Molecular Chaperones

In most cases proteins fold spontaneously under physiological conditions and do not require any external assistance. This has become evident since the experiments on ribonuclease A conducted by Anfinsen in ‘50s and ‘60s. The list of such “successful” folders has grown since and now includes many two-state proteins catalogued in *Folding & Design* 3, R81. These proteins fold rapidly and reliably to their native states without being trapped in intermediates. Yet protein folding is a very delicate process, which occurs in a narrow window of external conditions. As a result there are many factors, which may potentially impede folding by increasing the roughness of energy landscape of proteins. Due to particular sequence composition or adverse change in external condition (cellular stress), folding funnel may have significant energy ‘ripples” separated by deep minima, which are distinctive from the native state. As a result proteins may be trapped in non-native misfolded structures, which can be quite stable, because they have low energy comparable with the native one. The problem is further aggravated by the fact that in such misfolded conformations proteins often have patches of hydrophobic residues exposed to solvent. Because cellular environment is crowded with other molecules, misfolded proteins may interact with those and irreversibly aggregate into large assemblies. Aggregation often leads to dreadful consequences for cells and organism as a whole. Alzheimer’s, Parkinson’s, “mad cow” diseases are just few examples of deadly protein aggregation.

Cells have developed various strategies to fight irreversible protein aggregation. One of the most interesting is the evolvement of special cell molecules called chaperones, whose primary function is to find and rescue misfolded proteins called substrates. Chaperone molecules are now found in all compartments of a cell and even in extracellular medium (Recommended reading: Walter and Buchner, *Angewandte Chemie Int. Ed.* **41**, 1098 (2002)).

1. Types of chaperones

Hsp70-Hsp40-GrpE machinery: Hsp70 and Hsp40 proteins are involved in blocking aggregation of misfolded proteins by binding to their hydrophobic segments. The best studied system is DnaK Hsp70 protein from *E.coli*, the structure of which is known. Hsp70 consists of ATP-binding N-terminal domain and peptide binding C-terminal domain. ATP hydrolysis switches off and on the binding ability of C-terminal domain. A special hydrophobic groove formed by α-helices and β-strands provides the docking site for hydrophobic segments of misfolded proteins. Due to geometrical consideration the substrate has to be in a relatively open unstructured conformation. Furthermore, Hsp70 works in tandem with Hsp40 co-chaperone.
One of Hsp40 chaperones is DnaJ (75-residue protein), which interacts with DnaK (a Hsp70 chaperone) and assists in capturing substrate proteins.

The cycle of this chaperone system is shown in Fig. 1. The cycle starts, when DnaJ binds a hydrophobic peptide (actually, a segment of misfolded protein) and transports it to the peptide-binding groove of DnaK. DnaJ then promotes hydrolysis of ATP at N-terminal domain of Hsp70 that locks substrate in a groove. Then DnaJ dissociates and the complex waits for the signal to release the substrate. This is done by binding GrpE chaperone to nucleotide-free state of Hsp70. Upon hydrolysis of ATP GrpE and substrate are released. GrpE essentially performs a timing function in the chaperone cycle. It appears that DnaK system with its co-chaperones does not change the structure of misfolded protein nor does it provide a “safe” enclosed environment for a substrate to complete folding (the volume of the groove is simply too small to fit the entire substrate). The main function of Hsp70 machinery may be understood as follows. Proteins in a cell may experience partial unfolding due to variety of factors, such as temperature increase, pH change etc. Some proteins may also fail to reach their native states after synthesis. As result such proteins adopt aggregation–prone states. To prevent this Hsp70 binds to such proteins and act as a general “safe keeper” for misfolded proteins.

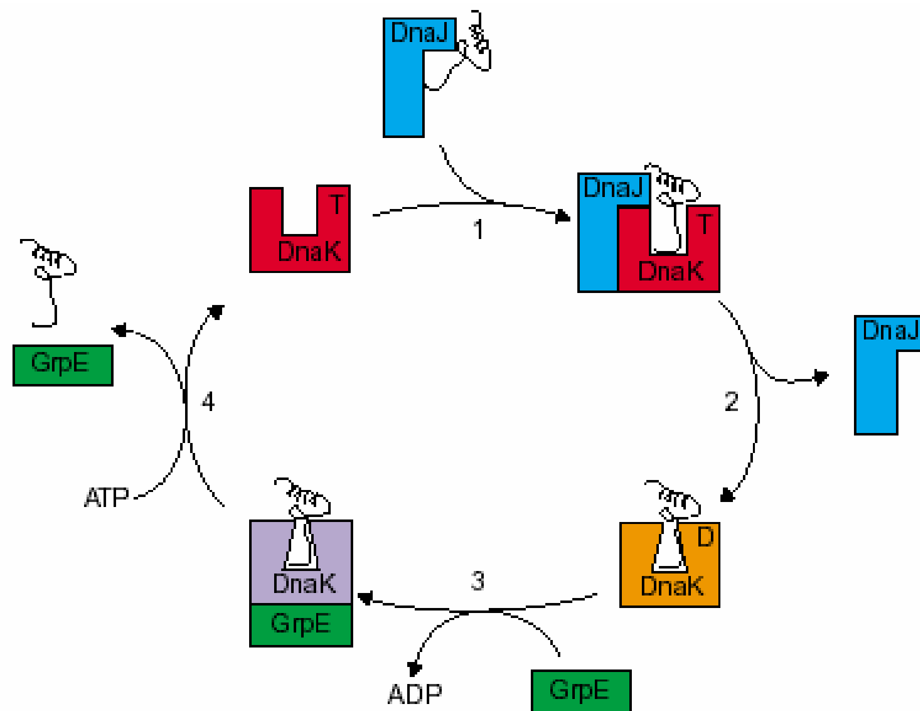

Fig. 1 DnaK chaperone cycle involves complex interaction between several chaperonin partners, Hsp70 DnaK, Hsp40 DnaJ, and GrpE (from *Angewandte Chemie Int. Ed.* **41**, 1098 (2002)).

Small heat-shock proteins: Small heat-shock proteins (sHsp) are found in almost all organisms. They form large oligomeric complexes consisting from 12 to 42 individual units. The oligomer is shaped in the form of a hollow sphere, which has hydrophobic

interior surface and hydrophilic exterior. The example of sHsp structure from one of archaeon microorganism is shown in Fig. 2.

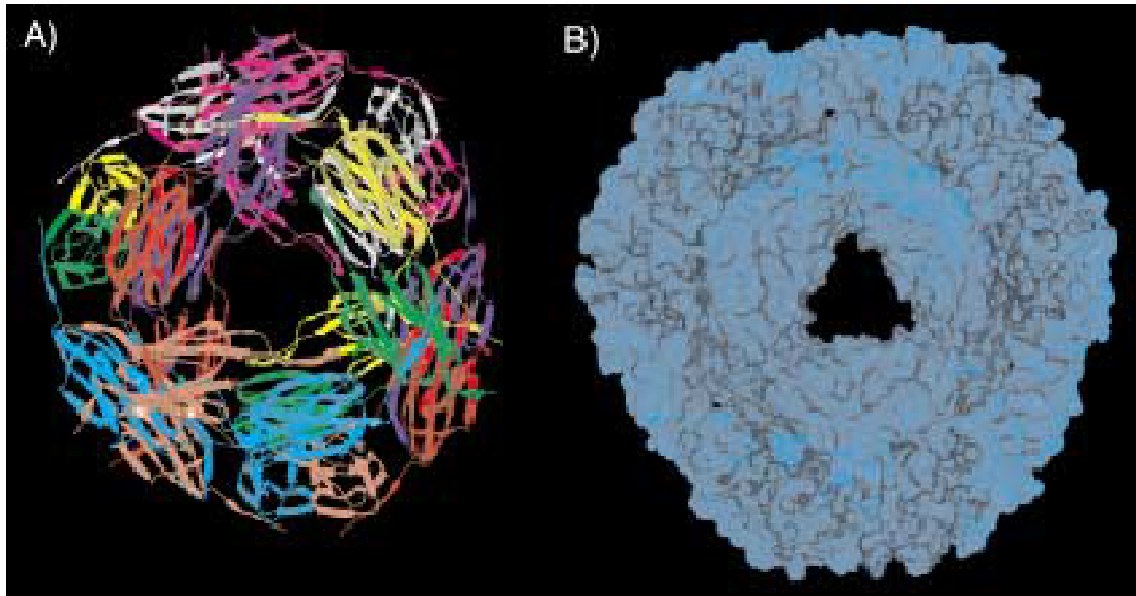

Fig. 2 Structure of a heat-shock protein oligomer appears as a hollow sphere consisting of 24 individual units. The diameters of the sphere and cavities are about 120 and 65 Å, respectively (from *Angewandte Chemie Int. Ed*. **41**, 1098 (2002)).

A sHsp oligomer may bind many substrate non-native proteins of different sizes at once that suggests a non-specific action of this chaperone. Furthermore, the formation of the complexes between sHsp and substrates is a highly cooperative process. Some details of the mechanism of substrate binding are known for the yeast Hsp26 chaperone. At elevated temperatures (> 40°C) the oligomer dissociates into individual units. Their exposed hydrophobic regions bind non-native proteins and form stable complexes with them. No timing mechanism for the release of substrate is discovered, but it was speculated that sHsp proteins are teamed with Hsp70 chaperones. For example, experiments show that addition Hsp70 chaperones and ATPs to sHsp increases the yield of native proteins. Furthermore, the dissociation of sHsp oligomer is reversible and its reassembly is observed as temperature lowered. Thus, it appears that sHsps act as a non-specific "cleaning machine", which controls the concentration of non-native proteins under stress conditions (e.g., "heat shock").

GroEL/GroES chaperone complex: GroEL and GroES chaperones is the best studied system, whose function is relatively well understood.

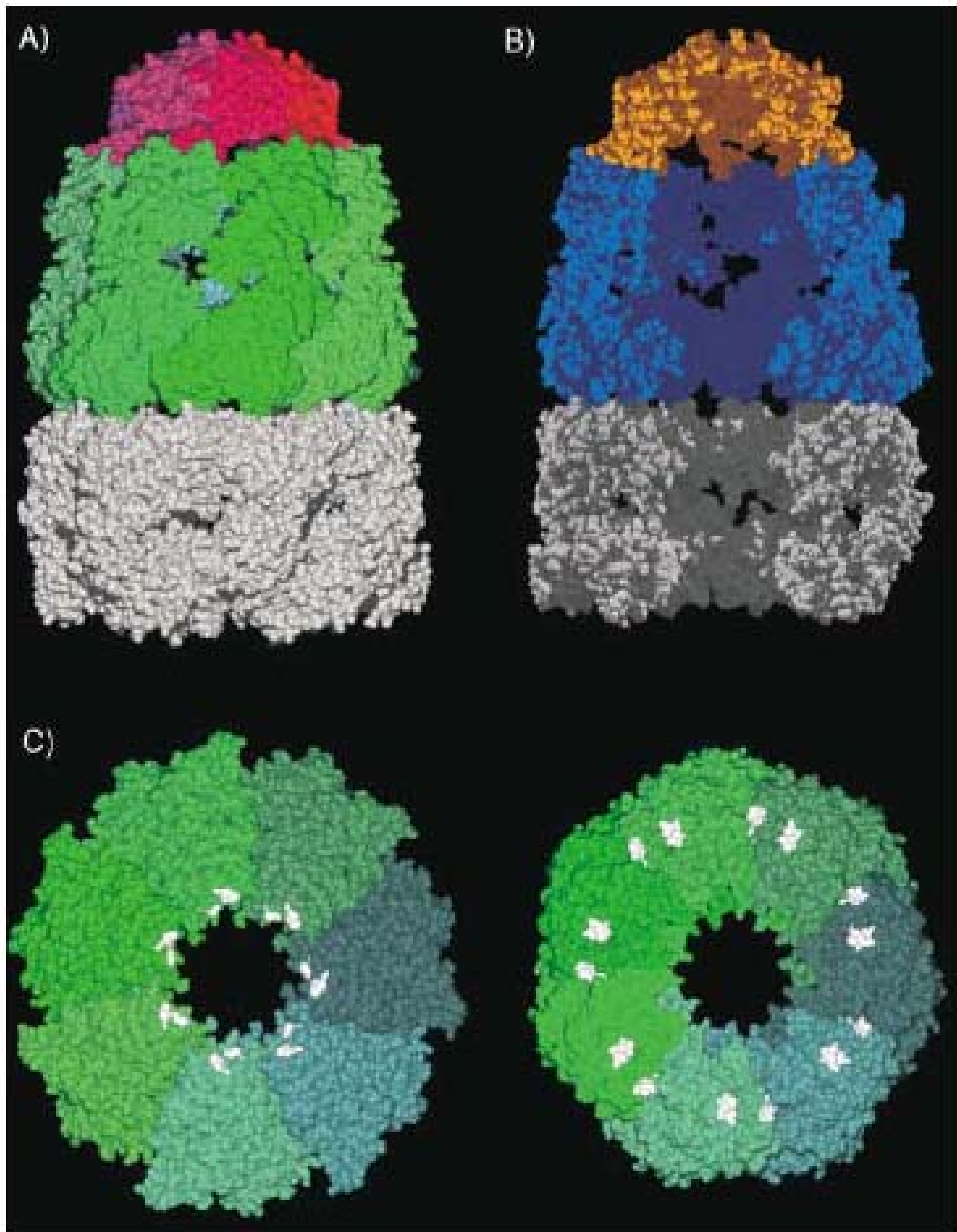

Fig. 3 (A) GroEL/GroES structure from *E.coli*. Two rings (*cis* and *trans*) are shown in green and grey, respectively. GroES is shown in red. (B) and (C) display the cross-sections of GroEL/GroES complex and single GroEL ring, respectively. (C) also shows the structural changes in GroEL ring upon concerted rotation of apical domains. It is estimated that about half of proteins in *E.coli* can interact with GroEL/GroES (*Angewandte Chemie Int. Ed*. **41**, 1098 (2002)).

GroEL chaperone consists of two rings – *cis* (or proximal) upper ring and *trans* (or distal) lower ring. GroES co-chaperone binds to both GroES rings (In Fig. 3 binding to the *cis* ring is shown). Each of the rings consists of seven identical units shown in the lower panel of Fig. 3 in the shades of green and blue. The units are arranged in a circular manner and form a cavity. The cavities of both rings are not connected. GroES upon binding to GroEL serves as "lid", which encapsulates the volume inside the cavity. Each of the GroEL ring units is composed of three domains – *apical* (the upper one), *intermediate*, and lower *equatorial*. The role of *equatorial* domain is to bind and hydrolyze ATP and to transmit signals for concerted function of *cis* and *trans* rings. *Apical* domains contain binding sites for GroES docking and capturing substrate protein.

Function of GroEL/GroES complex is based on a remarkable cycle, which involves spectacular, very coordinated global changes in the structure of chaperone molecules. Consider the cycle for a single GroEL ring and GroES displayed in Fig. 4 (Recommended reading: Thirumalai and Lorimer, *Annual Review of Biophysics and Biomolecular Structure* **30**, 245 (2001)).

A. <u>Capture (T state)</u>: The cavity of GroEL is open and GroES is not yet bound. Exposed hydrophobic sites in *apical* domains of GroEL non-specifically bind a non-native protein. Generally, only few of *apical* domains out of seven are actually involved in substrate capture. Also no sequence specificity is evident in binding substrate proteins; although due to geometrical considerations only relatively small (< 30 kDa) proteins can be captured. It is also important to remember that only non-native proteins can be bound, while native proteins are not "recognized". The volume of the cavity at this point is 85,000 $Å^3$. In T state GroEL has a high affinity toward substrate and the lining of GroEL cavity is largely hydrophobic.

B. <u>Encapsulation (R and R' states):</u> Binding of ATP by *equatorial* domain causes several structural changes. T→R transition is accompanied by a concerted clockwise rotation of all *apical* domains. Consequently, the distance between their binding sites increases from 25 to 33 Å. Because substrate is bound to these sites, their rotation exerts a stretching force on a protein by presumably partially unfolding it. Binding of GroES encapsulates substrate protein and also causes further structural rearrangement of *apical* domains. Each of them rotate upwards by 60° and twist by 90° with respect to *equatorial* domains. These motions effectively displace substrate deeper into the cavity and break its interactions with the binding sites, which are now engaged in the interactions with the GroES molecule. All the conformational movements are highly cooperative and concerted. For example, introduction of a single tether in one of the domains blocks the movements of *all* domains. Cooperativity of domain motions was also revealed in molecular dynamics simulations (Ma, Singler, and Karplus, *Journal of Molecular Biology* **302**, 303 (1998)). Full encapsulation of substrate protein and binding of GroES create R' state, in which the volume of the cavity is about 175,000 $Å^3$ and its radius is about 33 Å. More importantly, the wall of GroEL cavity in R' state turns hydrophilic. There is evidence that a substrate experiences

additional force-induced unfolding during R→R’ transition. In accord with this, tritium exchange experiments indicate unfolding of a substrate protein in the cavity (about 4 sec after the start of the cycle). Furthermore, experiments further show that successful folding of substrate proteins may take place, while they are encapsulated by the GroEL/GroES complex. GroES is attached to GroEL for a total of about 7 sec.

C. ATP hydrolysis (R’’ state) and substrate release: ATP hydrolysis in the *equatorial* domain of *trans* ring serves as a timing event, which sends a signal to *cis* ring to release substrate (R’→R’’ transition). The process is started with the initial release of GroES followed by the release of substrate itself. As GroEL returns to the T state, the cavity shrinks back to the volume of 85,000 Å and its walls become again hydrophobic. These events are accompanied by binding substrate protein to the lower *trans* GroEL ring. Without the lower ring the cycle cannot be completed and substrate would remain encapsulated in the *cis* ring. Therefore, two-ring GroEL system works as a two-stroke engine.

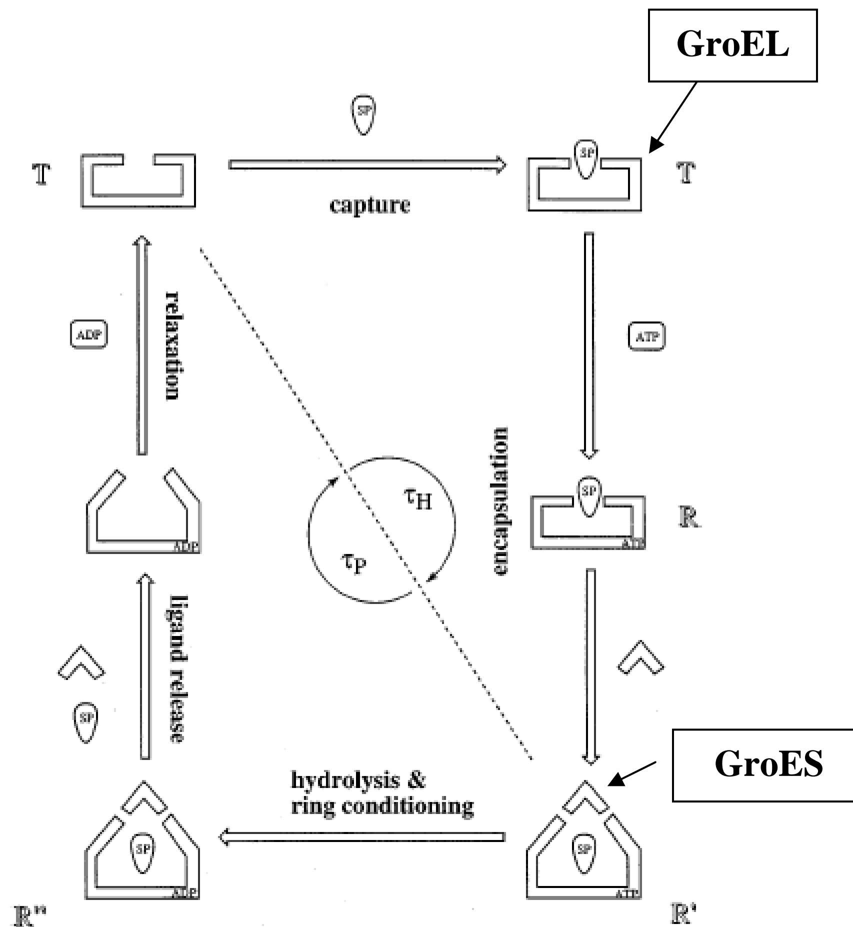

Fig. 4 Illustration of GroEL/GroES function cycle. The total time of the cycle is about 15 sec. SP denotes a substrate protein. (from *Annual Review of Biophysics and Biomolecular Structure* **30**, 245 (2001)).

GroEL/GroES can be viewed as an iterative annealing machine. This chaperone complex actively participates in conformational change in substrate protein as opposed to Hsp70 chaperones, which merely bind to substrate proteins. Let us consider a protein, which occasionally becomes trapped in misfolded states. Assume that the probability of rapid folding without intermediates with the time scale τ_{FAST} is Φ. This implies that the initial pool of unfolded proteins partitions into the fraction Φ, which folds fast, and the faction (1-Φ), which becomes misfolded. Assuming biexponential folding kinetics, we write the fraction of unfolded proteins at the time t is

$$P_u(t) = \Phi e^{-\frac{t}{\tau_{FAST}}} + (1-\Phi)e^{-\frac{t}{\tau_{SLOW}}}, \quad (1)$$

where τ_{SLOW} is the timescale for slow folding through intermediates. Let the time interval during which the wall of the cavity is hydrophilic be τ_P and consider the most relevant case, when $\tau_{FAST} << \tau_P << \tau_{SLOW}$. It follows from Eq. (1) that at $t \approx \tau_P$ $P_u \approx 1-\Phi$ or the yield of the native state $\psi = \Phi$.
After n iterative applications of GroEL/GroES cycles, one can show that the yield increases as

$$\Psi = 1-(1-\Phi)^n \quad (2)$$

If Φ=0.05, then the native yield reaches 0.6 after n=20 iterations of the cycle. From this calculations of iterative action, it is clear that GroEL/GroES machine works only for the proteins, for which 0<Φ<1. Two-state proteins do not require GroEL/GroEL assistance, because for them $\tau_{FAST} \approx \tau_{SLOW} << \tau_P$ and all unfolded proteins rapidly reach native state. It is important to remember that all structural transitions in GroEL/GroEL complex are highly concerted. The origin of this cooperativity lies in the tight packing of GroEL system and steric interactions.

The iterative annealing theory of GroEL/GroES machine is not the only one, seeking to explain chaperone function. “Anfinsen case” theory suggests that even passive sequestering of a non-native protein in inert cavity is enough to ensure successful folding, because protein is spared from aggregation. While confinement of a protein to a restricted volume does increase folding rates, this viewpoint does not explain the effective unfolding of misfolded proteins, which would require additional energy (in the form of ATP).

General outlook on chaperone machinery: Diversity of biological functions requires diversity of protein sequences. Not all sequences are equally well optimized for fast folding, and some proteins, especially large ones, fold slowly through intermediates. Furthermore, proteins are marginally stable in their native states and often their biological functions involve large conformational fluctuations. Since these factors increase, at least, transiently the exposure of hydrophobic residues, they also raise the risk of aggregation. To cope with this problem evolution forced cell to develop molecular chaperones, which

assist proteins to reach their native states. Chaperones not only assist monomeric folding, but also increase solubility of already formed protein aggregates, perform polypeptide degradation (such as Hsp100 chaperones), or participate in trafficking proteins in a safe "protected" form (SecB chaperones). All these functions require a source of energy, which is provided by the hydrolysis of ATP. It is also important to note that chaperones frequently form a network. For example, not only Hsp70 works in tandem with Hsp40 and GrpE, but there is a link between this chaperone and Hsp100 and GroEL as well. Therefore, chaperones are vital for living organisms and mutations affecting their function can be frequently fatal.