ANNALS OF THE NEW YORK ACADEMY OF SCIENCES
Volume 672

ENZYME ENGINEERING XI

Edited by Douglas S. Clark and David A. Estell

The New York Academy of Sciences
New York, New York
1992

SP
Printed in the United States of America
ISBN 0-89766-763-8 (cloth)
ISBN 0-89766-764-6 (paper)
ISSN 0077-8923

ANNALS OF THE NEW YORK ACADEMY OF SCIENCES

Volume 672
November 30, 1992

ENZYME ENGINEERING XI[a]

Editors and Conference Organizers
DOUGLAS S. CLARK and DAVID A. ESTELL

CONTENTS

[a]The papers in this volume were presented at the Eleventh International Enzyme Engineering Conference, held in Keauhou-Kona, Hawaii, on September 22–27, 1991, sponsored by the Engineering Foundation (USA).

Part II. Improved Biocatalysts through Genetic Engineering

Part III. Enzyme Stability and Stabilization

Part V. Enzymes under Unusual Conditions

Part VI. Applications of Enzymes in Synthesis

The Influence of Chaperonins on Protein Folding

A Mechanism for Increasing the Yield of the Native Form

S. G. BURSTON, R. SLEIGH, D. J. HALSALL, C. J. SMITH, J. J. HOLBROOK, AND A. R. CLARKE

Molecular Recognition Center and
Department of Biochemistry
University of Bristol
School of Medical Sciences
Bristol BS8 1TD, England

Molecular chaperones are proteins that promote the correct folding, assembly, and transport of other protein molecules.[1] The most widely studied of the molecular chaperones are the chaperonins; these form a subgroup found in prokaryotic cells, in mitochondria, and in plastids.[2–5] They are oligomeric proteins of high molecular weight, which work in conjunction with a smaller, also oligomeric, coprotein. One such chaperonin is cpn60 from *Escherichia coli,* the product of the *gro-EL* gene locus. This has a subunit mass of approximately 60,000 daltons, and the whole protein exists as an assembly of 14 subunits in a "double-doughnut" structure, each ring comprising 7 subunits. The coprotein is produced by the *gro-ES* locus and is termed cpn10 by virtue of its 10,000-dalton subunit molecular weight. Cpn10 exists as a "single-doughnut" structure of 7 subunits.[6] In conjunction, these proteins have been shown to aid the folding and/or assembly of bacteriophages, multimeric and monomeric protein molecules.[7,8]

They function by binding to unfolded or partially folded protein molecules and, by transducing the energy of ATP hydrolysis, increasing the yield of the natively assembled form.[9] The mechanistic detail of this process remains undefined.

Here we report on preliminary kinetic experiments that explore the interactions of cpn60 with (a) the folding intermediates of *Bacillus stearothermophilus* lactate dehydrogenase (LDH), (b) with ATP and an unreactive analogue (AMP-PNP), and (c) with the coprotein cpn10. From these results with LDH we suggest a mechanism that explains the ability of chaperonins to improve the efficiency of protein folding in general.

THE BINDING OF cpn60 TO PROTEIN-FOLDING INTERMEDIATES

When bacterial LDH is unfolded to equilibrium in solutions of the denaturant guanidinium chloride (Gdm-Cl), four structural states of the protein can be isolated.[10] The native state (NN) is a dimer which dissociates to give, at 1.0 M Gdm-Cl, an inactive monomer (N). This unfolds to give, at 2.2 M Gdm-Cl, an expanded

monomer (G) which retains 60% of its alpha-helix content but is freely penetrated by solvent. At 4.0 M Gdm-Cl, the protein is fully unfolded (U).

Previous results have shown that when LDH is allowed to renature from increasingly unfolded states, cpn60 can retard refolding when initiated from the unfolded state (U) but not from the "molten globule" (G), or the monomer (N).[11] A kinetically identifiable, but structurally undefined, intermediate (INT) which occurs on the conversion from U to G is also bound by cpn60. A summary of the folding intermediates of LDH is presented in FIGURE 1.

From the above observations we conclude that cpn60 associates most strongly with the least-folded forms of the protein and not with its molten globule state.

To seek a fuller description of this process, we measured the return of LDH activity over a continuous time course when the fully denatured LDH refolded in the presence and absence of cpn60 (see FIGURE 2a). LDH folding alone proceeds by two, slow, unimolecular steps (U → INT → G) and an apparent bimolecular step (2G → NN). To describe refolding in the presence of cpn60, further steps must be introduced to account for the interactions of these folding intermediates with the

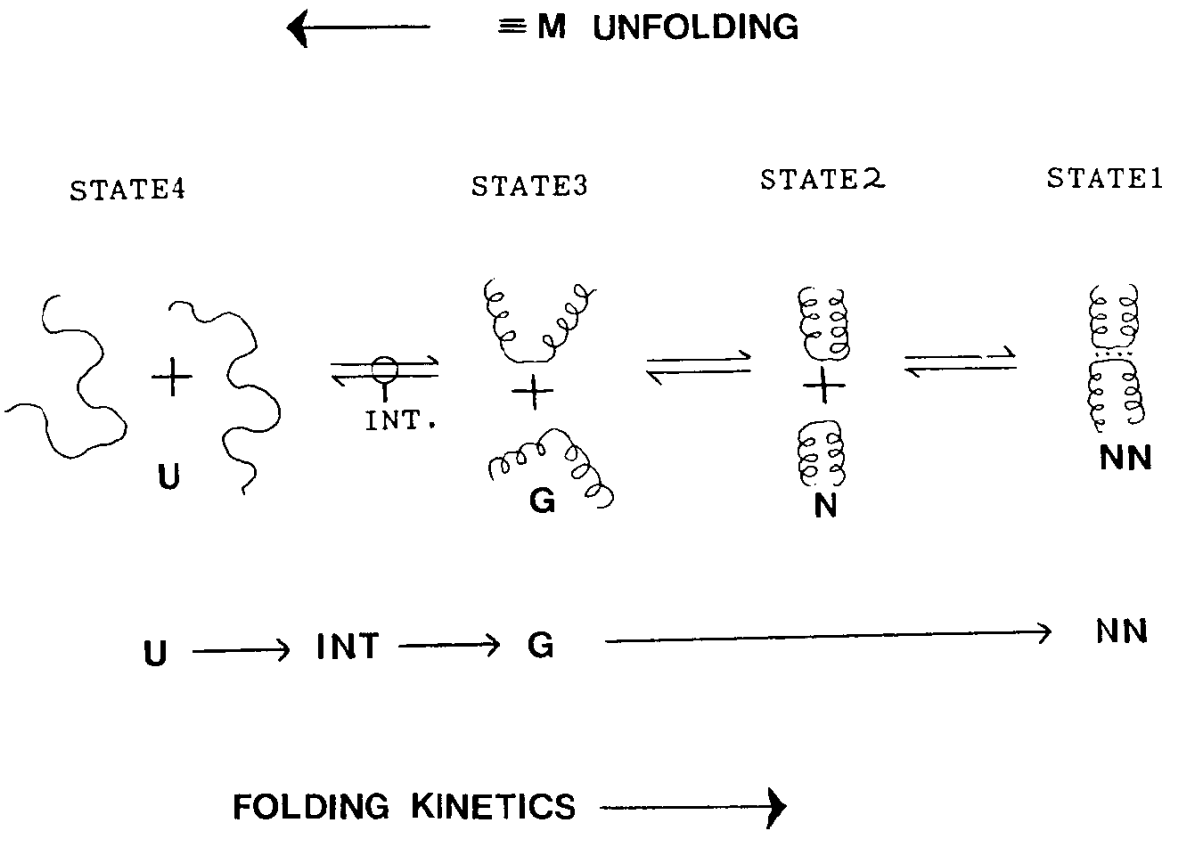
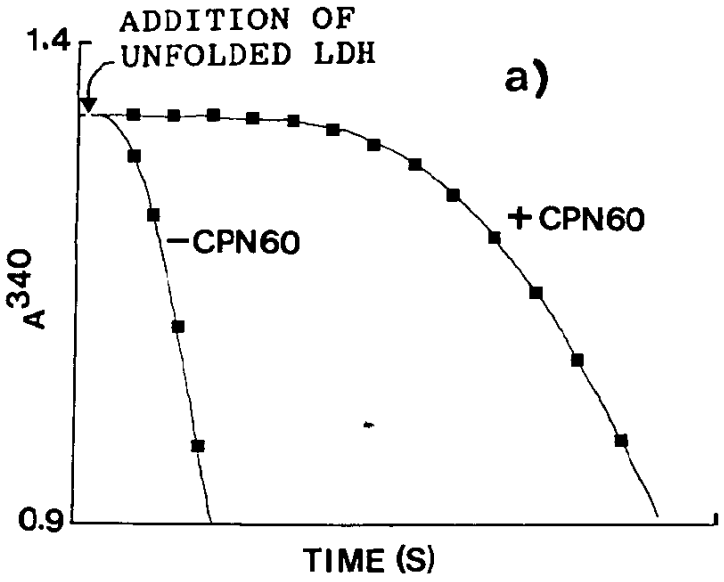


FIGURE 1. The folding mechanism of lactate dehydrogenase. The equilibrium (≡M) and kinetic states identifiable in LDH folding are summarized. In the former experiments,[10] state 1 (NN) is native, state 2 (N) predominates at 1.0 M Gdm-Cl, state 3 (G) at 2.2 M Gdm-Cl, and state 4 (U) at and above 4.0 M Gdm-Cl. In the kinetics of refolding,[11] the protein passes from U to G in two, slow unimolecular steps (3×10^{-3} and 15×10^{-3} seconds^{-1}, see FIGURE 2b). As protein refolding is performed at low protein concentration, the conversion of G to NN is measured as a single bimolecular step with an apparent rate constant of 5×10^{5} M^{-1} seconds^{-1} (see FIGURE 2b).



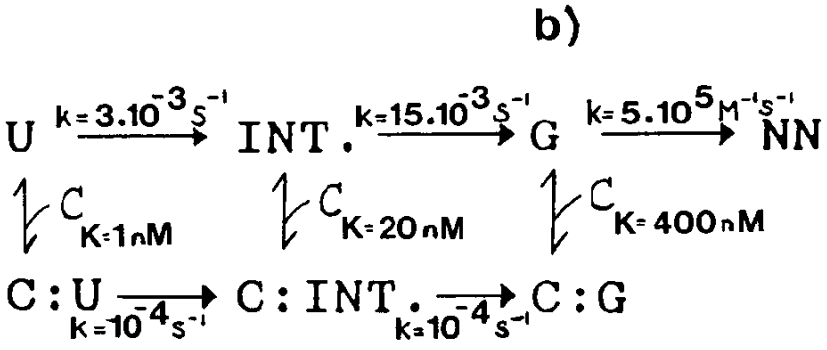


FIGURE 2. The association of cpn60 with LDH during refolding. (**a**) Aliquots of unfolded LDH (in 4.0 M Gdm-Cl/2 mM dithiothreitol) were added to an enzyme assay mix (0.2 mM NADH/10 mM pyruvate/2 mM dithiothreitol/50 mM triethanolamine-HCl pH 7.0). As the enzyme (final subunit concentration 20 nM) refolded, the turnover of NADH was measured by continuously recording absorbance at 340 nm. The experiment was performed in the presence and absence of 200 nM cpn60 (14-mer). The points on the curve represent the best fit of the model described below to the experimental data. (**b**) Model summarizing the results of fitting these curves by numerical integration using the FACSIMILE program.[12] The folding kinetics in the absence of cpn60 (C) are represented on the top line and those steps necessary to describe the interactions of intermediates with the chaperonin are included underneath (binding is described by a dissociation constant K). The analysis demonstrates that the more folded the LDH becomes, the less is its propensity to bind to cpn60.

chaperonin; this goes beyond what is feasible with standard analytical methods, so numerical integration was used to fit these continuous time courses of refolding.

The model shown in FIGURE 2b is the simplest that will fit the experimental observations; it confirms that cpn60 binds most tightly to the least structurally organized forms of LDH and shows that folding can occur on the surface of the chaperonin, but at a much retarded rate. This latter point is supported by the experimental observation that the presence of cpn60 increases the yield of active LDH by two- to threefold when allowed to fold from the fully denatured state.[11] The implication is that, by virtue of its slow folding on the chaperonin surface, the

early-folding intermediates of LDH are sequestered from the bulk phase and are therefore protected from irreversible aggregation.

THE INFLUENCE OF NUCLEOTIDE BINDING IN THE ABSENCE OF HYDROLYSIS

Cpn60 is known to bind and hydrolyze ATP, this process mediating release of bound protein substrates.[9] To begin to investigate this aspect of chaperonin function, we have recorded continuous LDH-folding curves in the presence of cpn60 and at increasing concentrations of the nonhydrolyzable ATP analogue AMP-PNP. The results are shown in FIGURE 3, together with the analysis by numerical integration. This requires the introduction of new binding steps describing the interaction of the nucleotide with cpn60. The result is striking in that the cpn60/LDH complex gets weaker as the protein folds, but the binding of nucleotide to cpn60 gets stronger. The apo-cpn60 has a K_d for nucleotide of 2 μM (the same as that predicted for the unstable cpn60/molten globule complex), whereas in the stable cpn60/unfolded LDH complex the K_d rises to 200 μM. We conclude that cpn60 can adopt at least two structural forms: one associates preferentially with nucleoside-triphosphates, one with unfolded proteins; nucleotide hydrolysis is not a prerequisite of protein release.[11]

ANALYSIS OF FOLDING TIMES AND EFFICIENCY—THE EFFECT OF ATP HYDROLYSIS AND THE BINDING OF THE COPROTEIN CPN10

A simplified analysis of the rate of refolding can be made by measuring the lag time for the regaining of LDH activity (defined in FIGURE 4b) in the presence of ligands and cofactors. The influence of AMP-PNP and ATP on the lag time for refolding in the presence of cpn60 is shown in FIGURE 4a. In both cases their effects on LDH release are half-saturated at a concentration of 100–200 μM; this is in accord with the model presented in FIGURE 3. However, at saturating concentrations, ATP is able to reduce the lag time further than is AMP-PNP. This shows that either ATP itself or its turnover products, ADP · P_i, are better at displacing LDH from cpn60 than is AMP-PNP. In view of the data presented in FIGURE 4, the latter seems more likely.

FIGURE 4a also shows the effect of cpn10 on LDH displacement by AMP-PNP. The result demonstrates that cpn10 alone promotes displacement, but inhibits the binding/displacement effect of AMP-PNP. This latter effect then appears to be cooperative. This observation is, at first sight, puzzling, but can be explained if we accept that the hydrophobic cpn10[13] binds to the same form of cpn60 as unfolded LDH, thus further reducing its nucleoside-triphosphate affinity. This interaction may lead to a degree of cpn10/LDH competition for hydrophobic binding faces on cpn60.

The bar chart in FIGURE 5 summarizes these effects on lag times for folding and illustrates the important point that even in optimal conditions for LDH release from

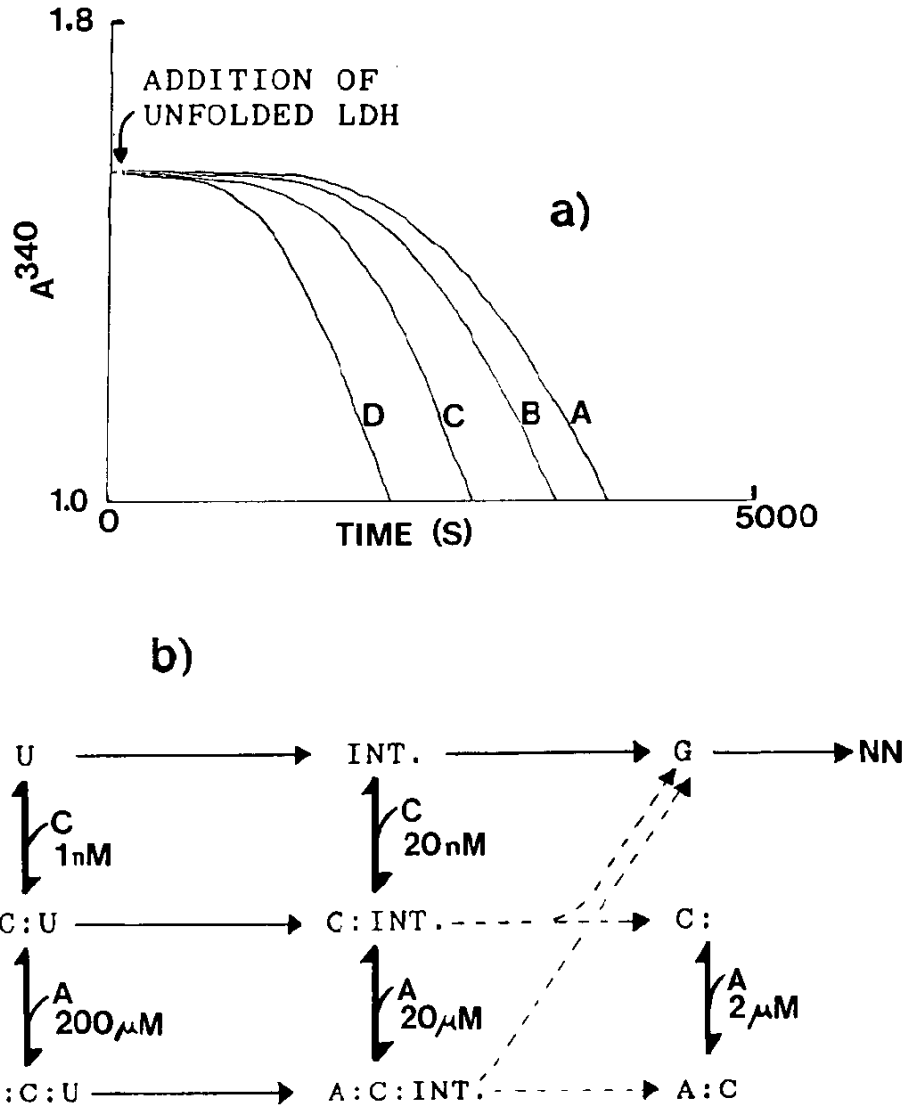


FIGURE 3. The interaction of cpn60 with AMP-PNP. (**a**) Unfolded LDH (20 nM) was added to an assay medium containing 200 nM cpn60 as described for FIGURE 2. AMP-PNP was omitted for curve A and present at 0.05 mM for B, 1.25 mM for C, and 3.75 mM for D. (**b**) This scheme summarizes the apparent binding affinities of AMP-PNP (A) for the chaperonin: protein complexes (see also FIGURE 2). The numerical values represent dissociation constants. The result emphasizes that the nucleotide binds most tightly to apo-cpn60 and most weakly to the most stable cpn60-protein complex (C:U). In these conditions the interaction of G with the chaperonin is sufficiently weak that the C:G and A:C:G complexes never accumulate. This is indicated by the dashed lines showing C:INT and A:C:INT decaying to products at a rate limited by the folding rather than the dissociation rate.

cpn60 (high [ATP] plus cpn10), the *rate* of refolding is not enhanced. In these conditions, however, the *yield* is improved dramatically and over a short time period. This result is presented in full in FIGURE 4b and shows that at high ATP concentrations cpn60 can increase the yield of active LDH over a time period approaching that of "unaided" folding. The inclusion of cpn10 accelerates this effect still further. In the natural cellular environment the chaperonin system comprises cpn60, cpn10, and ATP; we must therefore extend the model shown in FIGURE 3 to account for this.

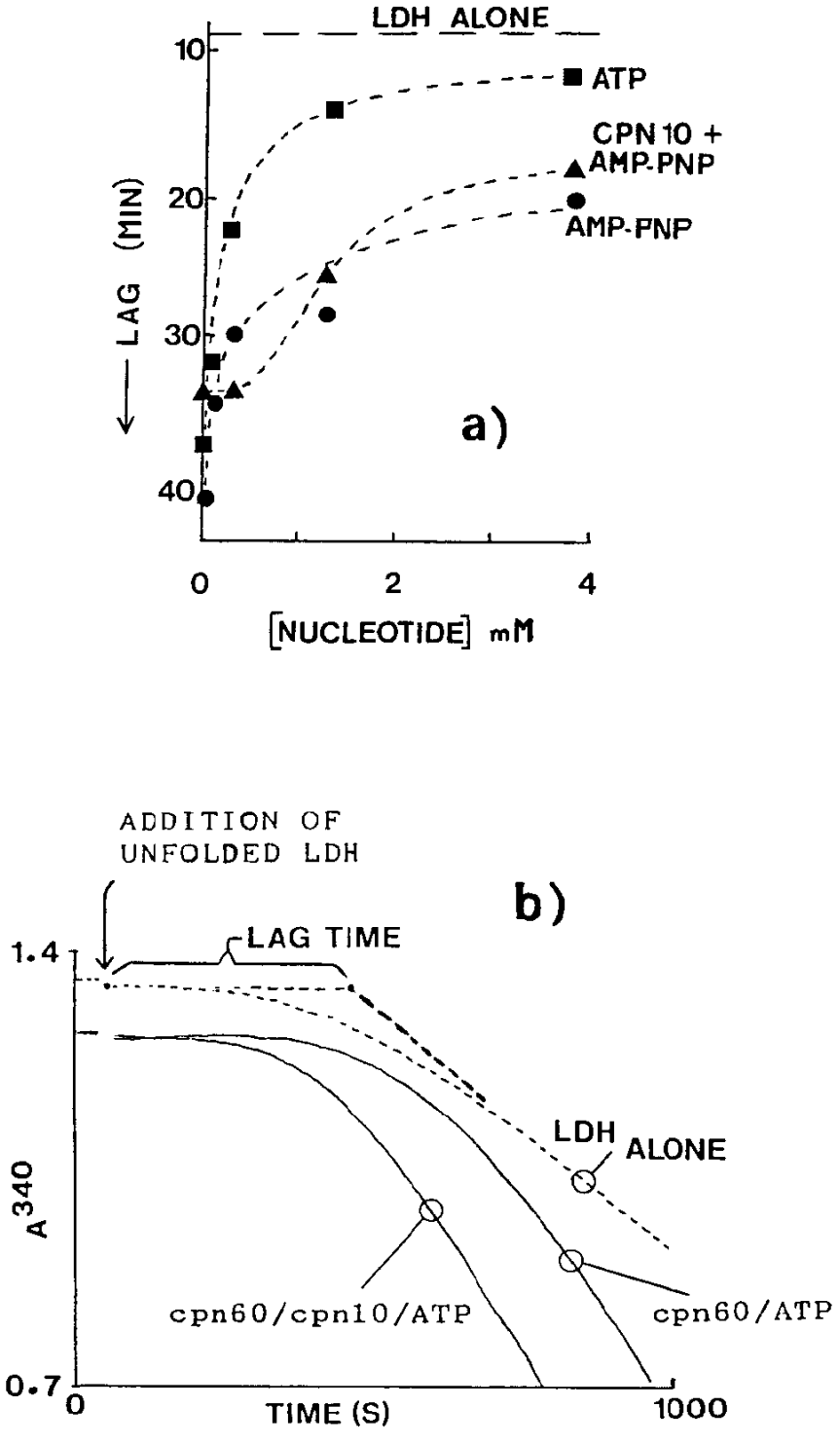


FIGURE 4. Lag times for folding; the concentration dependence of nucleotides and the effect of cpn10. Assays measuring the regaining of LDH activity were performed as described in FIGURES 2 and 3, and the lag time for renaturation recorded. **Plot a** shows the dependence of lag times on nucleotide concentration. In the case of experiments in the presence of cpn10, the concentration of 7-mer was 300 nM. Three progress curves are shown in **b**: LDH refolding in the absence of chaperone (giving the definition of the lag time) and in the presence of cpn60/3.75 mM ATP and cpn60/cpn10/3.75 mM ATP. The latter are included to show the improvement of yield in these conditions which occurs over a period comparable to that of unchaperoned folding.

A PRELIMINARY MECHANISM TO EXPLAIN THE ACTION OF CHAPERONINS IN IMPROVING THE EFFICIENCY OF PROTEIN FOLDING

Several properties of cpn60 are shown by the results presented here. They are:

1. The protein binds most tightly to the most unfolded forms of LDH, thereby stabilizing these states preferentially. We propose that the affinity is dictated by the degree of exposure of hydrophobic residues.
2. Release of the substrate protein is facilitated by the binding of AMP-PNP. This effect is more pronounced in the case of ATP; possibly due to the formation of a cpn60/ADP · P_i complex through hydrolysis.

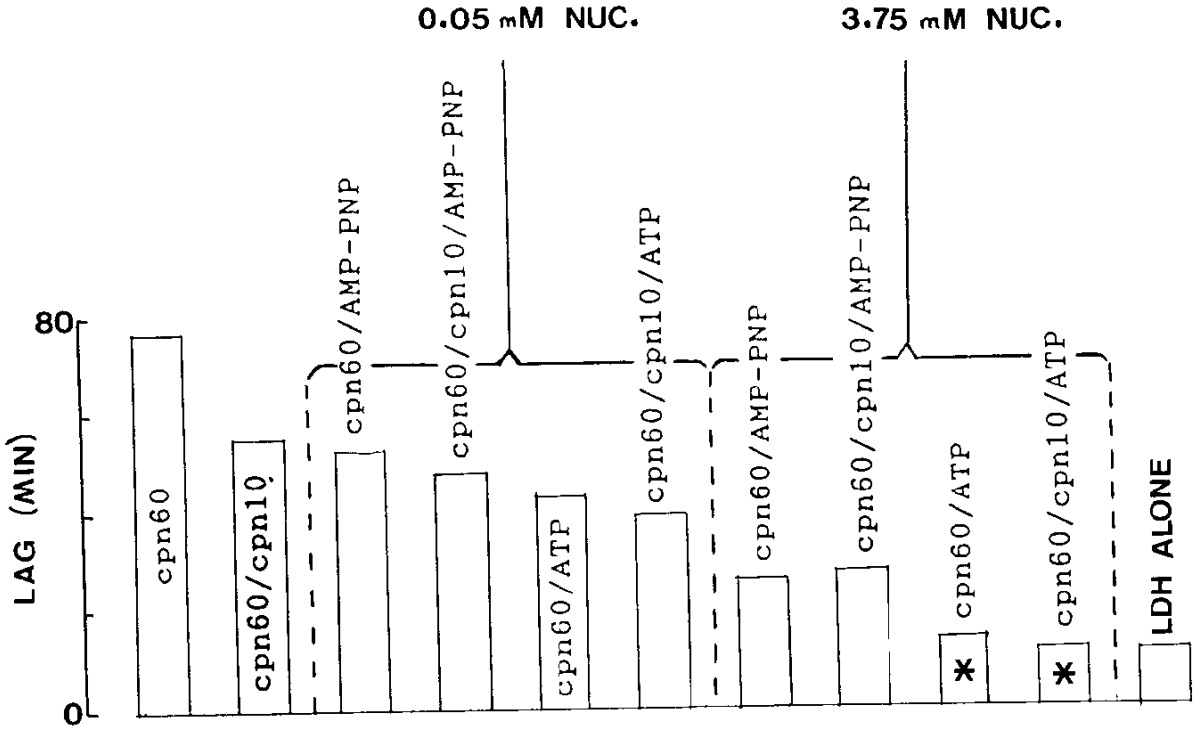


FIGURE 5. Summary of folding times; the influence of cpn60, cpn10, and nucleotides. The lag times for the regaining of LDH activity were measured as described for FIGURES 2, 3, and 4 and are represented here as a histogram. The asterisks mark conditions in which the yield is improved.

3. The binding of cpn10 to cpn60 stabilizes the form that associates strongly with unfolded proteins, not the one that binds AMP-PNP. In doing this, it competes for the hydrophobic binding surface.
4. In optimal conditions (i.e., with cpn10 and ATP), the chaperonin is able to increase the yield of folded protein over a short time period. We see no increase in the rate of folding. These observations demonstrate the ability of the chaperonin either to prevent misfolding or to convert misfolded protein to the native state. We suggest that the latter is more likely.

The previous results of Chandrasekhar *et al.*[13] show that:

5. Cpn10 associates with cpn60 after prolonged incubation with ATP, suggesting

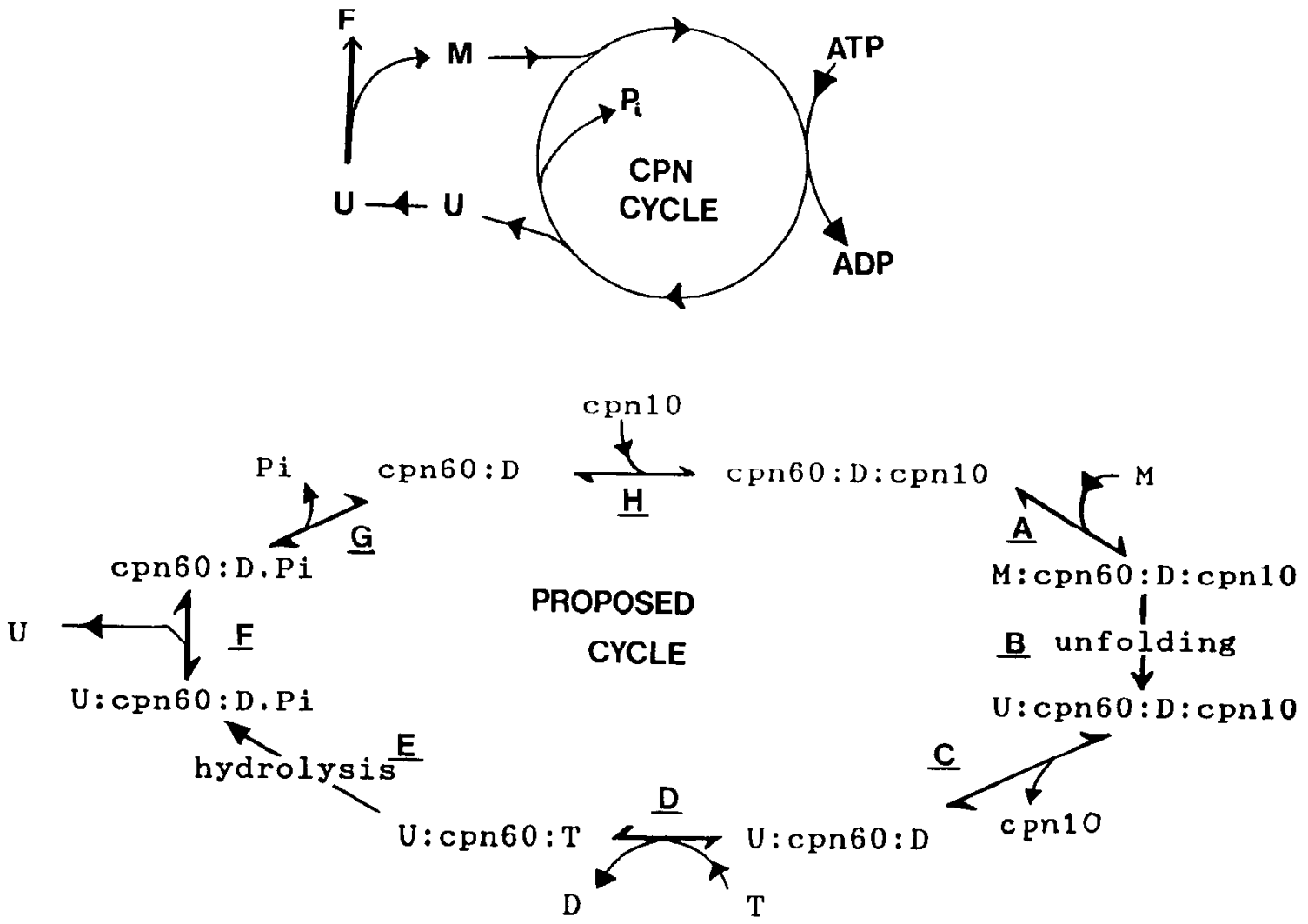


FIGURE 6. Mechanism for the enhancement of protein folding efficiency by chaperonins. Two schemes are shown, the first representing the coupling between folding and ATP hydrolysis where U can fold to its native state F or to a misfolded state M. In net terms, the energy of hydrolysis is coupled to protein unfolding. The second shows the chaperonin cycle in more detail, the steps of which are described in the text. The abbreviation T represents ATP and D represents ADP.

that hydrolysis is required. It does not associate in the presence of nonhydrolyzable analogues. This accords with number 3 above.

6. The rate of ATP hydrolysis by cpn60 is enhanced by association with unfolded protein and inhibited by cpn10.[9,13] We suggest that cpn10 stabilizes a product of ATP hydrolysis in the cpn60 nucleotide site, preventing constant turnover. This product is unlikely to be ADP · P_i as this would lock the chaperonin in a state that has a low affinity for protein substrates. For the mechanistic model, we tentatively suggest that this product is ADP.

The experimental observations are consistent with the mechanism presented in FIGURE 6. When LDH folds from the completely denatured state, a minority (10–15%) folds correctly. Misfolded structures, by virtue of their exposed hydrophobic surfaces, bind to the cpn60/cpn10/ADP complex (step A), this induces the substrate protein to unfold (step B). Cpn10 is then competitively displaced (step C), allowing release of ADP and association of ATP (step D). This reduces the affinity of cpn60 for the unfolded protein, and the hydrolysis of ATP to give ADP · P_i (step E) amplifies this effect. The unfolded protein is then released (step F), followed by P_i (step G), and the reassociation of cpn10 (step H) returns the system to its original state.

The net result of this cycle is to use the energy of hydrolysis of ATP to drive the unfolding of misfolded protein. Chaperonins therefore prevent irreversible aggregation of misfolded structures by recycling them. This allows proteins to refold to their native conformation through self-assembly and requires no specific recognition event.

REFERENCES

1. ELLIS, R. J. & S. M. VAN DER VIES. 1991. Annu. Rev. Biochem. **60:** 321–347.
2. GEORGOPOULOS, C. P. & B. HOHN. 1978. Proc. Natl. Acad. Sci. USA **75**(1): 131–135.
3. HENDRIX, R. W. 1979. J. Mol. Biol. **129:** 375–392.
4. HEMMINGSEN, S. M., C. A. WOOLFORD, S. M. VAN DER VIES, K. TILLY, D. T. DENNIS, C. P. GEORGOPOULOS, R. W. HENDRIX & R. J. ELLIS. 1988. Nature **333:** 330–334.
5. GOLOUBINOFF, P., J. T. CHRISTELLER, A. A. GATENBY & G. H. LORIMER. 1989. Nature **342:** 884–889.
6. SAIBIL, H., Z. DONG, S. WOOD & A. AUF DER MAUER. 1991. Nature **353:** 25–26.
7. ELLIS, R. J. & S. M. HEMMINGSEN. 1989. Trends Biochem. Sci. **14:** 339–342.
8. WICKNER, W. 1989. Trends Biochem. Sci. **14:** 280–283.
9. MARTIN, J., T. LANGER, R. BOTEVA, A. SCHRAMEL, A. L. HORWICH & F. ULRICH-HARTL. 1991. Nature **352:** 36–42.
10. SMITH, C. J., A. R. CLARKE, W. N. CHIA, L. I. IRONS, T. ATKINSON & J. J. HOLBROOK. 1991. Biochemistry **30:** 1028–1036.
11. BADCOE, I. G., C. J. SMITH, S. WOOD, D. J. HALSALL, J. J. HOLBROOK, P. LUND & A. R. CLARKE. 1991. Biochemistry **30:** 9195–9200.
12. CURTIS, A. R. & W. P. SWEETENHAM. 1985. United Kingdom Atomic Energy Authority (Unclassified) HMSO Report AERE-R 11771. Harwell, Oxon.
13. CHANDRASEKHAR, G. N., K. TILLEY, C. A. WOOLFORD, R. HENDRIX & C. P. GEORGOPOULOS. 1986. J. Biol. Chem. **261:** 12414–12419.