Introduction

I. Tuberculosis
II. Latency in Tuberculosis
III. Stress Response in Pathogenic Bacteria
IV. Mycobacterial Stress Response
V. Protein Folding and the role of Molecular Chaperones
Tuberculosis

I. Tuberculosis

TB: A Threat to Mankind

Tuberculosis (abbreviated as TB for tubercle bacillus or Tuberculosis) is a common and deadly infectious disease caused by mycobacteria. The bacillus causing tuberculosis, Mycobacterium tuberculosis, was identified and described on March 24, 1882 by Dr. Robert Koch. In the past, tuberculosis was known by several names. TB was called consumption, because it seemed to consume people from within, with a bloody cough, fever, pallor, and long relentless wasting. Other names included phthisis (Greek for consumption) and phthisis pulmonalis; scrofula (in adults), affecting the lymphatic system and resulting in swollen neck glands; and white plague, because sufferers appear markedly pale. Tuberculosis most commonly attacks the lungs (as pulmonary TB) but can also affect the other parts of the body, like the central nervous system, the lymphatic system, the circulatory system, bones, joints and even the skin.

Over one-third of the world's population has been exposed to the TB bacterium, and new infections occur every second (Fig. 1). Not everyone infected develops the full-blown disease; asymptomatic, latent TB infection is most common. However, one in ten latent infections progresses to active TB disease, which, if left untreated, kills more than half of its victims. The mortality and morbidity statistics include more than 14.6 million chronic active TB cases, 8.9 million new cases, and 1.6 million deaths, mostly in developing countries. In addition, a rising number of people in the developed world are contracting tuberculosis because their immune systems are compromised by immunosuppressive drugs, substance abuse, or HIV/AIDS. The World Health
Organization declared TB a global health emergency in 1993, and the “Stop TB Partnership” developed a Global Plan to Stop Tuberculosis aiming to save 14 million lives between 2006 and 2015.

The Bacterium
The primary cause of TB, *Mycobacterium tuberculosis*, is an aerobic bacterium that divides every 16 to 20 hours, an extremely slow rate compared with other bacteria, which usually divide in less than an hour (For example, one of the fastest-growing bacteria is a strain of *E. coli* that can divide roughly every 20 minutes). Since M. TB has a cell wall but lacks a phospholipid outer membrane, it is classified as a Gram-positive bacterium. However, if a Gram stain is performed, M. TB either stains very weakly Gram-positive or does not retain dye due to the high lipid & mycolic acid content of its cell wall. M. TB is a small rod-like bacillus (Fig. 2) that can withstand weak disinfectants and survive in a dry state for weeks. In nature, the bacterium can grow only within the cells of a host organism, but *M. tuberculosis* can be cultured in vitro.

Using certain histological techniques on expectorate samples from phlegm (also called sputum), scientists can identify mycobacteria under a regular microscope. Since M. TB retains certain stains after being treated with acidic solution, it is classified as an acid-fast bacillus (AFB). The most common staining technique, the Ziehl-Neelsen stain, dyes AFBs a bright red that stands out clearly against a blue background. Other ways to visualize AFBs include an auramine-rhodamine stain and fluorescent microscopy.

There is much less information on how the bacterium survives and grows in the lung (Fig. 3). It is known that infected macrophages in the lung, through the production of chemokines, attract inactivated monocytes, lymphocytes, and neutrophils, none of which kill the bacteria very efficiently. Then, granulomatous focal lesions composed of these macrophages and lymphocytes begin to form (Fig. 3). This process is generally effective in containing the spread of the bacteria.

A granuloma is actually a non-specific type of inflammatory response which may be triggered by diverse antigenic agents or by inert foreign materials. The antigenic triggering agents cause activation of the cellular immune system (T lymphocytes
Studies of tissues from infected asymptomatic individuals have shown viable *M. tuberculosis* in primary lesions in the lung, and also in lesion-free areas of lung and lymph nodes. Although primary lesions can occur anywhere in the lung, post-primary disease most commonly develops in the apical regions.

Granulomatous reactions to inert foreign bodies are generally considered to be non-immunologic in origin. The tuberculous granuloma (caseating tubercle) is a central caseous necrosis bordered by giant multinucleated cells (Langhans giant cell), and surrounded by epithelioid cells aggregates, lymphocytes and fibroblasts. Granulomatous tubercules tend to confluence. Multinucleated giant cell (Langhans giant cell) are 50-100 micrometres, consisting of numerous small nuclei (over 20) disposed at the

---

*Fig. 3. A) Studies of tissues from infected asymptomatic individuals have shown viable *M. tuberculosis* in primary lesions in the lung, and also in lesion-free areas of lung and lymph nodes. Although primary lesions can occur anywhere in the lung, post-primary disease most commonly develops in the apical regions. B) Acid fast stain showing numerous bacilli in giant cell. C) Necrotising Granuloma of M. TB.*
Tuberculosis

The periphery of the cell (crown or horseshoe). It results when activated macrophages merge. Epithelioid cells are activated macrophages resembling epithelial cells. They are elongated, with finely granular, pale eosinophilic (pink) cytoplasm and central, ovoid nucleus. They have indistinct shape contour and form aggregates. At the periphery are the lymphocytes (T cells) and rare plasma cells and fibroblasts. Caseous necrosis is a central area, amorphous, finely granular, and eosinophilic. If recent, it may contain nuclear fragments. The caseum is the result of the accumulated destruction of giant cells and epithelioid cells. With increase in cellular immunity, macrophages loaded with bacilli are killed, and this results in the formation of the caseous center of the granuloma, surrounded by a cellular zone of fibroblasts, lymphocytes, and blood-derived monocytes.

Treatment of TB

Although M. tuberculosis bacilli are postulated to be unable to multiply within the caseous tissue due to its acidic pH, the low availability of oxygen, and the presence of toxic fatty acids, some organisms may remain dormant but alive for decades.

Treatment for TB uses antibiotics to kill the bacteria. The two antibiotics most commonly used are rifampicin and isoniazid. However, instead of the short course of antibiotics typically used to cure other bacterial infections, TB requires much longer periods of treatment (around 6 to 12 months) to entirely eliminate mycobacteria from the body. Latent TB treatment usually uses a single antibiotic, while active TB disease is best treated with combinations of several antibiotics, to reduce the risk of the bacteria developing antibiotic resistance. People with these latent infections are treated to prevent them from progressing to active TB disease later in life. However, treatment using Rifampin and Pyrazinamide is not risk-free. The Centers for Disease Control and Prevention (CDC) notified healthcare professionals of revised recommendations against the use of rifampin plus pyrazinamide for treatment of latent tuberculosis infection, due to high rates of hospitalization and death from liver injury associated with the combined use of these drugs.

Drug resistant tuberculosis is transmitted in the same way as regular TB. Primary resistance occurs in persons who are infected with a resistant strain of TB. A patient with fully-susceptible TB develops secondary resistance (acquired resistance) during TB therapy because of inadequate treatment, not taking the prescribed regimen appropriately, or using low quality medication. Drug-resistant TB is a public health issue in many developing countries, as treatment is longer and requires more expensive drugs. Multi-drug
resistant TB (MDR-TB) is defined as resistance to the two most effective first line TB drugs: rifampicin and isoniazid. Extensively drug-resistant TB (XDR-TB) is also resistant to three or more of the six classes of second-line drugs\textsuperscript{12}. Thus, due to these difficulties in treatment, especially its resistance to drugs, the disease becomes even more deadly to mankind.

References:
II. Latency in Tuberculosis

- Dormancy or latency
- Oxygen Tension
- Studies in human
- Models of persistence
  - Cornell model
  - Bacteriological models
  - Wayne model
- Persistence mutants

Dormancy or Latency

The remarkable success of M. TB as a pathogen is closely associated with its ability to persist in humans for extended periods without causing disease. It is estimated that one-third of the world population, or about 1.9 billion people, harbors latent M. TB infections, which can last for years or decades. These individuals carry a 2-23% lifetime risk of developing reactivation TB.

When M. tuberculosis infects the host, an obvious outcome is the infection to get transformed to acute TB. However, in some cases it may so happen that the disease does not take shape because the bacteria are contained inside the granuloma. This enormous reservoir of latent disease greatly complicates efforts at tuberculosis control.

Latent TB is a clinical syndrome that occurs after an individual has been exposed to M. tuberculosis, the infection has been established, and an immune response has been generated to control the pathogen and force it into a quiescent state. In contrast to patients with active TB, individuals with latent TB do not transmit the disease and do not pose a public health risk. The diagnosis of latent TB is determined by the absence of clinical disease and either a positive tuberculin skin test using intradermal purified protein derivative (PPD) inoculation or a chest radiograph that demonstrates scars indicative of old TB.

Despite significant effort in recent years, progress has been slow in understanding the natural history of latent tuberculosis and reactivation. Important unresolved questions include the metabolic state of bacilli during latency, the role that metabolically dampened M. TB may play in lengthening the time necessary for effective chemotherapy, and the nature of the bacterial genetic program and host responses that underlie long-term persistence. The failure of bacteria to increase in numbers during latency, the lack of clinical sequel and the enhanced resistance of latent TB to chemotherapy argue that the bacilli may be metabolically dormant. However, there is no direct evidence from the genome or the
Latency in Tuberculosis

M. tuberculosis enters the host within inhaled droplets. Three outcomes are possible. (i) Immediate eradication of M. tuberculosis by the pulmonary immune system. This alternative is rare to absent. (ii) Infection transforms into tuberculosis. This frequently occurs in immunodeficient individuals, with the notable example of HIV infection increasing the risk of developing tuberculosis 800-fold. (iii) Infection does not transform into disease because M. tuberculosis is contained inside granulomas. In the diseased individual, M. tuberculosis is no longer contained because caseation of the lesion results in dissemination and transmission of M. tuberculosis. After inhalation, M. tuberculosis is engulfed by alveolar macrophages and DCs. In draining lymph nodes, these cells present mycobacterial antigens to different T cell populations. Antigen presentation probably involves cross-priming, allowing transfer of mycobacterial antigens from infected macrophages to dendritic cells. Antigen-specific CD4+ T cells, CD8+ T cells, T cells and CD1-restricted T cells participate in protection. Most importantly, macrophages are activated by IFN-γ and TNF-α. In addition, T cells may kill mycobacteria present in macrophages by means of perforin and granulysin.

Laboratory studies have shown that M. TB is capable of a truly dormant, spore-like state. In addition, chemotherapy can reduce the rate of reactivation in persons with latent TB, and immunotherapy can protect against reactivation in mice. It is hard to see how these therapies would have any effect in the complete absence of mycobacterial metabolism.

Oxygen Tension

Oxygen tension is one factor frequently associated with the establishment and maintenance of latent TB. In vivo, the
Latency in Tuberculosis

Fig. 5. Outcomes associated with exposure to *Mycobacterium tuberculosis*. Following close contact, 30% of individuals become infected, with about 40% of these individuals developing primary active tuberculosis (TB) and 60% developing latent infection. Among this 60%, only 2-23% of immunocompetent patients with latent TB will reactivate at a later date, while patients with HIV develop reactivation TB at an alarmingly faster rate of ~5-10% per year.

number of bacilli in a lesion generally correlates well with the degree of oxygenation, suggesting that O₂ supply may limit M. TB growth during infections. Also, inhibition of M. TB growth in *vivo* is associated with the formation of hard, fibrous, hypoxic granulomas. Replication of M. TB requires oxygen, but bacilli show a remarkable ability to survive for years without oxygen in *vitro* (Corper and Cohn model described later). M. TB maintained under anaerobic conditions in *vitro* lose their acid-fast character and human studies have associated latent TB with tubercle bacilli that were no longer acid fast.

Studies in human
Little is known regarding the location and state of *M. tuberculosis* in latently infected individuals. Following the inhalation of virulent micro-organisms, the bacilli enter and proliferate intracellularly within pulmonary alveolar macrophages. The intracellular survival of the pathogen is mediated by its ability to disrupt phagosome acidification within infected macrophages and prevent phagosome-lysosome fusion. During this stage of the initial infection process, which generally occurs within a period of 2-4 weeks, host immunity may be inadequate, and tubercle bacilli may escape from infected macrophages, causing a transient bacteremia with propagation of viable bacilli throughout the body. Despite the events in the alveoli and the transient bacteremia, patients are asymptomatic during the initial infection and control stages. In most immunocompetent individuals, T cells and macrophages are recruited, a secondary immune response is mounted, and the infection is controlled. Several features may alter the course of latent TB and influence the development of active disease. HIV infection, malnutrition, drug-use, cancer, diabetes, chronic renal insufficiency and immunosuppressive drug therapy are each believed to increase the risk of reactivation TB. Epidemiological studies of latent tuberculosis patients show that HIV infection increases the risk of reactivation from 2-23% per lifetime to an alarming rate of 5-10% per year (Fig. 5).

![Diagram](image-url)
Latency in Tuberculosis

Various models of persistence

Cornell model

An animal model for latent TB was developed at Cornell University, NY, USA by Walsh McDermott and colleagues in the 1950s and 1960s (Fig. 6). In this model outbred mice were infected intravenously with \( \sim 10^5 \) colony-forming units of the H37Rv strain of \( M. \) tuberculosis. These were then immediately treated with the antimycobacterial drugs INH and pyrazinamide (PZA)\(^{26,27} \) for a period of 12 weeks. The mice showed no evidence of cultivable tubercle bacilli for a period of 4-6 week after the withdrawal of INH and PZA. McDermott referred this phenomenon as the ‘sterile’ state. Nevertheless, \( \sim 12 \) weeks after INH and PZA treatment was withdrawn, one-third of the mice developed full-blown active TB, with nearly two-thirds displaying the disease after 24 weeks.

The Cornell team performed an extensive evaluation of the ‘sterile’-state mice available one month after the cessation of drug therapy. Homogenates of mouse lung, liver and spleen were sterile by bacteriological culture techniques. To their surprise, the tubercle bacilli could not be grown even when the team passed the ‘sterile’ state mouse tissues into additional mice or guinea pigs – a very sensitive test for infectious bacilli. In addition, smears from mice in the sterile state revealed that only very rare acid-fast forms could be detected after careful examination of thousands of high power fields. Cornell team estimated that there were\(<10^4\) acid-fast micro-organisms per ml of tissue. This inference was based upon limits of detection of their microscopic surveys. Of course, if the bacteria were in a non-acid fast form, a higher tissue burden of bacilli would be possible in the sterile-state tissues.

![Latency in Tuberculosis Diagram](image)

Fig. 6. Schematic diagram of the Cornell model of dormant tuberculosis (TB). a) Outbred mice are inoculated with \( \sim 10^5 \) colony forming units of H37Rv strain of \( M. \) tuberculosis. b) After treatment with the antimycobacterial drugs isoniazid (INH) and pyrazinamide (PZA) for a period of 12 weeks, there is no evidence of disease. c) 4 weeks after withdrawing INH and PZA, the mice appear to be well, and the TB infection cannot be demonstrated by any known cultivation methods. This is termed as sterile state of Latent TB. d) Eight weeks later however, \( \sim 35\% \) of the mice develop culture-positive active TB.

Bacteriological models

The ability of the tubercle bacillus to survive environmental hardship in culture was well documented by Corper and Cohn in a study published in 1933\(^{26} \). In 1920, they inoculated a culture bottle with...
Latency in Tuberculosis

*M. tuberculosis.* This was then incubated at 37 °C for a period of 12 years, after which it was re-examined. The old culture yielded viable micro-organisms in the sediment of the bottle.

**Wayne model**

Lawrence Wayne conducted pioneering studies of the bacteriological persistence of *M. tuberculosis.* In the Wayne model, cultures of *M. tuberculosis* were subjected to gradual oxygen withdrawal by incubation in sealed containers with controlled agitation. Growth under such conditions leads to two non-replicating persistent (NRP) states. The first being a microaerophilic state associated with induction of glycine dehydrogenase activity (NRP1), and the later an anaerobic state (NRP2) in which glycine dehydrogenase activity declines, and alterations in drug susceptibility are observed. Wayne also demonstrated that the reintroduction of oxygen to an aerobic bacilli leads to a resumption of growth, which progresses in a synchronous manner. At the initiation of replication, RNA is produced but there is a lag period before the production of DNA occurs. One interpretation of this finding is that the persistent form of the bacilli grown under anaerobic condition is diploid; in other words, the chromosome has already been replicated and microorganisms are awaiting cell division.

In addition, *M. tuberculosis* derived from the Wayne model reveals a different antibiotic susceptibility pattern: microorganisms in the exponential phase are sensitive to ciprofloxacin and resistant to metronidazole, whereas NRP-state microorganisms are resistant to ciprofloxacin and sensitive to metronidazole. This suggests that changes in metabolism during the NRP states permit the activated form of metronidazole to accumulate, resulting in bacterial killing, that the bacilli alter the supercoiling states of DNA, causing insensitivity to DNA gyrase inhibitors such as ciprofloxacin.

The Wayne model has clinical correlations with studies of human anaerobic lung cavities that have documented microscopically visible acid-fast bacilli but inability to cultivate the microorganisms from anaerobic body sites.

**Persistence mutants**

Transcriptional regulation in response to environmental changes encountered during infection is a common theme in bacterial pathogenesis, and similar concepts can be applied to persistent *M. tuberculosis.* Survival in poorly oxygenated lesions was likely to be an important element of persistence, encouraging several groups to investigate non-replicating persistence of mycobacteria in a low-oxygen environment. An important
metabolic adaptation to hypoxic survival involves a switch to metabolism of lipid substrates using the enzymes isocitrate lyase and malate dehydrogenase to bypass the complete tricarboxylic-acid cycle by combining acetyl CoA with oxaloacetate by the glyoxylate shunt. A similar adaptation is observed in activated macrophages, and a mutant strain of *M. tuberculosis* H37Rv lacking isocitrate lyase was characterized by a defect in the persistent phase of infection in the C57BL/6 model. This indicates that the persistent bacteria use the abundant fatty acids of the host cells as a carbon source. Detailed analysis of the response of *M. tuberculosis* to hypoxia using microarray-based whole-genome profiling has identified an important response regulon. This includes a two-component regulator (DosR/DosS) that is responsible for the control of a set of genes, which include a prominent protein antigen, the α-crystallin chaperone. Increased expression of α-crystallin is seen under microaerophilic conditions, in stationary-phase cultures and in the C57BL/6 model. An important feature of the hypoxia regulon is that it can also be activated by exposure to nitric oxide. Hypoxia and nitric oxide are key elements of the inflammatory response, and an ability to respond to changes in these signals could allow *M. tuberculosis* to tune changes in phenotypic properties to changes in host immunity. It is possible that mycobacteria exploit additional mechanisms to allow a coordinated response to environmental changes. In *Micrococcus luteus*, secretion of a resuscitation-promoting factor, or Rpf protein, synchronizes the regrowth of bacteria after nutrient deprivation, and the presence of analogous genes in *M. tuberculosis* indicates that a similar phenomenon could contribute to the reactivation of TB.

Wayne has pioneered the use of hypoxic culture conditions to generate non-replicating persistent bacilli in vitro as a model for latency. Variants of this model have been used to identify *M. TB* genes potentially important for the development or maintenance of the latent state. Studies by Clifton Barry and colleagues using 2D protein gels have shown that at least six proteins are upregulated in stationary-phase cultures of *M. tuberculosis*. One of the most prominent stationary phase induced gene is acr (also known as hspX, Rv2031), which encodes alpha-crystallin. *M. TB Acr* is a dominant antigen in vivo, recognized by most TB patient sera. Acr is a member of the small heat shock protein family that forms high-molecular-weight aggregates and has chaperone activity in vitro. Under hypoxic conditions, during latency, Acr expression is dramatically and rapidly increased. Several other persistence mutants have been identified. These include additional transcriptional
### Table 1.

<table>
<thead>
<tr>
<th>Genes</th>
<th>Name</th>
<th>Function</th>
<th>Association with persistence</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Experiments with knockout mutants</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rv0467</td>
<td><em>icl</em></td>
<td>Isocitrate lyase</td>
<td>Metabolic defect, reduced persistence in mice</td>
</tr>
<tr>
<td>Rv0470c</td>
<td><em>pcaA</em></td>
<td>Cyclopropane synthase</td>
<td>Altered cell wall, reduced persistence in mice</td>
</tr>
<tr>
<td>Rv0981</td>
<td><em>mprR</em></td>
<td>2-component regulator</td>
<td>Reduced persistence in mice</td>
</tr>
<tr>
<td>Rv0353</td>
<td><em>hspR</em></td>
<td>Transcriptional repressor</td>
<td>Over-expression of heat shock protein, reduced persistence in mice</td>
</tr>
<tr>
<td>Rv1908c</td>
<td><em>katG</em></td>
<td>Catalase peroxidase</td>
<td>Reduced persistence in mice</td>
</tr>
<tr>
<td>Rv3223c</td>
<td><em>sigH</em></td>
<td>RNA polymerase sigma factor</td>
<td>Reduced pathology, time-two-death phenotype in mice</td>
</tr>
<tr>
<td>Rv3416</td>
<td><em>whiB3</em></td>
<td>Transcriptional regulator</td>
<td>Reduced pathology, time-two-death phenotype in mice</td>
</tr>
<tr>
<td>Rv1161-4</td>
<td><em>narGHJ1</em></td>
<td>Nitrate reductase</td>
<td>Reduced persistence of <em>M. marinum</em> in frog granuloma</td>
</tr>
<tr>
<td>Rv1651c/3812</td>
<td><em>mag24-1</em></td>
<td>PE-PGRS</td>
<td>Reduced persistence of BCG in mouse model</td>
</tr>
<tr>
<td>Rv0126</td>
<td><em>treS</em></td>
<td>Trehalose synthesis</td>
<td>Time-two-death phenotype in mice</td>
</tr>
<tr>
<td>Rv3286c</td>
<td><em>sigF</em></td>
<td>RNA polymerase sigma factor</td>
<td>Time-two-death phenotype in mice</td>
</tr>
<tr>
<td>Rv3825c</td>
<td><em>pks2</em></td>
<td>Sulpholipid synthesis</td>
<td>Time-two-death phenotype in mice</td>
</tr>
<tr>
<td>Rv2946/7c</td>
<td><em>pks1/15</em></td>
<td>Phenolic glycolipid synthesis</td>
<td>Time-two-death phenotype in mice</td>
</tr>
<tr>
<td>Rv2958c</td>
<td><em>relA</em></td>
<td>GTP-pyrophospho kinase</td>
<td>Reduced survival under hypoxia and nutrient stress</td>
</tr>
<tr>
<td>Rv3132c</td>
<td><em>devR/dosR</em></td>
<td>2-component regulator</td>
<td>Increased lethality in immuno-deficient mice</td>
</tr>
<tr>
<td>Rv3764/5c</td>
<td><em>tcrY</em></td>
<td>2-component regulator</td>
<td>Increased lethality in immuno-deficient mice</td>
</tr>
<tr>
<td>Rv1032c</td>
<td><em>treS</em></td>
<td>2-component regulator</td>
<td>Increased lethality in immuno-deficient mice</td>
</tr>
<tr>
<td>Rv1027/8c</td>
<td><em>kdpDE</em></td>
<td>2-component regulator</td>
<td>Increased lethality in immuno-deficient mice</td>
</tr>
<tr>
<td><em>Gene expression studies</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rv3132/3c</td>
<td>devR/S/dosR/S</td>
<td>2-component regulator</td>
<td>Increased expression as part of hypoxia regulon</td>
</tr>
<tr>
<td>Rv2031c</td>
<td><em>hspX</em></td>
<td>α-crystallin</td>
<td>Increased expression during hypoxia/mouse model</td>
</tr>
<tr>
<td>Rv1736v</td>
<td><em>natX</em></td>
<td>Fused nitrate reductase</td>
<td>Increased expression during hypoxia and in human granulomas</td>
</tr>
<tr>
<td>Rv1737c</td>
<td><em>natK2</em></td>
<td>Nitrate/nitrite-transporter</td>
<td>Increased expression during hypoxia</td>
</tr>
</tbody>
</table>

regulators, *mpcA* and *whiB3*⁴⁹ which are involved in as-yet-undefined adaptations required for persistence⁵⁰ and changes in cell-wall components⁵¹.⁵². Deletion of
**hspR**, a regulatory gene that is involved in the control of expression of heat-shock proteins, resulted in a persistence defect that is thought to be mediated by an enhanced immune response. Many more genes have been implicated in persistence on the basis of elevated expression levels or changes in the survival time of mice infected with mutants. The recent application of modern biochemical and molecular techniques to bacteriological persistence, such as the use of 2-D protein gels and subtractive cDNA libraries, has already led to the identification of gene products that are specifically regulated during bacterial persistence.

The information obtained from various sources using models that appear to closely resemble latency indicates the operation of regulated pathway(s) that ensures successful manifestation of dormancy. There are multiple players involved in this process leading to the onset of dormancy. The transcriptional repressor **HspR** is one of the many.

**References:**

Latency in Tuberculosis

Latency in Tuberculosis


III. Stress Response in Pathogenic Bacteria

- Temperature Stress
  - Induction of virulence genes
  - Induction of heat-shock genes
- Oxygen Stress
- Osmotic Stress
- Metal ion Stress

Pathogenic bacteria, alternate between free living and host associated states. Not only are the physico-chemical parameters encountered by the bacteria in these two states very dissimilar, but due to this difference they exert different demands and stresses on the bacterial cell. Bacterial pathogens have evolved highly complicated mechanisms for sensing external conditions, which assist them to cope up with various adverse situations. These pathogens respond by altering the pattern of gene expression with activation of a set of genes whose products assist in survival and turning off those products which are not necessary in a particular environment. These sensor-activator systems allow the bacteria to monitor environmental parameters which distinguish host from external environment, thus allowing them to adjust gene expression accordingly, particularly by induction of virulence factors\(^1\). The expression of virulence genes is controlled by regulatory systems in such a manner that the virulence factors are expressed at different stages of the infection process. This is dictated by the changing micro-environment of the host as a consequence of the pathophysiology of infection. Accordingly, mutations in some of the regulatory systems attenuate virulence of several bacterial species\(^3\). Both at the level of transcription as well as translation, the environmental control of regulatory mechanisms are mediated by a variety of intricate processes. Moreover, stress conditions, like changes in the osmolarity of the growth medium, anaerobiosis and temperature which pathogenic bacteria encounter upon entry into the host, can control gene expression by inducing changes in DNA topology which can provide an overlap between response to different environmental stimuli\(^4\).

Temperature Stress

**Induction of Virulence genes**

An increase in temperature from that of the environment to the physiological temperature of the human body, which is 37 °C, is probably the first signal to an invading bacterium on its entry into the host. In several pathogenic organisms, the virulence gene cassettes get switched on at this physiological temperature. In many of these pathogens, the virulence determinants are under the control of
transcriptional activators which respond to fluctuations of growth temperature leading to an enhanced expression of virulence genes at 37 °C. Such an example can be seen in a gram-positive facultative intracellular pathogen, *L. monocytogenes*, where the activation of virulence genes is under the control of a transcriptional activator PrfA. The gene encoding PrfA is transcribed from its own promoter as a monocistronic transcript at 30 °C. At 37 °C, in addition to the monocistronic transcript, the prfA gene is also transcribed from a different promoter as a bicistronic message comprising of the prf and the pic (phosphatidylinositol-specific-phospholipase C) genes. Hence the prfA gene is transcribed from two different promoter regions at higher growth temperatures and it is the bicistronic transcript that is thermally regulated. At 30 °C none of the virulence genes under the control of prfA are transcribed, although at 37 °C all of them are expressed. The repression of transcription of the PrfA regulated virulence genes at low temperatures might be due to either limitation of other yet unknown cellular components which may act in concert with PrfA or temperature dependent conformational change in the PrfA protein.

Alternative mechanisms for activation of the transcription factors of virulence genes have been described in *Y. enterolytica* and *V. cholerae*. In *Y. enterolytica*, the *virF* gene is the transcriptional activator of the virulence genes. The transcription of this gene itself is regulated by temperature. Evidences suggest that it is due to the alterations of DNA superhelicity that with increasing temperature there occurs an increase in relative abundance of *virF* message. A surprising phenomenon reported is that, in contrast to other pathogenic organisms, the virulence genes in *V. cholerae*, *Yersinia pseudotuberculosis*, *Yersinia enterocolitica* are optimally expressed at 30 °C and reduced at 37 °C.

It has been proposed that the divergent transcription of a htpG-like heat shock gene in *V. cholerae* leads to a proportionate decrease in the expression of *toxR*, coding for a trans-membrane DNA-binding protein that positively regulates transcription of the genes for cholera toxin and other virulence determinants. Several gene functions, including some necessary for chemotaxis and motility are repressed by ToxR. Motility is involved in establishing infection; hence it is possible that ToxR remains repressed at an early stage of infection. It has been proposed that once the bacterium reaches the site of colonization, environmental signals at the surface of the mucosal epithelium may activate ToxR leading to the expression of virulence genes.

**Induction of Heat-Shock genes**

Apart from the regulation of the virulence genes as described earlier, temperature
Stress response in pathogenic bacteria also plays a vital role in pathogenic organisms by inducing the ubiquitous heat shock response involving the expression of a set of proteins known commonly as the heat shock proteins (HSPs). Induction of HSPs occurs primarily at the level of transcription and in *E. coli*, the product of the *rpoH* (*hsrR*) gene, a σ-factor (σ^32_), is required for the transcription of heat shock genes. σ^32_ binds to the core RNA polymerase and recognizes the heat shock promoters which differ significantly from regular promoters with respect to −35 (consensus TCTCnCCCTTGAA) and −10 (consensus CCCCCATnTA) regions and the length of the spacer (13–17 nucleotides) between these regions. At higher temperatures a greater number of level in the cell does maintain a low level of transcription of most of the heat shock genes. After upshift of temperature from 30 to 42 °C the concentration of σ^32_ increases by about 20-fold within the first few minutes. Following which, it decreases to a new steady state level characteristic of the elevated temperature, within 20-30 min. This increase is primarily due to enhanced rate of translation of the *rpoH* mRNA and increased stability of the protein at high temperature. The increase in the rate of translation has been attributed to an alteration in the secondary structure of *rpoH* mRNA which favours efficient ribosome binding and initiation of translation.

**Fig. 7.** At low temperatures, σ^32_ interacts with DnaKJE and the complex is cleaved by FtsH, making σ^32_ unstable. At high temperatures, the DnaKJE chaperone machinery binds to unfolded polypeptides, and is involved in the folding process, and σ^32_ remains free. The free σ^32_ interacts with RNA-polymerase and the cleavage by FtsH is hindered. The half life of σ^32_ gets increased at 42 °C.

σ^32_ associate with core RNA polymerase. Hence the heat inducibility of the heat shock genes is due to this availability of an excess of σ^32_ at increasing temperatures. At low temperature, σ^32_ present at a basal σ^32_ interacts with DnaK-DnaJ-GrpE chaperones and the complex is susceptible to cleavage by the protease FtsH. This makes σ^32_ highly unstable (half life of 1 min at 30°C) at low temperatures (Fig. 7).
At high temperatures, the DnaK chaperone machinery is sequestered by binding to unfolded polypeptides that accumulate within the cell following temperature upshift leaving free $\sigma^2$. FtsH protease cleaves free $\sigma^2$ \textit{in vitro} in the presence of ATP\textsuperscript{11}. It is possible that at high temperature, the free $\sigma^2$ has greater affinity towards core RNA polymerase and the cleavage by FtsH is hindered, thereby increasing the half life of $\sigma^2$ to about 4–5 min at 42 °C\textsuperscript{12}. During the adaptation period, the level of DnaK increases with the concomitant increase in the amount of $\sigma^2$ complexed with DnaK which are eventually cleaved by FtsH reducing the half life of $\sigma^2$. Under extreme heat stress conditions, another factor $\sigma^E$ ($\sigma^{24}$), a 24 kDa protein is induced which controls the expression of a set of genes including \textit{btrA}, \textit{btrC} and \textit{tpoH}\textsuperscript{13,14,15}.

Temperature regulation of expression of virulence genes is distinct from the heat shock response in the character of the induction process. In heat shock response, the initial large increase in transcription of the heat shock genes is transient and is followed by the adaptation phase when the level of induction falls to a lower steady state value characteristic of the new elevated temperature\textsuperscript{10}. Induction of the virulence genes is more directly coupled to temperature and does not decrease unless temperature is lowered. Thus, the mechanism of negative regulation of the heat shock response postulated to account for the decrease in expression of the heat shock genes after the initial increase\textsuperscript{12} is not applicable to temperature dependent virulence gene expression, although there appears to be a basic similarity, in principle, in the mechanism of positive regulation of these two processes, i.e., activation of a transcriptional factor by increased temperature which recognizes certain features present in the promoters of the genes it controls. Moreover, the thermometer, i.e., the basic temperature sensing system in the cell has not been identified in either case, although there is evidence suggesting that DnaK may be the cellular thermometer\textsuperscript{16}.

Major HSPs from bacteria to humans are a highly conserved class of proteins and represent a significant proportion of the total protein content of all living cells. Many of the HSPs are constitutively present in the cell even under no-stress situation and perform important housekeeping functions\textsuperscript{17}. The major HSPs
(DnaK, DnaJ, GrpE, GroEL, GroES) are molecular chaperones that assist in correct folding and assembly of proteins and are involved in diverse cellular processes including DNA replication, RNA transcription, flagella synthesis and UV mutagenesis17(Fig. 8). GroEL together with GroES facilitates protein translocation across membrane barriers and possibly also secretion. The fundamental functions of these HSPs are to prevent protein denaturation and to reactivate partially denatured proteins. Non-repairable denatured proteins are degraded by another class of HSFs which represent either an ATP dependent protease (Lon, La) or a catalytic (ClpP, Ti) or regulatory (ATPase) subunit (ClpB) of another protease Clp17. Under mild heat shock conditions (37-42 °C) DnaK, GroEL and GroES analogous of M. tuberculosis are induced18. When the temperature was raised to 48 °C, the amount of the 65 kDa GroEL was drastically reduced and several HSPs of molecular masses ranging from 90-15 kDa were induced. The pattern of HSP synthesis is dependent on the severity of the heat stress with two distinct phases being discernible in the heat shock response.

**Oxygen Stress**

The expression of adherence and invasion factors of several pathogenic bacteria is regulated by oxygen concentration. High oxygen usually represses whereas low oxygen induces invasiveness. During switch from aerobic to anaerobic growth condition, a set of genes are induced and some genes are repressed19. In *E. coli*, two regulatory mechanisms have been identified which control the expression of these genes. One regulatory network is the Fnr (fumarate-nitrate reductase)-dependent control in response to anaerobiosis. Fnr activates the transcription of several respiratory genes such as fumarate reductase *(frd)*20,21, dimethyl sulphoxide-triethylamine-N-oxide reductase *(dmox)*22 and nitrate reductase *(nar)*23,24 and represses the expression of cytochrome d *(cyd)*25,26. The *frd* and *nar* gene products are required for the reduction of fumarate and nitrate so that they can serve as alternative electron acceptors for oxidative phosphorylation even in the absence of oxygen. Under semi-anaerobic conditions, the cytochrome d oxidase, which has low *Km* value for oxygen is induced under the control of ArcA-ArcB, a two component sensor regulator system responsive to a shift from aerobic to anaerobic growth. ArcA-ArcB-dependent regulation represses transcription of several genes involved in aerobic metabolism.

**Osmotic Stress**

For a pathogenic bacterium which passes from environmental waters to the human body for infection, osmolarity is an
important criterion to distinguish between the external and host associated environments. Osmolarity of an aqueous environment is thought to be no greater than that equivalent to 0.06 M NaCl while in the intestinal lumen the osmolarity is much higher (equivalent to 0.3 M NaCl) and in the blood stream the bacteria encounters an osmolarity equivalent to about 0.15 M NaCl. Thus, an increase in osmolarity is associated with expression of virulence factors in many pathogenic organisms. An example of osmotic stress is seen in *S. flexneri*. In *S. flexneri*, expression of the plasmid located *vir* genes which are necessary for invasion of epithelial cells is markedly enhanced under conditions of high osmolarity (0.15 M NaCl). The two component regulatory system OmpR-EnvZ encoded by the OmpB locus is responsible for sensing and responding to the signal and controls the expression of the *vir* genes as well as the chromosomal ompF-ompC genes encoding the osmoregulated porins which are probably necessary for intracellular survival. Thus, ompB deletion mutants of *S. flexneri* are defective with respect to both invasion and survival within the host tissues.

Again in *V. cholerae* too, expression of virulence factors is dependent on osmolarity of the growth medium and it is maximum at osmolarity equivalent to about 60 mM NaCl. This value is much lower than the osmolarity of the intestinal lumen (equivalent to about 0.3 M NaCl), the site of infection by *V. cholerae*. The C-terminal periplasmic domain of ToxR, the transcriptional activator of virulence genes in *V. cholerae*, is probably the osmosensor, since replacement of this part of the protein (by fusion with PhoA) rendered the protein constitutively active even under high salt conditions.

### Metal ion Stress

Free iron is extremely limited in the tissues and fluids of mammalian systems. Iron is an essential element for bacterial growth and many pathogenic bacteria have evolved highly efficient iron scavenging systems which are regulated in response to the iron status of the environment. In addition, low iron concentration leads to the increased synthesis of virulence determinants in several pathogenic bacteria including shiga-like toxin of enteropathogenic *E. coli*, shiga toxin of *S. dysenteriae*, diptheria toxin of *Corynebacterium diptheriae*, exotoxin A of *P. aeruginosa* and so on. The molecular mechanism of iron regulation of gene expression has been thoroughly studied in *E. coli*. The coordinate expression of iron regulated genes involves the Fur protein as repressor and iron as co-repressor. The repressor-co-repressor complex binds to operator sites within the promoters of the iron regulated genes. A consensus DNA binding site for the Fur protein consists of
a 21 base pair dyad symmetric sequence\textsuperscript{34}. This sequence, present upstream of the gene encoding the shiga-like toxin (\textit{ild}) in \textit{E. coli is} responsible for the Fur dependent repression of shiga-like toxin at high iron concentration\textsuperscript{30}. A large number of pathogenic bacteria including \textit{S. typhimurium}, \textit{Serratia marcescens}, \textit{Y. pestis}, \textit{V. cholerae} have Fur like iron regulatory systems\textsuperscript{35,36,37,38}. The transcription of \textit{igA} gene, coding for a major iron regulated outer membrane protein of \textit{V. cholerae}, is repressed at high iron concentration by the Fur protein of \textit{V. cholerae}. The ability of \textit{V. cholerae} to acquire iron from the host is linked to virulence and disruption of either heme utilization or vibriobactin uptake system reduces the ability of the organism to colonize the intestine\textsuperscript{39}.

References:
Stress Response in pathogenia bacteria


Hypoxic stress

The DosR-DosS regulatory system

In an effort to mimic the environment during the persistent phase of infection, Sherman et al. compared *M. tuberculosis* H37Rv gene expression during growth under reduced oxygen levels with that of H37Rv grown at normal oxygen levels. More than 100 genes were altered in their expression in response to hypoxic growth conditions. Not surprisingly, many genes that encode proteins involved in biosynthetic pathways and aerobic metabolism were repressed during growth with reduced oxygen levels. Many of the genes whose expression was induced during hypoxic growth code for proteins of unknown function, but among these was an apparent operon that includes the genes encoding the dosR-dosS (also called devR-devS, RV3133c/RV3132c) two-component regulatory system that Kinger & Tyagi had shown to be expressed at higher levels in *M. tuberculosis* H37Rv than in H37Ra. In bacteria, two-component response regulator systems are an important means by which a variety of environmental signals are transduced into a phenotypic response. These systems typically consist of a membrane-bound sensor kinase and soluble response regulator that is activated by a histidine-aspartate phosphorelay to bind upstream of specific genes and alter their expression. dosR and dosS may form a signaling system involved in the initial adaptation of bacilli to conditions within the host. It is seen that DosR binds upstream of hypoxic response genes. Further, nearly all MTB genes rapidly upregulated in response to low doses of NO or by hypoxia require DosR for their induction. DosS (Rv3132c) is a functional kinase of the two-component class and that it can transfer phosphate to DosR in vitro.

Additional experiments in which the operon including dosR-dosS was disrupted indicated that normal regulation of expression of the hypoxic response gene acr (hspx) was eliminated in this mutant, suggesting that the dosR-dosS two-component system controls gene expression during growth under hypoxic conditions.
acr
One MTB gene induced under hypoxia is acr1 (also known as hspX, Rv2031), which encodes α-crystallin (Acrl). The availability of the genome sequence allows a detailed analysis of the acr1 gene locus.

As shown in Fig. 9, this locus is organized in the form of a divergent operon with acr1/hspX gene being transcribed leftwards and a gene (Rv2032) having homology with nitoreductase being expressed rightwards. M. TB Acrl was originally described as a dominant antigen, frequently recognized by TB patient sera6,7. Further evidence of its role in TB pathogenesis comes from studies showing that Acrl is needed for the growth of M. TB in cultured macrophages8. Acrl localizes to the inner surface of the cell membrane9 and is a member of the small heat shock protein family that forms high molecular weight aggregates and has chaperone activity in vitro8. Under hypoxic conditions, expression of acr1 is dramatically increased8,10. In addition to its potential role in M. TB latency, the powerful regulation of Acrl under reduced O₂ tension may provide insight into the nature and control of the whole genetic response by which M. TB adapts to hypoxic microenvironments of the host.

Response to reactive nitrogen intermediates
Nitric oxide formation is believed to have originated in metazoan cells as an ancient first-line defense against intracellular parasites11. The role of the high-Ca²⁺-independent, inducible NO synthase (iNOS) and reactive nitrogen intermediates (RNI) in the control of Mycobacterium tuberculosis by the mouse macrophage has been reasonably well established12,13,14, albeit with some debate concerning the magnitude15,16 and apparent strain dependence16,17 of associated effects. If NO and its metabolites contribute to the control of M. tuberculosis in the host, it is reasonable to assume that, in turn, this organism may have evolved ways to respond to RNI challenges.

In contrast to the apparent paucity of Mycobacterium tuberculosis response to reactive oxygen intermediates, this organism has evolved a specific response to nitric oxide challenge. Exposure of M. tuberculosis to NO donors induces the synthesis of a set of polypeptides that have been collectively termed Nox. The most prominent Nox polypeptide, Noxl6, was identified by immunoblotting and by N-terminal sequencing as the α-crystallin-
related, 16-kDa small heat shock protein, sHsp16. A panel of chemically diverse donors of nitric oxide, with the exception of nitroprusside, induced sHsp16 (Nox16). Nitroprusside, a coordination complex of Fe²⁺ with a nitrosonium (NO⁺) ion, induced a 19-kDa polypeptide (Nox19) homologous to the nonheme bacterial ferritins. Thus, the NO response in *M. tuberculosis* is dominated by increased synthesis of the α-crystallin homolog sHsp16, previously implicated in stationary-phase processes. It has been found to be a major *M. tuberculosis* protein induced upon exposure to reactive nitrogen intermediates.

A Stationary-phase stress response sigma-factor, Sig F
Sigma factors are subunits of bacterial RNA polymerase and confer promoter specificity to the holoenzyme complex. The unique affinity of each sigma factor for its promoter consensus sequence is an essential component in many gene regulation systems. Alternative RNA polymerase sigma factors are a common means of coordinating gene regulation in bacteria. Alternative sigma factors have been shown to mediate adaptive responses to environmental conditions in many bacterial species. Many studies have implicated sigma factors in the regulation of virulence gene expression by *M. tuberculosis*. The *M. tuberculosis* sigF gene was discovered by degenerate PCR and is a close homologue of sporulation sigma factors in *Streptomyces coelicolor* and *Bacillus subtilis* as well as stress response sigma factors in *B. subtilis*, *Staphylococcus aureus*, and *Listeria monocytogenes*. RNase protection assay (RPA) using an *M. tuberculosis* sigF-specific probe showed that the *M. tuberculosis* sigF open reading frame is a transcribed gene. Transcription was maximal during stationary phase, cold shock, and nitrogen depletion. Weaker RPA signals were present during other stress conditions, such as oxidative stress, alcohol shock, and microaerophilic stress. No evidence of transcription was seen during exponential-phase growth. RPA is highly sensitive and can detect mRNA at the femtogram level. These findings show that the *M. tuberculosis* sigF encodes a stationary-phase/stress-response sigma factor.

As stated earlier, sigF is strongly induced during the stationary phase of growth and under certain stress conditions, such as nitrogen depletion, cold shock, and exposure to certain antibiotics. No marked change of sigF mRNA expression in *M. tuberculosis* H37Rv can be seen after a short 2-h exposure to a variety of stresses in culture, but expression is upregulated during growth within macrophages. Recently, *M. tuberculosis* sigF has been shown to be expressed during nutrient starvation, which may be a model of the non-growing drug-resistant state that
mimics the persistence of *M. tuberculosis* in vivo\(^3\). The *M. tuberculosis* sigF mutant strain, in which the sigF gene is deleted and replaced by a hygromycin resistance gene, has been shown to be less virulent in mice by time-to-death analysis\(^3\).

*M. tuberculosis* can survive for relatively long periods in expectorated sputum. Survival outside the human host would require adaptation to oxidative stress, low nutrient levels, and low temperature. The biochemical and genetic alterations permitting the organism to survive under these conditions are unknown. All of these conditions, in particular cold shock, induce *M. tuberculosis* sigF transcription. It is possible that SigF is involved in survival outside of the host. Alternatively, *M. tuberculosis* SigF might be involved in the adaptation of the organism during latent infection.

**The MprAB two component system**

Apart from dosR-dosS, the MprAB two-component system is also involved in stress management. *mprA* and *mprB* encode the cytoplasmic response regulator, MprA and its partner the histidine sensor kinase, MprB. The protein pair responds to environmental stimuli by initiating adaptive transcriptional programs. Polyphosphate kinase 1 (PPK1) catalyses the synthesis of polyphosphate (poly P) which is a linear polymer composed of several orthophosphate residues. MprB autophosphorylates itself with poly P serving as the phosphate-donor. The phosphorylated MprB-P phosphorylates MprA via phosphotransfer reactions. There is also evidence that MprB-P functions as a MprA-P (phosphorylated MprA) phosphatase. MprA-P binds the promoter of the mprAB operon to initiate transcription. A positive feedback loop is functional in the network of interactions as the production of MprA brings about further MprA synthesis. The mprAB operon has a basal level of gene expression irrespective of the operation of the positive feedback loop. Once the mprAB operon is activated, MprA-P regulates the transcription of the alternate sigma factor gene SigE\(^32,33\), which in turn controls the transcription of *rel*. Rel plays an important role in stress response by producing (p)pGpp\(^34,35\). The stringent response utilizes hyperphosphorylated guanine [(p)pGpp] as a signaling molecule to control bacterial gene expression involved in long-term survival under starvation conditions.

Expression of mprAB is induced in *M. tuberculosis* during growth within granulomas. It is also induced in response to nutrient starvation and in the presence of SDS. Microarray based investigations revealed that the expression of a number of genes including the stress responsive sigma factors SigB and SigE are regulated by MprAB\(^36\). One particular gene that
deserves specific mention in the context of the present investigation is \(\text{acr2}\), which codes for a alpha-crystallin class small molecular weight heat shock protein\(^{37}\). \(\text{acr2}\) is like \(\text{acr1}\), which encodes another alpha crystalline family protein known as HspX, have been found to be involved in mycobacterial pathogenesis although the mechanism is likely to be different. \(\text{acr2}\) has been termed as HrpA for heat-stress ribosome binding protein A. Unlike \(\text{acr1}\) which is induced under hypoxia, \(\text{acr2}\) is induced under heat shock, oxidative stress and SDS stress. MprAB has been found to regulate \(\text{acr2}\) expression. There are three binding sites for MprAB, one of which overlaps with the predicted sites for SigE, SigH and HspR, the regulatory protein which is the subject of the present investigation. How so many regulators bind to the \(\text{acr2}\) promoter and execute their functions in a regulated manner is an interesting issue that needs to be looked into in future.

**Heat Shock response of *Mycobacterium tuberculosis***

The classical heat shock proteins (HSPs) are stress-inducible molecular chaperones which represent the most conserved proteins in cellular life. In prokaryotes and eukaryotes their main role is to maintain a correctly folded and assembled protein component of the cell\(^{38}\). What makes the expression of HSPs so interesting to the pathogen biologist is that pathogen HSPs are also recognized by specific receptors on host immune cells, triggering an inflammatory immune response\(^{39,40,41,42}\). In addition to this, the conservation between host and pathogen HSPs means that pathogen HSPs can be utilized by the host to shuttle peptides into the HSP-mediated antigen presentation pathway\(^{43,44}\). Thus, while the pathogen needs to increase expression of its HSPs in response to the stresses induced by host defenses\(^{45,46,47,48}\), it must temper this need so as not to alert the host immune response to its presence.

The infection profile of *Mycobacterium tuberculosis*, is critically governed by a dynamic relationship with the host immune response. The vital importance of HSP regulation to *M. tuberculosis* has been demonstrated by generating a mutant strain lacking the HspR repressor protein, thus effecting dysregulation of the Hsp70 response\(^ {49}\). The mutant strain constitutively overexpressed Hsp70 and associated HSPs, and during murine infection its survival was emphatically reduced. The underlying cause for this attenuation was enhanced immune recognition of the bacterium.

Of course, heat shock is a complex response which varies with both time and temperature, but this snapshot of the response provides a reference with which to compare transcriptomic changes in
defined regulatory mutants. One general observation on the heat shock transcriptional profile is that the response is not simply the elevated transcription of the known HSPs but encompasses genome-wide changes in gene expression.

To dissect this response, two likely heat shock repressor proteins in *M. tuberculosis*, HspR and HrcA, which were identified by homology to regulators in *Streptomyces* and *Bacillus*, were knocked out. Mutants of both of these regulators using gene replacement system were made. Comparison of the expression profiles of strains lacking these regulators with wild-type, combined with identification of repressor binding sites in promoter regions, established the HspR and HrcA regulons. All members of these regulons were also found to be upregulated during heat shock and amongst them were many of the classical heat shock chaperones. Comparative studies examining the regulation of other arms of the mycobacterial stress response reveals a high degree of crosstalk and overlap between the different stress regulons. For example, the Hsp70 regulon forms a central element of the heat shock response and is under negative regulation by HspR in complex with Hsp70 itself. However, the Hsp70 operon is also under control of the heat-inducible alternative sigma factor σH, which in addition promotes transcription at the stress response sigma factors σB and σE.

Further to this, the functional activity of these sigma factors is under the control of anti-sigma factor pathways. Understanding the regulatory mechanisms behind mycobacterial HSP expression may allow the development of novel strategies for the treatment of tuberculosis. The dysregulation of the *M. tuberculosis* Hsp70 response allows the host to mount a more effective immunological response against the bacterium. Thus, drugs that disrupt HSP regulation by interfering with specific regulators make an attractive mechanism by which to enhance host immunity.

**HspR, a heat shock response regulon**

Heat shock is both a convenient and robust laboratory model for studying transcriptional regulation as well as an important element in the pathogenesis of *M. tuberculosis*. Heat-shock proteins are induced following uptake of *M. tuberculosis* by host cells, presumably contributing to bacterial survival in the phagocytic environment. They are also used by the host cell as peptide chaperones to enhance MHC class I presentation and as a signal
to the immune system. Directly because of these immunological roles of heat shock proteins, the constitutive overexpression of bacterial Hsp70 proteins results in reduced survival of \textit{M. tuberculosis} during prolonged infection\textsuperscript{49}. By understanding the regulation of the heat-shock response, we may be able to design interventions capable of enhancing resistance to persistent infection with tuberculosis.

Exposure of \textit{M. tuberculosis} to increased temperature results in elevated transcription of heat-shock genes and expression of the corresponding proteins\textsuperscript{57,58}. The regulatory mechanisms involved have not been characterized. Two general mechanisms for heat-shock regulation have been identified in bacteria.

Induction of the response in \textit{E. coli} involves transcriptional activation, with increased levels of an alternative \(\sigma\)-factor, \(\sigma-32\), directing RNA polymerase towards genes proceeded by a consensus heat-shock promoter sequence\textsuperscript{59}. In contrast, in \textit{Bacillus subtilis} the heat-shock response is regulated by transcriptional repression\textsuperscript{60}. Inspection of the genome sequence of \textit{M. tuberculosis}\textsuperscript{61} indicates repression as the probable mechanism of heat-shock regulation. Open reading frame (ORF) Rv2374c encodes a homolog of the HrcA repressor, whereas Rv0353 encodes a protein similar to HspR, a repressor identified in \textit{hsp70} regulation in \textit{S. coelicolor}\textsuperscript{62} and in \textit{Helicobacter pylori}\textsuperscript{63}.

The \textit{M. tuberculosis} \textit{hspR} is the fourth gene in the DnaK operon, comprising of the \textit{hsp70} or \textit{dnaK}, followed by genes encoding GrpE and DnaJ, HSPs that have functional interactions with Hsp70\textsuperscript{64} and are known as their co-chaperones (Fig. 10).

HspR is a DNA-binding protein related to the \textit{MerR} family (Fig. 11), and contains an HTH-motif. It recognizes either of two inverted repeat sequences, HAIR (HspR Associated Inverted Repeats), in the promoter region of the \textit{hsp70} operon, and acts as an autoregulator by reducing the level of transcription in unstressed conditions. HAIR sequences, or binding regions of HspR, are not only identified upstream of the Hsp70 operon but can
Microarray analysis of an hspR deletion mutant of M. tuberculosis confirms and extends previous studies of Hsp70 regulation. In the absence of HspR, there is release of transcriptional repression, and the genes of the Hsp70 operon are upregulated. Other than the genes of the Hsp70 operon, a further of 46 different genes have been identified which are seen to have significantly elevated transcription (Table 2).

Within the ΔhspR-upregulated ORF set, the hsp70 and acr2 operon genes were upregulated during heat shock along with bfrB, groES and Rv3654c. The bacterioferritin gene, bfrB, and Rv3654c, encoding an 8 kDa protein with unknown function, are not preceded by obvious HspR binding sites, but their co-regulation with HAIR-associated genes in both heat shock and the mutant suggest an indirect link to HspR. The majority of genes upregulated in the mutant were neither associated with HAIR sequences nor upregulated during heat shock. The induction of these genes is a consequence of the physiological changes associated with overexpression of the HspR-regulated proteins and may not be directly relevant to the normal heat-shock response. A surprising omission from the ΔhspK upregulated list is clpB, which encodes another probable molecular chaperone. This is probably because the clpB mRNA is of a sufficiently short half-life to preclude detection of the ΔhspR-associated transcriptional increase. The detection of substantially increased clpB mRNA in the wild-type after heat shock at 45 °C is explained by upregulation of clpB
### Table 2

<table>
<thead>
<tr>
<th>Gene Designation</th>
<th>Gene name</th>
<th>P Value</th>
<th>Fold Change in ΔhspR compared with wild-type</th>
<th>Fold Change in the complimented strain, ΔhspR (pSMT168), compared with wild-type</th>
<th>Regulatory sequence</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rv0004</td>
<td></td>
<td>9.3E-03</td>
<td>2.77</td>
<td>1.19</td>
<td>CH</td>
<td></td>
</tr>
<tr>
<td>Rv0108c</td>
<td></td>
<td>4.12E-04</td>
<td>2.09</td>
<td>1.16</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>Rv0249c</td>
<td></td>
<td>1.06E-03</td>
<td>2.00</td>
<td>1.10</td>
<td>HAIR</td>
<td>Membrane protein</td>
</tr>
<tr>
<td>Rv0250c</td>
<td></td>
<td>2.93E-05</td>
<td>2.26</td>
<td>1.15</td>
<td>HAIR</td>
<td>CH</td>
</tr>
<tr>
<td>Rv0251c</td>
<td></td>
<td>2.14E-07</td>
<td>2.54</td>
<td>1.11</td>
<td>HAIR</td>
<td>Chaperone</td>
</tr>
<tr>
<td>Rv0268c</td>
<td></td>
<td>5.46E-03</td>
<td>1.96</td>
<td>1.23</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>Rv0291</td>
<td></td>
<td>7.7E-03</td>
<td>2.93</td>
<td>0.84</td>
<td>Membrane protease</td>
<td></td>
</tr>
<tr>
<td>Rv0350</td>
<td>dnaK</td>
<td>2.49E-25</td>
<td>4.50</td>
<td>2.23</td>
<td>HAIR</td>
<td>Chaperone</td>
</tr>
<tr>
<td>Rv0351</td>
<td>pepE</td>
<td>1.58E-28</td>
<td>4.99</td>
<td>2.12</td>
<td>HAIR</td>
<td>Chaperone</td>
</tr>
<tr>
<td>Rv0352</td>
<td>dnaJ</td>
<td>1.40E-19</td>
<td>4.01</td>
<td>1.60</td>
<td>HAIR</td>
<td>Chaperone</td>
</tr>
<tr>
<td>Rv0455c</td>
<td></td>
<td>3.89E-06</td>
<td>2.32</td>
<td>1.25</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>Rv0469</td>
<td>mmaA1</td>
<td>2.15E-04</td>
<td>2.06</td>
<td>0.98</td>
<td>Mycolic acid synthesis</td>
<td></td>
</tr>
<tr>
<td>Rv04516c</td>
<td></td>
<td>1.19E-04</td>
<td>2.16</td>
<td>1.15</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>Rv0569</td>
<td></td>
<td>1.06E-03</td>
<td>1.98</td>
<td>0.62</td>
<td>CH</td>
<td></td>
</tr>
<tr>
<td>Rv0635</td>
<td></td>
<td>1.50E-04</td>
<td>2.13</td>
<td>1.05</td>
<td>CH</td>
<td></td>
</tr>
<tr>
<td>Rv0643c</td>
<td>mmaA3</td>
<td>5.09E-04</td>
<td>2.04</td>
<td>0.88</td>
<td>Mycolic acid synthesis</td>
<td></td>
</tr>
<tr>
<td>Rv0655</td>
<td></td>
<td>9.26E-03</td>
<td>1.82</td>
<td>1.11</td>
<td>ABC transporter</td>
<td></td>
</tr>
<tr>
<td>Rv0701</td>
<td>rpsC</td>
<td>7.34E-03</td>
<td>1.95</td>
<td>1.27</td>
<td>Ribosomal protein</td>
<td></td>
</tr>
<tr>
<td>Rv0714</td>
<td>rpsN</td>
<td>2.55E-03</td>
<td>2.00</td>
<td>1.20</td>
<td>Ribosomal protein</td>
<td></td>
</tr>
<tr>
<td>Rv0715</td>
<td>rpsX</td>
<td>2.01E-03</td>
<td>2.03</td>
<td>1.10</td>
<td>Ribosomal protein</td>
<td></td>
</tr>
<tr>
<td>Rv0909</td>
<td></td>
<td>1.90E-03</td>
<td>1.98</td>
<td>-</td>
<td>CH</td>
<td></td>
</tr>
<tr>
<td>Rv1078</td>
<td>prsA</td>
<td>4.08E-04</td>
<td>2.00</td>
<td>1.09</td>
<td>CH</td>
<td></td>
</tr>
<tr>
<td>Rv1172c</td>
<td></td>
<td>2.27E-03</td>
<td>1.99</td>
<td>1.42</td>
<td>PE protein Family</td>
<td></td>
</tr>
<tr>
<td>Rv1174c</td>
<td></td>
<td>2.90E-03</td>
<td>1.93</td>
<td>1.10</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>Rv1298</td>
<td>rpsE</td>
<td>4.64E-04</td>
<td>2.11</td>
<td>1.30</td>
<td>Ribosomal protein</td>
<td></td>
</tr>
<tr>
<td>Rv1738</td>
<td></td>
<td>6.42E-07</td>
<td>2.46</td>
<td>0.89</td>
<td>CH</td>
<td></td>
</tr>
<tr>
<td>Rv1872c</td>
<td>ilvD2</td>
<td>1.02E-03</td>
<td>1.93</td>
<td>1.14</td>
<td>Oxidoreductase</td>
<td></td>
</tr>
<tr>
<td>Rv2005c</td>
<td></td>
<td>5.43E-04</td>
<td>2.03</td>
<td>0.83</td>
<td>Universal stress protein</td>
<td></td>
</tr>
<tr>
<td>Rv2007c</td>
<td>fbcA</td>
<td>2.04E-14</td>
<td>3.17</td>
<td>0.86</td>
<td>Ferredoxin</td>
<td></td>
</tr>
</tbody>
</table>
transcription by the heat-inducible sigma factor, σH, as well as release of HspR repression.\textsuperscript{54}

The HspR protein interacts tightly with Hsp70 \textit{in vitro} \textsuperscript{49,52} and DnaK functions as a transcriptional co-repressor by binding to HspR at its operator sites in \textit{Streptomyces coelicolor} (Fig. 13). A system where this heterodimer forms the functional repressor unit with feedback achieved by titration of Hsp70 away from the HspR complex in the presence of unfolded polypeptides represents an attractive model for regulation\textsuperscript{52,65}. DnaK activates HspR repressor, rather than inactivating an activator (such as σ\textsuperscript{32}). Despite this reversal of function, the end-result is the same, the molecular chaperones negatively

<table>
<thead>
<tr>
<th>Rv2069</th>
<th>sigC</th>
<th>7.07E-03</th>
<th>1.94</th>
<th>1.16</th>
<th>Sigma factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rv294c</td>
<td>τττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττττtauA</td>
<td>2.34E-04</td>
<td>2.17</td>
<td>1.15</td>
<td>Protein translocase</td>
</tr>
<tr>
<td>Rv2137c</td>
<td>tatA</td>
<td>9.95E-04</td>
<td>1.97</td>
<td>1.09</td>
<td>CH</td>
</tr>
<tr>
<td>Rv2185c</td>
<td>daE</td>
<td>7.65E-04</td>
<td>2.07</td>
<td>1.02</td>
<td>CH</td>
</tr>
<tr>
<td>Rv2193</td>
<td>chpE</td>
<td>2.05E-04</td>
<td>2.02</td>
<td>1.22</td>
<td>Cytochrome c Oxidase</td>
</tr>
<tr>
<td>Rv2288</td>
<td>cdb</td>
<td>6.64E-04</td>
<td>2.09</td>
<td>1.08</td>
<td>Unknown</td>
</tr>
<tr>
<td>Rv2289</td>
<td>cdh</td>
<td>6.63E-03</td>
<td>1.87</td>
<td>1.03</td>
<td>CDP diglyceride hydrolase</td>
</tr>
<tr>
<td>Rv2764c</td>
<td>dhaA</td>
<td>8.29E-03</td>
<td>1.83</td>
<td>1.10</td>
<td>Thymidylate synthase</td>
</tr>
<tr>
<td>Rv2816c</td>
<td>ciaA</td>
<td>3.93E-10</td>
<td>2.84</td>
<td>1.59</td>
<td>CH</td>
</tr>
<tr>
<td>Rv2876</td>
<td>ciaA</td>
<td>1.89E-03</td>
<td>1.90</td>
<td>1.20</td>
<td>Unknown</td>
</tr>
<tr>
<td>Rv2960c</td>
<td>ciaA</td>
<td>1.03E-03</td>
<td>2.04</td>
<td>1.04</td>
<td>Unknown</td>
</tr>
<tr>
<td>Rv3101c</td>
<td>ftsX</td>
<td>3.61E-03</td>
<td>1.92</td>
<td>1.04</td>
<td>Cell division protein</td>
</tr>
<tr>
<td>Rv3128c</td>
<td>ftsX</td>
<td>2.12E-03</td>
<td>1.96</td>
<td>1.02</td>
<td>CH</td>
</tr>
<tr>
<td>Rv3407</td>
<td>ftsX</td>
<td>1.44E-04</td>
<td>2.10</td>
<td>1.15</td>
<td>CH</td>
</tr>
<tr>
<td>Rv3418c</td>
<td>ftsX</td>
<td>4.45E-03</td>
<td>1.92</td>
<td>1.06</td>
<td>Chaperone</td>
</tr>
<tr>
<td>Rv3508</td>
<td>ftsX</td>
<td>9.24E-03</td>
<td>1.90</td>
<td>-</td>
<td>PE-PGRS protein family</td>
</tr>
<tr>
<td>Rv3654c</td>
<td>ftsX</td>
<td>1.34E-27</td>
<td>5.08</td>
<td>1.53</td>
<td>Unknown</td>
</tr>
<tr>
<td>Rv3679</td>
<td>ftsX</td>
<td>2.57E-03</td>
<td>1.96</td>
<td>1.12</td>
<td>Anion transporter</td>
</tr>
<tr>
<td>Rv3680</td>
<td>ftsX</td>
<td>5.91E-03</td>
<td>1.85</td>
<td>1.45</td>
<td>Anion transporter</td>
</tr>
<tr>
<td>Rv3841</td>
<td>ftsX</td>
<td>7.64E-04</td>
<td>1.98</td>
<td>1.02</td>
<td>Bacterioferritin</td>
</tr>
</tbody>
</table>

![Fig. 13. In \textit{Streptomyces coelicolor}, DnaK forms a complex with HspR and this acts as the functional repressor unit.](image-url)
Mycobacterial Stress Response

regulate their own synthesis by modulating the activity of their respective transcription factors. Thus, although mechanisms controlling molecular chaperone gene expression vary widely, homeostatic feedback control by a key chaperone is emerging as a common theme. In DnaK-mediated regulation of the *S. coelicolor* dnaK operon, under normal growth condition, native HspR binds DnaK, and this complex binds avidly to the dnaK promoter region to repress transcription of the operon efficiently. In the (transient) absence of DnaK, HspR does not bind with high affinity to its DNA target. Thus, during heat shock, when DnaK is sequestered by denatured or partially unfolded proteins, the HspR protein is unable efficiently to repress transcription and the operon is induced.

References:
gene cluster with similarities to the Bacillus subtilis sigG and sigB operons. Tuber. Lung Dis. 78: 3-12.
Mycobacterium tuberculosis sigF gene: implications for chemotherapy against mycobacterial persistors.
synthesized in response to phagocytosis by human macrophages by selective capture of transcribed
nutrient starvation model of Mycobacterium tuberculosis persistence by gene and protein expression
characterization of a Mycobacterium tuberculosis mutant lacking the alternate sigma factor gene, sigF.
The stringent response of Mycobacterium tuberculosis is required for long-term survival. J Bacteriol 182:
4889-98.
36. He, H., Hovey, R. Kane, J. Singh, V. and T. C. Zahrt. (2006) MprAB is a stress-responsive two-
component system that directly regulates expression of sigma factors SigB and SigE in


V. Protein Folding and the role of molecular chaperones

The Role of Major Heat Shock Proteins as Molecular Chaperones
What began as a molecular curiosity in fruit flies over 30 years ago, the so-called heat shock or stress response, now constitutes an active area of research in cell biology and biochemistry. The heat shock proteins (hsp70), one of the most highly conserved group of proteins so far characterized, are being implicated as essential components in a number of diverse biological processes. Although referred to as heat shock proteins, most of these proteins in fact are expressed at rather significant levels in all cells maintained under normal growth conditions and are essential for cellular growth at all physiologically relevant temperatures. Much of the current interest in the hsp70 follows from recent studies demonstrating their role as molecular "chaperones", being intimately involved in various steps of protein maturation. Members of the Hsp70 (DnaK) and Hsp60 (GroEL) families, for example, participate in protein folding, protein translocation, and perhaps higher ordered protein assembly. Yet other members of the heat shock protein family, such as Hsp90, play important roles in the regulation of certain transcription factors and protein kinases.

The Heat Shock Proteins
For largely historical reasons, the hsp70 from different organisms are referred to by a variety of different names. Unfortunately this often leads to some confusion when discussing the structure and function of a particular member of the heat shock protein family. Following the nomenclature first used in *Drosophila*, the various hsp70 in animal cells are referred to on the basis of their mode of
induction, and apparent molecular mass on SDS gels. Hence, their designation as heat shock protein 70 (Hsp70) and heat shock protein 90 (Hsp90), for example, refers to heat-inducible proteins of 70 or 90 kDa, respectively. In eukaryotic cells, another class of stress-induced proteins, distinct from those induced by heat, were first observed to exhibit an increased expression in cells starved of glucose, and therefore were called the glucose-regulated proteins (grps)\textsuperscript{2}. Subsequent work revealed increased grps expression in cells under anoxic conditions, treatment with inhibitors of glycosylation, exposure to agents that perturb calcium homeostasis (e.g., calcium ionophores or EGTA), or exposure to different reducing agents (e.g., DTT, BME)\textsuperscript{3}. The major grps, with apparent masses of 75, 78, and 94 kd, now are known to be related to the heat shock protein family and, like most of the other hsps, are also expressed in cells maintained under normal growth conditions.

In bacteria, the nomenclature used to refer to the hsps is based on earlier genetic studies, examining bacterial host functions that were essential for bacteriophage growth. Here, the genes coding for the constitutively expressed hsps were found to map within specific operons and shown to participate in various aspects of bacteriophage DNA replication and morphogenesis\textsuperscript{4,5,6}. For example, the DnaK protein (prokaryotic homologue of Hsp70) is required for bacteriophage DNA replication, while the GroEL/GroES proteins (prokaryotic homologues of Hsp60/Hsp10) were shown to be necessary for the assembly of the bacteriophage head and tail components. Finally, in S. cerevisiae, the hsps are referred to by genetic nomenclature, or sometimes by their apparent size. In Table 3, a summary of the nomenclature being currently used to refer to the two major stress protein families known to function as molecular chaperones in bacteria, S. cerevisiae, and higher eukaryotes.

**The Problem of Protein Folding**

The pioneering work of Anfinsen\textsuperscript{7} (1973) with the \textit{in vitro} refolding purified ribonuclease A left the long-lasting impression that the folding of a newly synthesized polypeptide was an intrinsic feature of its primary structure, independent of other factors\textsuperscript{8}. Most of the \textit{in vitro} protein refolding experiments are usually carried out by first denaturing a given purified polypeptide and then removing the denaturant. Under these conditions, most polypeptides quickly collapse into a compact structure, usually called "molten globule", thought to possess extensive secondary structure, but still exposing hydrophobic groups, which may lead to aggregation. The probability that a given unfolded polypeptide will fold properly increases at relatively low protein
Table 3.

<table>
<thead>
<tr>
<th>Family</th>
<th>Name</th>
<th>Locale</th>
<th>Size (kD)</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hsp70</td>
<td>DnaK</td>
<td>Cytosol</td>
<td>70</td>
<td>Mutants grow poorly; constitutive/inducible</td>
</tr>
<tr>
<td>Prokaryotes</td>
<td>Ssa</td>
<td>Cytosol</td>
<td>70</td>
<td>Multiple members: essential family</td>
</tr>
<tr>
<td>Yeast</td>
<td>Ssb</td>
<td>Cytosol</td>
<td>70</td>
<td>Multiple members: mutants cold sensitive</td>
</tr>
<tr>
<td></td>
<td>Ssc</td>
<td>Mitochondria</td>
<td>70</td>
<td>Essential for viability</td>
</tr>
<tr>
<td></td>
<td>Kar2</td>
<td>ER</td>
<td>70</td>
<td>Essential for viability</td>
</tr>
<tr>
<td>Higher eukaryotes</td>
<td>Hsc70</td>
<td>Cytosol/nucleus</td>
<td>70</td>
<td>Constitutive</td>
</tr>
<tr>
<td></td>
<td>Hsp70</td>
<td>Cytosol/nucleus</td>
<td>70</td>
<td>Stress inducible</td>
</tr>
<tr>
<td></td>
<td>Grp75</td>
<td>Mitochondria</td>
<td>75</td>
<td>Most homology to DnaK</td>
</tr>
<tr>
<td></td>
<td>Hsp70</td>
<td>Chloroplasts</td>
<td>70</td>
<td>Constitutive</td>
</tr>
<tr>
<td></td>
<td>Grp78(BiP)</td>
<td>ER</td>
<td>78</td>
<td>ADP-ribosylated</td>
</tr>
<tr>
<td>Hsp60</td>
<td>GroEL</td>
<td>Cytosol</td>
<td>60</td>
<td>Essential for viability</td>
</tr>
<tr>
<td>Prokaryotes</td>
<td>Mif4</td>
<td>Mitochondria</td>
<td>60</td>
<td>Essential for viability constitutive/inducible</td>
</tr>
<tr>
<td>Yeast</td>
<td>TCP-1 α</td>
<td>Cytosol</td>
<td>60</td>
<td>Essential for viability</td>
</tr>
<tr>
<td></td>
<td>TCP-1 β</td>
<td>Cytosol</td>
<td>60</td>
<td>Essential for viability</td>
</tr>
<tr>
<td>Higher eukaryotes</td>
<td>Hsp60</td>
<td>Mitochondria</td>
<td>60</td>
<td>Constitutive/inducible</td>
</tr>
<tr>
<td></td>
<td>Rubisco binding protein</td>
<td>Chloroplasts</td>
<td>60</td>
<td>Additional members?</td>
</tr>
<tr>
<td></td>
<td>TCP-1</td>
<td>Cytosol</td>
<td>60</td>
<td></td>
</tr>
</tbody>
</table>

concentrations (which limit inter-polypeptide aggregation) and low temperature (which attenuates hydrophobic interactions). However, the relatively high protein concentrations in the cytosol [or the specialized organelles, i.e. endoplasmic reticulum (ER), mitochondria, chloroplasts] may subject the growing polypeptide chains, as they emerge from the ribosomes (or as they enter various cellular compartments), to premature interactions with other intra- or inter-polypeptide domains, thereby leading to misfolding and aggregation. To deal with these sorts of problems, a set of proteins, collectively called chaperones, has been identified and whose primary function is to ensure that polypeptides will fold or assemble properly in the cell. These chaperone proteins act primarily by
binding to the reactive surfaces of polypeptides (such as the hydrophobic surfaces exposed by the molten globule intermediates of various proteins). In doing so, chaperones sequester these reactive sites from the rest of the reactive surfaces present in their vicinity, thus effectively preventing aggregation and favoring the proper folding pathway. The chaperone proteins act without covalently modifying their polypeptide substrates and without being part of the finished product. Because high temperatures tend to favor both protein unfolding on the one hand and hydrophobic interactions on the other, there is an extra need for chaperone protein function to prevent protein aggregation in vivo. This is most likely the reason why many protein chaperons are expressed at higher levels in cells after heat shock. A detailed discussion is given on the properties of the two major stress proteins that function as molecular chaperones, which are the Hsp70 (DnaK) and the Hsp60 (GroEL) families of stress proteins.

Activity of Hsp70 involves Cycles of Polypeptide Binding and Release
Hsp70 chaperones, with their co-chaperones, comprise a set of abundant cellular machines that assist a large variety of protein folding processes in almost all cellular compartments. Historically, they were identified by induction under conditions of stress, during which they are now known to provide an essential action of preventing aggregation and assisting refolding of misfolded proteins. But they also play an essential role under normal conditions, including (1) assisting folding of some newly translated proteins; (2) guiding translocating proteins across organelar membranes through action at both the cis and trans sides; (3) disassembling oligomeric protein structures; (4) facilitating proteolytic degradation of unstable proteins; and in selected cases, (5) controlling the biological activity of folded regulatory proteins, including transcription factors. All of these activities rely on the ATP-regulated association of Hsp70 with short hydrophobic segments in substrate polypeptides, which prevents further folding or aggregation by shielding these segments. In Hsp70-assisted folding reactions, substrates undergo repeated cycles of binding/release, frequently at a stoichiometry of a single Hsp70 monomer per substrate molecule. Hsp70 binding does not appear to induce global conformational changes in the substrate but, rather, appears to act locally. Substrates released from the chaperone undergo kinetic partitioning between folding to native state, aggregation, rebinding to Hsp70, and binding to other chaperones or proteases as part of a
multidirectional folding network. Hsp70 proteins all consist of same working parts: a highly conserved (Fig. 14) NH₂-terminal ATPase domain of 44 kDa and a COOH-terminal region of 25 kDa, divided into a conserved substrate binding domain of 15 kDa and a less-conserved immediate COOH-terminal domain of 10 kDa.

**E. coli** Hsp70, DnaK
The *Escherichia coli* 70-kDa molecular chaperone, DnaK, folds, transports, and assembles other proteins in an ATP dependent activity cycle that is regulated by the cochaperones, GrpE and DnaJ.\(^{18,19,20}\). DnaK cycles from an ADP bound, high-affinity state that tightly binds unfolded substrate to an ATP-bound, low-affinity state that weakly binds substrate. GrpE is the nucleotide exchange factor that catalyzes the release of ADP from ADP-DnaK complexes\(^{21,22}\). ADP release in turn permits ATP binding which induces the high-to-low affinity transition. The rate of GrpE-facilitated ADP/ATP exchange is known to be controlled by temperature\(^{23}\). It has been reported that with increasing temperature, melting of the long helix pair in the GrpE dimer decreases the efficacy of GrpE in catalyzing nucleotide exchange. DnaJ
Fig. 15. Domains of DnaK. DnaK consists of a 44 kDa N-terminal ATP-ase domain and a 25 kDa C-terminal peptide binding domain. The peptide binding domain comprises of the helices A to E along with the random tail.

promotes the reverse transition\textsuperscript{24,25}.\textit{E. coli} DnaK is composed of three domains: the ATPase domain, the substrate-binding domain, and the lid comprise residues 1-388, 389-508, and 509-638 respectively. The ATPase domain is a bilobed structure that contains a deep channel between the two lobes,\textsuperscript{26,27} nucleotide binds at the base of the channel. The substrate-binding domain consists of a uniquely folded \(\beta\)-sandwich subdomain followed by an \(\alpha\)-helical domain that consists of five antiparallel \(\alpha\)-helices (Fig. 15)\textsuperscript{28,29}. The substrate peptide is bound in an extended conformation between two loops that protrude from the \(\beta\)-sandwich subdomain (L1,2 and L3,4), and are in turn buttressed by two other loops (L4,5 and L5,6) (Fig. 16). The helical subdomain of the Substrate Binding domain D acts like a lid over the \(\beta\)-sandwich subdomain. The \(\alpha\)-helices A and B are packed onto the \(\beta\)-sandwich, and helix B covers the substrate binding cavity. The distal part of helix B, together with helices C, D, and E, builds up a hydrophobic helical core that acts as a “lid”\textsuperscript{29,30}. Although interdomain coupling occurs in the absence of the lid and the bound peptide contacts the \(\beta\)-structure but not the \(\alpha\)-helical subdomain, its presence has been confirmed to be essential in stabilizing Hsp70-substrate complexes\textsuperscript{31,32}, especially at stress temperatures\textsuperscript{33}. A network of hydrogen bonds and a salt bridge links the lid noncovalently to the \(\beta\)-sandwich. The salt bridge Asp\textsuperscript{526} - Asp\textsuperscript{445} (Fig. 16) plays a pivotal role in the proper positioning of the lid. The bound peptide interacts with the \(\beta\)-sandwich but not the lid\textsuperscript{29}. The role of the lid in the chaperone activity cycle is
still not precisely understood. Interdomain coupling occurs in the absence of the lid \cite{34,35}, although ATP-induced peptide dissociation is significantly accelerated \cite{31} compared to the wild-type protein. The 33 C-terminal residues of DnaK, which constitutes a flexible, mobile region of the protein, has no known function \cite{36}. There are intriguing hints that the lid region of DnaK and other Hsp70s interacts with DnaJ \cite{37}. Deletion of the bulk of DnaK’s lid, residues 518-638, increases the rate of ATP induced peptide release.

The peptide Binding domain of DnaK and the consensus motif of DnaK-interacting peptides

Studies with peptides have indicated that DnaK binds with maximum affinity to short hydrophobic segments in extended conformation \cite{29,38}. In order to qualify as a substrate, the minimal requirement of a protein is that it should expose a single recognizable segment, either through local unfolding or as an intrinsically unfolded structural element, such as a loop. The crystal structure of the COOH-terminal substrate binding domain of DnaK complexed with a heptapeptide substrate, NRLLTTG \cite{29}, reveals a β sandwich composed of two sheets of four strands each, followed in the primary structure actions by two α helices, A and B, which span back over the sandwich (Fig. 17A). Along with loops flanking them at either side, they are stabilized by critical contacts with the overlying helix B. This helix B may function as a lid in permitting the entry and release of substrates (without directly contacting it) (Figs. 17A and 17B). The peptide in the substrate binding pocket pierces the narrow dimension of the domain and is contacted by DnaK only through its central five residues (Figs. 17C and 17D). The hydrophobic side-chain contacts between the loops and it has been seen that the three central leucines in the peptide (L3-5, Fig. 17D) makes the most extensive contacts. The central-most leucine side chain dwells in a hydrophobic pocket in the floor of the channel. It is surmounted by a hydrophobic “arch” that also interacts with the hydrophobic side chains of the neighboring leucine residues in the peptide. In between the main chain of the peptide, and the main chain of DnaK, a total of seven hydrogen bonds can be observed. These contacts, together with the hydrophobic specificity-determining pockets, dictate the requirement for an extended conformation of the bound peptide. The interactive surface of DnaK at the ends of the hydrophobic channel is negatively charged. This negatively charged interactive surface favors the presence of basic residues at the end positions of the peptide (e.g., arginine at position 2). In order that the peptide binds DnaK or is enclosed within it, the interacting polypeptide segment should be separated from the remainder of the
Fig. 17. Substrate Binding Domain of DnaK in Complex with a Peptide Substrate. Ribbon diagram and space-filling models of the substrate binding domain in standard view (A and B) and views rotated 90° counterclockwise around the vertical axis of the standard view (C and D). In (A), the peptide substrate is shown in blue, the strands (1, 2, 4, and 5) and loops (L1,2, L3,4, L4,5, and L5,6) of DnaK constituting the upper β sheet in green, and the hinge-forming residues Arg536–Gln538 of helix B, which may allow ATP-dependent opening of the helical lid (arrows in [A]), in orange. Helices, strands, and loops are labeled; the large N and C indicate the NH2- and COOH-termini of the DnaK fragment. (B) shows the buried nature of the bound peptide substrate in space-filling representation: peptide in blue, DnaK in green, and side chains contributing to the interaction of helix B with the loops of the upper β sheet in pink. (C) is the rotated version of (A), colored similarly; here, N and C refer to the NH2 and COOH termini, respectively, of the bound peptide. (D) shows details of the peptide binding cavity of DnaK: peptide backbone in blue, side chains of the peptide as yellow sticks, DnaK backbone in interacting region in green, and side chains of DnaK residues interacting with the peptide as red van der Waals spheres. (Modified from Rüdiger et al., 1997a).

substrate protein by 10 Å or more, implying that the bound region of the polypeptide must be significantly unfolded. A consensus motif recognized by DnaK in substrate polypeptides has been identified by screening a library of peptides derived from known protein sequences. This consensus motif comprises of a hydrophobic core of 4-5 residues flanked by basic residues. The hydrophobic cores of individual peptides recognized by DnaK contain typically 2-4 hydrophobic residues, with Leu the most common, present in ~90% of recognized peptides. Acidic residues are excluded from the cores and disfavored in the flanking regions. Such a motif occurs frequently within protein sequences, every 36 residues on average, and localizes preferentially to buried β strands of the corresponding folded proteins. The motif identified corresponds remarkably well with the observed features of interaction of the NRLLLTG peptide from the structural study.

Association between the ATP-ase Cycle and Substrate binding of DnaK

ATP binds the NH2-terminal domain of Hsp70. This ATP bound amino terminal domain is used to drive conformational changes in the carboxy-terminal peptide
binding domain of DnaK. The ATP binding is known to alter the affinity of Hsp70 for its substrates (Fig. 18). The binding of ATP increases the dissociation constant for Hsp70–substrate complexes between 5- and 85-fold as a result of increases in $k_{off}$ of 2-3 orders of magnitude, coupled with increases in $k_{on}$ of ~50-fold. The ATPase cycle of Hsp70 can thus be viewed, in its simplest form, as a swinging between two states: the ATP-bound state, which has a low affinity and fast exchange rates for substrates. The ATP bound state has the substrate binding pocket open. The second is the ADP-bound state, with high affinity and slow exchange rates for substrates. In this case, the substrate binding pocket remains closed. The rapid association of Hsp70 with substrates can only occur in the ATP state, because substrate binding to the ADP state is too slow on the time scale of folding reactions. Dissection of ATP binding reveals that it occurs in two steps. The first step is the rapid formation of a weak complex, followed by a slower structural rearrangement, leading to an overall $K_d$ for ATP in the sub-micromolar range. The second step probably reflects the precise locking in of the nucleotide in the binding pocket which is essential for hydrolysis. This is kinetically coupled to the release or exchange of a previously bound poly-peptide. The subsequent conversion of the ATP-
peptide-Hsp70 ternary complex to the ADP state as a result of ATP hydrolysis, then stabilizes the chaperone-peptide interaction. The importance of this step has been demonstrated by an interesting finding where it has been shown that mutant Hsp70 proteins, arrested in the ATP bound state due to defects in hydrolysis, are completely deficient in chaperone activities^{45}, which proves that in order that DnaK functions as a chaperone the ATP hydrolysis step is indelible.

Hydrolysis of ATP is the rate-limiting step in the ATPase cycle of Hsp70 proteins in isolation^{24,41,46,47} and this likely results in dramatic conformational changes in Hsp70 that convert it to the high affinity, slow exchange state, which sequesters substrate protein. The final step in the ATPase cycle is the release of ADP and Pi from ATP. This does not induce any detectable conformational changes but allows the subsequent rapid binding of ATP, which restabilises the starting point. Although nucleotide exchange is 10-20-fold faster than the rate of hydrolysis in the unstimulated cycle^{41,47}, it can become rate-limiting when the hydrolysis step is stimulated by co-chaperones.

Structure of the ATPase Domain of Hsp70
ATPase domain of DnaK consist of two large, globular subdomains (I and II), separated by a deep central cleft. The central cleft is connected by two crossed \(\alpha\) helices (Figs. 18B and 18C). Both the subdomains as well as the connecting helices contribute to forming the binding pocket for nucleotide and the required Mg\(^{2+}\) and K\(^+\) ions at the bottom of the cleft^{26}. The nucleotide interacts with two phosphate-binding loops (loop 1 and loop 2 shown in Figure 18C) and a hydrophobic adenosine binding pocket^{26}, together with contacts with the Mg\(^{2+}\) ion, which is coordinated by several side chains. Such an interaction positions it in the active site. Structural rearrangement of Hsp70 during ATP binding leads to adjustment of the position of the \(\gamma\)-phosphate and a bidentate complex is formed between the \(\beta\)- and \(\gamma\)-phosphate oxygens and Mg\(^{2+}\), permitting an in-line attack by a water (or OH\(^-\)) that is hydrogen-bonded to Lys-71 (Fig. 18C). Precise geometry of the nucleotide and the surrounding residues requires the correct positioning of the Mg\(^{2+}\) ion, established in part by the binding of two K\(^+\) ions nearby. This accounts for the absolute requirement of K\(^+\) for ATP hydrolysis and chaperone activity of Hsp70 proteins^{40}.

### Coupling between the ATPase and Peptide Binding Domains

The molecular mechanism by which the chemical energy of ATP is used to perform mechanical work, in the form of the opening and closing of the substrate
binding pocket, is not yet clearly understood. The available atomic structures of the ATPase domain do not indicate large-scale motions in response to either nucleotide binding or hydrolysis. There are two possibilities: either there are subtle conformational changes in the ATPase domain that are amplified to produce dramatic changes in the rest of the chaperone or the crystallized ATPase domain fragments do not reflect conformational changes occurring in the full-length protein in response to the nucleotide. Supporting the later possibility, biochemical demonstration of nucleotide-dependent conformational changes in Hsp70 proteins requires the physical connection of the NH2-terminal ATPase domain with the adjacent substrate binding domain. In particular, several studies suggest that ATP binding triggers an association of the ATPase domain with the substrate binding domain and that this causes further conformational changes within the substrate binding domain itself, although the precise changes that open the substrate binding pocket remain unknown. Hendrickson and coworkers proposed a structural basis for the ATP-induced opening of the binding pocket. One structure has a kink or a hook at residues 536-538 of the lid-forming \( \alpha \) helix B (indicated by orange segment in Fig. 17A), and consequently, the subdomain COOH-terminal to the substrate binding domain has rotated upwards by \( 11^\circ \), causing a loss of stabilizing contacts between \( \alpha \) helix B and the outer loops. It has been proposed that this represents the initial stages of release by a “latch” mechanism and that further movement of the “lid” opens up the substrate binding pocket and triggers substrate release.

Control of the Hsp70 ATPase Cycle by Co-chaperones

The steady-state turnover rate of the unstimulated Hsp70 ATPase is too slow (between 0.02 and 0.2 min\(^{-1}\)) to drive the chaperone activities of Hsp70, even in the presence of substrates, which typically stimulate the ATPase activity 2-10-fold. Therefore, the presence of certain regulatory mechanisms becomes necessary, which will increase ATP turnover and, hence, chaperone function (Fig. 18A). ATP hydrolysis is the prime target for regulation, mainly by members of the DnaJ family, found in all Hsp70-containing compartments of prokaryotic and eukaryotic cells, as well as in several tumor viruses. DnaJ proteins are a heterogeneous group of multi-domain proteins. They are defined by a highly conserved domain of ~80 amino acids, the J domain, often located near the NH2 terminus, which is essential for the stimulation of the Hsp70 ATPase activity.
Regulation of release of ADP and Pi from Hsp70 is also essential for some homologs, such as bacterial DnaK and mitochondrial Ssc1p, and is accomplished by members of the GrpE family. Association of GrpE with DnaK-ADP accelerates nucleotide exchange 5000-fold, reducing the affinity of DnaK for ADP by 200-fold. For the DnaK system, GrpE and DnaJ together stimulate the ATP turnover rate at least several hundred-fold at saturating conditions, which may be more than is necessary to support chaperone function. Hence it becomes necessary that the effects of DnaJ and GrpE be balanced to optimize the equilibrium between substrate binding and release; and this is achieved in vivo by coregulation of expression of their genes.

The chaperone cycle of Hsp70

The key to the entire functional cycle of Hsp70 proteins lies in the mechanism by which the action of DnaJ and GrpE proteins couples the regulated ATPase cycle of Hsp70 with productive substrate binding. Such coupling prevents stimulating futile ATPase cycles in the absence of substrate. In the case of the DnaK-DnaJ system, it has been observed that in addition to binding to DnaK, DnaJ itself associates with substrates of the DnaK system with kinetics fast enough to prevent their aggregation. This is achieved possibly by binding to a sequence motif similar to that recognized by DnaK. Furthermore, the efficiency with which DnaJ stimulates the ATPase activity of DnaK is strongly increased by the presence of polypeptide substrates. These data suggest a model of the functional cycle of the DnaK system: (Fig. 19) (1) The cycle starts with the transient and rapid association of DnaJ with a substrate, followed by transfer of the substrate to the ATP form of DnaK (K-ATP). This transfer is coupled to the locking-in of the substrate in the substrate-binding pocket of DnaK by ATP hydrolysis. Following substrate transfer, DnaJ leaves the complex, and GrpE associates with the DnaK-substrate complex to trigger ADP release from DnaK. This allows binding of ATP and subsequent release of GrpE and substrate from DnaK.
second is the transfer of the substrate protein from DnaJ to the open substrate binding pocket of DnaK-ATP. Both these steps together are required to stimulate ATP hydrolysis by DnaK, resulting in stabilization of the DnaK substrate complex and tightly coupling ATP hydrolysis to substrate binding by DnaK.

(3) Upon substrate transfer to DnaK and conversion of DnaK to the ADP state, the affinity of the DnaK–substrate complex for DnaJ gets reduced. This leads to the dissociation of DnaJ from the complex.

This step is reflected in observations that ternary DnaK–DnaJ–substrate complexes are unstable and that DnaJ acts catalytically in targeting DnaK to substrates (4) GrpE is the nucleotide exchange factor and hence it binds to the DnaK–ADP–substrate complex, thereby triggering the release of ADP. Consequently, (5) ATP binds rapidly to DnaK, which releases the bound substrate and GrpE and returns DnaK to its initial state.

The Hsp60 chaperonins
Chaperonins are the most fascinating group of cellular machinery which utilise ATP binding and hydrolysis to drive ordered conformational changes. These are fascinating group of devices made up of collective of double-ring assemblies that promote folding of proteins to the native state. This complex is composed of back-to-back rings of identical or closely related rotationally symmetric subunits (~60 kDa), and play an essential role in all cells, assisting a large variety of newly synthesized and newly translocated proteins to reach their native forms by binding them and facilitating their folding inside a large central channel within each ring. The central cavity of each heptameric ring, the worksite of the machine, functions in two major states. In the binding-active state, it is open at the end of the cylinder for ingress of...
Protein Folding

nonnative proteins, exposing a flexible hydrophobic lining that likely binds nonnative species through exposed hydrophobic surfaces (which will become buried to the interior in the native state) (Fig. 20). Binding is most likely multivalent in character, with the substrate protein contacted simultaneously by many of the chaperonin apical (end) domains surrounding the channel. Regardless of how it becomes bound, a captured nonnative protein is conferred the extraordinary opportunity to reach the native state after release from the binding sites when the machine proceeds to its other state, the folding-active state. The folding-active state is reached by conformational changes in the chaperonin, induced by the action of ATP binding and, for the organelar/bacterial chaperonins, by binding also of a lid-like co-chaperonin, GroES, itself a ring of rotationally symmetric subunits (each of ~10 kDa) that trigger folding of the substrate protein encapsulated in the central channel. The conformational change of the bacterial chaperonin, GroEL, upon binding its co-chaperonin, GroES, initially detected by Electron Microscope and recently resolved crystallographically at 3 Å resolution is a molecular spectacle, with dramatic en bloc movements of the seven apical domains in the ring bound by nucleotide and co-chaperonin, resulting in a global change of the shape and character of the central cavity. The contact with GroES is mediated through mobile β-hairpin loop segments, one from each GroES subunit, that extend downward and laterally to make contact with a corresponding GroEL apical domain. The contact is in part hydrophobic in character, involving interaction between several of the GroEL apical hydrophobic residues required for polypeptide binding and a sequence, Ile-Val-Leu, in one “edge” of the GroES mobile loop.

The cavity of the bound ring enlarges 2-fold in volume and is closed off at the open end by the dome-shaped GroES ring, encapsulating/sequestering the nonnative protein (Fig. 21A); the hydrophobic binding surface is elevated and twisted away from the polypeptide, releasing it into the cavity; and the aspect of the apical domains now forming the cavity surface is hydrophilic in character (Fig. 21B), favoring burial of hydrophobic surfaces in the folding substrate protein and exposure of its hydrophilic surface, thus acting to promote the native state. Both the seven-valent, subunit-to-subunit contacts between GroES and GroEL and the new supporting interfaces act to stabilize the opened-up, folding-active conformation of the apical domains. Although nucleotide binding alone, in the absence of GroES, has not been observed to produce this extent of apical movement, it must be capable of
transiently driving the full or nearly full extent of these changes, which are then stabilized by GroES binding.

**ATP Binding and hydrolysis in GroEL**

Given the asymmetry of the chaperonin machine in the presence of its ligands, it has seemed likely that ATP would have distinct actions in \textit{cis} (GroES-bound) and \textit{trans} rings. These roles in the formation and dissolution of a folding-active \textit{cis} complex (Fig. 22) have been ATP turnover experiments and the analysis of single ring mutants and mutant rings able to bind but not hydrolyze ATP. It was observed\textsuperscript{78,79} that GroES and bound \textit{cis} ADP were rapidly discharged upon exposure of such asymmetric complexes to ATP, indicating that ATP binding or hydrolysis in the \textit{trans} ring sends an allosteric signal that evicts the ligands from the \textit{cis} ring (Fig. 22, panel 6). Consistent with this interpretation, a single-ring version of GroEL bound GroES in the presence of ATP but could not release GroES or refolded protein, apparently resulting from the failure to receive a signal from the non-existent \textit{trans} ring\textsuperscript{69,71}. While the requirement for ATP in the \textit{trans} ring was recognized early, a specific requirement for ATP in the \textit{cis} ring has only recently been uncovered through observations that single-ring, obligatorily \textit{cis} versions of GroEL can productively fold “stringent” substrates, such as RUBISCO from \textit{Rhodospirillum

Protein Folding

rubrum or mitochondrial malate dehydrogenase (MDH), in GroES and ATP. In addition, earlier kinetic observations that ATP is the preferred nucleotide for GroES binding suggest that, under physiologic conditions, the majority of GroEL-bound protein is triggered to commence folding by formation of GroES-ATP \( \text{cis} \) ternary complexes (Fig. 22, panel 4). It was unclear whether \( \text{cis} \) folding was triggered by ATP-GroES binding or by subsequent \( \text{cis} \) ATP hydrolysis. This was resolved by study of a mutant single ring complex bearing the intermediate domain substitution, D398A, that reduces ATP hydrolysis to \( \approx 2\% \) of wild type. Conversion of bound substrate polypeptide to the native state occurred inside the D398A-GroES complex only in the presence of ATP (and not AMP-PNP), with the same kinetics and to the same extent as in hydrolysis competent single and double rings, despite the absence of ATP hydrolysis on this time-scale. Thus, ATP-GroES binding, and not hydrolysis, is required to produce the native state. Remarkably, folding-active \( \text{cis} \) ATP-GroES-GroEL is a very stable complex that even resists dissociation by chaotrope. Thus, the energy of the \( \gamma \)-phosphate and its contacts are employed to assure the stability of the folding-active environment. Paradoxically, however, this complex is so stable that the refolded substrate remains locked up inside it. It seemed likely that hydrolysis of \( \text{cis} \) ATP was the next step forward, based on the hypothesis that it would relax the high affinity interaction between GroES and GroEL (Fig. 22, panel 5). This was supported by observation of the reduced stability of de novo-formed \( \text{cis} \) ADP versus ATP complexes and the lability of stable ATP complex attendant upon hydrolysis of \( \text{cis} \) ATP. Thus, hydrolysis of \( \text{cis} \) ATP acts to ‘prime’ the \( \text{cis} \) ring for releasing GroES upon receiving the eviction signal sent by ATP in the \( \text{trans} \) ring (Fig. 22, panels 5 and 6). The observation that binding of ATP/GroES, but not AMP-PNP/GroES, triggered folding in the \( \text{cis} \) ring raised the issue of whether it is likewise ATP binding acting in \( \text{trans} \) that evicts the \( \text{cis} \) ligands (Fig. 22, panel 6). Once again, the hydrolysis-defective D398A ring was used, and binding of ATP, but not AMP-PNP, to such a \( \text{trans} \) ring evicted the \( \text{cis} \) ligands. Thus, binding of the nucleotide and not hydrolysis triggers the major work on the substrate protein: \( \text{cis} \)-bound ATP/GroES triggers release of substrate into the cavity followed by productive folding; \( \text{trans} \)-bound ATP triggers discharge of GroES and the substrate protein (Fig. 22). Hydrolysis steps are used to drive the machine, directionally, to the next state. \( \text{Cis} \) hydrolysis relaxes the high affinity interaction of bound GroES for GroEL, “priming” the \( \text{cis} \) ring for release of GroES and substrate protein, but the
Fig. 22. Model for a GroEL-GroES-Mediated Folding Reaction. The asymmetric GroEL-GroES complex is the likely polypeptide acceptor state in vivo and binds unfolded polypeptides (U) or kinetically trapped folding intermediates (Iuc) to form a trans ternary complex (second panel). This complex is highly dynamic with respect to GroES binding in the presence of ATP (third panel); two possible pathways of GroES rearrangement that lead to the cis complex are shown. When GroES binds to the ring containing polypeptide in the presence of ATP (forming the folding active cis intermediate, fourth panel), major conformational changes occur in the cis GroEL ring, polypeptide release from the apical binding sites is triggered, and folding commences. Hydrolysis of ATP in the cis ring weakens the GroEL-GroES interaction (fifth panel), priming GroES release, while polypeptide folding continues. Binding of ATP in the trans ring evicts GroES from the cis ring (last panel), giving polypeptide the opportunity to depart. The released polypeptide is either native (N) or committed to fold (Ic), or is in an uncommitted or kinetically trapped state (Iuc), which can rebind to the same or a different GroEL complex and undergo another cycle of folding. In the complexes, D designates ADP, T, ATP; and T-D, ATP hydrolysis.

precise action of trans hydrolysis is less clear at this point. One possibility is that ATP binding in trans may enable a second GroES molecule to bind to GroEL. If a polypeptide has been bound in the trans ring, then this would, at the same time as discharging the cis complex on the opposite ring, produce a new cis ternary complex. Such a step would reflect a “two-stroke” action of the GroEL machine, and ATP hydrolysis would always occur in the context of a cis complex, proceeding alternately on one ring, then the other. While such a model seems economical, several results argue against it. In one experiment, trans ternary complexes were formed in ADP with GroES and a substrate protein, ornithine transcarbamylase (OTC). Following addition of ATP and additional GroES, as well as a “trap” mutant of GroEL, able to bind but not release nonnative substrate, no OTC enzymatic activity was recovered. This implies that, instead of GroES binding to the OTC-containing ring and forming a productive cis complex, OTC was released and trapped. In a second experiment, the kinetics of folding by a mixed-ring GroEL complex, able to bind substrate and GroES on only one of its rings, was measured and found to be the same as wild-type GroEL. Here, if GroEL were functioning through a two-stroke mechanism, the wild-type would have
been expected to produce more rapid recovery of activity than the mixed-ring complex. The timing of the sequence of binding and hydrolysis of ATP in the *cis* and *trans* GroEL rings has obvious influence on the fate of substrate polypeptide (Fig. 22).

At 23 °C, polypeptide is released into the central channel within a second of the binding of ATP and GroES to form a *cis* complex. The substrate then has ~20 s in this favorable environment and, after *cis* hydrolysis, ~10 s in an ADP state, following which GroES is discharged by binding of ATP in the *trans* ring. From dynamic measurements of substrate fluorescence as well as activity, it seems that there is a seamless transition of the folding reaction between *cis*-ATP and ADP, with no discernible phases, even though in structural terms, the apical domains are likely to occupy different conformations. Thus, while the initiation of folding of stringent substrates cannot be triggered by ADP–GroES, the progression of folding, once release has been driven by ATP–GroES binding, does not appear to be affected by conversion to the ADP state. A lifetime for the folding-active state of ~20 s may be sufficient for many substrates to reach GroEL the native state, but for others, only a fraction of the molecules reach native form before the “timer” goes off and GroES is discharged from the *cis* ring, ending the lifetime of the favorable environment. Along with GroES departure, both native and nonnative molecules have been observed to leave the *cis* cavity (Fig. 22).

**A hypothesis for folding by Hsp70 and Hsp60 machinery**

Recent findings for protein folding in mitochondria may suggest a principle mechanism for the acquisition of tertiary structure *in vivo*. Division of labour between molecular chaperones of the Hsp70 and Hsp60 families appears to be an integral element. Although members of both groups of stress proteins interact with unfolded or incompletely folded polypeptides and utilize the energy of ATP hydrolysis for releasing the bound substrates, important functional differences are becoming apparent. Physiological folding of a newly synthesized polypeptide may occur by the following three-step reaction consisting of (a) protection and maintenance of folding competence by Hsp70, (b) ATP-dependent transfer from Hsp70 to Hsp60, and (c) ATP-dependent folding at Hsp60 (Fig. 23).

**a) Maintenance of Folding Competence by Hsp70**

Components of the Hsp70 family can associate with the folding polypeptide before those of the Hsp60 family, suggesting a hierarchy of interaction with
molecular chaperones. For example, the mitochondrial Hsp70 binds to the extended protein sequences reaching into the matrix while they are still in the process of membrane translocation. Similarly, cytosolic Hsp70 is thought to interact with nascent polypeptides emerging from the ribosome.

The main purpose of the interaction of a nascent polypeptide chain with Hsp70 would be to prevent premature (mis)folding and aggregation of proteins (or their independently folding domains) during synthesis. Binding of Hsp70 could prevent or retard hydrophobic collapse of the molecule, and the emerging chain would transiently be held in a relatively extended conformation with little stable secondary structure. The initial binding may require the cooperation between Hsp70 and a DnaJ-homologous component. Because of its slow, ATP-dependent release activity, bound Hsp70 would fall off the substrate perhaps with kinetics adjusted to the speed of protein synthesis. This would allow the internalization of hydrophobic regions to a molten globule-like state now stabilized by fewer Hsp70 molecules.

The interaction with Hsp70 alone appears not to be sufficient to promote the formation of ordered tertiary structure, at least in the case of proteins imported into the mitochondrial matrix. The role of Hsp70 in protein folding could mainly consist of stabilizing the newly synthesized chains in loose conformations competent for the folding process mediated by a further component(s) or for membrane translocation (so-called "folding-competent" and "translocation-competent" states). Alternatively, the stabilization of nascent chains followed by cycles of ATP-dependent release and rebinding of Hsp70 may be sufficient for the folding of a subset of total soluble proteins that might contain sequence elements supporting the productive interaction with Hsp70.

b) Transfer from Hsp70 to Hsp60
Folding in mitochondria appears to depend on the newly imported proteins being transferred from Hsp70 to Hsp60. Very little is known about this transfer reaction. In light of the finding that binding of Hsp70 is necessary for membrane translocation, while Hsp60 function appears to be of little importance at this stage, transfer to Hsp60 should occur when import of the protein (or its domains in case of larger polypeptides) is complete. Other mitochondrial components, perhaps similar to the N-Ethylmaleimide (NEM) sensitive factor contained in reticulocyte lysates, may have to cooperate with Hsp70 to ultimately allow its complete release from the imported protein. Natural candidates could be mitochondrial homologues of E. coli DnaJ and GrpE that are known to physically interact with the Hsp70 DnaK.
DnaK is dispensable at intermediate growth temperatures around 30 °C (although the cells grow very poorly) and becomes essential for growth at temperatures below and above 30 °C. At present, all results would be compatible with DnaK having a critical buffer function in protein folding, holding and protecting nascent chains until they can be taken over by GroEL. Under non-heat-shock conditions, cells would contain fewer GroEL molecules than nascent chains; the basal level of the GroEL 14-mer is about 15 μM, while the concentration of nascent chains is 30-50 μM and that of DnaK is 100μM.

Furthermore, GroEL may bind nascent chains only at a later stage of synthesis than DnaK and would therefore be unable to efficiently prevent their aggregation. This lack of protection would become more critical at higher growth temperatures. So it seems that DnaK, DnaJ, and GrpE cooperate with GroEL/GroES in the de novo folding of proteins.

c) Hsp 60-Mediated Folding

Evidently, Hsp70 cannot assume some specific functions of Hsp60 in protein folding. The chaperonins appear to mediate folding by a process of step-wise release while the folding polypeptide...
remains sequestered at their surface\textsuperscript{92,99}. Surface-mediated folding may critically depend on an oligomeric machinery in which the individual Hsp60 monomers contribute binding sites for polypeptide segments whose action would be coordinated in the oligomeric chaperonin complex. Although both mitochondrial Hsp60 and \textit{E. coli} GroEL/GroES fulfill essential functions under all growth conditions\textsuperscript{100}, it is unclear what percentage of total proteins depend on chaperonins for folding. Given the importance for protein folding assigned to Hsp60, one would expect that chaperonin-like components were present in every cellular compartment capable of de novo folding.

**Antibacterial Peptides**

One of the most serious emerging health concerns today is the appearance and spread of antibiotic-resistant bacterial strains\textsuperscript{101}. It is thus becoming increasingly important to identify antimicrobial compounds with novel modes of action for which the bacteria are unable to mount a quick response and to build resistance. Perhaps the most promising among these novel compounds is a family of antibacterial peptides originally isolated from insects\textsuperscript{102-106}. Antibacterial peptides are the effector molecules with innate immunity. Generally they contain 15-45 amino acid residues and the net charge is positive. The Table 4 lists the related peptide sequences. Some of these act in a stereospecific manner on a target bacterial protein\textsuperscript{104,107,108}. In contrast, studies on model membranes as well as on live bacteria have indicated that other types of antibacterial peptides provoke an increase in plasma membrane permeability\textsuperscript{109}. A direct correlation between antibiotic effect and membrane disruption has been found for defensins from mammals and insects, magainins, and cecropins\textsuperscript{110-113}. However, such antibacterial peptides that lyse bacterial membranes are potentially toxic to eukaryotic cells and are, therefore, unsuitable as a systemic drug. This highlights the potential of Drosocin, Apidaecin, and Pyrrhocoricin. None of these peptides lyse sheep erythrocytes, and in addition, Pyrrhocoricin also appears to be completely nontoxic to COS cells of primate origin\textsuperscript{114}. Significantly, Pyrrhocoricin and one of its analogues, Chex-pyrrhocoridn-Dap(Ac) designed to withstand protease cleavage, remain nontoxic \textit{in vivo} and protect mice from live \textit{Escherichia coli} challenge\textsuperscript{114}. Some antibacterial peptides are known to act as inhibitors of enzymes produced by the bacteria either by serving as a pseudo-substrate or by tight binding to the active site eliminating the accessibility of the native substrate\textsuperscript{109}. For example, histatins are capable of inhibiting a trypsin-like protease from \textit{Bacteriocides gingivalis} with an IC50 in the submicromolar range\textsuperscript{115}. Microbial serine protease(s) are also
inhibited by the equine version of the peptide NAP-2. Other peptides can control yet additional proteinases involved in inflammatory processes, such as the inhibition of furin by histatin 5. There are significant similarities in the mechanism of action between the latter case and the Pyrrhocoricin-Drosocin-Apidaecin family. The bioactive secondary structure of Drosocin has been suggested to comprise of two reverse turns, one at each terminal region, which constitute the binding sites to the target molecule. The general fold of native Pyrrhocoricin, as determined by nuclear magnetic resonance spectroscopy (NMR) and circular dichroism spectroscopy, is similar. Also for Pyrrhocoricin, reverse turns are identified as pharmacologically important elements at the termini, bridged by an extended peptide domain. The recognition mechanism of histatin 5 involves a well-defined bioactive conformation of this 24-mer peptide, as was demonstrated by its selectivity to other proprotein convertases. In addition, the all-D-peptide remained inactive, just like for Drosocin, Pyrrhocoricin, and Apidaecin. The antibacterial peptide Pyrrhocoricin, shown in Table 4, is a proline-rich antimicrobial peptide originally isolated from the insect pyrrhocoris apterus. It has been proposed that Pyrrhocoricin binds predominately to the multi-helical lid sub-domain of DnaK (Fig. 24), thereby preventing DnaK’s lid from opening and closing, a function that is deemed essential for its chaperon activity. This proposition that Pyrrhocoricin functions by binding to the lid domain has however been contested in another study, where it has been argued that

<table>
<thead>
<tr>
<th>Antibacterial peptides</th>
<th>Amino acid sequences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drosocin</td>
<td>GKPRYPSPRP TSHPRPIRV</td>
</tr>
<tr>
<td>Formaecin</td>
<td>GRPNPVNN KPTYPYPHL</td>
</tr>
<tr>
<td>Pyrrhocoricin</td>
<td>VDKGSYLRPR TPRPIYNRN</td>
</tr>
<tr>
<td>Dipterisin</td>
<td>DEKPKLIPLPT PAPPNLPO</td>
</tr>
<tr>
<td>Apidaecin</td>
<td>GNNRP VYIP QRPHPRI</td>
</tr>
</tbody>
</table>

![Fig. 24. Characteristic structures of Pyrrhocoricin and C terminal region of E. coli DnaK generated by flexible docking process. The blue-white domain corresponds to the peptide binding cavity, orange-green domain corresponds to the multi-helical lid, and the purple structure corresponds to Pyrrhocoricin.](image)
Protein Folding

Pyrrhocoricin interacts mainly with the substrate binding domain of DnaK. Despite such contradictory claims, there is no disagreement about the fact that Pyrrhocoricin does interact with DnaK, and that this interaction is linked at least partly to its bactericidal effects. The interaction of Pyrrhocoricin, Drosocin and Apidaecin, with the bacterial heat shock protein, DnaK, identifies DnaK as a convenient target for drug design. Considerable amount of light has been thrown on the structure and function of the DnaK protein. Significantly, DnaK is essential for bacterial growth at most temperatures. The C terminal 10 kDa region appears to carry an important regulatory function, with many individual amino acid residues being involved in the various bioactivities of the protein. Currently, the heat-shock family of proteins is considered as suitable targets for development against many diseases including cancer.

References:


