, in K_2SO_4 buffer based upon the amino acid composition of the proteins using the method of Cohn and Edsall (*49*). The data were fit assuming that this value did not change as a function of urea for the wild-type, K84L, and K84L/−11 aa proteins.

Analysis of Denaturation Data. The differences in the fluorescence and CD spectroscopic signals for wild-type protein and for −11 aa deletion protein were assumed to be proportional to the fraction of the native state present at a specific concentration of urea. The pre- and post-transition base lines were treated by linear extrapolation (*50*, *51*), and the intensity of the signal (Y), was converted to fraction unfolded using the equation:

$$F_u = \frac{Y_f - Y}{Y_f - Y_u} \quad (2)$$

where Y_f and Y_u are the signals of the folded and unfolded states. For the −11 aa protein, the data were fit to the equations described previously for the dimer to unfolded monomer transition (*34*) using a nonlinear least-squares program (Igor Pro).

The ultracentrifuge unfolding data for K84L were normalized in a similar manner:

$$M_{corr} = M_t - \left[M_{t,app}\left(\frac{M_d}{M_{d,app}}\right)\right] \quad (3)$$

where the molecular mass in native buffer was M_t for K84L and M_d for K84L/−11 aa, and the molecular mass apparent in different concentrations of urea was $M_{t,app}$ for K84L and $M_{d,app}$ for K84L/−11 aa. M_{corr} from the ultracentrifuge data was plotted as fraction of dimer (F_d), calculated by dividing the M_{corr} at each point by the value of M_{corr} at 6 M urea. This method of normalizing the data corrects for any changes in the value of $\bar{v}$ for the proteins at different concentrations of urea in the ultracentrifuge, assuming that this value is affected in the same manner for the dimeric and tetrameric proteins.

The data for K84L were fit to a two-state transition, $M_4 \leftrightarrow 2M_2$, where M_4 is the K84L tetramer and M_2 represents K84L dimer present in high concentrations of urea, and $K = [M_2]^2/M_4$. The relationship to determine fraction of dimer, F_d, can be expressed as

$$F_d = \frac{\sqrt{K^2 + 8K[P]} - K}{4[P]} \quad (4)$$

where [P] is the total protein concentration in dimer and K is the equilibrium constant for dissociation of K84L from tetramer to dimers. The free energy change for this transition in the absence of denaturant was determined by fitting the data using a nonlinear least-squares program (Igor Pro) to the relationship $K = K^\circ \exp\{-(m[\text{urea}])/(RT)\}$ with base line correction, where m is the dependence of K on denaturant concentration, R is the gas constant (1.987 cal deg^{-1} mol^{-1}), and T is the temperature in degrees kelvin. The free energy change for the transition (ΔG°) in the absence of urea could be determined based upon the relationship $K^\circ = \exp(-\Delta G^\circ/RT)$.

Wild-type tetramer spectroscopic data were fit to two models, each of which potentially could describe the unfolding process. The first model was $M_4 \leftrightarrow 4U$ where M_4 is the native tetrameric protein, U is the unfolded monomer, [P] is the total protein concentration in monomer, and $K = [U]^4/M_4$. Since this model involved a fourth-order dependence on protein concentration, an iterative loop was incorporated into the NonLin (*52*) fitting routine to provide a numerical solution of the fourth-order equation. The data were fit to fraction unfolded using NonLin (Figure 2A):

$$F_u = [U]/[P] \quad (5)$$

It is also possible that the wild-type protein unfolds via a three-state model of $M_4 \leftrightarrow U_4 \leftrightarrow 4U$ where the fluorescence and CD data monitor only the $M_4 \leftrightarrow U_4$ transition. If the $M_4 \leftrightarrow U_4$ transition is sufficiently separated from the $U_4 \leftrightarrow 4U$ transition, the spectroscopic data can be fit as a first-order process. Therefore, normalized data were analyzed by the following equation for fraction unfolded:

$$F_u = K/(1 + K) \quad (6)$$

where K in this equation represents the equilibrium constant for monomer unfolding. The data were fit to fraction

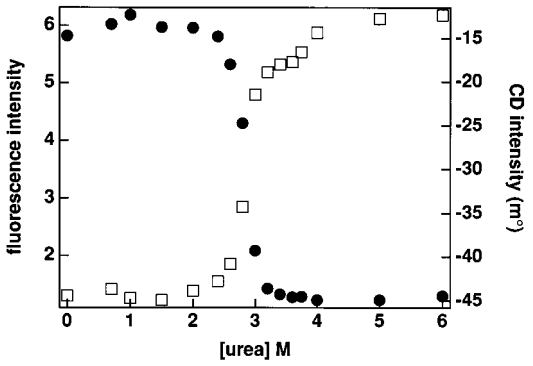

FIGURE 1: Denaturation of wild-type repressor protein. The protein concentration was 4 μM monomer in a buffer of 0.1 M K_2SO_4, 0.01 M Tris-HCl (pH 7.4) with various concentrations of urea. The spectroscopic signals were monitored by the fluorescence intensity at 340 nm (●) and the CD signal at 222 nm (□).

unfolded using Igor Pro (Figure 2B). Because the spectroscopic signal from which the $\Delta G°$ values were derived corresponds to the monomer transition within the tetramer, the free energy value for this transition was multiplied by 4 to determine the value for the tetrameric species.

RESULTS

Wild-Type Lactose Repressor Denaturation. The spectral changes associated with the denaturation of *lac* repressor were monitored by circular dichroism (CD) at 222 nm and tryptophan fluorescence at 340 nm. As shown in Figure 1, the changes in these signals were highly correlated and exhibited an apparent cooperative transition centered at ~2.8 M urea. A small deviation between fluorescence and CD data at higher urea concentrations was observed for both wild-type and mutant dimeric proteins near the end of the transition and may represent residual helical content in the largely unfolded protein. Previously, studies with a dimeric *lac* repressor mutant (−11 aa) established the species at the end point of the CD and fluorescence transitions to be unfolded monomeric *lac* repressor (*34*).

Because the fluorescence and CD signal intensities monitor the loss of secondary and tertiary structure of the repressor protein, information about the oligomeric state of the protein must be determined by other techniques. Equilibrium sedimentation experiments utilizing different centrifuge speeds and protein concentrations demonstrated that *lac* repressor sediments in the absence of urea with the expected molecular mass of 150 000 Da (data not shown). In these experiments, wild-type protein was tetrameric at concentrations of urea in the regions of the pretransition base line for the fluorescence and CD experiments. However, above 2.5 M urea, reproducible scans could not be collected in the ultracentrifuge experiments. Similarly, sedimentation data also could not be collected for the −11 aa repressor protein at concentrations of urea beyond the fluorescence and CD pretransition base lines. Therefore, the process causing this variation cannot be ascribed to the presence of the dimer−dimer interface.

Because no intermediate dissociation to dimer was detected in the ultracentrifuge in the pretransition range, the spectroscopic data for wild-type repressor were fit to a model of folded tetramer to four unfolded monomers ($M_4 \leftrightarrow 4U$). Due

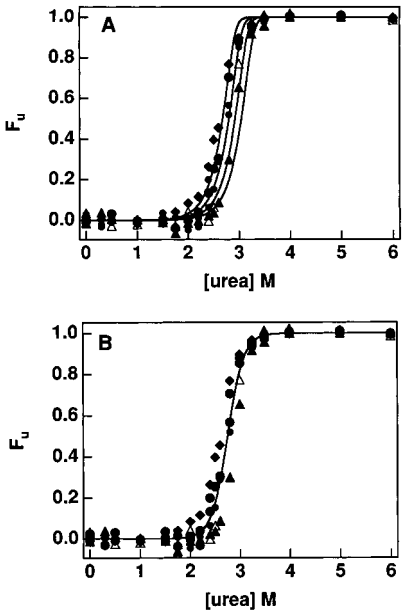

FIGURE 2: Concentration dependence of wild-type repressor unfolding. The intensity of the fluorescence signal was monitored as a function of [urea] at four different monomeric concentrations of wild-type repressor: 4 μM (Δ), 2 μM (▲), 1 μM (●), and 0.5 μM (◆). These data were fit to two different models. (A) The data were fit globally to the model $M_4 \leftrightarrow 4U$ using Non-Lin (*52*) as explained under Materials and Methods. (B) All data points were fit simultaneously as described under Materials and Methods to the first transition of the second model, $M_4 \leftrightarrow U_4$, using Igor Pro.

to the fourth-order nature of this process, this model predicts that the folding process will display concentration dependence. Figure 2A shows simulations of unfolding curves representing data expected for the $M_4 \leftrightarrow 4U$ transition. These curves are overlaid with the experimental data, which do not precisely display the concentration dependence predicted by the model. Therefore, the data were fit to an alternative model of folded tetramer to unfolded tetramer ($M_4 \leftrightarrow U_4$) (Figure 2B), where the unfolded tetramer (U_4) represents four unfolded monomers still associated at the C-terminus to form a tetramer. This reaction is not concentration dependent because no dissociation event(s) occur(s). These data provide the first indication that dissociation to unfolded monomers may not be monitored by fluorescence and CD during wild-type repressor denaturation experiments. Further unfolding experiments with mutant repressor proteins are used to better define which model best represents wild-type repressor unfolding events.

Reversible Denaturation of Wild-Type Lactose Repressor. To analyze the thermodynamic parameters for *lac* repressor unfolding, the process must be reversible. In contrast to the refolding behavior observed for dimeric *lac* repressor mutants (*34*), dilution of wild-type *lac* repressor from high to low concentrations of urea results in significant precipitation (*53*). Because dimeric repressors generated by both point mutation (e.g., L349A) and deletion at the C-terminus exhibit the same refolding behavior (data not shown), this precipitation likely results from the presence of a functional dimer−dimer interface rather than simple aggregation of this hydrophobic region. A refolding protocol for wild-type repressor which incorporates the incubation of the protein at intermediate concentrations of urea (2.5−3 M) results in regeneration of >80% of the original fluorescence signal (Figure 3). The

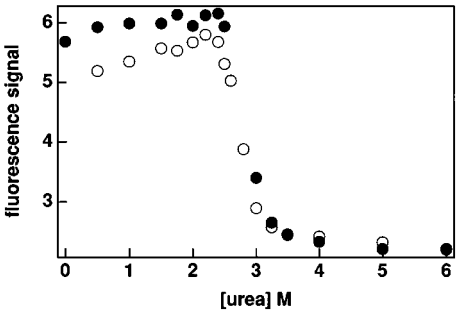

FIGURE 3: Reversibility of wild-type repressor denaturation. The intensity of the fluorescence signal as a function of [urea] was determined for denaturation (●) and renaturation (○) of wild-type repressor. Wild-type protein was renatured by diluting samples from 5.25 M urea to 2.5 M urea and incubating them for 2 h prior to dilution to lower concentrations of urea. The final protein concentration in these assays was 0.5 μM monomer.

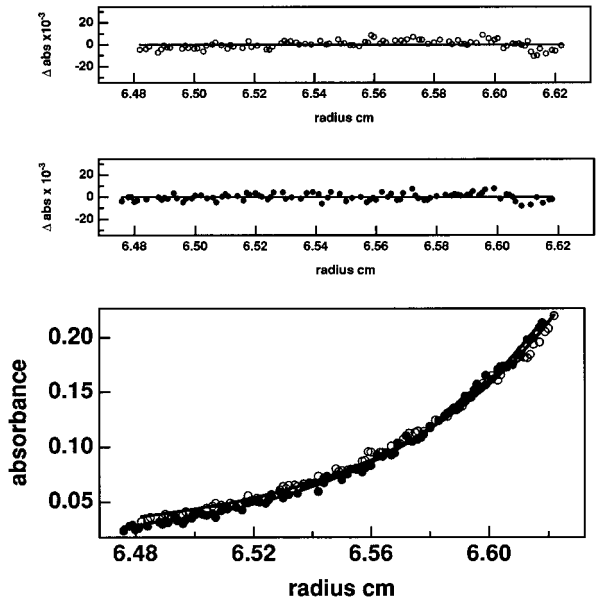

FIGURE 4: Sedimentation equilibrium analysis of refolded wild-type repressor. Wild-type protein was renatured as described under Materials and Methods. The control (○) and renatured (●) wild-type repressor proteins were centrifuged at multiple speeds (12 000, 14 000, and 16 000 rpm) at a single concentration of 4 μM monomer until equilibrium was reached. These data sets for the control and renatured samples were each independently fit to a molecular mass of 150 000 Da in a nonassociating model. For clarity, only the scans for one rotor position of the 12 000 rpm data are shown for each sample. The residuals for the data set compared to the fitted values are also included.

structural and functional characteristics of the refolded protein were tested to determine whether the restored spectroscopic signal correlates with properly folded, wild-type *lac* repressor. Sedimentation equilibrium experiments using analytical ultracentrifugation demonstrate that the refolded protein is tetrameric with an apparent molecular mass of ~150 000 Da (Figure 4). The operator binding activity of the refolded protein was assessed to determine if the secondary and tertiary structural elements of the protein were restored. Specific recognition of operator DNA depends on proper folding of the monomer−monomer interface to align the N-terminal DNA binding domains in the correct position to contact the operator sites. Figure 5 demonstrates that the renatured protein binds operator specifically, with no loss of affinity compared to wild-type controls. Together,

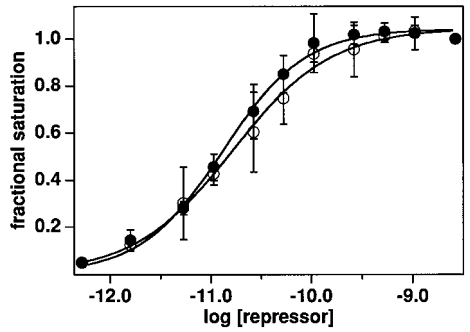

FIGURE 5: Operator binding to refolded wild-type repressor. Wild-type protein was renatured as described under Materials and Methods. Nitrocellulose filter binding assays were used to measure operator binding for renatured wild-type protein (●), and wild-type control (○). The curves were generated using Igor Pro by fitting the data as described previously (*44*).

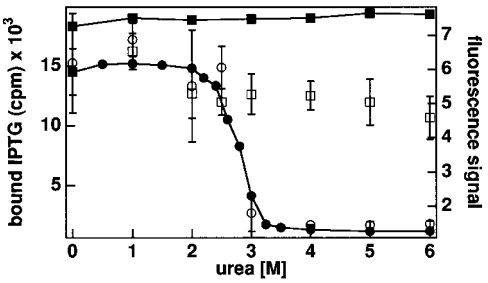

FIGURE 6: Denaturation of K84L mutant repressor. The fluorescence intensity for K84L (■) and wild-type (●) repressors was monitored as a function of [urea]. Protein concentration for both repressor proteins was 4 μM monomer. Lines are drawn through the points for comparison purposes only. The IPTG binding activity for repressor proteins at different concentrations of urea was determined using an ammonium sulfate assay. These data for K84L (□) and wild-type (○) protein are shown in comparison to the fluorescence data. Error bars represent the standard deviation of three experiments used to determine these values.

these experiments demonstrate that both interfaces are correctly assembled in the renatured protein.

Denaturation of Mutant Repressor with a Strengthened Monomer−Monomer Interface. Wild-type *lac* repressor protein unfolds via a single transition when monitored by fluorescence or CD spectroscopy even though multiple processes (monomer−monomer interface dissociation, monomer unfolding, and dimer−dimer dissociation) are potentially occurring. The tetramer to dimer transition for wild-type *lac* repressor appears to be transparent to both fluorescence and CD signals. It is therefore possible that this transition occurs concomitantly with monomer−monomer dissociation and monomer unfolding. To dissect these processes, a repressor mutant, K84L, with a significantly strengthened monomer−monomer interface was employed (*42*, *47*). Fluorescence signal intensity for this mutant repressor does not decrease at urea concentrations as high as 6 M (Figure 6; *42*). Furthermore, the K84L repressor binds inducer effectively in 6 M urea, indicating that the monomer species remains properly folded under these conditions (Figure 6). Since the K84 residue is located in the monomer−monomer interface, the ability of the leucine substitution to stabilize the protein presumably derives from stabilizing the dimer structure (*8*, *16*, *42*, *47*). Using this mutant *lac* repressor protein is required to generate information regarding the dimer−dimer subunit interface; however, the inherent assumption in this approach is that the K84L mutation at the monomer−

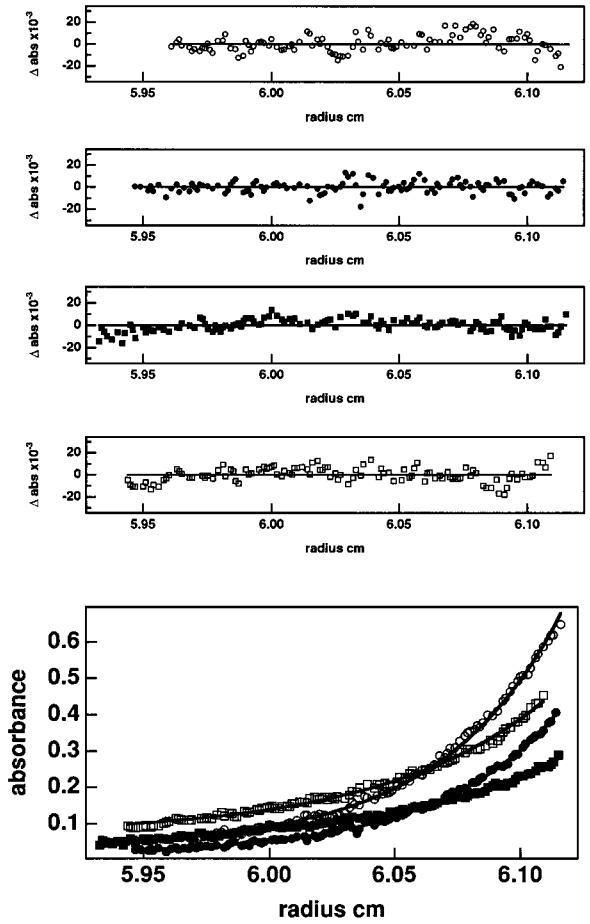

FIGURE 7: Sedimentation equilibrium analysis for K84L mutant repressor. Two speeds (12 000 and 14 000 rpm) and two concentrations (4 and 8 μM) were fit simultaneously to a molecular mass of 150 000 Da using a nonassociating model. The residuals for each fit versus the data points for each curve shown are plotted. The different conditions are represented by the following symbols: 4 μM at 12 000 rpm (■), 4 μM at 14 000 rpm (●), 8 μM at 12 000 rpm (□), and 8 μM at 14 000 rpm (○).

monomer interface does not influence the stability of the dimer−dimer interface.

The transition from tetramer to dimer for K84L repressor protein was determined by monitoring the molecular mass of the protein at different concentrations of urea in the analytical ultracentrifuge. Under nondenaturing conditions, sedimentation equilibrium experiments using multiple speeds and concentrations demonstrated that the apparent molecular mass of K84L is ~150 000 Da (Figure 7). The molecular mass of a dimeric form of the K84L mutant, K84L/−11 aa, was also determined (*42*). This mutant repressor provided a control for the ultracentrifuge experiments because the molecular mass of the dimer should remain constant throughout the concentrations of urea used. In sedimentation equilibrium experiments in the absence of urea, this repressor mutant fit well to an apparent molecular mass of ~72 800 Da (data not shown), corresponding to the loss of 22 amino acids due to the −11 aa mutation at the C-terminus of both monomers. At high concentrations of urea, K84L and K84L/−11 aa repressor proteins each have an observed molecular mass ~65 000 Da, which corresponds to the dimeric species. At intermediate concentrations of urea (between 3 and 5.5 M), the K84L protein displays a decrease in apparent molecular mass (Figure 8A). In Figure 8B, this

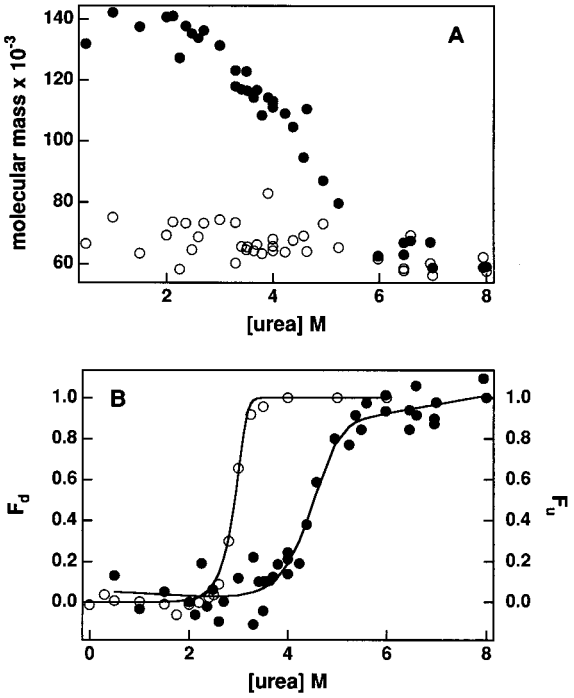

FIGURE 8: Tetrameric to dimeric transition for K84L mutant repressor. (A) The apparent molecular mass for K84L (●) and K84L/−11 aa (○) in different concentrations of urea was determined using sedimentation equilibrium experiments as described under Materials and Methods. (B) The apparent molecular mass for K84L (●) was converted into fraction of dimer, F_d, and these data points were fit to a two-state model of $M_4 \leftrightarrow 2M_2$ using Igor Pro as described under Materials and Methods. The fraction unfolded for wild-type repressor protein (○) determined by fluorescence spectroscopy is shown for comparison purposes.

transition is compared to the fluorescence data for the wild-type protein. The midpoint of the transition between tetrameric and dimeric species for K84L corresponds to ~4.5 M urea, while the midpoint for the transition from folded to unfolded protein for wild-type repressor is ~2.8 M urea. The overall reaction for K84L repressor was assumed to be tetramer to dimer ($M_4 \Leftrightarrow 2M_2$), and the ultracentrifuge data were fit to a two-state model. The free energy value of 20.4 kcal/mol for dissociation of tetramer into dimer determined from these data indicates that the dimer−dimer interface is remarkably stable, and information about dissociation of this interface can be obtained directly when monomer−monomer dissociation is prevented by K84L mutation.

Integration of Thermodynamic Analyses of Dissociation and Unfolding. Mutant *lac* repressor proteins have been used previously to determine the stability of the monomer−monomer interface and the monomeric species (*34*). In this study, we have expanded the analysis to include the wild-type tetrameric repressor protein and a mutant that allows the dimer−dimer interface to be assessed independently. The composite results of these experiments are summarized in Figure 9. A monomeric *lac* repressor mutant, Y282D, was used previously to determine the contribution of the unfolding of monomeric *lac* repressor ($\Delta G°_{M-U} \approx 4.8$ kcal/mol) to the overall transition for the dimeric species (*34*). The energy associated with the dissociation of the monomer−monomer interface determined previously for −11 aa dimeric repressors was $\Delta G°_{M_2-2M} \approx 9.7$ kcal/mol (*34*), and the same value determined in the present studies was $\Delta G°_{M_2-2M} \approx 10.0$ kcal/mol. The free energy of the tetramer to dimer dissociation

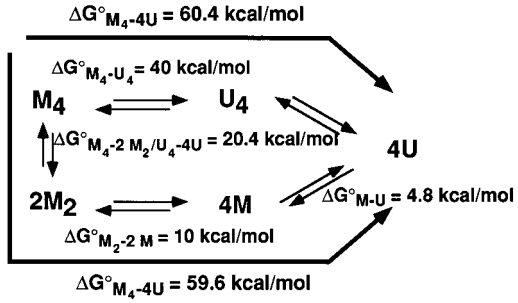

FIGURE 9: Wild-type repressor dissociation and unfolding pathways. Two different pathways by which wild-type repressor tetrameric structure can dissociate and unfold are shown. The free energies associated with each of the different steps in these pathways were determined.

transition obtained from the present experiments using the K84L mutant is $\Delta G°_{M_4-2M_2} \approx 20.4$ kcal/mol. The transition of tetrameric to unfolded monomeric *lac* repressor can be described by a sequence of these transitions observed for mutant proteins. The total free energy change can be described as $\Delta G°_{M_4-4U} = \Delta G°_{M_4-2M_2} + 2\Delta G°_{M_2-2M} + 4\Delta G°_{M-U}$, which yields an overall calculated $\Delta G°_{M_4-4U} \approx 59.6$ kcal/mol. However, the *lac* repressor mutant proteins used to generate these values may not necessarily reflect the stability of these interfaces or the energy of the folded monomer in the wild-type protein.

The wild-type fluorescence and CD unfolding data could be fit to a variety of models. These spectroscopic techniques only detect a single transition and not the multiple equilibria between oligomeric states that might occur during the unfolding of wild-type *lac* repressor. In choosing the models for fitting these data, the following two assumptions were made: (1) the loss of CD signal correlates with unfolding of monomer; and (2) no tetramer to dimer dissociation occurs prior to the unfolding transition. The latter assumption was confirmed by analytical ultracentrifugation. Two models meeting these criteria are (1) $M_4 \leftrightarrow 4U$, and (2) $M_4 \leftrightarrow U_4 \leftrightarrow 4U$, where U_4 represents unfolded monomers that remain associated as a tetramer through the four-helical coiled-coil structure which forms the dimer−dimer interface. As shown in Figure 2, simultaneous fitting of the unfolding curves for wild-type repressor at multiple concentrations indicated that the first model, $M_4 \leftrightarrow 4U$, does not describe the data as well as the second model, $M_4 \leftrightarrow U_4 \leftrightarrow 4U$. The difference between these two models is the concentration dependence predicted for the unfolding process. While either model could be interpreted to fit the data, additional support for the $M_4 \leftrightarrow U_4 \leftrightarrow 4U$ model is provided by a comparison of the wild-type and K84L unfolding curves shown in Figure 8B. The transition of K84L from tetramer to dimer occurs well after the unfolding of wild-type protein detected by spectroscopy, consistent with the interpretation that the dimer−dimer interface remains associated after the remainder of the wild-type protein has unfolded. With this additional information from the K84L mutant protein regarding the stability of the dimer−dimer interface, we conclude that the second model with a tetrameric intermediate best describes the behavior of the protein. Using the $M_4 \leftrightarrow U_4 \leftrightarrow 4U$ model, the derived $\Delta G°$ for monomer−monomer dissociation and unfolding of 10 kcal/mol corresponds to the transition for a single monomer; therefore, the total $\Delta G°$ for unfolding/dissociation of the tetramer is 40 kcal/mol. Assuming the value for the free energy corresponding to $U_4 \leftrightarrow 4U$ is identical to that for the tetramer to dimer transition in K84L, given that these are identical interfaces, the free energy for the complete dissociation and unfolding of the tetrameric lactose repressor is $\Delta G°_{M_4-4U} \approx 60.4$ kcal/mol.

DISCUSSION

An essential aspect of the mechanism by which *lac* repressor diminishes transcription of the *lac* operon genes is the ability of the tetrameric protein to interact with DNA at two sites and produce loops (*26*, *28*). Two different interactions must occur in this process: (1) the monomeric protein must be assembled into dimers that bind to single operator sites; and (2) the dimeric protein must be assembled into tetramers that bring two operator sites in separate regions together via a single protein molecule. These processes therefore encompass protein−protein as well as protein−DNA interactions. Although *lac* repressor is an exceptionally stable tetramer, detailed information about the overall stability of the protein has been indirect.

Determining the unfolding pathway for *lac* repressor is important for analyzing the energetics of these interactions. Mutant *lac* repressors have provided crucial information in dissecting the unfolding pathway. Previous experiments demonstrated that denaturation of dimeric mutant repressor proteins involved a concerted dimer to unfolded monomer transition based on concentration dependence and simultaneous loss of spectroscopic signal intensity and inducer binding capacity (*34*). In the present experiments, the $\Delta G°$ for this transition for −11 aa protein to unfolded monomer was measured to be 19.6 kcal/mol, in good agreement with the previously reported value of 19.3 kcal/mol (*34*). The monomeric Y282D mutant protein unfolds with a $\Delta G°$ of ∼4.8 kcal/mol (*34*); using this value for monomer unfolding, the $\Delta G°$ for dissociation of the monomer−monomer interface is ∼10 kcal/mol. The dissociation of the dimer−dimer interface in K84L was monitored by analytical ultracentrifugation at multiple urea concentrations to calculate a $\Delta G°$ of ∼20.4 kcal/mol. The dimer−dimer interface in K84L is distant from the site of mutation and is presumed to be identical to the corresponding interface in the wild-type protein. The overall $\Delta G°$ value for the transition of tetramer to unfolded monomer can be calculated from these individual measurements to be ∼59.6 kcal/mol. The correspondence between the value estimated based on individual transitions for mutant proteins and that for the wild-type tetrameric protein of 60.4 kcal/mol is striking.

A combination of the unfolding data for *lac* repressor wild-type and mutant proteins indicates that *lac* repressor remains assembled as a tetramer after the monomer−monomer interface has dissociated and the monomeric protein unfolded. No stable monomeric intermediate is detected because in the absence of the dimeric interface, the monomer would not remain folded. The physical separation of the monomer−monomer and dimer−dimer interfaces, determined by mutagenic studies and evident in the crystal structures (*8*, *11*−*19*), allows one interface to remain assembled after the other interface has dissociated. However, the unfolded but still assembled protein could not be detected by equilibrium ultracentrifugation in our studies or in previous measurements by size-exclusion column chromatography (*33*). In this

context, difficulty in determining the oligomeric state of folding intermediates for other tetrameric proteins has also been noted (*54*). In folding studies of tetrameric β-galactosidase, the oligomeric state of the folding intermediate could not be established using size-exclusion column chromatography or light-scattering experiments (*54*). The combined data on *lac* repressor wild-type and mutant proteins lead to the proposed model for denaturation in which wild-type repressor exists in an intermediate as an unfolded but still associated tetrameric protein due to the strength of the tetramer (dimer−dimer) interface.

As more folding studies of proteins with higher order oligomeric states are conducted, an increasing number of stable, intermediate structures have been detected in these folding and assembly pathways (*39*−*41*). Essentially two types of folding pathways are typical of dimeric proteins (reviewed in *40*): (1) a single two-state transition, or (2) multiple transitions with the presence of intermediates of assembly and/or folding. In general, the oligomeric states of the intermediates tend to be monomeric (e.g., *55, 56*), but there are also reports of stable, dimeric intermediates (e.g., *57*−*59*). Folding studies of tetrameric proteins have also identified folding intermediates with different oligomeric structures (*41*). Tetrameric β-galactosidase folds via monomeric intermediates that associate to form dimers, although these intermediates were not detected in the unfolding reaction (*54*). The unfolding pathway for pyruvate decarboxylase from *Z. mobilis* involves cofactor dissociation, tetramer dissociation to monomer, and unfolding of the monomer species (*60*). R67 dihydrofolate reductase dissociates to a stable dimeric structure before a transition to unfolded monomers (*61*). *lac* repressor is unique in that the dimeric structure unfolds before the dimer−dimer interface that forms the tetramer dissociates.

A combination of data for the K84L and wild-type repressors allows us to assign a $\Delta G°$ value for the transition from tetramer to unfolded monomer of ~60 kcal/mol. The equilibrium dissociation constant derived from the free energy value for the dimer−dimer interface indicates that the protein does not dissociate at concentrations above $\sim 10^{-15}$ M monomer. This equilibrium constant can be compared to those estimated using other techniques. Previously, the unfolding transition for *lac* repressor was monitored by fluorescence polarization in pressure denaturation experiments (*29, 30*). Unfolding of wild-type and modified repressors results in an estimated equilibrium dissociation constant for what was presumed to be a tetramer to dimer transition for the wild-type protein of 1×10^{-8} to 1×10^{-9} M (*30*). These values result in a much lower free energy estimation than values obtained from the present urea denaturation experiments and are also contradicted by operator binding experiments (see below; *31, 35*). In the pressure denaturation experiments, treatment of protein prior to the experiment, assumptions inherent in extrapolating the data to atmospheric pressure, and the unknown oligomeric state of the transition end point could all contribute to the lower apparent free energy values.

Operator binding studies provide the most conclusive evidence that *lac* repressor is a tightly assembled tetramer. Brenowitz et al. (*31*) used footprint titration analysis and electrophoretic mobility shifts to correlate double-site occupancy with looped or tandem DNA−protein complexes. These data could be fit to any value for tetramer−dimer dissociation between 1×10^{-10} and 1×10^{-18} M. Finzi and Gelles (*27*) monitored the formation of loops by *lac* repressor for DNA fragments with two operator binding sites. Their conclusion was that a tetramer to dimer equilibrium did not contribute to DNA unlooping (*27*), and therefore the protein−protein interactions must be stronger than the protein−operator interactions ($K_d < 10^{-11}$ M). Fickert and Müller-Hill (*32*) determined the fraction of tetrameric wild-type protein present in solution by indirectly detecting the amount of protein associated with two operator fragments at saturating concentrations of DNA. They concluded that at 3×10^{-11} M protein, more than 80% of the complexes were tetrameric, while at 3×10^{-12} M only 30−50% were still tetrameric. Levandoski et al. (*35*) extensively modeled filter binding experiments over a wide concentration range of ligand and protein to detect the potential contribution of dimeric *lac* repressor protein to operator DNA binding. A thermodynamic model incorporating the presence of dimer−operator, tetramer−operator, and tetramer−operator$_2$ species for fitting these extensive binding isotherms did not indicate the presence of dimer−operator binding even in the femtomolar protein concentration range (*35*). These experiments set a concentration limit for *lac* repressor tetramer dissociation and indicate that *lac* repressor is tetrameric in the concentration range in which it interacts with DNA. Our present estimate of $\sim 10^{-15}$ M for the tetramer to dimer transition based on the free energy of dissociation for the dimer−dimer interface measured using the K84L mutant protein is consistent with these results. This correspondence further justifies using this value to determine the contribution of the dimer−dimer interface derived from the K84L ultracentrifugation data to estimate the overall stability of the repressor protein.

lac repressor proteins with mutations in the subunit interfaces allow the energetic contribution of each interface to be estimated and permit speculation about how the *lac* repressor assembles both in vitro and in vivo. The free energy values determined were assigned to each step of the pathway for *lac* repressor folding and assembly (Figure 9). An important note is that the dimeric structure is never actually present in the transition that occurs for the wild-type protein because the monomer−monomer association free energy is lower than that for the dimer−dimer association. These values explain why *lac* repressor dimers are not detected in operator binding experiments even at low concentrations in vitro (*31, 33, 35*; Moraitis and Matthews, unpublished results). The results of this study also demonstrate the difficulties of refolding proteins in vitro that possess two interfaces with different stabilities. In vivo, the N→C-terminal synthesis of proteins may prevent this assembly problem. Attachment to the ribosome may allow the monomer to fold and the dimer to form before the C-terminal region is present to form a tetramer. Therefore, in the bacterium, the reaction would be second order and less complicated than the fourth-order reaction that must occur in vitro. The overall stability and strength of the tetramer interface indicate that *lac* repressor under most circumstances does not exist as a dimer in the cell and confirm that the tetramer to dimer equilibrium does not play a regulatory role in modulating formation of looped DNA structures.

ACKNOWLEDGMENT

We thank Markos Moraitis for assistance with data analysis and simulations and Drs. J. S. Olson and M. S. Hargrove for discussions of the data and their interpretation.

REFERENCES

1. Miller, J. H., and Reznikoff, W. S. (1980) *The Operon*, 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
2. Gilbert, W., and Müller-Hill, B. (1966) *Proc. Natl. Acad. Sci. U.S.A. 56*, 1891−1898.
3. Riggs, A. D., and Bourgeois, S. (1968) *J. Mol. Biol. 34*, 361−364.
4. Müller-Hill, B. (1971) *Angew. Chem., Int. Ed. Engl. 10*, 160−172.
5. Adler, K., Beyreuther, K., Fanning, E., Geisler, N., Gronenborn, B., Klemm, A., Müller-Hill, B., Pfahl, M., and Schmitz, A. (1972) *Nature 237*, 322−327.
6. Jovin, T. M., Geisler, N., and Weber, K. (1977) *Nature 269*, 668−672.
7. Chuprina, V. P., Rullmann, J. A. C., Lamerichs, R. M. J. N., van Boom, J. H., Boelens, R., and Kaptein, R. (1993) *J. Mol. Biol. 234*, 446−462.
8. Lewis, M., Chang, G., Horton, N. C., Kercher, M. A., Pace, H. C., Schumacher, M. A., Brennan, R. G., and Lu, P. (1996) *Science 271*, 1247−1254.
9. Pfahl, M., Stockter, C., and Gronenborn, B. (1974) *Genetics 76*, 669−679.
10. Files, J. G., and Weber, K. (1976) *J. Biol. Chem. 251*, 3386−3391.
11. Schmitz, A., Schmeissner, U., Miller, J. H., and Lu, P. (1976) *J. Biol. Chem. 251*, 3359−3366.
12. Miller, J. H. (1979) *J. Mol. Biol. 131*, 249−258.
13. Kleina, L. G., and Miller, J. H. (1990) *J. Mol. Biol. 212*, 295−318.
14. Chakerian, A. E., and Matthews, K. S. (1991) *J. Biol. Chem. 266*, 22206−22214.
15. Markiewicz, P., Kleina, L. G., Cruz, C., Ehret, S., and Miller, J. H. (1994) *J. Mol. Biol. 240*, 421−433.
16. Friedman, A. M., Fischmann, T. O., and Steitz, T. A. (1995) *Science 268*, 1721−1727.
17. Chakerian, A. E., Tesmer, V. M., Manly, S. P., Brackett, J. K., Lynch, M. J., Hoh, J. T., and Matthews, K. S. (1991) *J. Biol. Chem. 266*, 1371−1374.
18. Alberti, S., Oehler, S., von Wilcken-Bergmann, B., Krämer, H., and Müller-Hill, B. (1991) *New Biol. 3*, 57−62.
19. Chen, J., and Matthews, K. S. (1992) *J. Biol. Chem. 267*, 13843−13850.
20. Daly, T. J., and Matthews, K. S. (1986) *Biochemistry 25*, 5474−5478.
21. O'Gorman, R. B., Rosenberg, J. M., Kallai, O. B., Dickerson, R. E., Itakura, K., Riggs, A. D., and Matthews, K. S. (1980) *J. Biol. Chem. 255*, 10107−10114.
22. Whitson, P. A., and Matthews, K. S. (1986) *Biochemistry 25*, 3845−3852.
23. Whitson, P. A., Olson, J. S., and Matthews, K. S. (1986) *Biochemistry 25*, 3852−3858.
24. Hsieh, W.-T., Whitson, P. A., Matthews, K. S., and Wells, R. D. (1987) *J. Biol. Chem. 262*, 14583−14591.
25. Krämer, H., Niemöller, M., Amouyal, M., Revet, B., von Wilcken-Bergmann, B., and Müller-Hill, B. (1987) *EMBO J. 6*, 1481−1491.
26. Eismann, E. R., and Müller-Hill, B. (1990) *J. Mol. Biol. 213*, 763−775.
27. Finzi, L., and Gelles, J. (1995) *Science 267*, 378−380.
28. Oehler, S., Eismann, E. R., Krämer, H., and Müller-Hill, B. (1990) *EMBO J. 9*, 973−979.
29. Royer, C. A., Weber, G., Daly, T. J., and Matthews, K. S. (1986) *Biochemistry 25*, 8308−8315.
30. Royer, C. A., Chakerian, A. E., and Matthews, K. S. (1990) *Biochemistry 29*, 4959−4966.
31. Brenowitz, M., Pickar, A., and Jamison, J. (1991) *Biochemistry 30*, 5986−5998.
32. Fickert, R., and Müller-Hill, B. (1992) *J. Mol. Biol. 226*, 59−68.
33. Chen, J. (1993) Ph.D. Dissertation, Rice University.
34. Chen, J., and Matthews, K. S. (1994) *Biochemistry 33*, 8728−8735.
35. Levandoski, M. M., Tsodikov, O. V., Frank, D. E., Melcher, S. E., Saecker, R. M., and Record, M. T., Jr. (1996) *J. Mol. Biol. 260*, 697−717.
36. Dill, K. A. (1990) *Biochemistry 29*, 7133−7155.
37. Creighton, T. E. (1990) *Biochem J. 270*, 1−16.
38. Kim, P. S., and Baldwin, R. L. (1990) *Annu. Rev. Biochem. 59*, 631−660.
39. Jaenicke, R. (1987) *Prog. Biophys. Mol. Biol. 49*, 117−237.
40. Neet, K. E., and Timm, D. E. (1994) *Protein Sci. 3*, 2167−2174.
41. Jaenicke, R. (1996) *Curr. Top. Cell. Regul. 34*, 209−314.
42. Nichols, J. C., and Matthews, K. S. (1997) *J. Biol. Chem. 272*, 18550−18557.
43. O'Gorman, R. B., Dunaway, M., and Matthews, K. S. (1980) *J. Biol. Chem. 255*, 10100−10106.
44. Barry, J. K., and Matthews, K. S. (1997) *Biochemistry 36*, 15632−15642.
45. Riggs, A. D., Bourgeois, S., Newby, R. F., and Cohn, M. (1968) *J. Mol. Biol. 34*, 365−368.
46. Bourgeois, S. (1971) *Methods Enzymol. 21*, 491−500.
47. Chang, W.-I., Olson, J. S., and Matthews, K. S. (1993) *J. Biol. Chem. 268*, 17613−17622.
48. Kawahara, K., and Tanford, C. (1966) *J. Biol. Chem. 241*, 3228−3232.
49. Cohn, E. J., and Edsall, J. T. (1943) *Proteins, Amino Acids, and Peptides as Ions and Dipolar Ions*, p 372, Van-Nostrand-Reinhold, New York.
50. Greene, R. F., Jr., and Pace, C. N. (1974) *J. Biol. Chem. 249*, 5388−5393.
51. Pace, C. N. (1975) *CRC Crit. Rev. Biochem. 3*, 1−43.
52. Johnson, M. L., and Frasier, S. G. (1985) *Methods Enzymol. 117*, 301−342.
53. Schnarr, M., and Maurizot, J.-C. (1981) *Biochemistry 20*, 6164−6169.
54. Nichtl, A., Buchner, J., Jaenicke, R., Rudolph, R., and Scheibel, T. (1998) *J. Mol. Biol. 282*, 1083−1091.
55. Cheng, X., Gonzalez, M. L., and Lee, J. C. (1993) *Biochemistry 32*, 8130−8139.
56. Malecki, J., and Wasylewski, Z. (1997) *Eur. J. Biochem. 243*, 660−669.
57. Clark, A. C., Sinclair, J. F., and Baldwin, T. O. (1993) *J. Biol. Chem. 268*, 10773−10779.
58. Couthon, F., Clottes, E., Ebel, C., and Vial, C. (1995) *Eur. J. Biochem. 234*, 160−170.
59. Mei, G., Di Venere, A., Buganza, M., Vecchini, P., Rosato, N., and Finazzi-Agro, A. (1997) *Biochemistry 36*, 10917−10922.
60. Pohl, M., Grotzinger, J., Wollmer, A., and Kula, M. R. (1994) *Eur. J. Biochem. 224*, 651−661.
61. Zhuang, P., Eisenstein, E., and Howell, E. E. (1994) *Biochemistry 33*, 4237−4244.

BI9900727