Chapter 1

Heat Shock response and its regulation
1.1 Heat shock response (HSR)

When cells are shifted from normal growth temperature to high temperature, the newly induced proteins due to this stress are commonly known as heat shock proteins (HSPs). The heat-shock response (HSR) was first identified by Ritossa (Ritossa F., 1963), who observed that exposure to heat leads to transient changes in the puffing pattern of salivary chromosomes and induction of specific proteins in Drosophila; Tissieres et al. (Tissieres et al., 1974) demonstrated that these changes result in the transient induction of several proteins. Initially, function of heat shock protein was not clear; however, experiments in several organisms revealed that many hspS were chaperones that promote protein folding (Pelham HR, 1986; Gaitanaris et al., 1990; Skowyra et al., 1990; Beckmann et al., 1990). These studies not only suggested that a major function of the HSR is to maintain the protein folding process of the cell, but also indicated that some of these chaperones, such as Hsp90 and Hsp70, are present in all organisms (Bardwell & Craig, 1984; 1987). Thus, both the HSR and some hspS are universally conserved among organisms.

1.1.1 Induction of heat shock proteins

A complex control system regulates expression of heat shock genes, guaranteed that they can counter even to small changes in cellular environment. In *Escherichia coli*, heat shock gene regulation is interceded by rpoH, which encodes the $\sigma^{32}$ transcription factor (Connolly et al., 1999; Arsene et al., 2000). $\sigma$ factors are bacterial-specific initiation factors that recruit RNA polymerase to particular classes of promoters (Gruber & Gross, 2003). $\sigma^{32}$ is regulated at multiple levels and use inputs both directly from some stresses and also from the protein folding state of the cell. Increases in temperature directly increase translation of rpoH mRNA by destabilizing RNA structural element overlapping the translation start point (Morita et al., 1999a; 1999b). The protein folding state of the cell regulates both the degradation and activity of $\sigma^{32}$. In response to an increased need for protein folding agents (e.g., immediately after temperature upshift), $\sigma^{32}$ is transiently
stabilized (Straus et al., 1987). In response to the decreased requirement for protein folding agents (e.g., during the recovery phase following temperature upshift or immediately after temperature downshift), the activity of $\sigma^{32}$ decreases (Straus et al., 1989). Additionally, cells use these feedback systems to constantly monitor their folding status during growth under steady-state conditions so that the cellular folding environment remains optimal.

### 1.1.2 Regulation of heat shock response

The expression of heat-shock genes in *E.coli* is positively controlled by the product of the *rpoH* gene, which encodes the sigma-32 transcription factor (Connolly et al., 1999; Arsene et al., 2000). The DnaK chaperone machine, consisting of DnaK (Hsp70 homolog), DnaJ (Hsp40 homolog), and GrpE (nucleotide exchange factor), is implicated in regulation of $\sigma^{32}$ activity and $\sigma^{32}$ degradation (Tilly et al., 1983; 1989). Sigma factors are bacterial-specific initiation factors that recruit RNA polymerase to particular classes of promoters (Gruber & Gross, 2003). Stress condition, such as a sudden temperature upshift, induces transient heat-shock gene expression until the cells have adapted to the applied stress. Sigma32 is necessary for induced expression as well as uninduced basal expression of the heat-shock genes. The cellular concentration of sigma32 is very low under steady state conditions (10-30 copies /cell) at $30^0C$ and is limiting for heat-shock gene transcription.

![Fig. 1.1: Schematic representation of $\sigma^{32}$ post-translational modulation](image)

The heat-shock response is induced as a consequence of a rapid increase in sigma32 levels and stimulation of its activity. The shut-off of the response occurs as a consequence of a decline in level of sigma32 and inhibition of its activity. Under heat stress, the increase in sigma 32 is mediated by increases in the translation, stabilization
and activation of sigma32 by sequestration of the DnaK chaperone and its DnaJ co-chaperone, which form a complex with sigma32.

Translation of rpoH mRNA is reserved at steady-state conditions and is rapidly de-repressed upon shift of the cells from 30 to 42°C. At 2-4 min after temperature upshift, a 12-fold higher translation level compared to the pre-heat shock condition is reached. Then, translation becomes increasingly repressed during the shut-off phase of the heat-shock response to reach a new steady-state level. Translation of sigma32 is mediated through distinct mechanisms involving three cis-acting elements of the rpoH coding sequence, termed regions A, B and C. The region A is a positive regulatory element comprising the initiation codon and the downstream 20 nucleotides. This region is complementary to the 3' region of the 16S rRNA and acts as a translational enhancer (Morita et al, 1999). The region B is a negative regulatory element located within nucleotides 110 and 247. Extensive base pairing between regions A and B has been substantiated by mutational analysis (Morita et al, 1999). The formation of this secondary structure is hypothesized to mediate repression of rpoH translation at steady-state conditions, by preventing translation initiation, because of the unaccessibility of the Shine-Dalgarno sequence and initiation codon. The thermal induction of translation results from partial melting of the mRNA secondary structure due to the increased temperature, enhancing the entry of ribosome and translational initiation (Morita et al, 1999). Region C is a negative regulatory element located within nucleotides 364-433 of the rpoH coding sequence, involved in repression of rpoH translation during the shut-off phase of the heat-shock response. It is suggested that region C acts at the protein level to mediate translational repression of rpoH (Nagai et al., 1994).

During steady-state growth, sigma32 has an extremely short half-life or less than 1.0 min (Nagai et al., 1994; Straus et al., 1987). Upon temperature upshift from 30 to 42°C, it becomes transiently stabilized at least 8-fold until the beginning of the shut-off phase of the heat-shock response (Straus et al., 1987). One protease responsible for its degradation is the ATP-dependent metalloprotease FtsH, which is an integral cytoplasmic
membrane protein with the active site located in a cytosolic domain (Tatsuta et al., 1998; Tomoyasu et al., 1998). The housekeeping sigma factor sigma70 is more abundant at 30°C and therefore prevents the binding of sigma32 to the RNA polymerase. The interaction of sigma32 with the RNA polymerase core enzyme prevents degradation of sigma32 which indicates that FtsH and RNA polymerase compete for binding to sigma32 (Tomoyasu et al., 1998).

The DnaK chaperone system has role in controlling sigma32 activity. DnaK and DnaJ are also involved in the FtsH-dependent degradation of sigma32. It is supposed that the chaperones affect the stability of sigma32 by preventing its binding to the RNA polymerase, and therefore increasing sigma32 degradation (Blaszczyk et al., 1999). The homeostatic regulation model, also called the "unfolded protein titration model" (Straus et al., 1990; Craig & Gross, 1991; Bukau, 1993; Arsene et al., 2000), for regulation of sigma32 proposes that unfolded proteins and sigma32 compete for binding to DnaK chaperone system; DnaK binding, not only makes sigma32 inactive for transcription, but also facilitates sigma32 degradation by the protease FtsH. Therefore, when unfolded proteins are lesser in amount relative to DnaK, the inactive DnaK-sigma32 complex predominates, sigma32 is rapidly degraded, and heat shock gene expression is low. However, when unfolded proteins are higher in amount relative to DnaK, DnaK is titrated away from sigma32; the active, stable, chaperone-free state of sigma32 predominates and heat shock genes are induced. Thus the induction of heat-shock response after stress treatment relies on the sequestration of the DnaK chaperone system through binding to misfolded proteins accumulating during stress. According to this model, a common feature of inducers of the E. coli heat-shock response is their potential to generate misfolded proteins. Additionally, cells use these feedback systems to constantly monitor their folding status during growth under steady-state conditions so that the cellular folding environment remains optimal. A few recent findings indicate that in addition to the role of DnaK/J/E chaperone system, GroEL/S system may also have important role in regulation of the stability and activity of sigma32.
An important strategy to control expression of environmentally responsive genes is to regulate the stability of key regulators by proteolysis. At least two *E. coli* sigma factors are prone to proteolytic attack during exponential growth. The starvation sigma factor $\sigma^5$ (RpoS) is degraded by the ClpXP protease, whereas the heat-shock sigma factor $\sigma^{32}$ (RpoH) is degraded by FtsH and other proteases (Kanemori et al., 1994; Herman et al., 1995; Tomoyasu et al., 1995). The half-lives of both sigma factors are very short (in the range of 1–2 min) because of their rapid turnover during normal growth in exponential phase. A Complex regulatory circuit prevents the degradation of these transcription factors under stress conditions. As a consequence, the stability of $\sigma^{32}$ and $\sigma^5$ is increased $\approx$ 10-fold during heat stress and in stationary phase, respectively (Straus et al., 1987; Tilly et al., 1989; Lange & Hengge-Aronis, 1994).

*E. coli* FtsH is a 70kDa membrane-anchored, ATP-dependent metalloprotease and is a member of the highly conserved and widely distributed AAA (ATPases associated with diverse cellular activities) family (Tomoyasu et al., 1995; Tomoyasu et al., 1993). Apart from controlling the cellular level of $\sigma^{32}$, FtsH degrades a number of functionally and structurally unrelated proteins. Degradation of $\sigma^{32}$ by FtsH is a very complicated process. To control the heat-shock response in *E. coli*, FtsH protease works with the DnaK machinery in close co-operation. Binding of the DnaK and DnaJ chaperones to $\sigma^{32}$ influences both the activity and stability of the sigma factor (Tomoyasu et al., 1998). The sequestration of the sigma factor probably prevents its association with RNA polymerase and thus impedes transcription of heat-shock genes (Blaszczak et al., 1995; Gamer et al., 1996). Here, the DnaK machinery may be considered as being functionally equivalent to an antisigma factor. As RNA polymerase-bound $\sigma^{32}$ is protected from degradation by FtsH (Tomoyasu et al., 1998; Blaszczak et al., 1999), active displacement by the DnaK system was proposed as a prerequisite for sigma factor turnover. It is not clear at present condition whether DnaK also actively presents $\sigma^{32}$ to the protease. The equilibrium between RNA polymerase-bound sigma factor and DnaK provides the molecular basis for a postulated homeostatic heat-shock control mechanism in *E. coli* (Tomoyasu et al., 1998; Gamer et al., 1996; Craig & Gross, 1991). After implication of heat shock, the DnaK
machinery interacts with accumulated unfolded proteins leading to the release of $\sigma^{32}$. The free sigma factor can now interact with RNA polymerase core enzyme to initiate transcription of heat-shock genes. One of these heat-shock genes is $\text{ftsH}$ itself, which participates in returning $\sigma^{32}$ to low levels in the shut-off phase when DnaK and DnaJ become available again to recruit the sigma factor.

### 1.1.3 Important *Escherichia coli* heat shock proteins

Two types of heat shock proteins are induced during heat shock or at stress condition one is called chaperones and another is called protease. Cellular chaperones, not only hold or help denatured proteins to for their proper conformation at stress conditions, they also help in maturation of newly translated proteins at steady state condition. They are Hsp90, Hsp70, hsp60 and small heat shock proteins. Lon and ClpP play the main role of protease.

<table>
<thead>
<tr>
<th>Molecular Weight (in kD)</th>
<th>Eukaryotic Homolog</th>
<th>Prokaryotic Homolog</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>Hsp104, Hsp110</td>
<td>ClpB, ClpA, ClpX</td>
</tr>
<tr>
<td>90</td>
<td>Hsp90</td>
<td>HtpG</td>
</tr>
<tr>
<td>70</td>
<td>Hsp70, Hsp72</td>
<td>DnaK</td>
</tr>
<tr>
<td>60</td>
<td>Hsp60</td>
<td>GroEL</td>
</tr>
<tr>
<td>40</td>
<td>Hsp40</td>
<td>DnaJ</td>
</tr>
<tr>
<td>20</td>
<td>Hsp27</td>
<td>GrpE</td>
</tr>
<tr>
<td>10</td>
<td>Hsp10</td>
<td>GroES</td>
</tr>
</tbody>
</table>
Chapter 1

*Escherichia coli* heat shock Chaperons:

**Hsp70:**

The term chaperone was first used to describe an activity associated with nucleoplasm in *Xenopus* oocytes (Laskey *et al*., 1978). The term has been stretched to include the biochemical activities displayed by a broad collection of proteins that aid nascent polypeptide folding process, transit across cellular and organelle membranes, disassembly of macromolecular complexes or aggregates, targeting for proteolysis, and quality control, as well as the regulation of conformational changes that affect biological functions, such as signaling (Ellis, 1987; Young *et al*., 2004; Bukau *et al*., 2006). In-depth studies with the canonical *Escherichia coli* DnaK-DnaJ-GrpE Hsp70 machine initially furnished a wealth of insights into chaperone function. Crystal structure of DnaK substrate binding domain shown in Fig. 1.2.

![Fig 1.2 Crystal structure of substrate binding domain of DnaK](image)

In *E. coli*, the best studied Hsp70 is termed DnaK and its canonical partner JDP is termed DnaJ. The nucleotide exchange factor (NEF) that functions with DnaK is termed GrpE. The two DnaK partner proteins, GrpE and DnaJ, are referred to as co-chaperones, and the DnaK chaperone machine is usually abbreviated as KJE. The discovery of the chaperone genes, which encode DnaJ, DnaK and GrpE in the mid-1970s, is traced to the
first described missense mutant alleles that define each member, namely *dnaK756*, *dnaJ259* and *grpE280*. Only decades later did DNA sequence analysis pinpoint the corresponding amino acid changes (*dnaK756*: G32D, G468D and G455D; *dnaJ259*: H33Q; and *grpE280*: G122D) and their precise biochemical functions on what became known as a molecular chaperone machine. In the early studies that uncovered KJE, genetic selections were applied that specifically searched for host genes that, due to mutation blocked the propagation of various bacteriophages. Just after the initial description of these mutations, it was shown that both *dnaJ* and *dnaK* had a role in both bacteriophage and host DNA replication, resulting in the renaming of the original *groPC756* and *groPC259* alleles as *dnaK756* and *dnaJ259* respectively (Saito & Uchida, 1977; Saito et al., 1978). This finding was later extended to include *grpE* (Saito et al., 1978), but the gene name was not changed.

Members of the Hsp70 family are characterized by an N-terminal ATPase domain, the substrate binding domain, and a short C-terminal domain of largely unknown function that can interact with various partner proteins (for example Hop/p60/Sti1 in eukaryotes) to modulate chaperone function (Young et al., 2004; Mayer & Bukau, 2005). The open/closed state of the Hsp70 substrate binding domain, mediated by the repositioning of an alpha-helical lid over the substrate binding pocket, controlled by the nucleotide occupancy and status in the ATPase domain (Schmid et al., 1994; Palleros et al., 1993; Zhu et al., 1996; Mayer & Bukau, 2005). While the ATP-bound Hsp70 exhibits a low affinity and fast exchange rate for its substrates and the ADP-bound form is characterized by high affinity and low exchange rates. To prevent unrestrained cycling upon fake interaction with random substrates, Hsp70 is regulated by co-chaperones. The dual-purpose JDP co-chaperones help stimulate ATP hydrolysis and substrate delivery to Hsp70, both the events leading to an ADP-bound Hsp70 in complex with its substrate (Liberek et al., 1991; Karzai & McMacken, 1996; Laufen et al., 1999). Though the substrate specification of Hsp70 and JDP may be similar in some cases, the attachment of Hsp70 and JDP to distinct segments in polypeptide substrates that are close in space tends to enhance the rates of JDP stimulation of the Hsp70 chaperone cycle (Laufen et al., 1999; Han & Christen, 2003).
Chapter 1

There is general agreement that the Hsp70 chaperone machine binds to short extended hydrophobic polypeptide sequences present in its substrate proteins. Such regions of extended conformation can occur either during de novo protein synthesis or subsequent translocation through membranes, or following chemical or thermal stresses. In addition, that exposed peptide segments can be found within native macromolecular structure, resulting in Hsp70 orchestrating either oligomeric disassembly or macromolecular partitioning. The dissociation of ADP, followed by the subsequent binding of ATP that triggers substrate release which is mediated by another type of cochaperone GrpE. The apparent requirement for a NEF in some Hsp70s and not others relies on specific structural elements within the ATPase domain which control nucleotide dissociation rates (Brehmer et al., 2001; Kluck et al., 2002). Measurement of protein abundance of the Hsp70 chaperone machine members in E. coli under different physiological states suggested a wide range of regulation. It turned out that K-J-E is the only heat shock-induced machine. DnaK is present at approx. $10^4$ copies per cell under non-stressed conditions at 37°C, and this number doubles following a shift to 42°C. DnaJ levels mirror those of DnaK, but are approximately 20-fold less. The biochemical properties of the Hsp70 machines of E. coli depict functional divergence at several levels. As example, only DnaK seems to require a NEF for its chaperone cycle. In vitro biophysical measurements revealed that the GrpE protein undergoes a reversible conformational change at high temperatures, which changes its nucleotide exchange activities towards DnaK. Such abridged activity of GrpE will result in a transient increase in the fraction of ADP-bound DnaK-substrate, which may be advantageous for the prevention of protein aggregation during heat shock (Siegenthaler & Christen, 2005; 2006).

Hsp90:

Hsp70, i.e. DnaK chaperone family play the main role of chaperone in E. coli. On the basis of in vitro studies and homology considerations, some additional E. coli proteins have been proposed to participate in cellular protein folding. These include the Clp ATPases (ClpA, ClpB, ClpX, and ClpY), the Hsp90 homologue HtpG (Squires and Squires,
1992; Jakob & Buchner, 1994). Although ClpA, ClpB and HtpG are dispensable up to 45°C, their absence affects bacterial growth at very high temperatures (Bardwell and Craig, 1988; Squires et al., 1991; Katayama et al., 1988; Thomas and Baneyx, 1998) and deletions in *clpB* or *htpG* exacerbate the growth defects of cells containing mutations in the DnaK-DnaJ-GrpE system at 42°C (Thomas & Baneyx, 1998). Thus, these proteins may only play a significant role in folding events under stressful conditions and/or when the Hsp70 team becomes unable to handle this task efficiently. Crystal structure of nucleotide binding domain of ClpB shown in fig. 1.3.

Like *Saccharomyces cerevisiae* (Glover and Lindquist, 1998; Parsell *et al.*, 1994) and *T. thermophilus* homologues (Motohashi *et al.*, 1999), *E. coli* ClpB collaborates with the DnaK-DnaJ-GrpE system to reactivate large insoluble aggregates *in vitro* (Zolkiewski, 1999; Goloubinoff *et al.*, 1999) and thermally aggregated proteins *in vivo* (Mogk *et al.*, 1999). It is, however, unclear if, in addition to functioning as a disaggregase, ClpB also plays a vital role in the folding of newly synthesized proteins. Hsp90, the most abundant proteins in the mammalian cytosol and it is part of a super chaperone system (Buchner, 1999). The *E. coli* Hsp90 homologue, suppresses the aggregation of citrate synthase *in vitro* (Jakob *et al.*, 1995), but very little is known about its *in vivo* property, with the exception of a possible involvement in secretion (Ueguchi & Ito, 1992; Shirai *et al.*, 1996). Finally, although the ClpA ATPase is not a heat shock protein, it provides substrate

![Fig. 1.3 Crystal structure of nucleotide binding domain of ClpB](image-url)
specificity to the heat-shock protease ClpP which converts plasmid P1 RepA dimers into active monomers and protects certain proteins from heat inactivation in vitro (Gottesman et al., 1998; Wickner et al., 1994). So it is possible that ClpA also participates in cellular protein folding. ClpB and HtpG participate in cellular protein folding in heat-shocked E. coli cells and the main function of these chaperones in de novo folding is to expand the ability of the DnaK-DnaJ-GrpE team to interact with newly synthesized polypeptides. The chances that ClpB actively unfolds intermediates on a dead-end branch of their folding pathway, whether HtpG acts as a holder chaperone that transiently maintains a subset of newly synthesized proteins in a conformation that is accessible to DnaK (Freeman & Morimoto, 1996), could explain the relative importance of ClpB relative to HtpG in de novo folding events.

Hsp60:

GroEL chaperon is called hsp60 and GroES act as a cochaperone of it. Historically, the GroE proteins of Escherichia coli were the first chaperone proteins to be studied on a molecular level (Georgopoulos et al., 1973; Lorimer, 2001). In the early 1970s, temperature-sensitive, mutant E. coli strains were isolated that were unable to support the growth of bacteriophage λ (Georgopoulos et al., 1973). Further analysis revealed that apparently two host proteins, GroEL (57kDa) and GroES (10 kDa), were required for the correct assembly of the phage capsids. Both proteins were found to be essential for the growth of E. coli (Fayet et al., 1989). At that time, the cellular function of GroEL was unknown. The turning point came in the late 1980s, when George Lorimer and co-workers began to investigate whether the GroE proteins could assist in the biogenesis of Rubisco expressed in E. coli (Goloubinoff et al., 1989a). They observed that in wild-type cells the formation of active Rubisco was severely compromised. Upon over expression of both GroEL and GroES, however, active Rubisco was produced. It was suggested that GroE’s primary function was to prevent the aggregation of Rubisco during its folding (Goloubinoff et al., 1989b), as could be demonstrated later (Buchner et al., 1991; Höll-Neugebauer et al., 1991). In the past decade, a wealth of both biochemical and structural
data on the GroE chaperone has been accumulated, making it the most thoroughly investigated chaperone system so far (Fenton & Horwich, 1997; Sigler et al., 1998; Thirumalai & Lorimer, 2001; Grallert & Buchner, 2001; Walter & Buchner, 2002). According to the model emerging from this data, GroE-assisted protein folding is a three-step process. An aggregation-prone folding intermediate is first captured by GroEL and thereby becomes protected from aggregation. Upon binding of ATP and GroES to the GroEL/polypeptide complex, the polypeptide is ejected into a closed compartment formed by the GroE chaperone, where folding is initiated. After hydrolysis of ATP, both GroES and the polypeptide are released. While this basic mechanism of GroE action is now widely accepted, there are still a number of details that remain controversial and require further experimental investigation.

The mechanism of GroE-mediated protein folding is intimately related with the oligomeric structure of the chaperone (Weber et al., 1998). The GroEL molecule is a complex assembly comprising 14 identical 57-kDa subunits. The transitions between the different functional states of the chaperone are triggered by a set of domain movements which in turn are controlled by the binding of ATP and the co-chaperone GroES. A more detailed picture became available with the X-ray structure of GroEL (Braig et al., 1994), which confirmed that the GroEL molecule resembles a barrel with dimensions of 137 Å (diameter) and 146 Å (height). Its 14 subunits are arranged in two rings stacked back to back [fig. 1.4 (A)]. The two rings enclose two separate cavities (45 Å wide) that serve as folding compartments for polypeptide substrates [fig. 1.4 (B)]. Each GroEL subunit can be dissected into three distinct domains (Braig et al., 1994). The equatorial domains (residues 6–133 plus 409–523) constitute the central part of the cylinder and consist mainly of alpha-helices. They serve as the foundations of the GroEL oligomer, since they mediate all inter-ring contacts, and most of the intra-ring contacts. They also contain the binding pockets for ATP [fig. 1.4 (C)], which face towards the inside of the central cavity.
Fig. 1.4: Structure of the GroEL chaperone of *E. coli*. (A) Side view of the GroEL tetradecamer. Subunits comprising the top ring are shown in color; subunits of the bottom ring are shown in gray. Each subunit can be dissected into three domains: apical (orange), intermediate (yellow) and equatorial (red). (B) Top view of the GroEL tetradecamer. The seven subunits of the ring are shown in shades of green. For one subunit, the apical and the intermediate domains are highlighted in orange and yellow, respectively. (C) Ribbon representation of a GroEL subunit. The equatorial domain (red) consists almost exclusively of helices and contains the nucleotide binding site, which is occupied by ATP (blue). The intermediate domain (yellow) serves as a molecular hinge that connects the equatorial domain with the apical domain (orange). Binding of GroES and polypeptides occurs in a hydrophobic groove formed by the two helices (white) facing the central cavity.

The co-chaperone, GroES is a dome-shaped heptamer with diameter of 75 Å and a height of 30 Å (Hunt et al., 1996). It consists almost exclusively of beta sheets. Residues 16–33 form the so-called mobile loops, flexible extensions that dangle from the GroES molecule like the tentacles of a jellyfish (Landry et al., 1993; Richardson et al., 2001). Binding of GroES occurs at the apical domains of the GroEL tetradecamer and requires that the nucleotide binding sites of the respective GroEL ring are occupied with either ATP or ADP (Saibil et al., 1991; Chandrasekhar et al., 1986; Jackson et al., 1993). Upon association, the mobile loops of GroES bind to the hydrophobic peptide binding groove of GroEL and become immobilized (Xu et al., 1997; Landry et al., 1993). Because of the common seven-fold symmetry of both proteins, binding is thought to be highly cooperative. Upon binding of its co-chaperone, the GroEL molecule undergoes major structural rearrangements that are central to its functional cycle (Xu et al., 1997; Roseman et al., 1996). First, the apical domains of the *cis* ring, i.e. the ring to which GroES binds, swing upward by ~60° and rotate outward by ~90°. As a result, the diameter of the
central cavity almost doubles, and its volume increases from 85,000 Å³ to 175,000 Å³. Second, the hydrophobic residues, which form the peptide binding site of GroEL, are moved away from the cavity surface and become buried within the wall. Thus, the surface of the cis cavity becomes largely hydrophilic. Third, GroES now blocks the exit of the cavity. As a result, the cis cavity is converted from an acceptor site for hydrophobic polypeptides into a closed microenvironment for protein folding.

**Small Heat Shock Proteins:**

Small heat-shock proteins (sHsps) are one of the families of molecular chaperones preventing the irreversible aggregation and assisting in the refolding of denatured proteins. The sHsps have been found in all organisms tested so far and are characterized by their low molecular mass (12–30 kDa), oligomeric structure and the presence of a conserved ‘α-crystallin’ domain (Jacob & Buchner, 1994). The sequence of this domain is highly similar to the vertebrate protein α-crystallin which prevents the formation of cataracts associated with protein aggregation in the eye lens (MacRea, 2000). According to the present model, ATP-independent sHsps bind and stabilize denatured proteins in a folding competent state, preventing their irreversible aggregation. The stabilized proteins are subsequently delivered to the Hsp70–Hsp40 chaperone system for ATP-dependent refolding (Ehrsperger et al., 1997; Veinger et al., 1998; Lee & Vierling, 2000). *Escherichia coli* has two members of the sHsps family, IbpA and IbpB proteins, sharing 50% amino acid homology and identified for the first time as proteins bound to heterologous proteins forming inclusion bodies in *E. coli* cells (Allen et al., 1992). Previously study found that IbpA and IbpB are major components of the S fraction which contains endogenous, heat-denatured *E. coli* proteins, separated from the membranes and soluble proteins by sucrose gradient ultracentrifugation (Kucharczyk et al., 1991; Laskowska et al., 1996a). The proteins IbpA and IbpB were so named because they were first identified as components of inclusion bodies. It is known that over expression of a gene product, say a recombinant protein, may result aggregation of that protein in cell cytosol as inclusion
bodies. Sometimes secreted proteins in *E. coli* can come out of the cell as inclusion bodies, whose formation depends on the protein synthesis rate and the amino acid sequence of the signal peptide. Inclusion bodies are such aggregates that are not homogeneous in protein composition. The S fraction appears in *E. coli* wild-type (WT) cells 15 min after the temperature change from 30 °C to 45 °C. During subsequent growth at 37 °C, the S fraction disappears due to refolding and degradation of the aggregates by heat-shock proteins. Removal of the S fraction is inhibited or retarded by mutations affecting the heat-shock proteases and chaperones Lon, HtrA, ClpP, ClpX, ClpA, ClpB, DnaK, DnaJ and GroEL/ES (Laskowska et al., 1996b; Kedzierska et al., 1999). The S fraction is also stable in *rpoH* cells, which lack active σ^{32}, the transcriptional activator of heat-shock genes (Kucharczyk et al., 1991).

It was found that IbpB reduced thermal aggregation of model substrates: citrate synthase and alcohol dehydrogenase (Shearstone & Baneyx, 1998) and formed a stable and soluble complex with MDH which was refolded subsequently by the DnaK/DnaJ and GroEL/GroES systems (Veinger et al., 1998). Compared to the considerable amount of data derived from studies on the *in vitro* systems, there is not much information available on the *in vivo* function of the IbpA/B proteins. (Kitagawa et al. 2000) found that *E. coli* strains overproducing IbpA or IbpB acquired higher levels of resistance to heat and oxidative stresses and accumulated lower amounts of aggregated proteins after exposure to 50 °C than the WT strain. Deletion of the *ibpA/B* operon affected neither the cell viability nor the intracellular protein aggregation at high temperature. The only phenotypic change observed in a Δ*ibpA/B* mutant was a slightly slower growth at 46 °C (Thomas & Baneyx, 1998).

It was demonstrated that IbpB assembles *in vitro* into oligomeric structures of heterogeneous size. The basic 600 kDa oligomers are roughly spherical and interact to form larger complexes (Shearstone & Baneyx, 1998). It is not clear either whether IbpA and IbpB are functionally equivalent. The fact that overproduced IbpA in Δ*ibpA/B* cells was partitioned between soluble and insoluble cellular fractions whereas overproduced
IbpB was located mostly as a soluble protein (Shearstone & Baneyx, 1998) suggests that *E. coli* sHsp may have a different affinity for unfolded proteins.

**Escherichia coli** heat shock proteases:

Many HSPs are molecular chaperones or ATP-dependent proteases. The DnaK-DnaJ-GrpE and GroEL-GroES chaperone teams play major roles in protein folding and assembly in many cellular processes, perhaps including protein turnover (Hartl, 1996). The HSPs Lon/La, ClpP, ClpX, and FtsH are all ATP dependent proteases or their subunits: Lon is a single polypeptide that forms homotetramers, whereas Clp is a large complex that consists of a catalytic subunit (ClpP) and a regulatory subunit (ClpA or ClpX) which confers substrate specificity (Goldberg, 1992; Gottesman & Maurizi, 1992; Maurizi, 1992, Gottesman, 1996). FtsH is a membrane-bound metalloprotease essential for growth and can degrade σ32 in vivo and in vitro (Herman, 1995; Tomoyasu, 1995). Degradation of cellular proteins is mostly energy dependent, since it is inhibited by 90% upon depletion of ATP. Lon and ClpAP/Ti proteases play major roles in proteolysis: disruption of both lon and clpA (or clpP) inhibits degradation of cellular proteins by 70 to 80% (Gottesman & Maurizi, 1992; Maurizi, 1992). Crystal structure of FtsH functional domain shown in fig. 1.4.

Fig 1.4: Crystal Structure of FtsH

The Lon protease family (Rawlings et al., 2004) is conserved in the prokaryotes and in eukaryotic organelles such as mitochondria and peroxisomes, is the most common family of ATP-dependent proteases. Prokaryotic Lons are the key enzymes responsible for intracellular proteolysis which contribute to protein quality and cellular homeostasis by
eliminating mutant and abnormal proteins and participating in rapid turnover of select short-lived regulatory proteins (Goldberg, 1992; Gottesman & Maurizi, 1992; Gottesman, 1996; Wickner et al., 1999). A number of direct and indirect observations indicate that bacterial Lons form rings consisting of six subunits (Lee et al., 2004; Park et al., 2006). Lons are divided into two subfamilies, LonB and LonA, based on differences in the number of domains and characteristic sequences within the domains (Rotanova et al., 2004, 2003). Both the subfamilies contain the ATPase (A) domains that include typical AAA⁺ modules, as well as the proteolytic (P) domains, but whereas LonA enzymes contain a large N-terminal (N) domain, the LonB enzymes have no N-domain but have a large transmembrane domain insertion within the AAA⁺ module between the Walker motifs A and B. It had been suggested that Lons are serine proteases (Chung & Goldberg, 1981; Waxman & Goldberg, 1982; Goldberg et al., 1994), although amino acid sequences of LonA and LonB P-domains show no homology with serine proteases containing the classical catalytic triad Ser-His-Asp (Amerik et al., 1991; Goldberg et al., 1994; Rotanova et al., 2004), or, to any other proteases. Active-site inhibitors of classical serine proteases, viz. sulfonyl fluorides, chloromethyl ketones, etc. are poor inhibitors of Lon, and none of these shown to modify the active site residues. Comparative analysis of primary structures of the Lon pool and site-directed mutagenesis of the full-length Escherichia coli Lon (EcLon) (Rotanova et al., 2004, 2003), along with the recently determined structure of the EcLon P-domain (Botos et al., 2004), established that the active sites of Lon proteases have a Ser-Lys catalytic dyad (Ser679 and Lys722 in the EcLon numbering). The crystal structures of the EcLonA P-domains have provided data favoring hexameric ring structures for these enzymes. In the EcLonA P-domain crystals (Botos et al., 2004), the asymmetric unit contained six molecules related by an approximate sixfold non-crystallographic symmetry (NCS) axis, revealing a hexamer in which individual monomers form a ring. Viewed from the side, the ring is dome-shaped, with a diameter of ~100 Å at the base and ~50 Å at the top. A solvent-accessible central pore ~32 Å long runs through the hexamer. It has a diameter of ~18 Å at the entrance from the proximal side and widens slightly toward the distal end. The pore entrance from the proximal surface (16–
18 Å) is significantly larger in size than the entry channel in *E. coli* ClpP (∼10 Å) (Wang et al., 1997) which could be expected to accommodate two folded α-helices or β-strands or unstructured loops with two polypeptide chains. A number of negatively charged and polar residues are located toward the distal part of the pore and making the distal end rather negatively charged, while the proximal part of the pore has several positively charged residues, which might be served as gating function for substrate entry.

Two types of Clp protease exist in *E. coli*, the first with ClpP subunits as the proteolytic core, and the second with ClpQ (or HslV). The ClpP proteolytic subunit associates with either of two ATPases, ClpA or ClpX (Katayama et al., 1988; Gottesman et al., 1993), whereas the ClpQ proteolytic subunit associates with the ClpY (or HslU) ATPase (Kessel et al., 1995; Rohrwild et al., 1996). ClpP and ClpQ are unrelated proteins in both amino acid structure and mode of proteolytic action (Maurizi et al., 1990a; Larsen & Finley, 1997). ClpP is classified as a serine-type protease (Maurizi et al., 1990b) and has been identified not only in a wide range of bacteria but also in plants and mammals. The *clpP* gene from *E. coli* encodes a predicted polypeptide of 207 amino acids, of which the first 14 residues are rapidly and autolytically cleaved *in vivo* to yield a mature protein of 21.5 kDa (Maurizi et al., 1990a). The reason for this self-processing by ClpP has proved elusive, although it may serve as a specific regulatory mechanism, as only the mature protein is proteolytically active (Thompson & Maurizi, 1994). ClpP has active sites characteristic of serine-type proteases, but alone is incapable of degrading proteins longer than six amino acids (Thompson et al., 1994). For full proteolytic activity, ClpP must associate with one of two related ATPase subunits, ClpA or ClpX, both of which are members of the Hsp100/Clp family of molecular chaperones (Schirmer et al., 1996). Of the two ATPase subunits in *E. coli*, ClpA was the first identified and, thus, the ClpAP protease is so far the best characterized. The proteolytic complex is composed of two central heptameric rings of ClpP flanked by one or two hexameric rings of ClpA, with ClpA having about twice the affinity for binding to ClpP as ClpX (Grimaud et al., 1998). The optimal ratio of ClpA hexamer to ClpP tetradecamer is 2:1 (Maurizi et al., 1998). The ClpAP protease degrades large proteins down to short peptides of seven to 10 amino acids.
acids without any apparent sequence specificity (Thompson & Maurizi, 1994; Thompson et al., 1994). The process requires both Mg$^{2+}$ and ATP hydrolysis; other nucleoside triphosphates cannot substitute for ATP (Katayama et al., 1988). ATP is bound at two distinct domains on the ClpA protein, of which ATP hydrolysis at one is sufficient for protein degradation by the ClpAP complex; ATP bound to the other domain is primarily required for ClpA oligomerization (Singh & Maurizi, 1994; Seol et al., 1995). ClpA also binds the protein targeted for degradation and, in steps requiring ATP, apparently unfolds the protein to increase its accessibility to the ClpP proteolytic active sites (Wang et al., 1997). One example of a native substrate for ClpAP is the MazE protein, which functions as an endogenous regulator of programmed cell death in *E. coli*. 