Chaperonins as protein-folding machines

Nidhi Bhutani and Jayant B. Udgaonkar*

National Centre for Biological Sciences, Tata Institute of Fundamental Research, GKVK Campus, Bangalore 560 065, India

The *Escherichia coli* chaperone GroEL epitomizes the group of chaperone proteins termed as chaperonins. The wealth of structural and functional information available for GroEL, and its accessory protein, the co-chaperonin GroES, has been of much value in deciphering the role of chaperonins in facilitating the folding of substrate proteins in the cell. The chaperonin machinery has a complex architecture which can undergo dramatic structural rearrangements upon interaction with GroES and nucleotides. Although the reaction cycle of GroEL binding with ATP/ADP as well as GroES is understood in some detail, the kinetic and structural details of the effect of the reaction cycle on the refolding polypeptide are only beginning to be understood. This review discusses the mechanism by which the GroEL/ES/ATP cycle can effect the correct folding of an unfolded substrate, or can even rescue some misfolded or kinetically trapped intermediates. The classical definition of chaperonin action is prevention of intermolecular aggregation, by binding partially-folded molecules, and allowing folding to take place in a 'cage of infinite dilution'. More recent studies have, however, indicated that the chaperonin machinery plays a more active role in facilitating the folding reaction, by altering the energy landscape available to the substrate polypeptide for folding. New data has been analysed, and discussed not only in terms of how they have improved our understanding of chaperonin function, but also in terms of the important new issues that have been raised.

THE first use of the term chaperone in biochemistry was to describe the property of an acidic nuclear protein 'nucleoplasmin'[1], which prevents off-pathway aggregation between histones and DNA during the correct assembly of nucleosomes. The definition was later extended to include the chloroplast Rubisco binding protein (RBP) which assists oligomerization of the protein ribulose bisphosphate carboxylase (Rubisco) by preventing inter-subunit aggregation[2]. RBP shows about 50% identity at the amino acid sequence level with the *Escherichia coli* protein GroEL[3], a protein that had been identified as being indispensable for bacteriophage *l* replication[4]. GroEL was later shown to assist in the correct assembly of the phage-encoded head and tail protein subunits[5]. These initial observations established the general concept of chaperones[6] as a class of unrelated proteins that help in intermolecular assembly, by transiently binding and preventing incorrect interactions, without being a part of the final assembly.

The requirement of such accessory proteins for folding was somewhat surprising because it is well accepted that the amino acid sequence of a polypeptide is the primary determinant of, and contains complete information for, its final folded conformation[7]. Thus, many denatured polypeptides are able to regain their native folded structure with 100% efficiency *in vitro*, upon removal of the denaturant. The cellular milieu is, however, distinguished from the *in vitro* environment in terms of molecular crowding[8]. The total macromolecular concentration (mainly protein, RNA and DNA) in the cell is at least 300 g/l[9], with the result that 20–30% of the cellular volume is occupied by these macromolecules. Due to an excluded volume effect, the reaction rates and equilibria of many macromolecular reactions in the cell are expected to be considerably different *in vitro*, where reactions are usually studied at low concentrations of reactants. In particular, intermolecular association constants are expected to be greatly increased in the cell. Another aspect of *in vivo* folding that is different is that the folding information in terms of the amino acid sequence does not become available all at once, because protein biosynthesis is vectorial, i.e. from the N-terminus to the C-terminus of the polypeptide. As a consequence, the nascent polypeptide chains being synthesized on the ribosomes are in danger of misfolding in the absence of complete folding information, as well as in danger of aggregating because of the proximity of other nascent polypeptides being synthesized on the large poly-ribosomal assemblies[10].

It is the presence of molecular chaperones that maintains a high fidelity of protein-folding reactions *in vivo*, in spite of the potential hazards outlined above. Chaperones may differ in their structure and mode of action, but they all recognize and bind non-native or thermodynamically unstable forms of proteins. Some chaperones are known to act co-translationally, by binding nascent polypeptides to maintain them in their folding– or translocation–competent forms, while others act post-translationally, by providing a cavity for the polypeptide to fold in a dilute environment. Different chaperones have also been known to act in a network, thereby directing a newly synthesized polypeptide towards its correct *in vivo* fate[11,12]. Thus, a good working definition of a molecular chaperone is 'a protein which transiently binds to and stabilizes an

*For correspondence. (e-mail: jayant@ncbs.res.in)

unstable conformer of another protein, and through regulated binding and release, facilitates its correct fate *in vivo*: be it folding (following *de novo* synthesis, transit across a membrane, or stress-induced denaturation), oligomeric assembly, interaction with other cellular components, switching between active and inactive conformations, intracellular transport, or proteolytic degradation, either singly or with the help of co-factors'[13]. Table 1 (refs 13 and 14) provides a list of the better-known cytosolic chaperones, with a brief outline of their functions, to emphasize the diversity of cellular processes in which these molecular machines are now implicated to play a role.

Chaperonins

A prominent and well-studied class of molecular chaperones are the chaperonins[3], grouped together on the basis of striking structural and sequence similarities. The chaperonins are characterized by ring-shaped structures made up of seven, eight or nine subunits enclosing a cylindrical cavity, and with two rings stacked back-to-back[15–17]. Chaperonins are further subdivided into two classes: group I and group II. The sequence identity between members within each group is about 50%, but is much lower between the two groups. The group II chaperonins are more heterogeneous in sequence and structure compared to the group I chaperonins[15,18–20].

Group I consists of chaperonins found in chloroplasts and other plastids, mitochondria and eubacteria, with the prominent examples being the *E. coli* chaperonin GroEL and its accessory co-chaperonin GroES, Rubisco binding protein (Chloroplast cpn60) and its co-chaperonin Chloroplast cpn10, as well as cpn60 and cpn10 from the mitochondria[16,21–23]. In these chaperonins, each ring consists of seven identical subunits. In the co-chaperonin, a single ring of seven subunits can form a cap-like structure atop the chaperonin cavity. Co-chaperonins play an essential role in chaperonin-assisted folding, as exemplified by the role of GroES in GroEL-mediated protein folding[24,25].

Group II consists of chaperonins found in the archaebacteria and the eukaryotic cytosol, and is exemplified by TCP-1, the thermosome complex and TF55. Group II chaperonins have eight or nine subunits per ring and appear to work without any co-chaperonins. In the thermosome complex, there is however, an in-built lid formed by a sequence insertion in the substrate binding domain, which forms a large protrusion towards the central cavity[26,27]. This protrusion has been suggested to be the functional equivalent of the co-chaperonin, although the group II chaperonin-mediated folding mechanism is as yet poorly understood. Another feature that distinguishes the group II chaperonins is that each ring may contain non-identical subunits: the thermosome contains two types of subunits that alternate within a eight-membered ring, while the eukaryotic cytosolic chaperonin ring is assembled from eight totally different proteins[28–30]. Biochemical studies

Table 1. Representative folding chaperones

Class	Selected member	Function
CCT	TF55, TriC	Folding of nascent polypeptides and misfolded proteins in eukaryotic cytosol.
DnaJ	DnaJ, Dj1A, CbpA, HscB, Hsp40, Ydj1, Sec63, Auxilin, CSP's, Mdj1, Hdj1, Hdj2	Co-chaperone of Hsp70
GimC	Gim 1–6, Prefoldin	Folding of actin and tubulins.
GrpE	GrpE, Mge1p	Co-chaperone of Hsp70 in bacteria, mitochondria and chloroplast.
Hsp10	GroES, Gp31, Hsp10, Cpn10	Co-chaperonin of Hsp60
Hsp47	Hsp47	Folding and assembly of collagen.
Hsp60	GroEL, Hsp60, Cpn60	Folding of nascent polypeptides and misfolded proteins in bacteria, mitochondria and plastids.
Hsp70	DnaK, HscA (Hsc66), hsc70, Hsp68, 70, 71 and 73, Bip, grp75, 78 and 80, KAR2, SSA1–4, SSB1, SSC, SSH1, LHS1, KAR2	Protein assembly and translocation across membranes in prokaryotes and eukaryotes.
Hsp90	HtpG, Hsp90, Grp94, ERp90, endoplasmin, Hsp108, gp96, Hsp83, 87	Prevention of aggregation and regulation of activity of steroid receptors and kinases.
Hsp100	C1pA, B, X and Y, Hsp104, Hsp78	Assist in proteolysis and disaggregation.
Prosequences	Pro-subtilisin, proα-lytic protease	Protease assembly and maturation
sHSP	IbpA, IbpB, Hsp16.5, Hsp12, Hsp42, $\alpha\beta$-crystallin	Prevention of aggregation
Trigger factor	TF	Associated with ribosomes, chaperones nascent chains and catalyses prolyl isomerization.

have suggested that this heterogeneity might contribute to the substrate specificity of the group II chaperonins, with each of the subunits recognizing specific motifs in the substrate proteins[31]. The subdivision of the chaperonins is consistent with the evolutionary view that chloroplasts and mitochondria have originated from eubacteria, while archaebacteria have contributed to the eukaryotic cytosol[10].

E. coli GroEL is the best-studied chaperonin, and its extensive characterization[16,17,21,22,32–34] has clearly resulted in delineating the role of chaperonins as folding machines that facilitate the protein folding reaction. This review examines the known aspects of the structure and function of the chaperonin machinery consisting of GroEL and GroES, from the viewpoint of how the interaction of a substrate protein with the chaperonin machinery affects both partners of the interaction.

E. coli chaperonin machinery

GroEL and GroES architecture

The structure of the GroEL/GroES machinery has been elucidated using an effective combination of X-ray crystallography and cryo-electron microscopy (cryo-EM). While it was a major achievement to solve the structure of such a large protein by X-ray crystallography, the large size and the symmetry in structure actually facilitated the electron microscopic studies. Cryo-EM studies have also been useful in providing structural information about the various conformational changes that take place in GroEL, upon binding different nucleotides, and GroES.

GroEL is a homo-tetradecamer composed of two seven-membered rings that are stacked back-to-back to form two discontinuous cavities. The unliganded GroEL complex is 142 Å in height and 140 Å in diameter with the central cavity within each ring being 45 Å in diameter[35]. Each subunit (57 kDa) consists of 547 amino acids and folds back upon itself like a 'U', to form three domains (Figure 1), as revealed by the crystal structure at 2.8 Å (ref. 35): (1) The *equatorial* domain (residues 6–133 and 409–523) contains the nucleotide-binding sites and provides most of the intra-subunit as well as inter-subunit contacts. The domain is largely helical and well-ordered. (2) The *apical* domain (residues 191–376) contains the residues involved in polypeptide binding, identified using mutagenesis[36]. The domain also contains the residues involved in binding of the co-chaperone GroES. (3) The hinge-like *intermediate* domain (residues 134–190 and 377–408) connects the apical and equatorial domains, besides containing structural elements that contact the apical domain on the adjacent subunit. The last 24 C-terminal amino acid residues of each subunit appear disordered in the crystal. From electron micrographs[37–39] and small-angle neutron scattering studies[40], it has been deduced that the coming together of the C-terminal segments of all seven subunits in a ring at the equatorial plane, causes the discontinuity between the two cavities, turning them into two separate folding chambers.

The crystal structure of GroES at 2.8 Å shows that it is a heptamer made up of seven 10 kDa subunits. GroES appears to have a dome-shaped structure, which is 75 Å in diameter and 30 Å in height[23,24]. Each of the seven subunits has a core *b*-barrel structure with two *b*-hairpin loops, one of which (residues 16–32) arches downwards

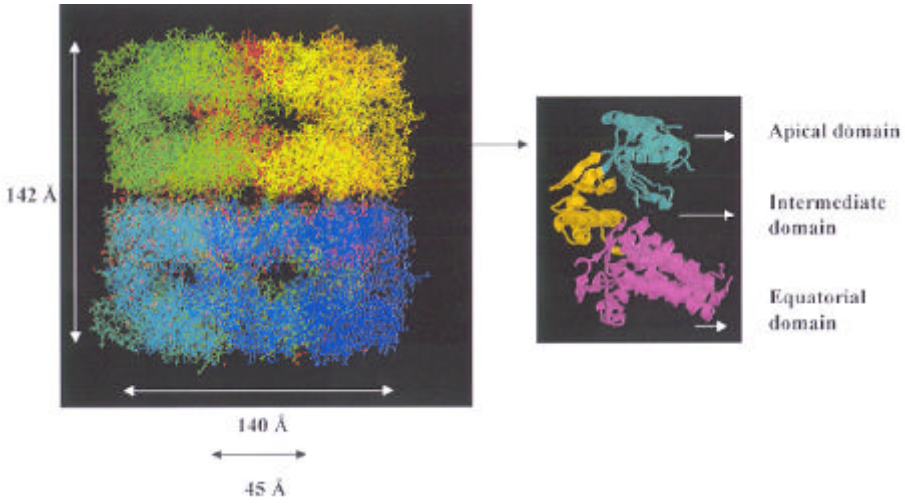

Figure 1. Structure of GroEL. (Left) The two heptameric rings are stacked back-to-back[61]. The figure has been generated using RASMOL[129] from the protein structure 1GR5 in the PDB, contributed by Ranson *et al.*[61]. (Right) The crystal structure of the GroEL oligomer at 2.8 Å[35] shows that each of the subunits is organized into three domains as described in the text. The figure has been generated using RASMOL[129] from the protein crystal structure 1GRL in the PDB, deposited by Braig *et al.*[35].

and is quite disordered and the other loop (residues 46–56) arches upwards and inwards to contribute to forming the roof of the dome. The core *b*-barrel (residues 60–97) in each subunit interacts with the first *b*-strand (residues 9–15) of the adjacent subunit for the formation of the dome-like oligomer.

Nucleotide and GroES-induced conformational changes in GroEL

Valuable insights into the facilitation of folding of a substrate polypeptide by GroEL have been provided by the determination of structural details of the GroEL conformations with the nucleotides (ATP or ADP) and GroES bound to it, while the sequence of events and their relevance in the folding cycle have been elucidated by many elegant biochemical studies. Early biochemical studies had established that each GroEL subunit has an ATP-binding site[41] and a K^+-dependent ATPase activity, which is cooperative with respect to ATP[42,43] as well as K^+-binding[44]. It was also evident that ATP binding and hydrolysis cause conformational changes in GroEL, alter its affinity for the substrate, and hence, drive the folding cycle[37,45–49]. In a crystal structure of GroEL, with bound ATP*g*S, determined at a resolution of 2.4 Å (ref. 50), it was observed that the ATP*g*S interacts with residues confined to the equatorial domain, and that it binds in a pocket facing towards the inside of the cylinder. The nucleotide was not observed to make any direct contact with either the apical or the intermediate domains.

Although early negative-stain EM images[37] showed some differences between GroEL and the GroEL–nucleotide complexes, a clearer understanding was obtained with the low-resolution structures of GroEL–nucleotide complexes determined by cryo-EM and single-particle analysis[39,51]. The arrangement of the GroEL, GroEL–ADP and GroEL–ATP structures reconstructed via cryo-EM images[51] showed that nucleotide binding elongates the GroEL oligomer and twists the apical domains to different extents. The apical domains appear to pivot on the hinge between the intermediate and apical domains, so that the residues initially facing the central cavity are now buried in the intersubunit contacts. In the case of GroEL–ADP, this pivoting movement is counterclockwise leading to a more opened up structure, while in the case of GroEL–ATP, the movement is mainly radially inward. The ADP-bound form of GroEL, therefore, appears to have a greater exposure of the substrate-binding apical domains. The structures corresponded well with the observations[52,53] that the substrate affinity of GroEL is reduced in the presence of ATP. The GroEL–GroES–nucleotide complexes showed the extended loops linking GroEL and GroES, and the GroEL–ES–ADP and GroEL–ES–ATP complexes were easily distinguished: the interface between the two rings being altered in the GroEL–ES–ATP form, so that it is more asymmetrical than the GroEL–ES–ADP form. These structures suggested that nucleotide binding and hydrolysis can induce a structural asymmetry between the two rings, in GroEL as well as in the GroEL–GroES complexes, which might be functionally significant.

In parallel studies, it was observed that GroEL displays nested cooperativity with respect to ATP binding, i.e. there is positive cooperativity for binding ATP to all the subunits within one ring, while there is negative cooperativity for it binding to subunits in the associated ring[54,55]. According to the nested cooperativity model, it was proposed that each GroEL ring is in equilibrium between two states: a tense (*T*) state with low affinity for ATP and high affinity for substrate, and a relaxed (*R*) state with high affinity for ATP but low affinity for substrate. Hence, the GroEL oligomer was postulated to exist in an equilibrium between three states: *TT*, *RT* and *RR*. Although this model is a considerable simplification, it provided a good framework for the inherent structural asymmetry between the two rings that was observed in the GroEL–nucleotide complexes. In the crystal structure[35], it was observed that Arg 197 in the apical domain of a subunit forms a salt bridge with Glu 386 in the intermediate domain of a neighbouring subunit. This salt bridge appears to be involved in the allosteric switching of GroEL, because the mutation of Arg 197 → Ala (R197A) reduces both positive and negative cooperativity of ATP hydrolysis by GroEL[56,57].

The crystal structure of the asymmetric GroEL–GroES–$(ADP)_7$ complex determined at 3 Å resolution[58] (Figure 2) gave further insights into the domain move-

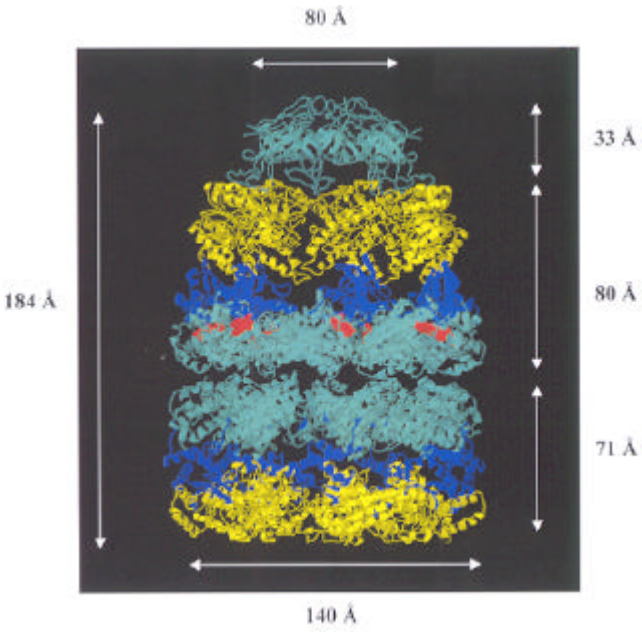

Figure 2. Crystal structure of the asymmetric GroEL–GroES–$(ADP)_7$ complex at 3 Å[58]. The GroEL–GroES contacts are revealed as are the domain rearrangements, brought about by GroES binding. The figure has been generated using RASMOL[129] from the protein crystal structure 1AON in the PDB, deposited by Xu *et al.*[58].

ments associated with nucleotide binding and their relevance for the folding cycle. The structure of the *cis* ring, which binds GroES, undergoes dramatic rearrangements while that of the *trans* ring deviates only slightly from the unliganded GroEL structure. The intermediate domains in the *cis* ring subunits swing downwards, creating new contacts with the equatorial domains of the same as well as neighbouring subunits, which impede the dissociation of the bound nucleotide. The central cavity volume is almost doubled as a result of the apical domains swinging up by 60° relative to the equator as well as twisting about the long axis of the domain. These movements also change the earlier hydrophobic nature of the central cavity lining to hydrophilic, by burying the initially exposed hydrophobic residues in either the newly formed inter-subunit contacts or in the GroEL–GroES interface. This hydrophilic cavity, capped by GroES, is 80 Å in diameter and 85 Å high with a volume almost doubled to 175,000 Å^3, and can easily accommodate a native, globular protein of a relative molecular mass of 70 kDa. Apart from the newly formed contacts with the intermediate domains, the equatorial domains of the *cis* ring undergo an *en bloc* inward tilt of around 4°, so that the inside of the *cis* ring is around 3 Å lower while the outside is 5 Å higher. It was postulated that the inter-ring subunit contacts are maintained by a complementary outward tilt of the *trans* ring, and that these *en bloc* movements are responsible for the negative cooperativity between the two rings[54,55]. In contrast to these dramatic rearrangements in the GroEL rings, the complexed GroES structure remains similar to its free structure[23,24], except that the mobile loops of all the subunits become structured as *b*-hairpins upon interacting with GroEL[59]. The importance of the mobile loops in GroEL–GroES interaction had been indicated by an earlier study which showed that mutations in the mobile loop disrupted GroES binding to GroEL[60]. GroEL residues that are involved in forming the GroEL–ES interface, as observed in the crystal structure[58], cluster on two helices of the apical domains in each subunit. Mutagenesis experiments[36] have shown that some of these interface residues are important for substrate polypeptide binding, suggesting that GroES binding helps in releasing the polypeptide into the GroEL cavity.

A recent study has provided a better molecular understanding of the nested cooperativity in GroEL using an elegant combination of cryo-EM, mutagenesis and X-ray crystallographic data[61]. By using an ATP hydrolysis-defective mutant, D398A, it has been possible to obtain cryo-EM images of an '*RT*' state of GroEL. The cryo EM images have been obtained at a resolution of 10 Å, and these images have been modelled by fitting domain structures from the X-ray crystal structure of GroEL, as rigid bodies. As observed earlier with the ADP-bound crystal structure of GroEL–GroES, the apical domain of the ATP-bound structure undergoes an upward tilt and twist with respect to the unliganded GroEL. The intermediate domain undergoes a 20° downward tilt, and this causes breaking of the salt bridge between the intermediate domain E386 and R197 present on the adjacent subunit apical domain. The breaking of this salt bridge is accompanied by the formation of a new salt bridge between E386 and K80 which lies on the adjacent equatorial domain. The E386–R197 salt bridge had already been implicated in the allosteric switching by mutagenesis as described above, as well as by theoretical predictions[62]. A mutation of K80 weakens positive cooperativity. Thus, the salt bridge rearrangement offers a plausible explanation for the intra-ring positive cooperativity. Also, helix D of the equatorial domain appears to change its inter-ring contacts upon ATP binding, and this manifests itself in increased separation between the two rings compared to the unliganded GroEL. This study thereby provides an alternative explanation for the inter-ring negative cooperativity, which is in contrast to the earlier model in which a mutual tilting of the two rings had been proposed[58].

The GroEL–ES reaction cycle

The GroEL–GroES reaction cycle has been elucidated[48,49,63–68] (Figure 3) in the presence as well as absence of the substrate polypeptide. Crucial to this elucidation was the use of a number of GroEL mutant proteins, e.g. mutant proteins defective in ATP hydrolysis, in binding the substrate polypeptide and GroES, as well as single-ring mutants that are association-incompetent[63,67]. These were further used for the generation of various mixed-ring combinations whereby one of the rings was mutant but the other was wild type, or for a combination of two mutant rings with similar or different defects. In this way, it has been possible to dissect out the precise sequence of events in the reaction cycle utilizing techniques such as stopped-flow fluorescence, anisotropy and fluorescence resonance energy transfer (FRET)[63,65,67,68]. In the presence of non-native substrate proteins, the GroEL rings alternate between folding-active *cis* and folding-inactive *trans* complexes[68]. The binding of a substrate polypeptide to a GroEL ring stimulates the binding of ATP as well as GroES. Binding of ATP molecules to each of the seven equatorial domains of the *cis* ring brings about conformational changes in the *trans* ring that decrease the affinity of the *trans* ring for the substrate as well as GroES. The binding of ATP and subsequently GroES, also brings about an increase in the *cis* cavity volume, and a change in the internal cavity surface from hydrophobic to hydrophilic[58]. The substrate protein is effectively displaced into the cavity by GroES, as a result of the stabilization of the GroEL–ES contacts that overlap with the substrate-binding sites, and is allowed to fold till the *cis* ring dissociates[49,63–65]. The hydrophilic lining of the cavity promotes sequestration of exposed hydrophobic patches in the substrate protein, causing it to refold. The *trans* ring cannot

bind a substrate till the ATP molecules bound to the *cis* ring are hydrolysed to ADP. Also, the *cis* GroEL–GroES–$(ATP)_7$ ring is stable and does not dissociate till the ATP molecules are hydrolysed. Hence, the hydrolysis of these ATP molecules is the rate-limiting step in the folding cycle[64] (~ 0.12 s^{-1}).

The hydrolysis of the *cis*-ring ATP molecules to ADP causes conformational changes that are transmitted to the *trans* apical domains, which then become competent to bind the substrate protein as well as GroES. It has also been demonstrated, using FRET with pyrene-labelled GroEL[64,68], that this hydrolysis is followed by a structural rearrangement of the ADP-bound *cis* ring of GroEL, which primes it for GroES dissociation. In the absence of a substrate, this structural rearrangement occurs at a rate of 0.042 s^{-1}, and would therefore be the rate-limiting step in the cycle[64]. Such a slow release of GroES is, however, not observed in the presence of substrate. In the presence of bound substrate on the *trans* ring, the GroES release step is accelerated to 1–2 s^{-1} (ref. 68). Since the binding of a substrate molecule to the *trans* ring is reasonably faster (k_B ~ 1–2 × 10^7 M^{-1} s^{-1}) than the rate of the structural rearrangement which leads to the dissociation of GroES from the *cis* ring, substrate-binding to the *trans* ring takes place before the dissociation of GroES from the *cis* chamber[68]. In a separate set of experiments, it was shown that the binding of ATP molecules to the *trans* ring[67], without the necessity of their hydrolysis, is sufficient to bring about the dissociation of such a primed *cis* GroEL–GroES–ADP complex.

The binding of GroES to a substrate-ATP-bound *trans* ring, and the dissociation of GroES from the *cis* ring are concerted events[68]. Hence, the *trans* ring first binds the substrate after ATP is hydrolysed to ADP in the *cis* ring, and then binds ATP. The binding of GroES to the *trans* ring concomitant to its dissociation from the *cis* ring then initiates the next reaction cycle. Thus, the GroEL machinery works efficiently as a 'two-stroke' machine[69]: the dissociation of GroES from the *cis*-folding active ring of GroEL is coupled to the conversion of the *trans* ring into the folding-active chamber. Apart from the hydrolysis of seven ATP molecules, no additional energy expenditure is required in one folding cycle, to clear all ligands from the GroEL rings.

GroEL-assisted folding

How does the reaction cycle assist the folding of the bound substrate polypeptide? Although the chaperonin

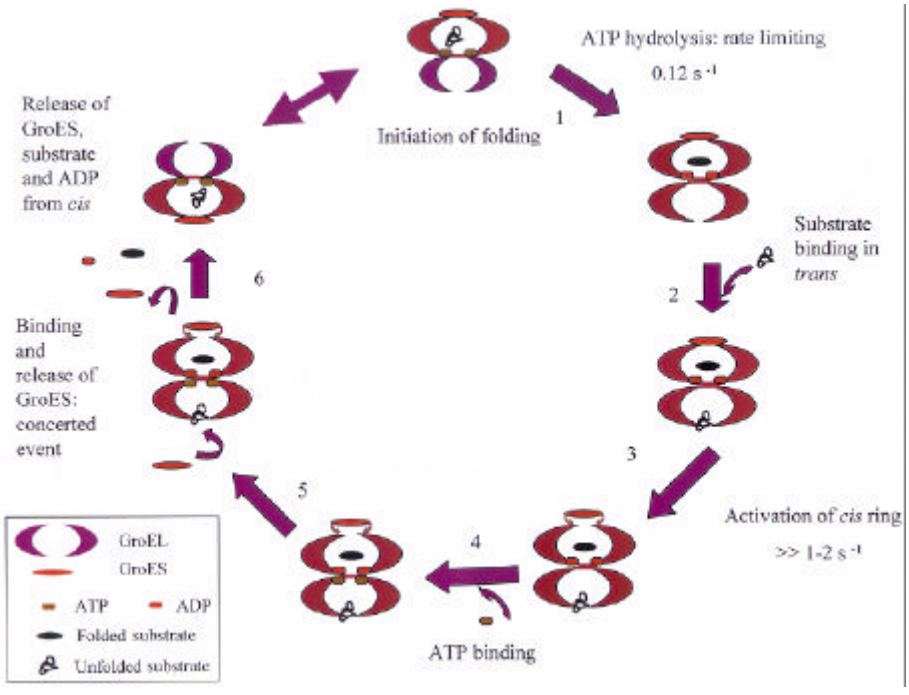

Figure 3. The GroEL–GroES reaction cycle. 1) Binding of substrate protein stimulates ATP and GroES binding in *cis*, which leads to the substrate protein being released in the cavity, and initiation of folding. 2) Substrate protein binds to the *trans* ring only after ATP hydrolysis takes place in the *cis* ring. 3) In the presence of substrate in the *trans* ring, there is a fast structural rearrangement in the ADP and GroES-bound *cis* ring that primes it for releasing GroES. 4) The binding of substrate protein in the *trans* ring stimulates ATP binding in *trans*. 5) The subsequent binding of GroES to the *trans* ring is simultaneous with the release of GroES from the *cis* ring. 6)The GroES- and ATP-bound *trans* ring causes structural rearrangements in the *cis* ring leading to release of ADP and substrate protein. Upon completion of one folding cycle, the next cycle is initiated in the alternate ring.

reaction cycle has been well delineated, it has remained unclear how it affects the kinetic and molecular events at the level of the refolding polypeptide. It has been difficult to study the role of GroEL in facilitating the folding of its natural substrate proteins *in vitro*, because their folding mechanisms are poorly understood. An additional drawback is that the partially-folded forms of these proteins, to which GroEL binds, are difficult to characterize because they are aggregation-prone. Hence, many of the studies pertaining to GroEL–substrate interaction have utilized proteins that are not natural GroEL substrates, but whose folding pathways are well-characterized.

Trying to understand the effect of GroEL on the folding of its substrates brings back into focus the fundamental question of whether protein folding *in vivo* is the same as that *in vitro*[70]. Does the crowded environment of the cell alter the way a polypeptide folds, or does the folding mechanism remain the same as in the dilute environment of a test-tube? Does folding in the GroEL cavity, which provides an environment of infinite dilution, occur in a manner similar to how it occurs outside the cavity in the crowded cytosol[71–73]? Does GroEL do more than just passively providing an effectively dilute environment in its cavity?

It has long been obvious that the chaperonin GroEL helps in preventing aggregation[74–76]. But can GroEL disaggregate or rescue a misfolded protein and if it can, what is the mechanism? Does the chaperonin assist folding only by utilizing the energy of its binding to the substrate protein? Is the binding energy utilized merely in a passive way to pull coupled equilibria in the direction of the most tightly bound form by a thermodynamic coupling mechanism? Or is it utilized in an active way to lower the activation energy for folding, thereby resulting in catalysis of the folding reaction? Can binding energy be similarly utilized by either an active or passive way to unfold a misfolded substrate protein? Is the energy from ATP hydrolysis used directly to assist folding or to rescue misfolded structures? The major changes in the substrate polypeptide conformation are expected to occur upon initial GroEL binding, on the release of the substrate in the cavity, and on subsequent refolding of the substrate before it is ejected out of the cavity. The rest of the review focuses on the information available on each of these events and how the available data can be interpreted in the light of the important outstanding questions.

GroEL binding to substrate

Early cryo-EM images suggested that the apical domains are involved in the binding of substrates to GroEL[39]. It was also shown that several mutations in the apical domain could completely abolish the binding of a known substrate OTC (Ornithine transcarbamylase) to GroEL[36]. These mutations – Y199E, Y203E, F204E, L234E, L237E, L259S, V263S and V264S (Figure 4) – were mainly in the hydrophobic residues, thus implicating hydrophobic interactions in the binding of a substrate molecule to GroEL.

Conformation of the bound substrate protein: The crystal structure of the isolated apical domain of GroEL (191–376 residues)[77] showed the same fold as intact GroEL (Figure 4), and the isolated apical domain was also shown to mimic oligomeric GroEL in binding substrate proteins. This isolated domain was hence termed a minichaperone, although later studies have shown that the intact oligomeric GroEL structure is required for full activity[78]. In any case, the isolated apical domain proved to be useful in determining the structure of the GroEL substrate-binding site. The crystal structure of the minichaperone, along with a 17-residue N-terminal peptide from the neighbouring apical domain, was solved at 1.7 Å resolution[79]. The structure showed that the N-terminal peptide was bound between two parallel α-helices of the apical domain, termed helix H (230–244 residues) and helix I (254–268 residues), and was in an extended conformation. This result was in good agreement with an earlier mutagenesis study that showed that point mutations in a few residues of helix H and helix I could abolish polypeptide binding to GroEL[36]. Residues within and near these helices were also implicated in substrate binding by studies in which discrete fluorescent labels that had been placed at various residues mutated to cysteines were monitored[80]. The crystal structure of the complex formed by the apical domain of GroEL with a high-affinity 12-mer peptide[81] that was selected out of a random 12-mer peptide library using phage-display, was solved to 2.1 Å resolution. This peptide, like the 17-residue N-terminal peptide, was also bound between helices H and I peptide, but it adopted a

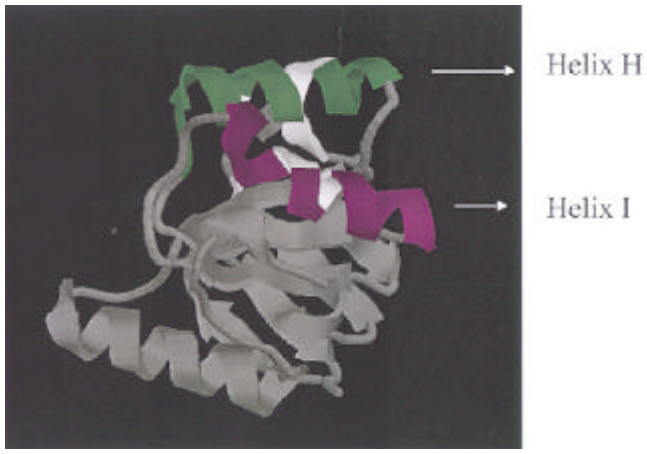

Figure 4. Structure of an isolated apical domain of GroEL[79]. Substrate-binding takes place between the two marked parallel helices: helix H (230–244 residues) and helix I (254–268 residues). (The figure also marks the residues, upon whose mutations polypeptide binding is abolished.) The figure has been generated using RASMOL[129] from the protein crystal structure 1JON in the PDB, deposited by Buckle and Fersht[72].

b-hairpin structure. Thus, different substrates seem to adopt different conformations upon binding GroEL: the neighbouring N-terminal peptide maintained an extended conformation, while the high-affinity 12-mer peptide formed a *b*-hairpin. Earlier NMR studies[82,83] utilizing transferred nuclear Overhauser effects, had also shown binding of *a*-helical polypeptides to intact GroEL. The 13-residue N-terminal peptide of the protein rhodanese was unstructured in solution, but appeared to assume an *a*-helical conformation upon binding GroEL[82]. A similar stabilization of *a*-helical structure was also observed with another short peptide, vsv-C[83].

The ability of GroEL to accommodate many different bound conformations reflects an inherent flexibility in the GroEL substrate-binding domain, which appears to form the basis of its interaction with a wide variety of substrates. More recently, it has been shown, using an elegant mutagenesis strategy[84], that substrate proteins can bind to multiple contiguous apical domains. A single-polypeptide chain GroEL was generated, which formed a fully functional covalent ring having all the seven subunits. Through mutagenesis, apical domains of different subunits could then be systematically rendered deficient in substrate binding. It was observed that while at least three contiguous wild-type subunits were required for binding malate dehydrogenase (MDH) or Rubisco, two contiguous domains were sufficient to bind rhodanese. Thus, the number of binding sites per GroEL ring varies from substrate to substrate. Moreover, because different contiguous and non-contiguous combinations of the apical domains are available for binding, GroEL might be able to bind many different conformations of the same substrate protein.

In addition to the structural studies described above, various kinetic studies have shown that GroEL can bind to many of the forms that accumulate on the folding pathways of proteins. A proper understanding of the nature of substrate-binding to GroEL, and the role of the GroEL reaction cycle in facilitating the folding reaction, has been difficult to obtain because of the transient nature of the folding intermediates, because of possible conformational heterogeneity of the bound forms, and because many of the proteins studied do not require or utilize the entire chaperone machinery to fold.

GroEL appears capable of binding unfolded forms as suggested for thermally unfolded *b*-lactamase[85], and then demonstrated more clearly for cyclophilin[86] and barnase[87] using equilibrium hydrogen–deuterium exchange coupled with NMR. Kinetic studies monitoring the intrinsic fluorescence of a substrate protein in the presence of GroEL, have shown that GroEL interacts with early unstructured intermediates in the case of lactate dehydrogenase[45], barnase[88] and barstar[89]. GroEL has also been shown to interact with a molten globule intermediate[90] as well as late structured intermediates of alpha-lactalbumin[91], and also with a late folding intermediate of a Fab antibody fragment having a defined quaternary structure[92]. GroEL can interact with the native state of barnase, and native-like structure is retained in human DHFR[93] when bound to GroEL. Proteins can fold while remaining bound to GroEL[88]; thus, it is not surprising that multiple conformations of the same protein can bind GroEL. As expected, GroEL binds to the various folding intermediates of different proteins with widely varying dissociation constants observed to be in the 10^{-6}–10^{-11} M range[32]. The binding event can be as fast as that of a diffusion-controlled reaction with a bimolecular rate constant greater than 10^8 M^{-1} s^{-1} (ref. 88), or it can be considerably slower, with the rate constant being in the range of 10^5–10^6 M^{-1} s^{-1} (ref. 45).

Due to the large substrate repertoire, no specific determinants in the sequence and structure of substrate proteins, which are recognized by GroEL, have been identified. *In vivo*, GroEL appears capable of binding up to 15% of *E. coli* proteins[94], as shown by anti-GroEL immunoprecipitation of newly synthesized polypeptides that were labelled using a pulse-chase strategy. Many of the proteins that bind to GroEL possess one or multiple *ab* domains[95], and it was suggested that such multi-domain proteins might need the mediation of chaperones because they have a greater tendency to misfold or aggregate.

Consequences of GroEL binding

Prevention of aggregation: One major action of the chaperonin machinery in assisting folding, that became obvious from the initial studies using substrates like rhodanese[74], citrate synthase[75] and DHFR[76], is to prevent aggregation which is a major side-reaction of protein folding[96]. These aggregation processes are expected to be more pronounced in the cellular context due to macromolecular crowding[8]. Various hypotheses have been put forth to explain the mechanism of action by which the chaperonin machinery can either prevent aggregation or even disaggregate preformed complexes. The original 'Anfinsen's cage' model[97] suggests that GroEL binds aggregation-prone intermediates and stabilizes them, allowing refolding only inside the cavity upon GroES-mediated release of the substrate. Further folding takes place in an environment of effectively infinite dilution, thereby preventing aggregation. Such a model was used to explain the effect of GroEL in preventing the aggregation of rhodanese[74] and citrate synthase[75].

Another model entails a catalytic role for the chaperonin machinery whereby it can bind to preformed aggregates and disrupt them, using multiple cycles of binding and release, thereby allowing them another chance at proper folding and oligomerization. Such a model predicts that substoichiometric amounts of GroEL and GroES should be capable of disrupting these aggregates. Mitochondrial MDH had been shown to undergo aggregation that competes with the folding process[98], and substoichiometric amounts of GroEL and GroES can catalyse a reversal

of the initial aggregation step but not the later irreversible steps[99,100]. The ability of GroEL to carry out disaggregation in this manner can explain its ability to increase the refolding yields as well as the apparent rates of refolding of proteins.

In recent studies on the thermal inactivation of rhodanese[101,102], it was observed that an intermediate is initially formed, which then forms dimers and small oligomers that show only a small increase in light scattering and fluorescence homotransfer of the label fluorescein. Larger aggregates are then formed by the association of these dimeric or oligomeric species. It was observed that GroEL binds to the early intermediate, but cannot recognize either the small oligomers or the large aggregates. GroEL appears, however, to be capable of unfolding the early intermediate so that it can refold upon addition of GroES and ATP. These observations reiterate the point that initial binding by GroEL can modulate the conformation of a substrate protein, before proper refolding proceeds inside the 'Anfinsen's cage'.

Unfolding action: Initial suggestions that binding to GroEL could lead to partial or complete unfolding of the substrate protein, originated from the observation that there was a loss of activity of folded pre-*b*-lactamase[103] and human DHFR (in the absence of ligand)[104] in the presence of GroEL. Further, the observation that a stable intermediate of MDH, which does not fold spontaneously to the native state, could fold to the native state in the presence of GroEL and GroES[105], suggested that partial unfolding by the chaperonin might provide a misfolded polypeptide another chance at refolding. It was also observed that a conformation of rhodanese that bound GroEL again after a round of binding and release, did not show any increased protection to protease action[65]. This observation suggested that the binding event pulls the protein back to the same unfolded form each time, until the protein achieves a folding-competent conformation that is no longer capable of binding GroEL.

Different mechanisms have been suggested for describing how a misfolded substrate protein might be unfolded by GroEL. According to the first active unfolding mechanism, the free energy of binding of substrate protein to GroEL is utilized to catalyse the partial or complete unfolding reaction. The only instance of a protein undergoing such catalysed unfolding in the presence of GroEL was observed by examining the deuterium exchange at the protected amide hydrogen sites of barnase[87]. In the presence of catalytic amounts of GroEL, native barnase was observed to undergo global unfolding, at least transiently, so that exchange is observed for the amide protons that exchange via global unfolding. Cyclophilin[86] appears to be another example of a protein that undergoes such an active unfolding action in the presence of GroEL; but alternative mechanisms cannot be ruled out because stoichiometric concentrations of GroEL had to be used for the effect to be seen. A number of other hydrogen exchange studies on stable binary complexes of GroEL with *a*-lactalbumin[106] and human DHFR[93,107] indicate that these proteins do not become globally exchanged in the complexes. Although partial, local unfolding is still possible in these and other substrate proteins, such unfolding has not been interpreted to occur through a catalytic mechanism.

According to the second passive unfolding or thermodynamic coupling mechanism, partial or complete unfolding of the substrate protein is accomplished by GroEL binding more tightly to a more unfolded form compared to any other co-existing form, thereby driving the equilibrium towards the former. The best evidence for this case was provided by an RNase T1 mutant[108] which can populate two non-native states, and the tighter binding of GroEL to the more unfolded state shifts the equilibrium towards that state, without affecting the observed microscopic conversion rates. The preferential binding of MDH monomers by GroEL has also been observed to shift the equilibrium, causing the disaggregation of low-order, reversible aggregates of MDH[100].

The acceleration of the refolding rates of barstar[89] could also be explained by a thermodynamic coupling mechanism. Unfolded barstar undergoes a rapid hydrophobic collapse[109] to an early intermediate that is spectroscopically indistinguishable from the unfolded state. The observed increase in the refolding rates in the presence of GroEL, could be explained by GroEL binding preferentially to the folding-competent intermediate state. The consequent increase in the concentration of the folding-competent forms, through thermodynamic coupling, was sufficient to account for the increase in the observed refolding rates[89].

A third type of unfolding action has been proposed for the case when the substrate polypeptide is released in the GroEL cavity only upon ATP and GroES binding[110], and is discussed in detail in a later section.

Facilitating domain docking/assembly: The process of binding to the polypeptide surface grants GroEL the first opportunity to influence the refolding process. Besides causing partial unfolding of kinetically trapped intermediates, GroEL has also been implicated in actively assisting the refolding of substrate proteins, as expected from the observation that proteins can fold while bound to GroEL, as described above.

When lysozyme, an *ab* protein, is refolded in the presence of GroEL, there is an acceleration in the rate of refolding[111]. Lysozyme refolding proceeds via three phases – a burst phase change forming a weakly protected *a*-domain, a subsequent fast phase forming a highly protected *a*-domain intermediate, and a final slow phase during which proper docking of the *a* and *b* domains takes place to form the native state. In the presence of GroEL, only the final slow phase is accelerated. Since the populations detected at various times of refolding, using pulsed hydro-

gen–deuterium exchange, were similar in the absence and presence of GroEL, it appears that the refolding mechanism is unaltered in the presence of GroEL. It was therefore suggested that GroEL is able to actively assist in the domain-docking step by reversing or rearranging the non-native contacts. Alternative explanations, such as the thermodynamic coupling mechanism that accounts for a similar effect seen in the case of barstar (see above), cannot, however, be ruled out, especially since the dependence of the rate of refolding of lysozyme on GroEL concentration is well described by a simple binding equation[111], as it is for barstar.

In vivo substrates of GroEL appear to predominantly contain one or more ab domains[95]. It is plausible that for these stringent ab substrates, GroEL can assist in the proper intra-molecular docking of these domains. Such a mechanism may also be applicable in the folding of oligomeric proteins, if GroEL can maintain aggregation-prone monomers in association-competent states, which may be released and properly assembled upon GroES binding.

Channelling the folding reaction along preferred pathways: Until recently, the effect of GroEL binding on the substrate protein was envisaged mainly as minimizing misfolding or rescuing misfolded proteins. In other words, it was believed that GroEL could modulate the energy landscape for folding, by allowing substrate protein molecules to refold without falling into kinetic traps. It has been suggested that the binding and hydrolysis of ATP can provide GroEL with the energy to alter and perhaps smoothen the energy landscape available to the substrate protein for folding[112–114]. It is possible for the energy landscape of the substrate to be modified even without ATP hydrolysis, because binding of GroEL can alter the kinetic barriers that separate various forms present at similar or lower energy levels in the folding funnel. In this manner, GroEL may be able to optimize folding by channelling the folding polypeptide along some preferred routes, when many routes are otherwise available. The feasibility of such a role for GroEL becames apparent from a study on the effect of GroEL on the parallel folding pathways of *E. coli* thioredoxin (Trx)[115].

Equilibrium unfolded Trx consists of at least three unfolded forms, presumably because of the presence of five proline residues that may be in *cis* or *trans* conformations[116–120]. These different unfolded forms progress to the native state via two channels: 10% molecules form the native state (*N*) via a fast channel, while the rest 90% follow a slow channel[120]. In the presence of a saturating concentration of GroEL, the refolding of Trx to *N* proceeds via only the slower of the two kinetic routes (Figure 5)[115]. Only the slowest-folding form, U_M, appears to be capable of folding further, when bound to GroEL. The two faster-folding forms, U_{VR} and U_R, are incapable of folding when bound. As the unfolded forms can inter-convert between themselves, a thermodynamic shift of equilibria replenishes the refolding U_M form, thereby channelling all the molecules via the slow route in the presence of GroEL.

Such a channelling mechanism may also explain the effect of GroEL on the kinetics of staphylococcal nuclease[121]. It might appear counter-intuitive that for both Trx and staphylococcal nuclease, the fast-folding route does not operate in the presence of GroEL. It is, however, plausible that the faster folding forms initially have more exposed hydrophobic area, compared to the slower folding forms; and hence, may bind more tightly to GroEL. Hydrophobic interactions are of primary importance in GroEL-substrate binding[36,80], and GroEL substrates tend to be aggregation-prone. Thus, binding to GroEL and slower folding, appears to be the cost of avoiding aggregation. Several proteins fold slower in the presence of GroEL, for example MBP[53], barnase[88], a-lactalbumin[90] and staphylococcal nuclease[121].

Theoretical studies have also predicted that the rates of substrate refolding can vary over a wide range in the presence of GroEL[112,114]. A theoretical model which considered that chaperone binding is biased, predicted that chaperone binding can either accelerate or slow down refolding, irrespective of which the folding yields are improved[112]. Another study in which a two-dimensional lattice model was used to simulate folding in the presence of chaperonin[114], also predicted that folding can be faster or slower in the presence of chaperonin.

Thioredoxin folding in the absence of GroEL

$$\begin{array}{l} U_R \rightarrow I_R \rightarrow I_S \rightarrow N \\ \Updownarrow \\ U_{VR} \rightarrow I_{VR} \rightarrow N \\ \Updownarrow \\ U_M \rightarrow I_M \rightarrow I_S \rightarrow N \end{array}$$

Thioredoxin folding in the presence of GroEL

$$\begin{array}{l} U_R.E \Leftrightarrow I_R.E \\ \Updownarrow \\ U_{VR}.E \Leftrightarrow I_{VR}.E \\ \Updownarrow \\ U_M.E \Leftrightarrow I_M.E \rightarrow I_S.E \rightarrow N + E \end{array}$$

Figure 5. Reaction scheme for thioredoxin folding in the absence and presence of GroEL[115]. In the absence of GroEL, thioredoxin folds to the native state via three parallel pathways, whereas all the unfolded molecules are channelled along a single kinetic route in the presence of equimolar concentrations of GroEL.

GroES-mediated substrate release in the cavity

It has been suggested that ATP and GroES-mediated conformational changes in GroEL can cause a forced unfolding of the bound substrate, whereby the substrate molecule is 'stretched on the rack'[110], by virtue of the outward opening of the apical domains of GroEL. The initial experimental observation that led to this hypothesis, came from studies with the stringent substrate protein Rubisco[110]. A metastable intermediate of Rubisco showed 12 amide protons that were highly protected from exchange with tritium, both in solution and when bound to GroEL. Upon addition of ATP and GroES, there was a rapid exchange of ten of these protons within 5 s. Earlier experiments monitoring the tryptophan anisotropy of Rubisco[67] in the presence of GroEL–GroES–ATP, had also shown an initial rapid decrease in anisotropy with a $t_{1/2} \sim 1$ s. These observations suggested that Rubisco unfolds upon ATP and GroES binding. It was proposed that the upward and outward movement of the apical domains, associated with GroES and ATP binding[38,51,56] could cause stretching of the bound substrate molecule before its release in the cavity. Such a forced mechanical unfolding mechanism was therefore suggested as the one by which GroEL can unfold misfolded substrate proteins before their release in the cavity, and in this manner can provide them with another chance at refolding.

Alternatively, the observed deprotection of amide hydrogens to exchange might be caused by the GroES-mediated release of the substrate[63]. Such a release would lead to breaking of any hydrogen bonds between the substrate and the binding site, resulting in the observed loss of protection to hydrogen exchange. This explanation was suggested by a study involving the folding of an MDH monomer in the presence of GroES, ATP and a single ring GroEL mutant (SR1-GroEL), which ensured that only the folding active *cis* ring is monitored[122]. Deuterium exchange was studied for a MDH–GroEL binary complex, and 45 protons were observed to be protected from exchange. A native-like tertiary fold appears to be conserved in the MDH–GroEL binary complex. After the addition of GroES and ATP to this binary complex, only 32 protons remained protected: a net deprotection of 13 protons occurred in the time frame of substrate release into the *cis* cavity. When the overall distribution of the deprotected residues was determined using tryptic digestion and mass spectrometry, the deprotection appeared to be broadly distributed. Only three out of the 13 deprotected residues were part of the protected native-like fold. This observation was used to argue against a 'stretching on the rack' mechanism, which should have led to a more widespread disruption of the native-like fold. The breaking of hydrogen bonds with the binding surface was therefore suggested as an alternative explanation for the observed partial deprotection. A partial 'stretching' mechanism cannot, however, be ruled out. Whatever be the mechanism, the substrate protein appears to unfold partially during its release into the cavity. Such partial unfolding upon release might occur because the intermediate is no longer stabilized by the energy of binding to GroEL. Partial unfolding of misfolded or kinetically trapped intermediates provides a mechanism by which they can be rescued, by providing them another chance to refold in an environment of apparent infinite dilution.

Folding in the cis chamber

How structured are the non-native forms that are released after each GroEL cycle? What conformational changes does the substrate protein undergo inside the *cis* folding chamber before it is ejected out? How much native structure actually develops inside the cavity, and what distinguishes the populations that proceed to the native state from those that do not? After the initial release of the substrate in the cavity, is there any subsequent interaction with the cavity walls? Is the GroEL cavity only a passive folding cage, or does it have a more active role in smoothening the energy landscape of folding?

Two recent papers have sought to address some of the above questions by reporting on the folding of stringent substrate proteins like MDH, Rubisco and rhodanese, using two different approaches[122,123]. The first paper has described a study of MDH folding in the presence of GroEL[122], using deuterium exchange, fragmentation into peptides by enzymatic cleavage, followed by mass spectrometry. As described above, deuterium-exchange experiments performed on a binary complex of MDH–SR1 GroEL (single-ring version) showed 45 protected protons that map out a native-like fold. An exchange-pulse was then provided at varying time intervals after the addition of ATP and GroES to this binary complex, to detect the folding of MDH inside this *cis* complex, from which GroES cannot dissociate because of the absence of the *trans* ring. Folding could be monitored in a SR1 GroEL–GroES–ADP complex before and after the first round of ATP hydrolysis. After 1s of ATP and GroES addition, only 32 protons remain protected and there was a net deprotection of 13 protons, as described in an earlier section. No additional protons were protected from exchange even after 9 s, which is the half-life of a wild type *cis* complex. This observation suggested that no substantial structure formation takes place in one round of the wild-type GroEL–GroES cycle. In the confined *cis* cavity**,** the native state of MDH started appearing only after 81 s, and around 74% of the molecules were fully folded by 6000 s. This observation emphasizes that native structure does form in the GroEL cavity. However, the study failed to provide any information on the extent of structure that is gained with each cycle, or the structural differences between a committed form that is no longer recognized by GroEL versus a form that can rebind.

The second study[123] made use of the rapid and very high affinity streptavidin–biotin interaction, to develop a methodology wherein the rapid rebinding of GroES and the non-native substrate protein to GroEL could be prevented after single or multiple rounds of cycling. A single-cysteine-containing mutant of GroEL (EL229C) was constructed in which all the three endogenous cysteines were replaced by alanines, and a cysteine was introduced at the position of Asn229, located close to, but not within the binding site for GroES and the substrate protein. The biotinylated version of EL229C was found to be structurally and functionally equivalent to wild-type GroEL. Streptavidin was shown to bind to biotinylated EL229C at a rate faster than the binding of GroES and the non-native protein, and in doing so it blocked the interaction as well as capped the cavity, so that the substrate remained inside. Even in conditions where spontaneous refolding was non-permissive, a 100% refolding yield of Rubisco as well as rhodanese was obtained in the presence of biotinylated EL229C, GroES and ATP, but the refolding could be blocked by the addition of streptavidin at any point of time. The blocking was complete in the case of Rubisco, showing that the folding of the unfolded form or an intermediate, to the native state of Rubisco, can take place only in the folding cage and not in free solution. In the case of rhodanese, 5–10% of the molecules continued to fold to the native state even after addition of streptavidin, irrespective of the time of its addition. Hence, at least in the case of Rubisco, folding inside the cage at infinite dilution is absolutely essential to prevent off-pathway aggregation.

The other important observation was that the refolding of Rubisco, under conditions where spontaneous refolding can occur, was accelerated in the presence of biotinylated EL229C, GroES and ATP. This acceleration was observed even when blocking by streptavidin was effected after a single round of ATP hydrolysis, so that only one GroEL cycle was allowed. This clearly showed that the mere confinement of the unfolded molecules in the folding cage was sufficient to prevent the partitioning of the polypeptide to a kinetically trapped intermediate, and multiple cycling rounds were not required. Acceleration of the intrinsic folding rates of various polypeptides in the presence of GroEL, GroES and ATP[44,105,124] had been observed previously, but had been explained by the ability of the chaperonin machinery to undergo multiple rounds of binding and partial unfolding of the trapped intermediates, resulting in an increase in the yield as well as rate of formation of the native state.

Folding within the GroEL cavity is distinguished from folding in solution, in that when the polypeptide is confined, it very likely interacts with the hydrophilic inner lining of the cavity. In fact, mere confinement of a folding reaction is expected to make it different from free folding in solution, with theoretical studies predicting a stabilization of secondary structures in such a case[125,126]. Little is known about the nature of the interactions, if any, of the substrate polypeptide with the cavity walls, after being released by the action of GroES and ATP. From fluorescence anisotropy measurements, it appears that the motion of native green fluorescent protein (GFP)[65] inside the SR1–GroEL cavity is significantly hindered. Free GFP has a rotational correlation time of 13.2 ns, as expected from its molecular weight of 27 kDa, and its rotational correlation time is 54 ns while sequestered under GroES in the SR1 cavity. This observation indicated that the tumbling of the native molecule in the cavity is slowed down due to some continuing interaction with the cavity walls. If the native state of a substrate protein like GFP can interact with the internal GroEL cavity surface, presumably so can various intermediates that accumulate during folding. Such interactions are expected to directly influence folding, and it is therefore important that they be understood better. Whether confinement in the GroEL cavity alone is enough to either restrict the number of or alter the nature of folding pathways that a polypeptide can sample, or whether this is dictated by continuing interactions with the cavity walls, still needs to be investigated in detail. Most of the information available on folding in the presence of GroEL is kinetic in nature, and detailed structural studies have been initiated only recently.

Folding in trans

Until recently, it was believed that enclosure of the substrate protein within the *cis* GroEL ring subsequent to ATP and GroES binding is required for productive folding. A recent study has, however, revealed that the GroEL chaperonin machinery can assist in the folding of yeast mitochondrial aconitase, an 82 kDa monomeric enzyme that is too large to be encapsulated in the GroEL cavity[127].

Earlier studies had shown that aconitase failed to fold functionally when imported into mitochondria deficient in either Hsp60 or Hsp10, the mitochondrial homologues of GroEL and GroES respectively[128]. The stringent requirement for GroEL, GroES and ATP was also demonstrated for the reconstitution of aconitase activity *in vitro*[127]. It was, however, observed that GroES failed to encapsulate the GroEL-bound aconitase, as shown by the lack of protection against proteinase K, which is observed for the smaller substrates that are encapsulated in the *cis* chamber. The observation that ATP or its transition analogues could not direct aconitase release in the absence of GroES, suggested that the conformational changes produced by the binding of GroES and ATP in the *trans* (with respect to the substrate) ring might be required for the release. This was also shown by the observation that negligible aconitase activity was recovered in the presence of SR1 or MR1 in the presence of GroES and ATP, where SR1 is a single-ring version of GroEL lacking a *trans* ring and MR1 is a mixed ring version, where one of

the rings is wild type while the other ring is incapable of binding GroES or polypeptide. Using a biotinylated version of GroEL as a trap that could be purified using streptavidin, it was shown that the binding of GroES and ATP in *trans* does direct aconitase release. Using an ATP-hydrolysis-deficient mutant, it was also observed that multiple rounds of binding and release are required for productive folding[127].

The alternative chaperonin cycle for large proteins might be coupled with the *cis* folding of a smaller substrate. As described in an earlier section, GroEL rings can alternate between folding-active *cis* and *trans* chambers[68], such that the release of GroES from the *cis* chamber and its binding in *trans* are simultaneous. It now appears that besides utilizing the *cis* cycles that alternate between the two rings, the chaperonin machinery may also utilize a *trans* cycle for a larger substrate, in combination with an ongoing *cis* cycle, and in this way further increase its efficiency in assisting folding.

Conclusions

It is clear that the working of the GroEL/ES/ATP chaperonin machinery has been characterized in exceptional detail, and the extensive studies have resulted in a better understanding of many of the structural and mechanistic details of its functioning. Recent investigations have focused on delineating the exact role of the machinery in assisting folding, particularly in terms of the effect on the refolding polypeptide. The observation that confinement in the folding cage can radically alter the energy landscape for folding, so that the polypeptide is channelled along a particular folding route to the native state, immediately leads to further questions. What is the structural basis for such a channelling? How does the hydrophillic *cis* cavity affect folding? Are there continuing interactions with the cavity walls? Can these interactions lead to a complete change in the refolding pathway?

The observation that the folding of larger substrates can be assisted by binding to the *trans* cavity also raises questions about how many substrate proteins can be assisted by this mode of chaperonin cycling, and whether the continued association of many large substrate proteins with GroEL *in vivo* is a reflection of a storekeeping function to maintain them in association-competent states[95]. Very little is known about how folding in the confined *cis* cavity is affected by the simultaneous participation of the *trans* ring in assisted folding. The *cis* and *trans* modes of assistance are likely to be very different mechanistically, and it will be interesting to determine which one of the two modes is more prevalent *in vivo*. Now, GroEL substrates can be classified on the basis of their sizes into *cis*-assisted and *trans*-assisted substrates. The role of GroEL in facilitating the assembly of multisubunit proteins remains largely unexplored, as does the possibility that GroEL might assist in the proper insertion of membrane proteins into membranes. The other functions of GroEL that remain poorly understood concern the extent and mechanism of the unfolding action of the chaperonin, and how exactly an initial unfolding action might help the chaperonin machinery in rescuing kinetically-trapped intermediates. These questions are under active investigation.

1. Laskey, R. A., Honda, B. M., Millis, A. D. and Finch J. T., *Nature*, 1978, **275**, 416–420.
2. Barraclough, R. and Ellis, R. J., *Biochim. Biophys. Acta*, 1980, **608**, 19–31.
3. Hemmingsen, S. M. *et al*., *Nature*, 1988, **333**, 330–334.
4. Sternberg, N., *J. Mol. Biol.*, 1973, **76**, 1–23.
5. Hendrix, R. W. and Tsui, L., *Proc. Natl. Acad. Sci. USA*, 1978, **75**, 136–139.
6. Ellis, R. J., *Nature*, 1987, **328**, 378–379.
7. Anfinsen, C. B., *Science*, 1973, **181**, 223–230.
8. Zimmerman, S. B. and Minton, A. P., *Annu. Rev. Biophys. Biomol. Struct.*, 1993, **22**, 27–65.
9. Zimmerman, S. B. and Trach, S. O., *J. Mol. Biol*., 1991, **222**, 599–620.
10. Ellis, J. in *The Chaperonins* (ed. Ellis, R. J.), Academic Press, San Diego, California, 1996, pp. 1–23.
11. Manning-Krieg, U. C., Scherer, P. E. and Schatz, G., *EMBO J*., 1991, **10**, 3273–3280.
12. Heyrovska, N., Frydman, J., Höhfeld, J. and Hartl, F. U., *Biol. Chem*., 1998, **379**, 301–309.
13. Agashe, V. R. and Hartl, F. U., *Semin. Cell Dev. Biol*., 2000, **11**, 15–25.
14. Bukau, B., *Molecular Chaperones and Folding Catalysts*, Haywood Academic Publishers, London, 1999.
15. Leroux, M. R. and Hartl, F. U., *Curr. Biol*., 2000, **10**, R 260–264.
16. Saibil, H., *Curr. Opin. Struct. Biol*., 2000, **10**, 251–258.
17. Gottesman, M. E. and Hendrickson, W. A., *Curr. Opin. Microbiol*., 2000, **3**, 197–202.
18. Kim, S., Willison, K. R. and Horwich, A. L., *Trends Biochem. Sci*., 1994, **19**, 543–548.
19. Willison, K. R. and Horwich, A. L., in *Cell Biology Series: The Chaperonins* (ed. Ellis, R. J.), Academic Press, London, 1996, pp. 107–136.
20. Gutshe, I., Essen, L. O. and Baumeister, W., *J. Mol. Biol*., 1999, **293**, 295–312.
21. Sigler, P. B., Xu, H., Rye, H. S., Burston, S. G., Fenton, F. A. and Horwich, A. L., *Annu. Rev. Biochem*., 1998, **67**, 581–608.
22. Bukau, B. and Horwich, A. L., *Cell*, 1998, **92**, 351–366.
23. Hartl, F. U., *Nature*, 1996, **381**, 571–579.
24. Hunt, J. F., Weaver, A. J., Landry, S. J., Gierasch, L. and Deisenhofer, J., *ibid*, 1996, **379**, 37–45.
25. Mande, S. C., Mehra, V., Bloom, B. R. and Hol, W. G., *Science*, 1996, **271**, 203–207.
26. Ditzel, L. *et al*., *Cell*, 1998, **93**, 125–138.
27. Klumpp, M., Baumeister, W. and Essen, L. O., *ibid*, 1997, **91**, 263–270.
28. Gao, Y., Thomas, J. O., Chow, R. L., Lee, G. H. and Cowan, N. J., *ibid*, 1992, **69**, 1043–1050.
29. Frydman, J., Nimmesgern, E., Erdjument-Bromage, H., Wall, J. S., Tempst, P. and Hartl, F. U., *EMBO J*., 1992, **11**, 4767–4778.
30. Lewis, V. A., Hynes, G. M., Dong, Z., Saibil, H. and Willison, K., *Nature*, 1992, **358**, 249–252.
31. Llorca, O. *et al*., *ibid*, 1999, **402**, 693–696.
32. Fenton, W. A. and Horwich, A. L., *Protein Sci*., 1997, **6**, 743–760.

33. Grantcharova, V., Alm, E. J., Baker, D. and Horwich, A. L., *Curr. Opin. Struct. Biol*., 2001, **11**, 70–82.
34. Saibil, H. R., Horwich, A. L. and Fenton, W. A., *Adv. Protein Chem.*, 2001, **59**, 45–72.
35. Braig, K. *et al*., *Nature*, 1994, **371**, 578–586.
36. Fenton, W. A., Kashi, Y., Furtak, K. and Horwich, A. L., *ibid*, 1994, **371**, 614–619.
37. Langer, T., Pfeifer, G., Martin, J., Baumeister, W. and Hartl, F. U., *EMBO J*., 1992, **11**, 4757–4765.
38. Braig, K., Simon, M., Furaya, F., Hainfeld, J. F. and Horwich, A. L., *Proc. Natl. Acad. Sci. USA*, 1993, **90**, 3978–3982.
39. Chen, S. *et al*., *Nature*, 1994, **371**, 261–264.
40. Thiyagarajan, P., Henderson, S. J. and Joachimiak, A., *Structure*, 1996, **4**, 79–88.
41. Viitanen, P. V., Lubben, T. H., Reed, J., Goloubinoff, P., O'Keefe, D. P. and Lorimer, G. H., *Biochemistry*, 1990, **29**, 5665–5671.
42. Gray, T. E. and Fersht, A. R., *FEBS Lett*., 1991, **292**, 254–258.
43. Bochkareva, E. S., Lissin, N. M., Flynn, G. C., Rothman, J. E. and Girshovich, A. S., *J. Biol. Chem*., 1992, **267**, 6796–6800.
44. Todd, M. J., Viitanen, P. V. and Lorimer, G. H., *Biochemistry*, 1993, **32**, 8660–8567.
45. Badcoe, I. G., Smith, C. J., Wook, S., Hallal, D. J., Holbrook, J. J., Lund, P. and Clark, A. R., *ibid*, 1991, **30**, 9195–9200.
46. Baneyx, F. and Gatenby, A. A., *J. Biol. Chem*., 1992, **267**, 11637–11644.
47. Jackson, G. S., Staniforth, R. A., Halsall, D. J., Atkinson, T., Holbrook, J. J., Clarke, A. R. and Burston, S. G., *Biochemistry*, 1993, **32**, 2554–2563.
48. Todd, M. J., Viitanen, P. V. and Lorimer, G. H., *Science*, 1994, **265**, 659–666.
49. Weissman, J. S., Kashi, Y., Fenton, W. A. and Horwich, A. L., *Cell*, 1994, **78**, 693–702.
50. Boisvert, D. C., Wang, J., Otwinowski, Z., Horwich, A. L. and Sigler, P. B., *Nature Struct. Biol*., 1996, **3**, 170–177.
51. Roseman, A., Chen, S., White, H., Braig, K. and Saibil, H., *Cell*, 1996, **87**, 241–251.
52. Staniforth, R. A., Burston, S. G., Atkinson, T. and Clarke, A. R., *Biochem. J*., 1994, **300**, 651–658.
53. Sparrer, H., Lilie, H. and Buchner, J., *J. Mol. Biol*., 1996, **258**, 74–87.
54. Yifrach, O. and Horovitz, A., *ibid*, 1994, **243**, 397–401.
55. Yifrach, O. and Horovitz, A., *Biochemistry*, 1995, **34**, 5303–5308.
56. White, H. E., Chen, S., Roseman, A. M., Yifrach, O., Horovitz, A. and Saibil, H. R., *Nature Struct. Biol*., 1997, **4**, 690–694.
57. Yifrach, O. and Horovitz, A., *J. Am. Chem. Soc*., 1998, **120**, 13262–13263.
58. Xu, Z. H., Horwich, A. L. and Sigler, P. B., *Nature*, 1997, **388**, 792–798.
59. Landry, S. J., Zeilstra-Ryalls, J., Fayet, O., Georgopoulos, C. and Gierasch, L. M., *ibid*, 1993, **364**, 255–258.
60. Zeilstra-Ryalls, J., Fayet, O. and Georgopoulos, C., *J. Bacteriol*., 1994, **176**, 6558–6565.
61. Ranson, N. A., Farr, G. W., Roseman, A. M., Gowen, B., Fenton, W. A., Horwich, A. L. and Saibil, H. R., *Cell*, 2001, **107**, 869–879.
62. Ma, J. and Karplus, M., *Proc. Natl. Acad. Sci. USA*, 1998, **95**, 8502–8507.
63. Weissman, J. S. *et al*., *Cell*, 1995, **83**, 577–587.
64. Burston, S. G., Ranson, N. A. and Clarke, A. R., *J. Mol. Biol*., 1995, **249**, 138–152.
65. Weissman, J. S., Rye, H. S., Fenton, W. A., Beechem, J. M. and Horwich, A. L., *Cell*, 1996, **84**, 481–490.
66. Hayer-Hartl, M. K., Weber, F. and Hartl, F. U., *EMBO J*., 1996, **15**, 6111–6121.
67. Rye, H. S., Burston, S. G., Fenton, W. A., Beechem, J. M., Xu, Z., Sigler, P. B. and Horwich, A. L., *Nature*, 1997, **388**, 792–798.
68. Rye, H. S., Roseman, A. M., Chen, S., Furtak, K., Fenton, W. A., Saibil, H. R. and Horwich, A. L., *Cell*, 1999, **97**, 325–338.
69. Lorimer, G., *Nature*, 1997, **388**, 720–721.
70. Fersht, A. R. and Daggett, V., *Cell*, 2002, **108**, 573–582.
71. Martin, J. and Hartl, F. U., *Proc. Natl. Acad. Sci. USA*, 1997, **94**, 1107–1112.
72. Van den Berg, B., Ellis, R. J. and Dobson, C. M., *EMBO J*., 1999, **18**, 6927–6933.
73. Martin, J., *Biochemistry*, 2002, **41**, 5050–5055.
74. Mendoza, J. A., Rogers, E., Lorimer, G. H. and Horowitz, P. M., *J. Biol. Chem.*, 1991, **266**, 13044–13049.
75. Buchner, J., Schmidt, M., Fuchs, M., Jaenicke, R., Rudolph, R., Schmid, F. X. and Kiefhaber, T., *Biochemistry*, 1991, **30**, 1586–1591.
76. Martin, J., Horwich, A. L. and Hartl, F. U., *Science*, 1992, **258**, 995–998.
77. Zahn, R., Buckle, A. M., Perrett, S., Johnson, C. M., Corrales, F. J., Golbik, R. and Fersht, A. R., *Proc. Natl. Acad. Sci. USA*, 1996, **93**, 15024–15029.
78. Weber, F., Keppel, F., Georgopoulos, C., Hayer-Hartl, M. K. and Hartl, F. U., *Nature Struct. Biol*., 1998, **5**, 977–985.
79. Buckle, A. M., Zahn, R. and Fersht, A. R., *Proc. Natl. Acad. Sci. USA*, 1997, **94**, 3571–3575.
80. Tanaka, N. and Fersht, A. R., *J. Mol. Biol*., 1999, **292**, 173–180.
81. Chen, S. and Sigler, P. B., *Cell*, 1999, **99**, 757–768.
82. Landry, S. J. and Gierasch, L. M., *Biochemistry*, 1991, **30**, 7359–7362.
83. Landry, S. J., Jordan, R., McMacken, R. and Gierasch, L. M., *Nature*, 1992, **355**, 455–457.
84. Farr, G. W., Furtak, K., Rowland, M. B., Ranson, N. A., Saibil, H. R., Kirchhausen, T. and Horwich, A. L., *Cell*, 2000, **100**, 561–573.
85. Zahn, R. and Pluckthun, A., *J. Mol. Biol*., 1994, **242**, 165–174.
86. Zahn, R., Spitzfaden, C., Ottiger, M., Wuthrich, K. and Pluckthun, A., *Nature*, 1994, **386**, 261–265.
87. Zahn, R., Perrett, S., Stenberg, G. and Fersht, A. R., *Science*, 1996, **271**, 642–645.
88. Gray, T. E. and Fersht, A., R., *J. Mol. Biol*., 1993, **232**, 1197–1207.
89. Bhutani, N. and Udgaonkar, J. B., *ibid*, 2000, **297**, 1037–1044.
90. Katsumata, K., Okazaki, A., Tsurupa, G. P. and Kuwajima, K., *ibid*, 1996, **264**, 643–649.
91. Hayer-Hartl, M. K., Ewbank, J. J., Creighton, T. E. and Hartl, F. U., *EMBO J*., 1994, **13**, 3192–3202.
92. Lilie, H. and Buchner, J., *Proc. Natl. Acad. Sci. USA*, 1995, **92**, 8100–8104.
93. Goldberg, M. S., Zhang, J., Sondek, S., Matthews, C. R., Fox, R. O. and Horwich, A. L., *ibid*, 1997, **94**, 1080–1085.
94. Ewalt, K. L., Hendrick, J. P., Houry, W. A. and Hartl, F. U., *Cell*, 1997, **90**, 491–500.
95. Houry, W. A., Frishman, D., Eckerskorn, C., Lottspeich, F. and Hartl, F. U., *Nature*, 1999, **402**, 147–154.
96. Fischer, G. and Schmid, F. X., *Biochemistry*, 1990, **29**, 2205–2212.
97. Ellis, R. J., *Curr. Biol*., 1994, **4**, 633–635.
98. Panda, M., Gorovits, B. M. and Horowitz, P. M., *J. Biol. Chem*., 2000, **275**, 63–70.
99. Hartman, D. J., Surin, B. P., Dixon, N. E., Hoogenraad, N. J. and Hoj, P. B., *Proc. Natl. Acad. Sci. USA*, 1993, **90**, 2276–2280.
100. Ranson, N. A., Dunster, N. J., Burston, S. G. and Clarke, A. R., *J. Mol. Biol*., 1995, **250**, 581–586.
101. Bhattacharyya, A. M. and Horowitz, P. M., *J. Biol. Chem.*, 2001, **276**, 28739–28743.
102. Bhattacharyya, A. M. and Horowitz, P. M., *Biochemistry*, 2002, **41**, 422–429.
103. Laminet, A. A., Ziegelhoffer, T., Georgopoulos, C. and Pluckthun, A., *EMBO J*., 1990, **9**, 2315–2319.
104. Viitanen, P. V., Donaldson, G. K., Lorimer, G. H., Lubben, T. H. and Gatenby, A. A., *Biochemistry*, 1991, **30**, 9716–9723.

105. Peralta, D., Hartman, D. J., Hoogenraad, N. J. and Hfj, P. B., *FEBS Lett*., 1994, **339**, 45–49.
106. Robinson, C. V. *et al*., *Nature*, 1994, **372**, 646–651.
107. Gross, M., Robinson, C. V., Mayhew, M., Hartl, F. U. and Radford, S. E., *Protein Sci*., 1996, **5**, 2506–2513.
108. Walter, S., Lorimer, G. H. and Schmid, F. X., *Proc. Natl. Acad. Sci. USA*, 1996, **93**, 9425–9430.
109. Agashe, V. R., Shastry, M. C. R. and Udgaonkar, J. B., *Nature*, 1995, **377**, 754–757.
110. Shtilerman, M., Lorimer, G. H. and Englander, S. W., *Science*, 1999, **284**, 822–825.
111. Coyle, J. E., Texter, F. L., Ashcroft, A. E., Masselos, D., Robinson, C. V. and Radford, S. E., *Nature Struct. Biol*., 1999, **6**, 683–690.
112. Gulukota, K. and Wolynes, P. G., *Proc. Natl. Acad. Sci. USA*, 1994, **91**, 9292–9296.
113. Todd, M. J., Lorimer, G. H. and Thirumalai, D., *ibid*, 1996, **93**, 4030–4035.
114. Chan, H. S. and Dill, K. A., *Proteins Struct. Funct. Genet*., 1996, **24**, 345–351.
115. Bhutani, N. and Udgaonkar, J. B., *J. Mol. Biol*., 2001, **314**, 1167–1179.
116. Kelley, R. F. and Stellwagen, E., *Biochemistry*, 1984, **23**, 5095–5120.
117. Kelley, R. F., Wilson, J., Bryant, C. and Stellwagen, E., *ibid*, 1986, **25**, 728–732.
118. Kelley, R. F., Shalongo, W., Jagannadham, M. V. and Stellwagen, E., *ibid*, 1987, **26**, 1406–1411.
119. Kelley, R. F. and Richards, F. M., *ibid*, 1987, **26**, 6765–6774.
120. Georgescu, R. E., Li, J., Goldberg, M. E., Tasayco, M. L. and Chaffotte, A. F., *ibid*, 1998, **37**, 10286–10297.
121. Tsurupa, G. P., Ikura, T., Makio, T. and Kuwajima, K., *J. Mol. Biol*., 1998, **277**, 733–745.
122. Chen, J., Walter, S., Horwich, A. L., Smith, D. L., *Nature Struct. Biol*., 2001, **8**, 721–728.
123. Brinker, A., Pfeifer, G., Kerner, M. J., Naylor, D. J., Hartl, F. U. and Hayer-Hartl, M., *Cell*, 2001, **107**, 223–233.
124. Fedorov, A. N. and Baldwin, T. O., *J. Mol. Biol*., 1997, **268**, 712–723.
125. Yee, D. P., Chan, H. S., Havel, T. F. and Dill, K. A., *ibid*, 1994, **241**, 557–573.
126. Zhou, H. X. and Dill, K. A., *Biochemistry*, 2001, **40**, 11289–11293.
127. Chaudhuri, T. K., Farr, G. W., Fenton, W. A., Rospert, S. and Horwich, A. L., *Cell*, 2001, **107**, 235–246.
128. Dubaquié, Y., Looser, R., Fünfschilling, U., Jenö, P. and Rosepert, S., *EMBO J*., 1998, **17**, 5868–5876.
129. Sayle, R. A. and Milner-White, E. J., *Trends Biochem. Sci*., 1995, **20**, 374.

ACKNOWLEDGEMENTS. This work was funded by the Tata Institute of Fundamental Research, the Department of Science and Technology, Government of India, and the Wellcome Trust. J.B.U. is the recipient of a Swarnajayanti fellowship from the Government of India.

Received 20 June 2002; accepted 9 August 2002

Ex situ conservation of rare and valuable forest tree species through seed-gene bank

Shyam S. Phartyal*, R. C. Thapliyal†, Nico Koedam and Sandrine Godefroid

Laboratory of Plant Science and Nature Management, Vrije Universiteit Brussel, Belgium
†Forest Tree Seed Laboratory, Forest Research Institute, Dehra Dun 248 006, India

There is growing concern throughout the world about the uncontrolled exploitation and depletion of the earth's natural resources, especially affecting the plant biodiversity of tropical forests. The extinction potential of a species is related to the degree of its biological vulnerability and the degree of threat by biotic and abiotic factors. Therefore, the need for conservation is exceptionally high and of paramount importance to preserve this plant heritage for posterity. One of the most effective biological techniques to conserve this biodiversity is the establishment of gene banks, i.e. *ex situ* conservation. Conventional seed storage is believed to be a safe, effective and inexpensive method of *ex situ* conservation of plant genetic resources, which not only maintains its viability but also its vigour without hampering the genetic makeup. The elucidation of various factors that regulate seed viability and vigour in storage is essential. An ideal condition to prolong the seed longevity is mainly dependent on seed moisture content, temperature and type of container used during storage. The optimum stage of seed maturity, seed-lot quality, their processing and harvesting techniques, germination eco-physiology and degree of dormancy too play a crucial role in maintaining seed longevity that need to be considered before large-scale seed storage is initiated. The present review is an attempt to discuss the importance of the aforementioned aspects of forest tree seeds in detail, to conserve their germplasm for *ex situ* conservation through seed-gene bank.

FORESTS, the biological diversity they contain and the ecological function they maintain, are our heritage. In tropical and subtropical regions, complex and species-rich ecosystems are being rapidly destroyed or altered, and in arid and semi-arid regions fragile environments are threatened by the increasing stress from human populations, domestic animals and fluctuating climates[1]. At

*For correspondence. (e-mail: shyamphartyal@yahoo.co.in)