

Protein folding: How the mechanism of GroEL action is defined by kinetics

(folding intermediates/aggregation/dihydrofolate reductase)

CARL FRIEDEN* AND A. CLAY CLARK

Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, St. Louis, MO 63110

Contributed by Carl Frieden, March 28, 1997

ABSTRACT We propose a mechanism for the role of the bacterial chaperonin GroEL in folding proteins. The principal assumptions of the mechanism are (i) that many unfolded proteins bind to GroEL because GroEL preferentially binds small unstructured regions of the substrate protein, (ii) that substrate protein within the cavity of GroEL folds by the same kinetic mechanism and rate processes as in bulk solution, (iii) that stable or transient complexes with GroEL during the folding process are defined by a kinetic partitioning between formation and dissociation of the complex and the rate of folding and unfolding of the protein, and (iv) that dissociation from the complex in early stages of folding may lead to aggregation but dissociation at a late stage leads to correct folding. The experimental conditions for refolding may play a role in defining the function of GroEL in the folding pathway. We propose that the role of GroES and MgATP, either binding or hydrolysis, is to regulate the association and dissociation processes rather than affecting the rate of folding.

The bacterial chaperonin GroEL is probably the best studied of all the proteins that are involved as accessory proteins for correct substrate protein folding. A member of the Hsp60 class of proteins, it is composed of 14 identical subunits arranged as two stacked heptameric rings. As determined from electron microscopy (1) and crystallographic (2, 3) studies, each ring contains a large cavity into which a substrate protein binds and subsequently folds. Even though the crystal structure is known and numerous site-directed mutants have been made (4), there is not much information about the molecular processes involved in GroEL-mediated folding. Thus, a series of questions remain unanswered. (i) Why do so many unfolded proteins bind to GroEL? A survey of the literature (as exemplified by the chaperonin web page[†]) indicates at least 40 proteins that form stable complexes with GroEL. Many of these proteins are not even from bacterial systems. For example, of those substrate proteins that have been examined most closely, several, like rhodanese and α -lactalbumin, are mammalian proteins whereas others, like barnase, are not from *Escherichia coli*. (ii) What is the mechanism of protein folding in the presence of GroEL? Many studies examining folding do so on the basis of recovery of enzymatic activity. Although these results show that the efficiency of activity recovery is increased in the presence of GroEL or that stable complexes may form with GroEL, they give little information about the mechanism of the folding process. A few studies (5–7) have examined the kinetics of refolding in the presence and absence of GroEL, but the conditions or proteins used were not always conducive to determining mechanism. (iii) What is the nature of complexes

formed between GroEL and substrate protein? Most studies discuss the formation of stable complexes, and although, for example, murine dihydrofolate reductase (DHFR) does form a stable complex, we have shown that (under the same solution conditions) the *E. coli* DHFR forms only a transient complex with GroEL (8). When stable complexes are formed they may be relatively unstructured or structured depending on which substrate protein is being studied. Recent results (9–11), for example, suggest that a substrate protein in such a complex can be significantly protected against exchange of amide protons when compared with other protein substrates such as cyclophilin (12), barnase (13), or α -lactalbumin (14) that show extensive amide proton exchange when bound to GroEL. (iv) Does GroEL interact with native protein or with some pre-existing isoform in equilibrium with the native form? Some controversy exists in the literature with Fersht and coworkers (e.g., ref. 15) suggesting the former mechanism while others propose the latter (e.g., ref. 16).

We have recently been studying the folding of several structurally homologous DHFRs both in the presence and absence of GroEL. DHFR has been used as a model protein for complex formation with the bacterial chaperonin GroEL (17), but an examination of the literature showed that it was the mammalian enzyme, rather than the *E. coli* enzyme that was used in these studies. Indeed, we found that the *E. coli* DHFR does not form a stable complex with GroEL at temperatures below 25°C (8). This observation led to investigation of the differences between the *E. coli* and mammalian forms of DHFR. Superimposition of the *E. coli* (18) and human (19) DHFR–folate binary complex structures suggested that although most of the structure was similar between the two enzymes, three external loops existed in the mammalian protein that were not present in the *E. coli* protein. By site-directed mutagenesis, we inserted each of two loops into the *E. coli* protein (20). This allowed the examination of four structurally similar proteins: the wild-type *E. coli* DHFR, the murine DHFR (which has essentially the same sequence as the human enzyme), and two mutant proteins, each containing a loop similar to that found in the mammalian protein but not in the *E. coli* protein. Although our comments may be biased because of the studies with DHFR, the results from these and other studies now allow us to make specific proposals about the mechanism of protein folding in the presence of GroEL.

Promiscuity

As mentioned above, numerous proteins, both prokaryotic and eukaryotic, form stable complexes with GroEL. A seemingly unrelated result may explain this observation. GroEL contains

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Copyright © 1997 by THE NATIONAL ACADEMY OF SCIENCES OF THE USA
0027-8424/97/945535-4\$2.00/0

PNAS is available online at <http://www.pnas.org>.

Abbreviation: DHFR, dihydrofolate reductase.

*To whom reprint requests should be addressed at: Department of Biochemistry and Molecular Biophysics, Box 8231, Washington University School of Medicine, St. Louis, MO 63110. e-mail: frieden@biochem.wustl.edu.

[†]The URL for the GroE chaperonin web page is <http://bioc09.uthscsa.edu/~seale/Chap/chap.html>.

no tryptophan residues (21) yet the usual preparations show significant fluorescence emission typical of tryptophan. This suggests that tryptophan-containing peptides are tightly bound to GroEL and are not removed during protein preparation. We have examined, by electrospray ionization mass spectrometry, the molecular weight distribution of species in a sample of protein that shows substantial fluorescence. Although there are species whose molecular weights are below 350, surprisingly, there are no observed masses between these low molecular weight species and that of the GroEL itself (~57,200) (22). This result shows the contaminating, presumably tightly bound, material to be very low molecular weight, certainly no larger than tripeptides[‡]. Although larger peptides (23, 24) and small amounts of numerous proteins (P. Horowitz, personal communication) have been shown to interact with GroEL, the most tightly bound species could be very small and unstructured segments of the substrate protein. Indeed, single amino acids have been shown to bind, although somewhat weakly (25), to GroEL. In this regard, it is of interest to calculate the concentration of a single molecule in the GroEL cavity. From crystallographic studies, the cavity within a single ring-like structure may be about 125,000 Å³ (2). Within a cavity this size, a single molecule may have a concentration of more than 10 mM. We propose, therefore, that the reason GroEL binds to many different proteins is that the binding site(s), which appears to contain many hydrophobic residues (26), recognize unstructured regions, but with some specificity toward hydrophobic residues such as tryptophan. It has been proposed on the basis of site-directed mutagenesis (4) and equilibrium binding studies (27) that hydrophobicity is an important factor in protein binding to GroEL. Many possible intermediates of a single protein substrate may form stable complexes because the apparent concentration within the cavity is so high. Conclusions from such experiments, however, may be misleading. In our studies of structurally homologous DHFRs, it is clear that it is the complex formed just prior to the final folding step that defines complex stability (8).

The Mechanism of Folding

Much of the literature refers to the action of GroEL as assisting or increasing the efficiency of folding and even catalyzing folding. We believe none of these concepts is strictly correct. Our studies on the refolding of the four structurally similar monomeric DHFRs show that the kinetics of the refolding process is the same in the presence or absence of GroEL (8). Thus the large cavity that exists in GroEL must be similar to bulk solution. Of course, since only a few (one to two) molecules are contained within this cavity, the chance of aggregation is slight. Consequently, the chaperonin increases the probability of the protein refolding correctly by decreasing the probability of aggregation that may occur in the bulk solution. The observation with respect to the DHFRs is not an isolated one: similar kinetics with and without GroEL have been observed by others (28). We propose, therefore that this is the general case, i.e., that folding in the cavity of GroEL occurs by the same mechanism as in bulk solution and that it occurs at the same rate until a stable complex is formed. Ellis (29), in describing the concept of the Anfinsen cage, has also suggested that folding within the GroEL cavity is the same as in bulk solution.

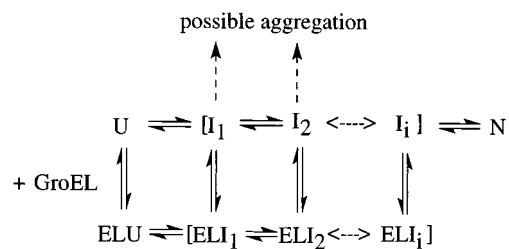
[‡]We note that many of these contaminating species, including essentially all the fluorescent ones, may be removed from GroEL by use of a Reactive Red 120-agarose column (20) and that this is important in the studies described herein with the DHFRs. The Reactive Red column or a similar procedure that removes low molecular weight contaminants should always be used in any quantitative study involving binding of proteins or peptides to GroEL.

The Nature of the Stable Complex

As mentioned above, there are conflicting data in the literature concerning the extent of protein substrate structure that is present in the GroEL-protein complex. Those results indicating a highly structured intermediate are consistent with our observations made with DHFR (8). In this case, since the kinetic behavior in the presence of GroEL is the same as in its absence, it is unlikely that a stable complex is formed early in refolding. That the protein substrate in the stable complex is an intermediate directly preceding the formation of the native structure is suggested both from the kinetic data and from the observation that the same complex is formed when starting with unfolded or native protein (8). In the latter case, the kinetic data show that complex formation is a consequence of GroEL binding not to the native structure itself but to a species that interconverts to the native structure (8). This indicates that at least one conformation or a region within the native structure may be similar to a partially unfolded form. However, in general, the nature of the substrate protein within the stable complex is really a function of a kinetic partitioning between the association and dissociation of the substrate protein with GroEL and the forward and reverse rate constants for formation of a specific intermediate. Thus the stability of a particular complex may be simply a matter of defining a slow step in the folding process as well as the affinity of an intermediate for GroEL. For protein substrates that are oligomeric in solution, but not within the GroEL cavity, the slow step in the formation of an active enzyme may well be the association of competently folded subunits and in such cases any stable complex would be with an intermediate close to the final folded form of the subunit.

Stability of the Complex and What it Means

With these thoughts in mind we need to consider only a portion of the folding process to define the nature of the stable complex. We can write a general scheme as follows.



Scheme I

The unfolded protein, U, may fold through the formation of one or more intermediates prior to forming the native state. In this mechanism the species I_i is the last intermediate state prior to forming the native structure. The species ELI_1 is written as if there is only a single molecule of protein substrate in the cavity, but it is certainly possible that the stoichiometry of substrate protein to GroEL may be greater than one. This will not affect the discussion, although the rate constants described in this case may be more complex than indicated in the above scheme. Additionally, the scheme does not include the possibility of conformational changes within GroEL that could affect association or dissociation rates without changing the folding mechanism. According to the scheme, intermediates may dissociate from GroEL at any point along the folding pathway, but if dissociation occurs early, the dissociated intermediates may aggregate, presumably due to exposed hydrophobic regions of the protein, rather than refold to a correct structure. This may be the case for a protein like rhodanese (30) or citrate synthase (31). In any case, the

stability of any of the complexes is clearly defined by the relative magnitude of rate constants leading into and out of the complex. Transient formation of the final complex may occur if binding of I_i to GroEL is weak and/or if the formation of the native structure from I_i is fast. This is what is observed for wild-type *E. coli* DHFR at temperatures below 25°C (8). Indeed, this is what one would want to happen for *E. coli* proteins under normal growth conditions. For if, under such conditions, a stable complex were formed, the cells would not continue to grow. This, in fact, is the heat shock response: an overproduction of GroE protects the viability of the cell by forming stable complexes with proteins.

The formation of a stable complex can also be defined in terms of these rate constants. If the binding of I_i to GroEL is reasonably tight and/or if the rate constant for the process $N \rightarrow I_i$ is large, the complex formed will be stable. The finding that many stable complexes are formed with nonbacterial proteins may simply be a question of the relative values of rate constants leading into or out of that complex, as shown in Scheme I. In this regard, it is of interest that mouse DHFR does not appear to form a stable complex with yeast Hsp60 at 25°C (32), whereas it does with GroEL at the same temperature (8). Wild-type *E. coli* DHFR does form a stable complex with GroEL as the temperature is raised above 30°C (unpublished data). This temperature dependence is exactly what would be expected for a heat shock response and has to do with the relative values of the rate constants shown in Scheme I.

The role of the kinetic partitioning between binding to GroEL and folding rates may play a part in the different results obtained for different protein substrates. Typical refolding experiments are, of course, performed by dilution of the chemical denaturant, either urea or guanidine hydrochloride. The rate of refolding is found to be dependent on the final denaturant concentration relative to the midpoint of the denaturation curve, increasing at lower final denaturant concentrations. On the other hand, the binding or dissociation of the substrate protein (or any intermediate) to GroEL is probably not very sensitive to the final denaturant concentration under refolding conditions. Thus the issue of the nature of the stable complex or the tendency of the substrate protein to aggregate may be a function of the conditions of the procedure used in the refolding experiment.

The Role of GroES and ATP

Some proteins require the presence of GroEL, GroES, and MgATP to fold properly (e.g., refs. 33 and 34) but others may require only MgATP in the presence of GroEL to form the final folded structure (17, 35–37). Since folding occurs within the GroEL cavity obviously neither GroES nor MgATP alone has chaperone activity. Rather, these components serve only to regulate the rate constants involved in the formation or dissociation of the GroEL–substrate protein complexes. This regulation may be quite complex (38–40) involving cooperativity between the two heptameric rings of GroEL and certainly related not only to MgATP binding but also to MgATP hydrolysis (34) as well as capping of the GroEL cavity (either *cis* or *trans* to the substrate protein) by GroES (41, 42). Even though formation of the substrate protein–GroEL complex may involve conformational changes within GroEL, the mechanism of the folding process is probably unchanged. We would expect that the observation of a temperature dependence of stable complex formation with wild-type *E. coli* DHFR reflects what would happen *in vivo* but that the temperature dependence would be regulated by the presence of GroES and ATP.

Conclusion

We propose that the mechanism and rate constants for folding of a protein are the same in the presence or absence of GroEL.

Preferential binding of small unfolded structural units is responsible for the lack of specificity in forming the initial complex with GroEL. At any point along the folding pathway the protein may dissociate from the GroEL, but if this occurs early, aggregation may result. The ability of GroEL to fold proteins correctly without aggregation is a consequence of the stability of the complex formed between GroEL and the protein conformation just prior to the final folding step. The stability of this complex is defined by rate constants involved in complex formation or dissociation compared with the rate constants for folding to or unfolding from the native protein as well as the experimental conditions used for the refolding experiment. Thus, rather than ask whether a given protein binds to GroEL, the relevant question would appear to be at what point in the folding pathway a given protein associates and dissociates from GroEL.

This work was supported by National Institutes of Health Grant DK13332.

- Harris, J. R., Zahn, R. & Pluckthun, A. (1995) *J. Struct. Biol.* **115**, 68–77.
- Braig, K., Otwinowski, Z., Hegde, R., Boisvert, D. C., Joachimiak, A., Horwich, A. L. & Sigler, P. B. (1994) *Nature (London)* **371**, 578–586.
- Hunt, J. F., Weaver, A. J., Landry, S. J., Gierasch, L. & Deisenhofer, J. (1996) *Nature (London)* **379**, 37–45.
- Fenton, W. A., Kashi, Y., Furtak, K. & Horwich, A. L. (1994) *Nature (London)* **371**, 614–619.
- Katsumata, K., Okazaki, A. & Kuwajima, K. (1996) *J. Mol. Biol.* **258**, 827–838.
- Ayling, A. & Baneyx, F. (1996) *Protein Sci.* **5**, 478–487.
- Sparrer, H., Lilie, H. & Buchner, J. (1996) *J. Mol. Biol.* **258**, 74–87.
- Clark, A. C. & Frieden, C. (1997) *J. Mol. Biol.* **268**, 512–525.
- Gervasoni, P., Staudenmann, W., James, P., Gehrig, P. & Pluckthun, A. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 12189–12194.
- Groß, M., Robinson, C. V., Mayhew, M., Hartl, F. U. & Radford, S. E. (1996) *Protein Sci.* **5**, 2506–2513.
- Goldberg, M. S., Zhang, J., Sonder, S., Matthews, C. R., Fox, R. O. & Horwich, A. L. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 1080–1085.
- Zahn, R., Spitzfaden, C., Ottiger, M., Wuthrich, K. & Pluckthun, A. (1994) *Nature (London)* **368**, 261–265.
- Zahn, R., Perrett, S., Stenberg, G. & Fersht, A. R. (1996) *Science* **271**, 642–645.
- Robinson, C. V., Groß, M., Eyles, S. J., Ewbank, J. J., Mayhew, M., Hartl, F. U., Dobson, C. M. & Radford, S. E. (1994) *Nature (London)* **372**, 646–651.
- Corrales, F. J. & Fersht, A. R. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 5326–5330.
- Walter, S., Lorimer, G. H. & Schmid, F. X. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 9425–9430.
- Viitanen, P. V., Donaldson, G. K., Lorimer, G. H., Lubben, T. H. & Gatenby, A. A. (1991) *Biochemistry* **30**, 9716–9723.
- Bolin, J. T., Filman, D. J., Matthews, D. A., Hamlin, R. C. & Kraut, J. (1982) *J. Biol. Chem.* **257**, 13650–13662.
- Davies, J. D., Delcamp, T. J., Prendergast, N. J., Ashford, V. A., Freisheim, J. H. & Kraut, J. (1990) *Biochemistry* **29**, 9467–9479.
- Clark, A. C., Hugo, E. & Frieden, C. (1996) *Biochemistry* **35**, 5893–5901.
- Hemmingsen, S. M., Woolford, C., van der Vies, S. M., Tilly, K., Dennis, D. T., Georgopoulos, C. P., Hendrix, R. W. & Ellis, R. J. (1988) *Nature (London)* **333**, 330–334.
- Clark, A. C., Ramanathan, R. & Frieden, C. (1997) *Methods Enzymol.*, in press.
- Landry, S. J., Jordan, R., McMacken, R. & Gierasch, L. M. (1992) *Nature (London)* **355**, 455–457.
- Rosenberg, H. F., Ackerman, S. J. & Tenen, D. G. (1993) *J. Biol. Chem.* **268**, 4499–4503.
- Richarme, G. & Kohiyama, M. (1994) *J. Biol. Chem.* **269**, 7095–7098.
- Fenton, W. A., Weissman, J. S. & Horwich, A. L. (1996) *Chem. Biol.* **3**, 157–161.

27. Lin, Z., Schwartz, F. P. & Eisenstein, E. (1995) *J. Biol. Chem.* **270**, 1011–1014.
28. Laminet, A. A., Ziegelhoffer, T., Georgopoulos, C. & Pluckthun, A. (1990) *EMBO J.* **9**, 2315–2319.
29. Ellis, R. (1996) *Folding Design* **1**, R9–R15.
30. Mendoza, J. A., Rogers, E., Lorimer, G. H. & Horowitz, P. M. (1991) *J. Biol. Chem.* **266**, 13044–13049.
31. Buchner, J., Schmidt, M., Fuchs, M., Jaenicke, R., Rudolph, R., Schmid, F. X. & Kiefhaber, T. (1991) *Biochemistry* **30**, 1586–1591.
32. Rospert, S., Looser, R., Dubaquié, Y., Matouschek, A., Glick, B. S. & Schatz, G. (1996) *EMBO J.* **15**, 764–774.
33. Martin, J., Langer, T., Boteva, R., Schramel, A., Horwich, A. L. & Hartl, F. U. (1991) *Nature (London)* **352**, 36–42.
34. Goloubinoff, P., Christeller, J. T., Gatenby, A. A. & Lorimer, G. H. (1989) *Nature (London)* **342**, 884–889.
35. Badcoe, I. G., Smith, C. J., Wood, S., Halsall, D. J., Holbrook, J. J., Lund, P. & Clarke, A. R. (1991) *Biochemistry* **30**, 9195–9200.
36. Fisher, M. T. (1992) *Biochemistry* **31**, 3955–3963.
37. Murai, N., Taguchi, H. & Yoshida, M. (1995) *J. Biol. Chem.* **270**, 19957–19963.
38. Todd, M. J., Viitanen, P. V. & Lorimer, G. H. (1994) *Science* **265**, 659–666.
39. Yifrach, O. & Horowitz, A. (1996) *J. Mol. Biol.* **255**, 356–361.
40. Hayer-Hartl, M. K., Weber, F. & Hartl, F.-U. (1996) *EMBO J.* **15**, 6111–6121.
41. Weissman, J. S., Hohl, C. M., Kovalenko, O., Kashi, Y., Chen, S. X., Braig, K., Saibil, H. R., Fenton, W. A. & Horwich, A. L. (1995) *Cell* **83**, 577–587.
42. Mayhew, M., Dasilva, A. C. R., Martin, J., Erdjumentbromage, H., Tempst, P. & Hartl, F. U. (1996) *Nature (London)* **379**, 420–426.