Articles

Unfolding and Disassembly of the Chaperonin GroEL Occurs via a Tetradecameric Intermediate with a Folded Equatorial Domain†

Jiwen Chen and David L. Smith*

Department of Chemistry and Eppley Institute for Research in Cancer and Allied Diseases, University of Nebraska, Lincoln, Nebraska 68588-0304

Received November 12, 1999; Revised Manuscript Received February 8, 2000

ABSTRACT: The chaperonin GroEL is a homotetradecamer in which the subunits (M_r 57 000) are joined through noncovalent forces. This study reports on the unfolding and disassembly of GroEL in guanidine hydrochloride and urea. Kinetic and equilibrium measurements were made using amide hydrogen exchange/mass spectrometry, light scattering, and size-exclusion chromatography. Hydrogen exchange in GroEL destabilized in 1.8 M GdHCl (the unfolding midpoint is 1.2 M GdHCl) shows that the apical and intermediate domains unfold 3.1 times faster than the equatorial domain. Light scattering measurements made under the same conditions show that disassembly of the native GroEL tetradecamer occurs at the same rate as unfolding of the equatorial domain. This study of the kinetics of GroEL unfolding and disassembly demonstrates the existence of an intermediate that was identified as a tetradecamer with the apical and intermediate domains unfolded. Although this intermediate was easily detected in dynamic unfolding measurements, its population in equilibrium measurements at the midpoint for GroEL unfolding was too small to be detected. This study of GroEL unfolding and disassembly points to features that may be important in the folding and assembly of the GroEL macroassembly.

The chaperonin GroEL, the most thoroughly characterized molecular chaperone, assists protein folding in vivo (*1*−*3*). GroEL is a homotetradecamer composed of two back-to-back seven-member rings. Each subunit consists of 547 residues and 3 distinctive structural domains (*4*, *5*). The equatorial domain (residues 5−132 and 408−522) is well-ordered and forms a solid foundation around the waist of the assembly. The apical domain (residues 190−375) is considerably less ordered and provides the binding site for the substrate protein. An intermediate domain (residues 133−189 and 376−407) links the equatorial and apical domains.

Little is known about the processes through which GroEL folds into these three structural domains, nor is much known about how GroEL subunits assemble to form the biologically active tetradecamer. There is increasing evidence suggesting that, in some cases, folding and unfolding involve similar paths on the folding energy landscape (*6*−*8*). When folding and unfolding do occur along similar paths, important features of the folding energy landscape can be deduced from studies of protein unfolding. This prospect led us to consider several previous studies of GroEL unfolding. Light scattering, circular dichroism (225 nm), and fluorescence measurements indicated that equilibration in 1.2−1.3 M GdHCl[1] induced major structural changes in GroEL (*9*, *10*). However, it was noted that GroEL lost ATPase activity at much lower concentrations of GdHCl (midpoint 0.5 M GdHCl) (*9*). Subsequent light scattering and fluorescence experiments also indicated that GdHCl induced an additional structural change when the concentration of denaturant was less than 0.5 M (*10*). Structural changes in GroEL induced by denaturants have also been studied by size exclusion chromatography and sedimentation velocity analysis. The size exclusion results indicated that GroEL remains a tetradecamer for GdHCl concentrations as high as 0.6 M (*10*). The sedimentation velocity results showed that GroEL is a tetradecamer in 2.0 M urea and a monomer in 2.3 M urea (*11*). Circular dichroism (CD) analysis of GroEL equilibrated in urea indicated a single major structural change with the midpoint at 2.2 M urea, in agreement with the sedimentation velocity measurements (*12*). Another report indicated that the midpoint for GroEL unfolding occurred at 3.3 M urea (*9*).

Collectively, these results show that low concentrations of GdHCl (<0.5 M) inhibit ATPase activity of GroEL and alter slightly its fluorescence and light scattering without appreciably altering its CD properties. However, these structural changes do not involve disassembly of the GroEL

† This work was supported by a grant from the National Institutes of Health (RO1 GM 40384) and by the Nebraska Center for Mass Spectrometry.

* Address correspondence to this author at the Department of Chemistry, University of Nebraska−Lincoln, Lincoln, NE 68588-0304. Tel: 402-472-2794, Fax: 402-472-9402, Email: dsmith7@unl.edu.

[1] Abbreviations: HX, hydrogen exchange; MS, mass spectrometry; CD, circular dichroism; HPLC, high-performance liquid chromatography; kDa, kilodalton(s); GdHCl, guanidine hydrochloride; GdDCl, deuterated guanidine hydrochloride; BSA, bovine serum albumin; Hepes, *N*-(2-hydroxyethyl)piperazine-*N*′-2-ethanesulfonic acid; Tris, tris(hydroxymethyl)aminomethane.

10.1021/bi992619n CCC: $19.00
Published on Web 03/23/2000

tetradecamer. Higher concentrations of denaturant (1.2 M GdHCl or 2−3 M urea) cause major structural changes (fluorescence and CD) and disassembly of the tetradecamer (light scattering and sedimentation velocity). In addition to these structural changes, residual structure in the apical domain has been detected by the fluorescence of bisANS when GroEL was equilibrated in urea (*12*). This structure, which was not detected by other spectroscopic methods, was lost when the urea concentration was greater than 3−4 M.

Results presented here provide a new view, based on amide hydrogen exchange (HX), of denaturant-induced unfolding and disassembly of GroEL. Isotope exchange was detected by analyzing both intact GroEL subunits and their peptic fragments by mass spectrometry (MS) (*13*−*16*). Analysis of multiple envelopes of isotope peaks in mass spectra showed that different regions of GroEL unfold at different rates. Furthermore, unfolding of the equatorial domain (based on HX) and disassembly of the GroEL tetradecamer (based on light scattering) occurred at the same rate, suggesting that they are closely coupled processes. The present results are combined with those of previous studies to develop a more detailed model for folding and assembly of native GroEL.

MATERIALS AND METHODS

Unfolding and Labeling. E. coli GroEL was overexpressed in *E. coli*, as described previously (*17*). Purified GroEL in Hepes, pH 7.4, was diluted 13.5-fold into a guanidine hydrochloride (GdDCl)/D_2O buffer (50 mM Hepes, pD 7.0) to initiate unfolding and subsequent deuterium exchange-in. All measurements of pH in solutions containing D_2O are presented here "as read" from the pH meter (*18*). The final concentration of GdDCl was 1.8 M, and the concentration of the GroEL tetradecamer was 1.1 μM. The temperature was 21 ± 2 °C. After various unfolding times (10, 22, 29, 45, 90, 180, and 360 s), isotope exchange was quenched by decreasing the pH to 2.4 and the temperature to 0 °C. The quench solution was 300 mM phosphate at pH 2.4. Samples labeled with deuterium were stored at −70 °C until analyzed by HPLC MS. Deuterium uptake in native GroEL in the absence of denaturant was determined from a "native" protein reference sample, which was prepared by incubating GroEL in the labeling buffer without GdDCl for 10 s. Deuterium loss during analysis was estimated by analyzing a sample that was completely exchanged in D_2O. This "unfolded" reference sample was prepared by incubating the protein in the labeling buffer containing 1.8 M GdDCl for 1 h. These reference samples, as well as samples from the time course, were analyzed by the same quench and HPLC MS procedures. About 400 pmol (subunit basis) of GroEL was used for each measurement.

Some measurements were made under equilibrium conditions. Preliminary measurements showed that approximately 4 h was required to achieve equilibrium populations of folded and unfolded forms of GroEL in 1.2 M GdHCl. The time required to achieve equilibrium, which depends on the concentration of denaturant, was considered in the design of all equilibrium experiments.

Analysis of Intact GroEL Subunits. Mass spectra of the intact GroEL subunits were acquired with a Micromass Platform II electrospray ionization mass spectrometer directly coupled to a reversed phase HPLC system. The injector and the column (Vydac C4 1 × 50 mm) were maintained at 0 °C, and both mobile phases (water and acetonitrile) contained 0.05% trifluoroacetic acid. An aliquot (50 μL) of sample was loaded using a 50 μL injection loop and a flow rate of 140 μL/min. Following a 1.5 min desalting period at 30% acetonitrile, the flow rate was decreased to 40 μL/min, and the protein was eluted using a 30−65% gradient in 4.5 min. The mass spectrometer was scanned over the mass/charge range 700−1800. Deuterium levels were determined from the centroids of peaks in the mass spectra reconstructed from the multiply charged state ions in the original spectra. The general procedures used for amide hydrogen exchange and mass spectrometry are described in more detail elsewhere (*13, 19*).

Pepsin Digestion and Fragment Analysis. The labeled protein (400 pmol of monomer) was digested with pepsin for 3.5 min (S:E mass ratio 1:1, 0 °C; pH 2.4) before analysis by HPLC MS. The HPLC injector and column (Vydac C4 1 × 50 mm) were maintained at 0 °C. An aliquot of the digest (50 μL) was loaded into a 50 μL loop at a flow rate of 40 μL/min. After desalting with 2% acetonitrile for 2 min, peptic fragments were eluted using a 2−38% acetonitrile gradient in 10 min and detected on-line by a Micromass AutoSpec magnetic sector mass spectrometer equipped for electrospray ionization. Ions in the mass-to-charge range 380−1350 were recorded using a focal plane detector. Although pepsin cleavage sites are difficult to predict from sequence alone, cleavage was reproducible under identical digestion conditions. All peptides were identified by collision-induced dissociation tandem mass spectrometry using a Finnigan LCQ ion trap mass spectrometer.

Data Analysis. Mass spectra of peptides exhibiting bimodal distributions of isotopes were deconvoluted into their constituent low- and high-mass envelopes using the computer program "Hxpro" (*20*). The areas of these envelopes were used to determine the populations of GroEL that were unfolded in specific regions. Unfolding rate constants were determined from these populations assuming consecutive first-order reactions:

$$\mathrm{N} \xrightarrow{k_1} \mathrm{I} \xrightarrow{k_2} \mathrm{U}$$

where N, I, and U refer to the populations of the native, intermediate, and unfolded species, respectively. Unfolding rate constants, k_1 and k_2, were calculated by fitting the following equations to the experimental data:

$$[\mathrm{N}]_t = [\mathrm{N}]_0 \exp(-k_1 t) \quad (1)$$

$$[\mathrm{U}]_t = [\mathrm{N}]_0 \left[1 + \frac{k_1 \exp(-k_2 t) - k_2 \exp(-k_1 t)}{k_1 - k_2}\right] \quad (2)$$

where $[\mathrm{N}]_0$ is the initial native protein population.

Light Scattering. Right angle light scattering (323 nm) measurements were performed on a SLM Aminco Bowman series 2 luminescence spectrometer. GroEL was diluted into the isotope labeling buffer (1.8 M GdDCl) to give a final protein concentration of 1.1 μM (tetradecamer).

Size Exclusion Chromatography. GroEL incubated in urea for 24 h was loaded on a Superose 12HR column (Pharmacia) and detected by UV absorbance at 229 nm. The mobile phase was 50 mM Tris, 100 mM KCl, 2.5 or 2.6 M urea, pH 7.8.

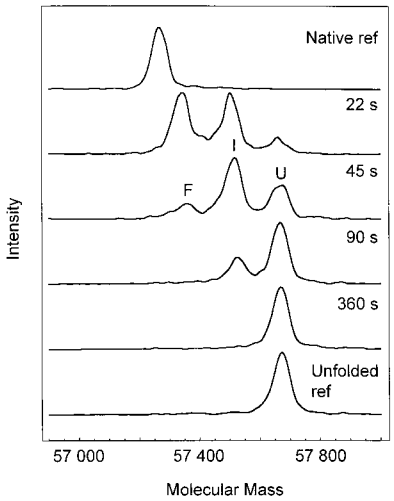

FIGURE 1: Electrospray mass spectra of intact GroEL monomers derived from native GroEL incubated in 1.8 M D_2O/GdDCl for 22, 45, 90, and 360 s. The native ref sample was prepared by exposing native GroEL to D_2O for 10 s in the absence of GdHCl. The unfolded ref sample represents GroEL that was completely exchanged in D_2O. These spectra were reconstructed from peaks for multiply charged ions in the original mass spectra. The area under each envelope of isotope peaks was used to determine the populations of folded, partially unfolded intermediate, and unfolded forms of GroEL.

To correlate the elution time with molecular mass, the elution times of native GroEL (M_r 800 000) and BSA (M_r 66 000) were determined using the same buffer but without urea.

RESULTS

Detection of a GroEL Unfolding Intermediate. Isotope exchange rates of hydrogen located at peptide amide linkages are highly sensitive to intramolecular hydrogen bonding and their access to the solvent (*21*−*25*). Since unfolding disrupts intramolecular hydrogen bonding and provides essentially complete access to the solvent, isotope exchange rates in unfolded proteins are much faster than in the folded forms. For example, exchange is about 85% complete in 1 s for an unfolded protein at pH 7.0 (*18*), but may require months to reach equilibrium in a folded protein. The high sensitivity of hydrogen exchange to the folding status of proteins has been used successfully in several studies of protein folding (*26*−*28*) and protein unfolding (*29*−*32*).

Amide hydrogen exchange was used in the present study to investigate the unfolding of GroEL in 1.8 M GdHCl. Results from previous CD and fluorescence studies showed that GroEL unfolding is complete in 1.8 M GdHCl (*9, 10*). In the present study, GroEL was unfolded in the presence of D_2O. Under these conditions, HX occurred slowly at most peptide amide linkages in folded molecules, proceeding rapidly only when molecules or specific regions within molecules unfolded. The large difference in HX rates in folded and unfolded GroEL is illustrated by the mass spectra of the native (M_r 57 266) and unfolded (M_r 57 671) reference samples (Figure 1). The average molecular mass measured for GroEL with a natural abundance of isotopes was 57 195 (calculated M_r 57 198). It is important to note that deuterium located in the side chains, as well as at the N- and C-termini, was replaced with protium during the HPLC step (*18, 33, 34*). As a result, the increased molecular mass of GroEL reflects only deuterium located at peptide amide linkages.

Table 1: Deuterium Levels Found in GroEL Monomers following Incubation in D_2O and 1.8 M GdDCl

	deuterium level[a]				
incubation time (s)	folded	inter-mediate	unfolded	intermediate−folded	unfolded−intermediate
native ref[b]	79				
10	136	332		196	
22	163	343	516	180	173
29	171	344	521	173	177
45	182	358	529	176	172
90		368	529		162
180			529		
360			530		

[a] The deuterium level was determined from the molecular masses of intact GroEL monomers (see Figure 1) and adjusting for deuterium losses during HPLC (*35*). [b] Native GroEL was exposed to D_2O with no GdDCl for 10 s.

The GroEL amino acid sequence used in this study, which is based on the scheme used by SwissProt (entry no. P06139), does not include Met 1, which was cleaved after translation. This form of GroEL, which has 547 residues (14 of which are proline), has 532 peptide amide linkages that could become deuterated. Assuming that HX was complete in the unfolded reference, GroEL should have 532 deuteriums located at peptide amide linkages. Finding only 476 indicates that 11% of this deuterium was lost during the HPLC step. This loss was used to adjust for deuterium losses during HPLC in other samples (*35*).

For a structurally homogeneous sample, deuterium is distributed randomly among all the molecules. Mass spectra of labeled proteins with a random distribution of deuterium have single envelopes of isotope peaks, as illustrated in Figure 1 for the reference samples. Mass spectra of proteins in which some of the molecules were folded and others were unfolded during isotopic labeling may exhibit bimodal isotope patterns (*28, 31, 32, 36*). A significant population of an intermediate state along the unfolding pathway in which only part of the molecule is folded may give an additional envelope of isotope peaks indicating a deuterium level intermediate between those of the native and unfolded references.

The deuterium uptake by GroEL was followed by mass spectrometry as a function of the time the protein was incubated in D_2O and 1.8 M GdDCl. After 10 s, two envelopes of isotope peaks were detected. The deuterium levels found in different forms of GroEL are given in Table 1. After 22 s, three envelopes of isotope peaks were evident (Figure 1), indicating the presence of three conformations of GroEL (F, I, and U). The deuterium level of ions comprising the high-mass envelope indicated that nearly all of the 532 amide hydrogens had been replaced with deuterium. It follows that this peak represents completely unfolded GroEL molecules, as sensed by hydrogen exchange. As the incubation time in D_2O/GdDCl increased to 360 s, the high-mass envelope became the dominant peak in the mass spectrum, indicating that GroEL was becoming increasingly unfolded. During this time, the deuterium level increased from 516 to 529 (see Table 1). Although this small increase in mass (0.02%) with exposure to D_2O /GdDCl may indicate some residual structure, it is close to the uncertainty of these measurements and may not be significant.

Native GroEL exposed to D_2O (no GdDCl) for 10 s had an excess of 79 deuteriums, indicating that HX was very fast at approximately 79 amide linkages (Table 1). The mass of the low-mass envelope for GroEL exposed to D_2O and GdDCl for 10 s indicated an excess of 136 deuteriums. The increase from 79 (native reference) to 136 indicates the loosening effect of GdDCl on GroEL that remained folded during the incubation time. As the incubation time increased from 10 to 45 s, the deuterium level in this population of molecules increased from 136 to 182 (Table 1). This increase reflects the increased time available for HX and possible continued loosening of folded GroEL. For incubation times greater than 45 s, this population was too small to be detected readily.

Three spectra in Figure 1 have a third peak representing partially unfolded GroEL monomers (I) with deuterium levels intermediate between those of the folded and unfolded forms. The molecular masses of this population indicated that their deuterium levels increased from 332 to 368 as the incubation time increased from 10 to 90 s. As described above for the population of folded GroEL, this increase is due to the increased time available for hydrogen exchange and possibly to increased loosening of the folded regions. The difference in the deuterium levels between two peaks (right columns in Table 1) gives the number of amide linkages that became deuterated following a particular unfolding step. Since exchange was complete at some amide linkages before unfolding occurred, the differences in the deuterium levels found between the three envelopes of isotope peaks represent only the lower limit to the number of residues involved in each unfolding domain. The results presented in Table 1 show that the first unfolding step must involve at least 196 residues and the second unfolding step must involve at least 177 residues.

Results of a separate pulse-labeling experiment confirmed that GroEL equilibrated in 1.8 M GdHCl was completely unfolded, as assessed by HX. Following equilibration in H_2O/GdHCl for 360 s, GroEL was exposed to D_2O/GdDCl for 10 s prior to quenching HX. The mass spectrum of this sample (data not presented) exhibited only one envelope of isotope peaks and a deuterium level of 532, indicating little protection against HX. The absence of peaks corresponding to either folded GroEL or the partially unfolded intermediate confirmed that these forms were much less stable than completely unfolded GroEL in 1.8 M GdHCl. These results are consistent with previous studies which reported that the midpoint for equilibrium unfolding of GroEL occurs at 1.2 M GdHCl (*9, 10*).

As discussed elsewhere (*19, 37*), peak intensities in continuous-labeling experiments represent populations that unfolded and underwent complete isotope exchange. This population includes molecules that were unfolded at the time that HX was quenched as well as those molecules that refolded prior to HX quenching. Because equilibrium in 1.8 M GdDCl strongly favored the unfolded protein, the population that unfolded and subsequently refolded prior to quench was small. Under these conditions, the peak intensities in Figure 1 indicate the populations of the folded, intermediate, and unfolded forms of GroEL at the time HX was quenched. The population of the folded form followed a single-exponential decay with increasing incubation time in D_2O/GdDCl. The population of the intermediate rose to a maximum and decayed to zero as the population of the unfolded form rose from 0 to 100%. Rate constants in eqs 1 and 2 were adjusted to optimize the fit between experimental and calculated populations. Unfolding rate constants of 0.048 and 0.015 s^{-1} gave the best fit for k_1 and k_2, respectively.

Identification of the Unfolding Domains in GroEL. To identify specific residues comprising the two unfolding domains detected by analysis of the intact GroEL monomers, the labeled protein was digested into peptic fragments whose deuterium levels were determined by HPLC MS. Except for addition of pepsin to the quench solution and a 3.5 min digestion time, the experimental conditions used to prepare samples for the two types of analyses were identical. For the experimental conditions used in this study, 54 peptic fragments covering 97% of the entire GroEL backbone were identified. A subset of these fragments covering 91% of the backbone was used in this study. The amino acid sequences of GroEL and the peptic fragments used in this study are presented in Figure 2.

Mass spectra of the segment including residues 214−221 derived from GroEL incubated in D_2O/GdDCl for 22, 45, and 90 s are presented in Figure 3. These spectra have two envelopes of isotope peaks (illustrated by broken lines), clearly demonstrating structural heterogeneity in this region of the GroEL backbone. The low-mass envelope of isotope peaks had an average mass similar to that of the native reference (top panel), indicating that the population of molecules represented by this envelope was folded. The high-mass envelope of the isotope peaks had an average mass similar to that of the unfolded reference (bottom panel), indicating that the population of molecules represented by this envelope was unfolded.

The relative areas of the low- and high-mass envelopes indicate the relative populations of molecules in which this segment was folded or unfolded, respectively (*19, 28*). With increasing incubation time in 1.8 M GdDCl, the population of molecules folded in this region decreased while the population of molecules unfolded in this region increased (Figure 3). In the backbone segment including residues 214−221, approximately 62% of the molecules were unfolded after 22 s of incubation in GdDCl. Approximately 89% of the molecules were unfolded after 45 s. By contrast, unfolding of the GroEL backbone segment including residue 419−425 was much slower. After 22 s, only 14% of the molecules were unfolded; after 90 s, 63% of the molecules were unfolded.

Similar analyses were performed on 38 peptic fragments of GroEL to determine the folding status of specific regions in GroEL as a function of the time it was incubated in D_2O/GdDCl. Unfolding rate constants for these regions, which were determined using eqs 1 and 2, are listed in Table 2. The unfolding rates of backbone regions corresponding to these 38 peptic fragments of GroEL form 2 distinct groups with unfolding rate constants of 0.049 and 0.016 s^{-1}. These unfolding rate constants correlate very closely with the unfolding rate constants determined by analyzing intact GroEL (Table 2). Peptides indicating fast or slow folding of GroEL are illustrated in Figure 2 as thick or thin lines, respectively. All of the fast unfolding peptides (20) map to the apical and intermediate domains, and all of the slow unfolding peptides (18) map to the equatorial domain (Figure 4).

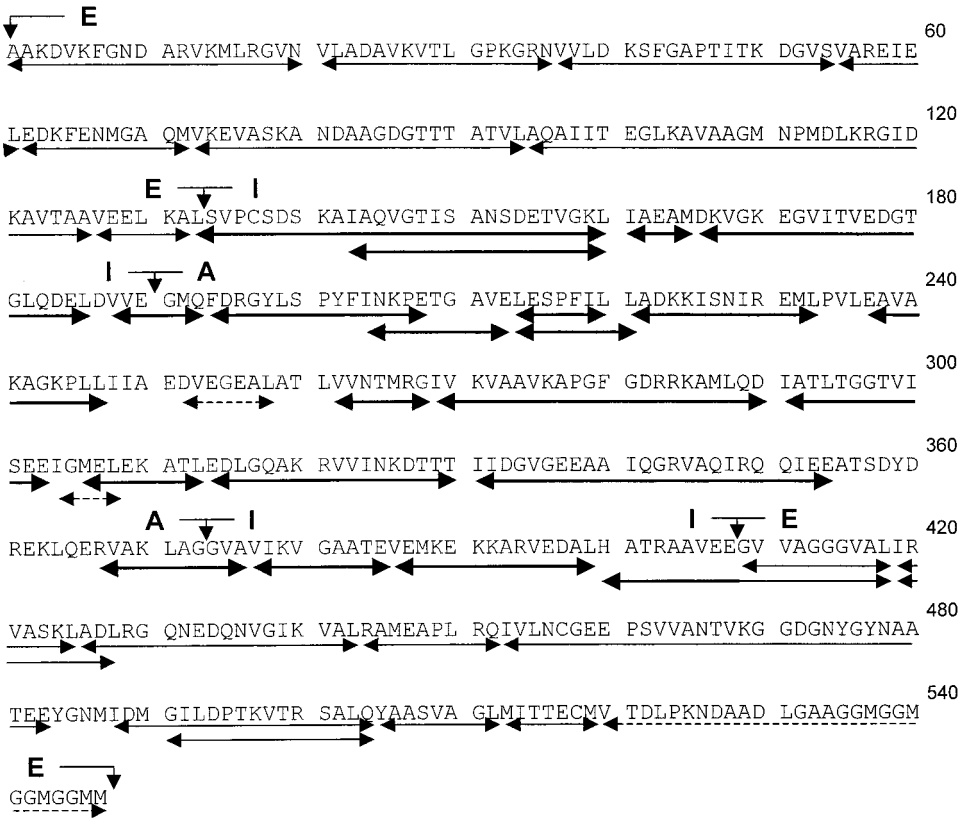

FIGURE 2: Amino acid sequence of GroEL (*46*). Arrows indicate the peptic fragments used in this study. Heavy and light arrows indicate segments derived from regions unfolding rapidly or slowly, respectively. Dashed lines indicate segments that did not exhibit bimodal isotope patterns. The apical (A), intermediate (I), and equatorial (E) structural domains were identified by X-ray crystallography (*4*, *5*).

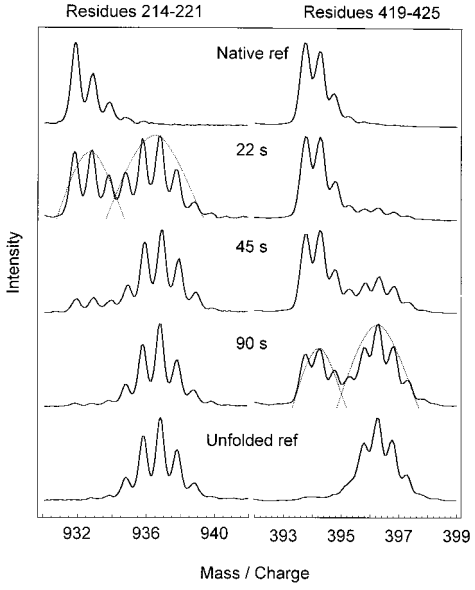

FIGURE 3: Electrospray mass spectra of two peptic fragments derived from GroEL after incubation of the native protein in D_2O and 1.8 M GdDCl for 22, 45, and 90 s. The native and unfolded references were described in Figure 1. The bimodal isotope distributions indicated by dashed lines were used to determine the populations of molecules that were unfolded in the region from which the peptides were derived.

Identification of unfolding domains by peptic fragments that exhibit similar unfolding rates can be extended to include comparison of the unfolding rates of the apical and intermediate domains. The mean unfolding rate constants for peptic fragments derived from the apical (14 peptides) and intermediate (6 peptides) domains are 0.048 and 0.053 s^{-1}, respectively, suggesting that the intermediate domain may unfold slightly faster than the apical domain. However, this difference is very near the confidence limit of these measurements and may not be significant. Although the exact unfolding mechanism cannot be determined from these results, it is evident that unfolding of the apical and intermediate domains of GroEL is a closely coupled process.

Table 2: Rate Constants for Unfolding of GroEL Incubated in 1.8 M GdHCl Determined from the Mass Spectra of Intact GroEL Monomers or Their Peptic Fragments

fast unfolding regions		slow unfolding regions	
residues	k (s^{-1})	residues	k (s^{-1})
133−160	0.051	1−20	0.017
143−160	0.054	21−36	0.015
161−165	0.052	37−54	0.016
166−186	0.058	55−61	0.016
188−193	0.051	62−72	0.017
194−208	0.047	73−94	0.016
204−213	0.051	95−126	0.011
214−220	0.049	127−132	0.015
214−221	0.049	409−418	0.018
221−233	0.043	419−425	0.017
237−247	0.047	419−428	0.017
262−268	0.043	426−443	0.018
269−290	0.048	444−452	0.016
291−303	0.045	453−483	0.018
306−313	0.051	488−504	0.018
314−330	0.045	491−504	0.017
331−354	0.052	505−512	0.015
367−376	0.053	513−519	0.015
377−385	0.055		
386−399	0.050		
av	0.049	av	0.016
intact monomer	0.048	intact monomer	0.015

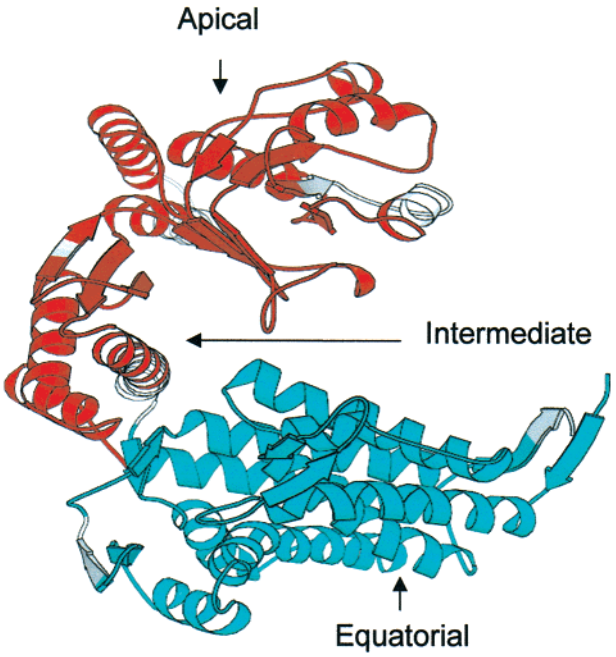

FIGURE 4: Structure of GroEL illustrating the apical, intermediate, and equatorial domains. Regions that unfold fast (apical and intermediate domains) or slow (equatorial domain) when GroEL was incubated in 1.8 M GdDCl are indicated by red and blue, respectively.

Most of the peptic fragments of GroEL analyzed in this study exhibited bimodal isotope distributions, suggesting that only the folded and unfolded forms were present at the time HX was quenched. Peaks due to structural forms with deuterium levels intermediate between deuterium levels found for the folded and unfolded forms were generally not detected. These results demonstrate the cooperative nature of the unfolding process. That is, these regions of the GroEL backbone were either folded or unfolded. The peptic fragment including residues 400−418 was an exception. Approximately half of this fragment is in the intermediate domain, and half is in the equatorial domain (see Figure 2). Mass spectra of this fragment had three envelopes of isotope peaks representing three populations of GroEL that differed by the folding status of residues 400−418. One population had all of these residues folded; another population had all of these residues unfolded; and another population had approximately half of these residues (as ascertained by the mass of the middle envelope) unfolded. These results demonstrate that the transition between the fast and slow unfolding domains was very close to the transition between the intermediate and equatorial structural domains (i.e., Glu 407).

Due to fast isotope exchange into the native form, several regions of the protein, including residues 252−258, 304−308, and 519−547, did not exhibit bimodal distributions of deuterium. Mass spectra of these peptic fragments had a single envelope of isotope peaks, the centroid of which indicated that the segments were completely deuterated. For example, after 10 s labeling in nondenaturing deuterated buffer, isotope exchange in the segment including residues 520−547 was essentially complete. This result indicates that this region in the native protein not only was solvent-accessible but also lacked stable secondary structure. These results are in agreement with the observation that residues 525−547 were not resolved in the crystal structure (*5*), as well as finding that the C-terminus protrudes from the equatorial domain into the GroEL central cavity to form a barrier between the two rings (*38*).

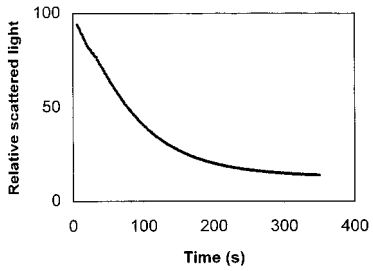

FIGURE 5: Change in intensity of scattered light (323 nm) with the time GroEL was incubated in 1.8 M GdDCl in D_2O.

Disassembly of GroEL Is Coupled to Unfolding of the Equatorial Domain. The hydrogen exchange results described above demonstrate that GroEL destabilized in GdHCl unfolded in two steps, where unfolding of the equatorial domain was the slower step. To determine whether either unfolding step correlates with the disassembly of the GroEL tetradecamer, the rate constant for disassembly of GroEL in GdHCl was determined by light scattering. This measurement was performed under the same conditions used to investigate unfolding of GroEL (D_2O/1.8 M GdDCl). The change in relative intensity of scattered light with the time GroEL was exposed to the denaturing conditions is presented in Figure 5. The magnitude of the intensity change from the start to the end of the time course was similar to that found for complete unfolding of native GroEL, indicating that the entire disassembly process was monitored in 1.8 M GdDCl. The decrease of scattered light intensity followed an exponential decay ($R^2 = 0.9993$) with a first-order rate constant of 0.014 s^{-1}. It may be of interest to note that performing the light scattering experiments under similar conditions but in H_2O gave similar results, but a disassembly rate constant of 0.044 s^{-1}. Previous studies have shown that many proteins fold faster in D_2O than in H_2O (*39*, *40*), suggesting that unfolding may be slower in D_2O. Finding that GroEL disassembles slower in D_2O is consistent with unfolding playing an integral role in the rate-limiting step for disassembly.

DISCUSSION

GroEL Unfolding. Analyses of intact GroEL monomers following equilibration of native GroEL in 1.8 M GdHCl showed that GroEL was unfolded, as sensed by amide hydrogen exchange. This low level of protection against hydrogen exchange demonstrated that most intramolecular hydrogen bonding was lost and that the amide hydrogens had ample access to the deuterated solvent. Mass spectra of intact GroEL monomers as well as peptic fragments derived from GroEL incubated in 1.8 M GdDCl show that unfolding involved at least two steps. In the first step, the apical and intermediate domains unfolded, either simultaneously or in rapid succession with a rate constant of 0.049 s^{-1}. Light scattering results show that native GroEL in 1.8 M GdDCl dissociated 3.1-fold slower ($k = 0.014$ s^{-1}). The excellent fit of the light scattering results to a single exponential term, as suggested previously (*11*), implies that the GroEL tetradecamer dissociates into monomers via a cooperative process or a series of closely coupled processes. These results suggest that the unfolding intermediate detected in 1.8 M

GdDCl is a tetradecamer in which the apical and intermediate domains of the 14 subunits are unfolded. Further analysis of the hydrogen exchange results shows that the unfolding rate of the equatorial domain ($k = 0.016\ s^{-1}$) is very similar to the disassembly rate of native GroEL under the same conditions. These results indicate that unfolding of the equatorial domain and disassembly of the tetradecamer are closely linked or possibly cooperative processes.

Previous studies show that the unfolding midpoint for GroEL, based on CD (220 nm), is 1.2 M GdHCl and that equilibrium strongly favors the unfolded form at 1.8 M (*9, 10*). The present hydrogen exchange results (both continuous labeling and pulse labeling) also show that GroEL is unfolded in 1.8 M GdHCl. Previous light scattering, ATPase activity, and fluorescence studies indicate that equilibration in GdHCl leads to two structural changes, one with a midpoint below 0.5 M and one with a midpoint at 1.2 M GdHCl (*9, 10*). Only the change at 1.2 M GdHCl was detected by CD, suggesting that this change was due to the loss of secondary structure of GroEL. Since loss of such secondary structure must involve loss of H-bonding, hydrogen exchange rates are expected to increase as α-helical structure is lost. Hence, the unfolding processes detected in the present hydrogen exchange experiments are likely the same structural changes detected by CD at 1.2 M GdHCl. Unfolding studies of two constructs of the apical domain suggest that the first unfolding step may be due to unfolding of helices 11 and 12 (*41*). In the present study, no evidence for this step was found in the unfolding rates of peptic fragments derived from these helices (Figure 2), perhaps because the concentration of GdHCl (1.8 M) was too high.

The present investigation of GroEL unfolding kinetics by hydrogen exchange shows that the apical and intermediate domains unfold prior to unfolding of the equatorial domains. Since all three domains of GroEL have high levels of α-helix character, the CD should change with the unfolding of each domain. However, equilibrium unfolding studies using CD indicated a single transition (*9, 10*), suggesting that the unfolding intermediate may be less stable than either the folded or the unfolded forms. As a result, its population might be too small to be detected. To test this hypothesis, additional hydrogen exchange experiments were performed in 1.2 M GdHCl, the apparent midpoint in the GdHCl-induced unfolding of GroEL. Following incubation of native GroEL in 1.2 M GdHCl for 0.8, 2.6, and 25 h, unfolded regions were labeled by exposing the protein to D_2O/1.2 M GdDCl for 10 s. Mass spectra of intact GroEL monomers (data not shown) exhibited bimodal isotope patterns indicating that approximately 71% was folded and 29% was completely unfolded (i.e., offered no protection to HX). There was no evidence of a third peak representing a partially unfolded population, as found when GroEL was incubated in 1.8 M GdDCl (Figure 1). These results show that the population of the partially unfolded intermediate must be very small in 1.2 M GdHCl, consistent with the CD results. Detection of the unfolding intermediate in 1.8 M GdDCl is possible because the unfolding rate constants for the apical and intermediate domains are much faster than the unfolding rate constant for the equatorial domain.

The unfolding of GroEL has also been studied in urea (*11, 12*). To relate results obtained when GroEL was destabilized in GdHCl or urea, the unfolding of GroEL in urea was also

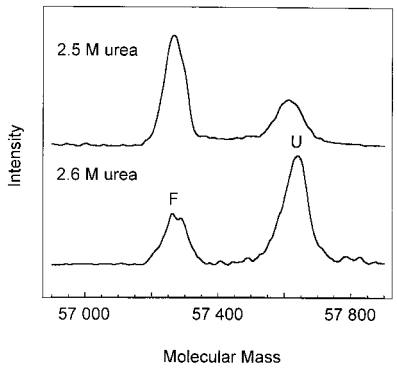

FIGURE 6: Mass spectra of intact GroEL monomers following 24 h equilibration in 2.5 or 2.6 M urea/H_2O and pulse labeling in urea/D_2O for 10 s.

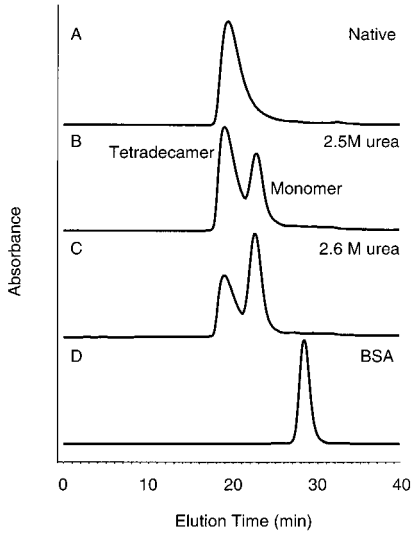

FIGURE 7: Size exclusion chromatography of GroEL following equilibration in 2.5 or 2.6 M urea for 24 h (panels B and C). Elution of molecular mass references, native GroEL and BSA (molecular mass 66 kDa), is presented in panels A and C. For panels B and C, the column was equilibrated with 50 mM Tris, 100 mM KCl, 2.5 or 2.6 M urea, pH 7.8; for panels A and D, the column was equilibrated with the same buffer but without urea.

investigated by hydrogen exchange. Mass spectra of native GroEL equilibrated in 2.5 or 2.6 M urea for 24 h followed by pulse labeling for 10 s (Figure 6) had two envelopes of isotope peaks, indicating that GroEL was either folded or unfolded. These results, which are consistent with those obtained in 1.2 M GdHCl, confirm that the equilibrium population of a partially unfolded intermediate is too small to be detected. To determine the aggregation state of GroEL under similar conditions, the same samples were analyzed by size exclusion chromatography. The chromatograms for GroEL equilibrated in 2.5 or 2.6 M urea (Figure 7) have two peaks corresponding to the GroEL tetradecamer and unfolded GroEL monomer. The relative intensities of the two peaks illustrate how the equilibrium populations of these two forms changed with small changes in urea concentration. These results show that the midpoint for unfolding and disassembly of GroEL is close to 2.5 M urea, consistent with previous studies using urea (*11, 12*).

GroEL Folding. Spontaneous folding of GroEL is inefficient (*9, 12*), suggesting that newly synthesized GroEL can fold only in the presence of functional GroEL (*42*). This hypothesis was confirmed by in vivo experiments (*43, 44*),

which demonstrated that one of the natural substrates for GroEL is GroEL itself. In vitro refolding of GroEL from low concentrations of denaturants (1 M GdHCl or 3 M urea) appeared to restore GroEL to its native structure (based on CD, fluorescence, and light scattering) and ATPase activity (*9, 12*). However, refolding from higher concentrations of denaturants gave non-native material with considerable secondary structure, but no ATPase activity or quaternary structure and an increased susceptibility to proteolysis. Investigations of denaturant-induced structural changes in GroEL by CD, velocity sedimentation, and light scattering suggested that the secondary and quaternary structures of GroEL were lost at the same concentration of denaturants. Binding of bisANS was maximum in 3 M urea, suggesting hydrophobic structure that was not detectable by fluorescence or CD (*12*). Photolabeling studies showed that bisANS binds to the region including residues 202−248 in the apical domain of GroEL. These results led Gorovtis et al. to propose a model for describing GroEL unfolding in denaturants. At low concentrations of denaturants, native GroEL dissociates to monomers. At intermediate levels of denaturants (3 M urea), these monomers retained some hydrophobic structure in the apical domain. When this residual structure was lost at higher concentrations of denaturants, GroEL would not refold to the native state. They also proposed that GroEL folding might start with formation of this hydrophobic core in the apical domain which facilitated folding of the N- and C-terminal domains to form the equatorial domain with eventual assembly of the native homotetradecamer.

Assuming that denaturants change the populations of various unfolding intermediates, but not the unfolding pathway, results obtained with different concentrations of denaturants may be combined to suggest an unfolding model for GroEL, as illustrated below. In the first step, the apical and intermediate domains unfold while the folded equatorial domains hold the tetradecamer together. The species $U(\underline{A}I)_{14}$ refers to a tetradecamer with all apical and intermediate domains unfolded (A, I, and E refer to apical, intermediate, and equatorial domains). This species was detected by amide hydrogen exchange when GroEL was incubated in 1.8 M GdHCl (Figures 1 and 3). The underlined A is used to indicate that some hydrophobic structure that binds to bisANS remains in residues 203−249 of the apical domain (*12*). GroEL monomers with the apical and intermediate domains unfolded, $U(\underline{A}I)$, are formed by a highly cooperative disassembly of the tetradecamer. This species has not been detected in any experiment, suggesting that it rapidly unfolds to give GroEL monomers which retain some hydrophobic structure in the apical domain. This hydrophobic structure, which has been detected only by binding of bisANS, is lost at high concentrations of denaturants.

$$\text{Native} \rightleftarrows U(\underline{A}I)_{14} \rightleftarrows 14\ U(\underline{A}I) \rightleftarrows 14\ U(\underline{A}IE) \rightleftarrows 14\ U(AIE)$$

Folding and assembly of native GroEL by this reaction scheme is possible only if each step is reversible. Equilibrium and kinetic studies of a GroEL construct including only the apical domain showed that unfolding of the apical domain is reversible (*41*). This result suggests that unfolding of native GroEL to give $U(\underline{A}I)_{14}$, which was detected in the present study under nonequilibrium conditions (1.8 M GdDCl), is likely reversible. Although GroEL monomers with the equatorial domains folded [i.e., $U(\underline{A}I)$] have not been detected, there is good evidence indicating that reactions joining Native and $U(\underline{A}IE)$ are reversible. GroEL equilibrated in 2.5 or 2.6 M urea for 24 h and analyzed by size exclusion chromatography (Figure 7) shows how the populations of Native and $U(\underline{A}IE)$ change with small changes in the urea concentration. Although the present results (Figures 6 and 7) show that paths joining Native and $U(\underline{A}IE)$ are reversible, the order in which the three domains fold cannot be determined. However, with the importance of the equatorial domain in stabilizing the tetradecamer, based on the present unfolding results as well as the X-ray crystal structure of GroEL, it is likely that the equatorial domain folds before the tetradecamer assembles. In addition, deletion mutagenesis results show that residues N-terminal to residue 522 are required for assembly of the tetradecamer (*45*), consistent with the requirement for folding of the equatorial domain before assembly. These results suggest that the failure of GroEL to assemble efficiently in vitro may due to misfolding of the equatorial domain. As proposed by Gorovtis et al., the residual hydrophobic structure in the apical domain may be required for proper folding of the equatorial domain in the absence of native GroEL. Alternatively, efficient folding of the equatorial domain may require the presence of native GroEL (*42*).

Previous studies of the unfolding of GroEL have used several different analytical methods, including tyrosine fluorescence, bisANS fluorescence, CD, light scattering, size exclusion chromatography, sedimentation velocity, and ATPase activity. Although these methods provide highly specific and valuable structural information, they provide no information on specific locations of structural changes. The present results demonstrate how amide hydrogen exchange, when combined with proteolytic fragmentation and mass spectrometry, can be used to locate structural changes in large protein complexes (i.e., determine which regions are unfolded) and to quantify populations of diverse unfolded structural forms. This experimental procedure will likely prove useful for folding and assembly studies of other large complexes.

ACKNOWLEDGMENT

We thank A. L. Horwich and S. G. Walter for helpful discussions and for providing the GroEL, M. A. Griep for use of his luminescence spectrometer, and Z. Zhang for providing software used for data analysis.

REFERENCES

1. Bukau, B., and Horwich, A. L. (1998) *Cell 92*, 351−366.
2. Ellis, R. J., and Hartl, F. U. (1999) *Curr. Opin. Struct. Biol. 9*, 102−110.
3. Sigler, P. B., Xu, Z., Rye, H. S., Burston, S. G., Fenton, W. A., and Horwich, A. L. (1998) *Annu. Rev. Biochem. 67*, 581−608.
4. Braig, K., Otwinowski, Z., Hegde, R., Boisvert, D. C., Joachimiak, A., Horwich, A. L., and Sigler, P. B. (1994) *Nature 371*, 578−586.
5. Braig, K., Adams, P. D., and Brunger, A. T. (1995) *Nat. Struct. Biol. 2*, 1083−1094.
6. Raschke, T. M., and Marqusee, S. (1997) *Nat. Struct. Biol. 4*, 298−304.

7. Xu, Y., Mayne, L., and Englander, S. W. (1998) *Nat. Struct. Biol. 5*, 774−778.
8. Li, R., and Woodward, C. (1999) *Protein Sci. 8*, 1571−1590.
9. Price, N. C., Kelly, S. M., Thomson, G. J., Coggins, J. R., Wood, S., and auf der Mauer, A. (1993) *Biochim. Biophys. Acta 1161*, 52−58.
10. Mizobata, T., and Kawata, Y. (1994) *Biochim. Biophys. Acta 1209*, 83−88.
11. Mendoza, J. A., Demeler, B., and Horowitz, P. M. (1994) *J. Biol. Chem. 269*, 2447−2451.
12. Gorovits, B. M., Seale, J. W., and Horowitz, P. M. (1995) *Biochemistry 34*, 13928−13933.
13. Smith, D. L., Deng, Y., and Zhang, Z. (1997) *J. Mass Spectrom. 32*, 135−146.
14. Zhang, Z., Li, W., Logan, T. M., Li, M., and Marshall, A. G. (1997) *Protein Sci. 6*, 2203−2217.
15. Resing, K. A., and Ahn, N. G. (1998) *Biochemistry 37*, 463−475.
16. Halgand, F., Dumas, R., Biou, V., Andrieu, J. P., Thomazeau, K., Gagnon, J., Douce, R., and Forest, E. (1999) *Biochemistry 38*, 6025−6034.
17. Weissman, J. S., Hohl, C. M., Kovalenko, O., Kashi, Y., Chen, S., Braig, K., Saibil, H. R., Fenton, W. A., and Horwich, A. L. (1995) *Cell 83*, 577−587.
18. Bai, Y., Milne, J. S., Mayne, L., and Englander, S. W. (1993) *Proteins: Struct., Funct., Genet. 17*, 75−86.
19. Deng, Y., Zhang, Z., and Smith, D. L. (1999) *J. Am. Soc. Mass Spectrom. 10*, 675−684.
20. Zhang, Z., personal communication.
21. Hvidt, A., and Nielsen, S. O. (1966) *Adv. Protein Chem. 21*, 287−385.
22. Woodward, C., Simon, I., and Tuchsen, E. (1982) *Mol. Cell. Biochem. 48*, 135−160.
23. Englander, S. W., and Kallenbach, N. R. (1984) *Q. Rev. Biophys. 16*, 521−655.
24. Miller, D. W., and Dill, K. A. (1995) *Protein Sci. 4*, 1860−1873.
25. Kragelund, B. B., Heinemann, B., Knudsen, J., and Poulsen, F. M. (1998) *Protein Sci. 7*, 2237−2248.
26. Roder, H., Elöve, G. A., and Englander, S. W. (1988) *Nature (London) 335*, 700−704.
27. Udgaonkar, J. B., and Baldwin, R. L. (1988) *Nature (London) 335*, 694−699.
28. Yang, H., and Smith, D. L. (1997) *Biochemistry 36*, 14992−14999.
29. Bai, Y., Sosnick, T. R., Mayne, L., and Englander, S. W. (1995) *Science 269*, 192−197.
30. Chamberlain, A. K., Handel, T. M., and Marqusee, S. (1996) *Nat. Struct. Biol. 3*, 782−787.
31. Heidary, D. K., Gross, L. A., Roy, M., and Jennings, P. A. (1997) *Nat. Struct. Biol. 4*, 725−731.
32. Deng, Y., and Smith, D. L. (1998) *Biochemistry 37*, 6256−6262.
33. Englander, J. J., Rogero, J. R., and Englander, S. W. (1985) *Anal. Biochem. 147*, 234−244.
34. Thévenon-Emeric, G., Kozlowski, J., Zhang, Z., and Smith, D. L. (1992) *Anal. Chem. 64*, 2456−2458.
35. Zhang, Z., and Smith, D. L. (1993) *Protein Sci. 2*, 522−531.
36. Miranker, A., Robinson, C. V., Radford, S. E., Aplin, R. T., and Dobson, C. M. (1993) *Science 262*, 896−900.
37. Engen, J. R., Smithgall, T. E., Gmeiner, W. H., and Smith, D. L. (1997) *Biochemistry 36*, 14384−14391.
38. Chen, S., Roseman, A. M., Hunter, A. S., Wood, S. P., Burston, S. G., Ranson, N. A., Clarke, A. R., and Saibil, H. R. (1994) *Nature 371*, 261−264.
39. Itzhaki, L. S., and Evans, P. A. (1996) *Protein Sci 5*, 140−146.
40. Parker, M. J., and Clarke, A. R. (1997) *Biochemistry 36*, 5786−5794.
41. Golbik, R., Zahn, R., Harding, S. E., and Fersht, A. R. (1998) *J. Mol. Biol. 276*, 505−515.
42. Lissin, N. M., Venyaminov, S., and Girshovich, A. S. (1990) *Nature 348*, 339−342.
43. Cheng, M. Y., Hartl, F. U., and Horwich, A. L. (1990) *Nature 348*, 455−458.
44. Ewalt, K. L., Hendrick, J. P., Houry, W. A., and Hartl, F. U. (1997) *Cell 90*, 491−500.
45. Burnett, B. P., Horwich, A. L., and Low, K. B. (1994) *J. Bacteriol. 176*, 6980−6985.
46. Burland, V., Plunkett, G., III, Sofia, H. J., Daniels, D. L., and Blattner, F. R. (1995) *Nucleic Acids Res. 23*, 2105−2119.

BI992619N