Research Communication

Mycobacterial Stress Regulation: The Dps "Twin Sister" Defense Mechanism and Structure-Function Relationship

Rakhi Pait Chowdhury*, Ramachandran Saraswathi* and Dipankar Chatterji
Molecular Biophysics Unit, Indian Institute of Science, Bangalore, India

Summary

In this work, we have tried to emphasize the connection between mycobacterial growth and regulation of gene expression. Utilization of multiple carbon sources and diauxic growth helps bacteria to regulate gene expression at an optimum level so that the inhospitable conditions encountered during nutrient depletion can be circumvented. These aspects will be discussed with respect to mycobacterial growth in subsequent sections. Identification and characterization of genes induced under such conditions is helpful to understand the physiology of the bacterium. Although it is necessary to compare the total expression profile of proteins as they transit from vegetative growth to stationary phase, at times a lot of insights can be deciphered from the expression pattern of one or two proteins. We have compared the protein expression and sigma factor selectivity of two such proteins in *M. smegmatis* to understand the differential regulation of genes playing diverse function in the same species. Some newer insights on the structure and function of one of the Dps proteins are also explained. **© 2009 IUBMB**
IUBMB *Life*, 62(1): 67–77, 2010

Keywords Dps; transcriptional regulation; DNA binding; interface cluster analysis; stress response.

INTRODUCTION

Mycobacterial Stationary Phase Physiology

Mycobacterium smegmatis shows a distinct phenotype of stationary phase adaptation. Previous work in this direction includes the establishment of the 0.02% glucose growth conditions as a starvation model (*1*). In this state *M. smegmatis* shows striking differences in colony morphology, growth rate and cellular appearance, as compared with the bacterium grown under nutrient-rich conditions. Figure 1 shows the alteration in the colony morphology and growth-curve upon starvation of *Mycobacterium smegmatis* (*1*). As shown in the figure, starved *M. smegmatis* cells form characteristic shiny smooth and small colonies. The rate of growth of the culture also reduces drastically when cells are starved as opposed to the cells grown in carbon-rich media containing 2% glucose as the carbon source.

Apart from the above directly visible changes, *M. smegmatis* also exhibits an increased resistance to acid, osmotic stress, and oxidative stress (*2*). The phenomenon of GASP (growth advantage in stationary phase), which has been well studied in *E. coli*, is also exhibited in *M. smegmatis*. The GASP mutations, as the name implies, confer a growth advantage under stationary phase conditions, to certain members of a population, which allows them a competitive edge over others within the same population (*3*). This also correlates with a view of stationary phase bacterial cultures as being dynamic and genetically diverse, as opposed to static and homogenous populations (*4*). Lately, similar observations have been made for *M. smegmatis* cultures. The diversity inherent in a stationary phase population seems to be an important mechanism enabling the bacterium to survive stress. An analysis of gene expression pattern and variable colony morphology in *M. smegmatis* cultures has shown that the cultures, in the stationary phase too, are physiologically heterogeneous and consist of subpopulations of cells that vary in their transcriptional profiles (*5*). The discovery of a new polar glycopeptidolipid in starved cultures of *M. smegmatis*, correlates with the smooth colony morphology upon starvation and points to a novel mechanism of stationary phase adaptation (*1*).

Stress-Induced Expression of Genes in Mycobacteria

In real conditions of growth bacteria often experience starvation (depletion of carbon source or nutrients) and hypoxia (lack of oxygen) invading the host macrophages. It is indeed important therefore to look at several mechanisms by which the bacilli manage the uptake of carbon source, mostly glucose,

Received 19 September 2009; accepted 6 November 2009
*Both the authors have contributed equally to the manuscript.
Address correspondence to: Dipankar Chatterji, Molecular Biophysics Unit, Indian Institute of Science, Bangalore-560012, India. Tel: +91-80-2293-2836. Fax: +91-80-2360-0535.
E-mail: dipankar@mbu.iisc.ernet.in

ISSN 1521-6543 print/ISSN 1521-6551 online
DOI: 10.1002/iub.285

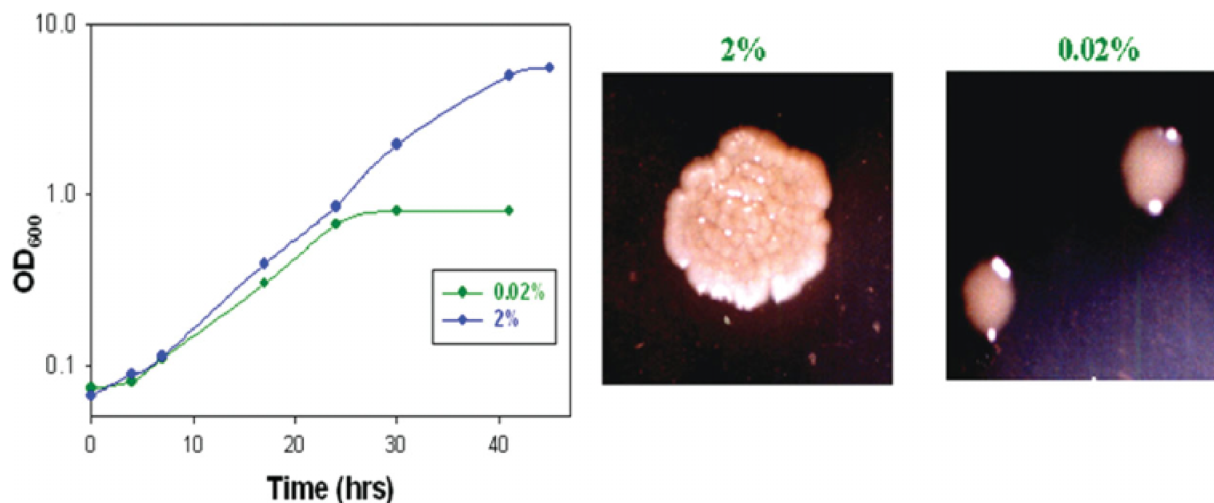

Figure 1. *Mycobacterium smegmatis* shows different colony morphology and growth pattern upon carbon starvation. The figure is taken from (*1*). [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

during the stationary phase growth. With the advent of total genome sequencing of mycobacterial species, DNA microarrays and transcriptome analysis, we are now at the verge of looking at this aspect of dormancy by exploring the expression and regulation of individual genes induced under explicit stress conditions. *Mycobacterium smegmatis*, the nonpathogenic homologue, has been taken as a universally accepted surrogate host to study *M. tuberculosis* gene expression and regulation. In this report our main aim will be to emphasize the expression and regulation of the genes in mycobacteria which may help in the survival of the organism. We have taken a stress induced protein Dps (DNA-binding Protein from Starved cells) from *M. smegmatis* expressed under glucose starvation condition as an example to address the question. Later with the discovery of a second Dps molecule we have found some differences with respect to induced transcription and sigma factor selectivity of the two *dps* genes in *M. smegmatis*. We, at this stage would also like propose a model depicting the differential regulation of the two *dps* genes from *M. smegmatis* and compare it with growth phase dependent sigma factor selectivity. We think that *dps* transcriptional analysis can be utilized as a model to understand the transcriptional regulation of genes in the mycobacterial stress response.

In the case of pathogenic mycobacteria, *M. tuberculosis,* the adaptation of the tubercle bacillus to the host environment is likely to involve a complex set of regulatory events and physiological changes. To identify *M. tuberculosis* components that may be responsible for successful intracellular survival of the bacteria, many individual genes inside the phagosome have been analyzed by microarray technique (*6*). Experiments have been carried out in broth culture, using conditions that may mimic the macrophage environment (*7, 8*). Dissection of regulation of gene expression at the molecular level has been possible due to the advent of tools by which the total transcriptome profile of the microroganism inside macrophage can be addressed (*9-14*). Recent transcriptome analysis of *M. tuberculosis* isolated from lungs of infected mice (*15*) and human tuberculosis patients (*16*) highlighted potential virulence factors expressed *in vivo*.

However, transcriptome analysis examines relative, not absolute, mRNA levels as a function of growth or extracellular conditions. Furthermore, genes whose level of expression does not alter significantly are completely ignored. Differential expression analysis thus provides an incomplete view of the transcriptome and determination of mRNA levels will address this question in a more elaborate way. *In vitro* experimental tools like RT-PCR and single round transcription assays can be of use in this regard (*17*).

Dps-DNA–Binding Protein from Starved Cells

The DNA-binding protein from starved cells is, as the name implies, a protein that is up regulated in cells that are starved of nutrition. It has the capacity to bind DNA. From its initial discovery in the starved stationary phase cultures of *E. coli*, this protein is now known to be primarily important in protecting the cells under conditions of oxidative stress (*18*). Subsequent to the original discovery in *E. coli*, Dps homologues have been found in almost all the bacterial groups, including archaebacteria. In some cases, more than one homologue of the protein is present in the same bacterium and each homologue performs a slightly different function inside the cell (*19*). The presence of this protein across various types of bacteria suggests an important role for this protein in the bacterial physiology. The DNA-binding ability and dodecameric structure are features which are common among many viral and cellular proteins like Beta, RecE, ICP8, and RAD52 (*20-22*). Bacterial Dps is one among them. However, in addition to the DNA-binding ability, Dps is

known to have ferroxidation activity to inhibit Fenton reaction mediated damage of the genomic material.

The binding of Dps to DNA was found to be independent of the sequence, length and topology of the DNA (*23*). Therefore Dps proteins were also classified as nucleoid-like proteins, which bind to DNA in a sequence independent fashion. Thus, Dps proteins seemed evolutionarily at the cross roads between nucleoid-like and ferritin-like proteins. Presently, they are considered to have evolved from the ferritin proteins, which do not bind to DNA. Dps later acquired DNA-binding property and the ability to form dodecamers (*24*). The structural analogy between Dps and ferritin will be discussed in a later section.

Essentially, the major functions of this protein relate to physical shielding and protection of the genome, as well as conferring resistance to the cell against oxidative stress, especially peroxide mediated oxidative damage. Due to its expression under conditions of stress, the Dps protein is considered to be a part of the stress response mechanism in bacteria. However, deletion of this protein does not lead to any significant deleterious effects on the cell. Cells without the *dps* gene show increased sensitivity to oxidative stress but no loss of viability under normal conditions (*23*).

Apart from its conventional role as an agent to protect cells from oxidative stress and DNA damage, various additional functions of Dps have been revealed lately. Dps is involved in a wide variety of roles, including cold shock response (*25*), antigenicity (*26*), biofilm formation, resistance to bacteriophage (*27*), and other functions. Often these other functions of Dps are specific to the organism in which the protein is found and are probably part of the unique adaptations of the particular organism. For instance, Dps is involved in the persistence inside the carrier 'ticks', of the pathogen *Burrelia burgdorferi* (*28*) that causes the Lyme disease. However, it is not present in another pathogen, *Mycobacterium tuberculosis*, which also needs to persist inside a host body. Instead *M. tuberculosis* is rich in ferritin protein. Dps is a glycosylated protein in *Salmonella enterica*, whereas in most other organisms it undergoes no such post-translational modifications (*29*).

Presence of More than One Dps Homologue

The Dps protein is nearly ubiquitous in various bacteria, although some bacteria do not have a Dps homologue. For instance there is no Dps homologue in *Mycobacterium tuberculosis*, which survives for long periods of time in the dormant state inside the host and therefore encounters prolonged starvation and oxidative stress (*30, 31*). However, intriguingly, there are also some species of bacteria which have no Dps at all. In the later case, the presence of multiple mechanisms of protection against oxidative agents and various general stress proteins probably compensate for the absence of Dps in bacteria that lack a Dps homologue. Indeed most bacteria have ferritin and bacterioferritin containing haeme, which aid in iron metabolism and homeostasis. However, the nucleoid-condensation in the bacterial stationary phase is uniquely attributed to Dps. The presence of an alternate protein to perform the same function in bacteria lacking Dps is not known so far. The bacteria that contain two copies of Dps include *Lactobacillus lactis* DpsA and DpsB (*32*), *Deinococcus radiodurans* Dps1 and Dps2 (*33, 34*), *Bacillus subtilis* MrgA and Dps (*35*), *Bacillus anthracis* Dlp-1 and Dlp-2 (*36*) and *Mycobacterium smegmatis* MsDps1 and MsDps2 (*37, 38*). Although these are the only proteins which have been experimentally characterized, the ORFs coding for two Dps in a single genome are found in many more bacteria. Out of 300 Dps molecules found in the bacterial sequences through bioinformatic analysis, 195 are single homologues. The rest of them are found as paralogs within the same organism (*38*). The sequence identity between the two homologues of Dps within the same organism is usually not very high. Therefore, these proteins are not duplicate versions of the same Dps molecule. However, certain biochemical differences do exist between the two Dps proteins from the same organism. For instance, the DpsA protein from *Lactococcus lactis* binds to DNA with greater affinity than the DpsB protein from the same organism. Also, the DpsA protein exhibits the presence of trimeric and dimeric forms apart from the dodecameric form as seen with the DpsB protein (*32*). In *Bacillus anthracis*, the BaDps1 protein utilizes O_2 as the substrate for the ferroxidation reaction. On the contrary, the other Dps homologue BaDps2 utilizes O_2 as well as H_2O_2 in oxidizing iron. It therefore provides greater protection to DNA as compared with BaDps1 (*39*). The BaDps1 and BaDps2 proteins share homology with the Dps and MrgA proteins, which are Dps homologues of *Bacillus subtilis*. BaDps1 is closer to Dps and BaDps2 is closer to MrgA by sequence similarity (*39*). In the case of the mycobacterial MsDps1 and MsDps2, the MsDps1 protein exists as a stable trimer which can convert into the dodecameric form, characteristic of the Dps proteins (*24*). The MsDps2 protein, on the other hand, only occurs as a dodecamer (*38*). From the sequence and structure of the two proteins, MsDps1 shows a long C-terminal tail, similar to the N-terminal tail in *E. coli* Dps, which is rich in positively charged amino acids and is involved in DNA binding (*40*). Interestingly, the MsDps2 protein binds to DNA but lacks any obvious N- or C-terminal DNA-binding extensions (*38*). Although the presence of two paralogous copies of a gene are not sufficient to propose independent functions for the genes, it is notable that a recent publication identified two forms of Dps in *Streptomyces coelicolor* (*19*) that play different functions in DNA metabolism (protection from DNA damage during vegetative growth and nucleoid formation).

Regulation of Dps Protein In Vivo

Dps was discovered as a starved stationary phase induced protein. Thus its expression is regulated to occur only under specific conditions. Upon starvation, two groups of proteins are found to be expressed. The class of *cst* genes or "cAMP -

dependent carbon starvation response" genes are induced through cyclic AMP. On the other hand, the cAMP independent genes are called *pex* or "post exponential genes." Dps is a pex protein, designated as PexB. Its expression in *E. coli* is dependent on the stationary phase specific sigma factor σ^{38} (also referred to as σ^{s}), under conditions of carbon starvation and osmotic stress. It is also transcribed by the housekeeping sigma factor σ^{70} during oxidative stress (*41*). Further studies confirmed that cAMP-CRP can affect σ^{70} levels inside the cells, but not Dps expression directly. The promoter of *dps* in *E.coli* also has an extended −10 element and no −35 element. The recognition of the *dps* promoter by both σ^{70} and σ^{s} is further regulated by OxyR and IHF (*3*). Thus, the expression of Dps is tightly regulated and involves more than one regulator. This implies that the precise roles played by Dps under specific conditions are important for the cellular physiology. Further complexities in the regulation of Dps expression have been observed. In addition to transcriptional control by OxyR and IHF, transcription from the *dps* promoter is also negatively regulated by the nucleoid proteins Fis and H-NS via novel mechanisms. Expression of Dps by the exponential phase sigma factor σ^{70} is negatively regulated by both Fis and H-NS (*42, 43*). This complex network of regulation ensures a tight control over Dps level. Dps expression is kept at minimum in the exponential phase and shoots up in the stationary phase (*42, 43*). Apart from the transcriptional level, Dps expression is also regulated at the post-translational level. Proteolytic cleavage by the ClpXP proteases controls the levels of Dps in *E.coli* in the exponential phase of growth. Parallely, the ClpA protease also enhances Dps expression at the post-translational level in the stationary phase, allowing for a large accumulation of the protein (*44*). An earlier report also showed that under nitrogen starvation conditions, Dps induction occurred at the post-transcriptional level (*41*).

Differential Regulation of the Two dps Genes in *M. smegmatis*

A large number of studies have been done on the structure as well as function of MsDps1 (*24, 37*). The expression profile of the protein inside wild type cells (*mc^{2}155*) as a function of its growth under carbon fed and depleted conditions has been well studied too (*45*). *In vivo* regulation of the protein has established the starvation response of this gene (*17*). Reconstitution of RNA polymerase, the key enzyme performing the transcription reaction, with sigma factor(s) and the core subunits ($\alpha_2\beta\beta'\omega$) from *E. coli* has been well studied (*46–48*). Recent studies through *in vitro* transcription using reconstituted polymerase composed of *M. tuberculosis* sigma factors and *M. smegmatis* core has been utilized in expanding promoter analysis in *M. tuberculosis* (*17*). This tempted us to assemble a mycobacterial heterologous transcription machinery *in vitro*, consisting of a reconstituted holo-RNA polymerase (combining *M. tuberculosis* sigma factors and *M. smegmatis* core enzyme) and a *M. smegmatis* promoter system, as a first attempt. We surmise that the *msdps1* 5′ region would be a suitable promoter system in this context, as its expression is inducible and conditional. This might further assist us to uncover the regulatory systems functional in mycobacteria under starvation. Similar experiments have been performed to expand our knowledge in understanding the regulation and sigma factor specificity of the second Dps molecule, MsDps2, which showed some distinct functional features compared with MsDps1. We, therefore, made an attempt in understanding any balance in the differential regulation of the two Dps proteins in *M. smegmatis*, resulting in divergent functions. The results obtained, so far, with *in vitro* transcription analysis on the two *msdps* promoters, revealed *msdps1* to be recognized by σ^{F} and σ^{H} exclusively (*17*), and not by σ^{A}; whereas *msdps2* was specific to σ^{A} and σ^{B} (unpublished data). Promoters being transcribed by dual sigma factors have been reported in case of *sigB* and *sigI* promoters from *M. tuberculosis* in recent studies (*49, 50*). *In vivo* promoter specific pull down assays also indicated the association of the same set of ECF sigma factors to *msdps1* promoter under starvation conditions at 72 hours of growth, a time-point where the protein expression profile was also maximum (*17*). At mid-log phase, however, *msdps1* fails to pick up σ^{A} from the *in vivo* lysate. Thus, we propose that both, when cells are grown in 2% glucose and when they are in the early stage of starvation, transcription at *msdps1* promoter is down regulated or is maintained at its basal level by σ^{F} and σ^{H}. Growth phase dependent expression of these two stress induced ECF sigma factors also follows a similar profile as MsDps1 itself, thereby suggesting an induced activated regulation of the *msdps1* gene during bacterial growth, shooting up dramatically when glucose depletion is encountered. As very late stationary phase of *M. smegmatis* mimics the latent state of *M. tuberculosis*, regulation of *msdps1* appears to be interesting. The regulatory sigma factors, identified as sigma F and H, play important roles toward the regulation of *msdps1* gene. The mechanism of binding of these two sigma factors at *msdps1* promoter region will be the immediate lead to be followed at the molecular level. While *msdps1* expression seems to be completely induced or conditional at both transcriptional and translational levels; the regulation of the *msdps2* promoter, specific to the *msdps2* gene, followed expression driven by the housekeeping sigma σ^{A} as well as the alternative sigma σ^{B}. Both *in vivo* promoter specific pull down assay (similar to *msdps1*) and *in vitro* single round transcription supports this contention. However, information on the expression analysis of MsDps2 protein *in vivo* is insufficient. Firstly, MsDps2 remains as a constitutively repressed protein throughout the microbial growth, and on the other hand, chances of some translational modifications keep the protein at very low levels that cannot be easily detected. The rigid dodecameric structure in solution, unlike MsDps1 trimers, makes MsDps2 distinct from MsDps1 from the point of stability. Bioinformatic analysis, in addition, also pointed toward a phylogenetically distant relationship between them, suggesting a possible origin by horizontal gene transfer (unpublished data). The difference in

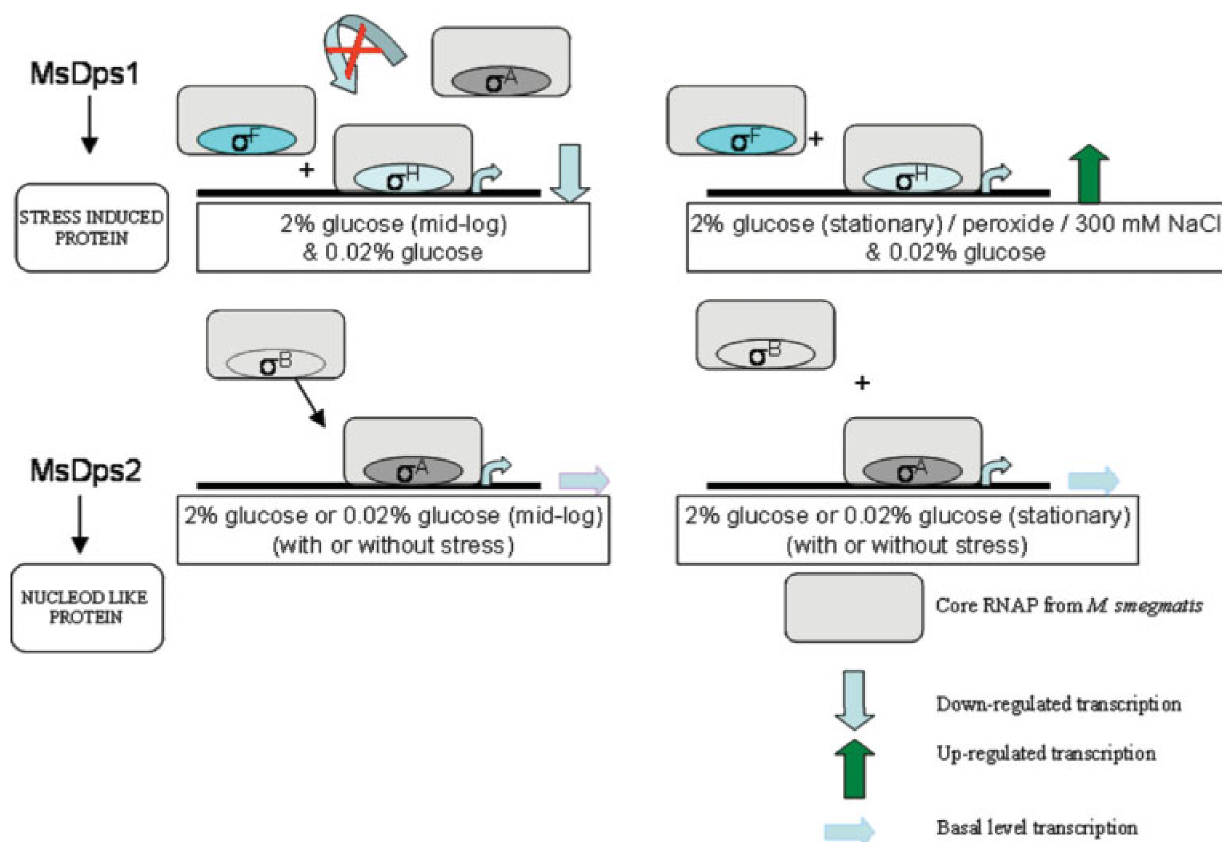

Figure 2. A model depicting the differential transcriptional regulation of the two *dps* genes in *M. smegmatis.* The regulation of MsDps1 and MsDps2 at transcription level is shown. MsDps1 protein expression is driven only by ECF sigma factors (sigma H and F) reconstituted RNA polymerases. However, MsDps2 expression gets initiated in the presence of housekeeping sigma factors (sigma A and B) reconstituted RNA polymerases. The condition of bacterial growth (as followed in growing cultured of *M. smegmatis* in our laboratory) is also added in the text boxes. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

sigma factor selectivity of their respective promoters at the transcription level further supports this discrimination. Additionally, there is quite a difference in the genome regions of the two genes as can be obtained from the TIGR database (www.tigr.org) (gene locus: MSMEG_6467 for *msdps1* and MSMEG_3242 for *msdps2*). In a recent report, the dual regulation of single *dps* gene in *E.coli* by two regulatory proteins, Fis and HNS has been attributed to sigma factor selectivity in a growth phase dependent manner (*42, 43*).

A Model for the Regulation of MsDps1 and MsDps2 In Vivo

All the aforementioned considerations put together, lead us to postulate a model which may eventually predict the stress induced regulatory network in the bacterium navigating the production of these two structurally similar DNA-binding proteins. However, the two proteins are distinct at the function and regulation level by control of the promoters by two separate classes of sigma factors (Fig. 2). This might show new insights on the transcriptional regulation of Dps proteins in bacteria in terms of sigma factor selectivity growing under uninduced and induced regimes. It is indeed interesting that so many regulatory mechanisms mediate the transcriptional, translational, and functional expression of Dps. This is especially so in light of the fact that Dps is a general stress protein that binds to DNA non-specifically and its deletion leads to no loss of viability for the cell. In this context, it appears that the bacterial gene regulation networks in general are more sophisticated than previously understood.

Structural Properties of the Dps Proteins

Structurally, the Dps protein is very similar to ferritin and is included in the ‘Ferritin-like’ superfamily. Both ferritin and Dps possess a monomeric structure consisting of a four helix bundle. The four helix bundle is made up of two α helices in tandem, followed by a loop region containing another short helix, and then the other two helices. Two subunits in Dps are related by a local 2-fold axis to form a dimer, which contains the ferroxidation site. The two subunits in a dimer are arranged adjacent to each other, such that the corresponding α helix in each subunit

is placed antiparallel to that in the other (*51*). Ferritins however have an additional 5th helix at the C-terminus, whereas Dps lacks the extra helix which is presumed to stabilize the tetracosameric ferritin structure. The absence of this helix is understood to make the 24-mer organization unstable thus possibly giving rise to a 12-meric structure of Dps. Additionally, Dps molecules have two different kinds of three-fold axes. One axis called "Ferritin-like," consists of a three-fold interaction having octahedral symmetry and occurs near the N-terminal region of the protein. The other, called "Dps-like" three-fold axis, occurs near the C-terminal region of the protein and is only present in the Dps proteins (*52, 53*).

Newer Insights on the Structure-Function Relationship of the MsDps1 Protein from Mycobacterium smegmatis

The dodecameric Dps protein is spherical with an external diameter of around 90 Å and inner diameter of 45 Å (*37*). Some of the Dps molecules have elongated N- and C-terminal regions outside the main body of the dodecamer. The C-terminal tail is known to confer the DNA-binding ability in MsDps1 (*40*). The N-terminal tails of *E. coli* (*54*) and *Dienococcus radiodurans* (*55*) Dps are implicated in their DNA-binding ability. From the crystal structure analysis, two modes of DNA packaging into the Dps protein layers have been proposed. In both the models, Dps and DNA complexes form a three-dimensional regularly organized structure. Evidence for both models exists, as in the case of the two *M. smegmatis* Dps proteins. The MsDps1 protein is postulated to pack DNA by the first model, similar to that found in *E. coli*, whereas the MsDps2 crystal structure provides evidence for the second kind of packing (*37, 38*). Structurally, the Dps proteins are conserved across various bacterial species, although their sequence similarity is usually low. The sequence identity can be as low as 17 % as seen with the Dps homologues Dps1 and Dps2, from the same organism *Deinococcus radiodurans* (*56*). On the other hand, there is a high degree of structural similarity among the Dps proteins (*55*). The structural conservation hints at the significance of the Dps architecture and the precise locations of the ferroxidation and irons storage sites. It also illuminates the fact that only a few residues, probably constituting the Dps "DNA-binding signature" (*52*), determine the overall scaffold of the Dps molecules and their further organization in the form of multi-layered arrays.

Surface Dynamicity as a Function of Oligomerization and DNA-Binding

The *M. smegmatis* Dps protein, MsDps1 has a disordered C-terminal tail, which is rich in positively charged amino acids (*37*). Previous work in our laboratory has shown that upon storage, this C-terminal tail gets spontaneously truncated to yield a non-DNA–binding version of the protein. The mechanism of spontaneous cleavage of MsDps1 is not very well characterized. However, it is speculated to occur through a cleavage-sensitive

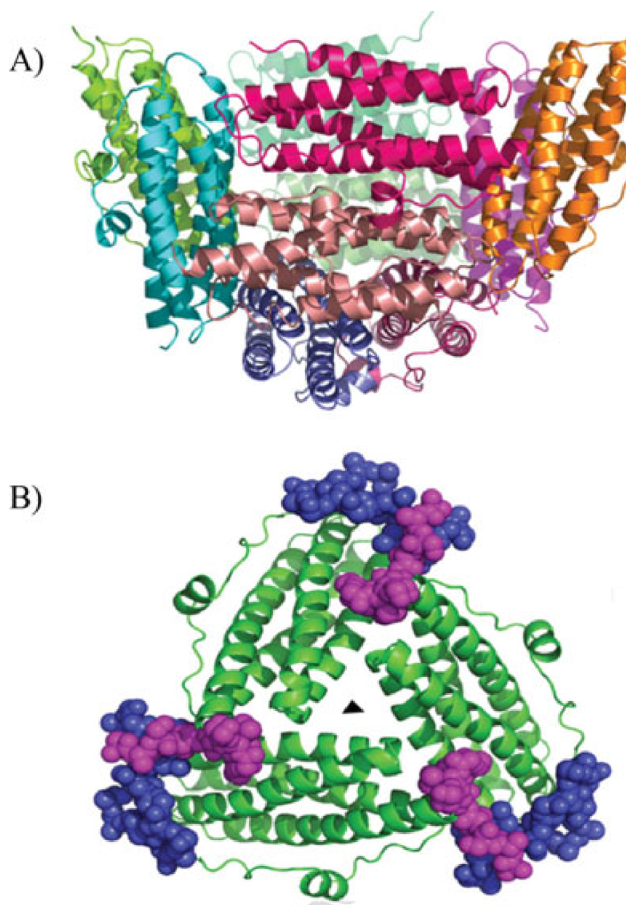

Figure 3. Role of N-C interaction in the dodecameric stability of MsDps1 (A) Crystal structure of MsDps1ΔC26 side view. Subunits are colored differently. Figure is taken from (*40*) (B) Interaction of the N-terminal stretch (blue) of one subunit with the C-terminal stretch (hot pink) of another subunit related by 3-fold symmetry in native MsDps1. Figure is taken from (*40*). [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

site at the beginning of the C-terminal tail. Nevertheless, it finally leads to the formation of a non-DNA–binding mutant of the protein. An N-terminal deletion interestingly, also forms trimer in solution and occurs in the dodecameric form under crystallization conditions. The N-terminal deletion mutant is found to exist as a stable trimer incapable of dodecamerizing in solution. This suggested that the N-terminus of the protein also determines the stability of the dodecamer. This data was considered in the context of the solved crystal structures of MsDps1 and its N and C-terminal deleted versions. The absence of either one of these regions destabilizes the dodecamer yielding a trimer in solution. Under the crystallization conditions involving high protein concentrations, the C-terminal deleted MsDps1 (MsDps1ΔC26) exists as an open bowl structure, as shown in Figure 3A (*40*). The C-terminal region of MsDps1 is thus involved in it DNA-binding ability. It additionally provides

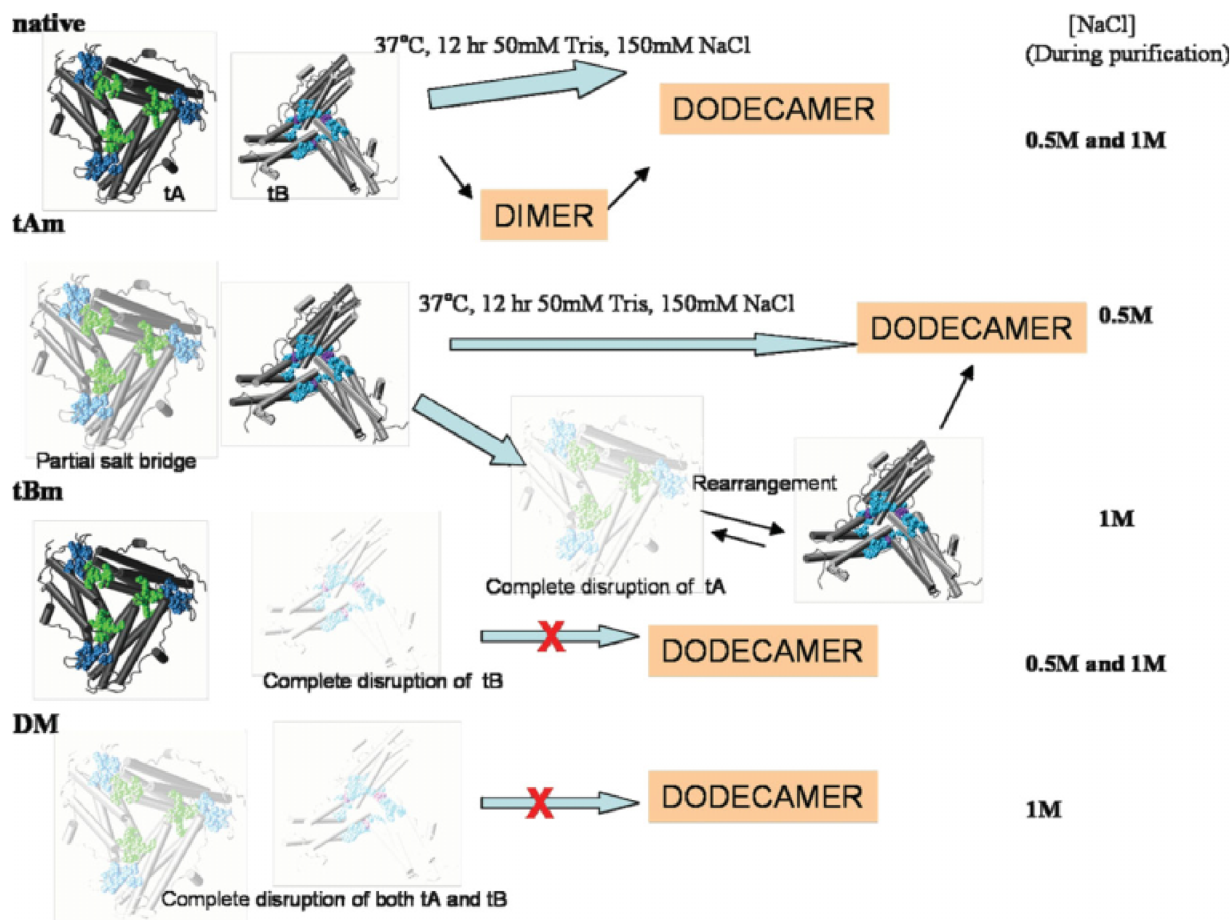

Figure 4. Model depicting the *in vitro* oligomerization of MsDps1 protein in solution. The trimeric interfaces are shown in VDW presentation for tA and tB. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

structural stability to the protein, through its interaction with the N-terminal domain (Fig. 3B) (*40*).

The interaction between N- and C-terminal tails was probed with Förster Resonance Energy Transfer (FRET) (*57*). The mutation of serine residue at 169th position to cysteine generated a single cysteine mutant of MsDps1, that was further labeled with a specific fluorophore 1,5-IAEDANS (5-({2-[(iodoacetyl)amino]ethyl}amino)naphthalene-1-sulfonic acid). From the crystal structure of the dodecamer (*37*), the intramolecular N–C_{169} distances are associated with magnitudes of 16.79, 48.34, and 65.08 Å (*52*). Fluorescence spectroscopic studies on C_{169}-IAEDANS labeled as well as both N-terminal-FITC(Fluorescein iso-Thiocyanate) -C_{169}-IAEDANS labeled MsDps1 protein showed that the 27 amino acid residues long C-terminal tail of the protein is responsible for the DNA-binding activity of the protein (*40*). The oligomeric switch from a trimeric species to a dodecamer does not lead to any significant change in overall protein conformation. However, after binding to DNA, as found from the FRET data, all the N–C_{169} distances in each monomer in a dodecameric species is confined to a distance of 50–55 Å. MEM (Maximum Entropy Method) analysis of the same phase-modulation data revealed averaging of these distances and two sets of stable distributions were obtained in the trimer as well as in the dodecamer (*57*). However, there was a definite change in the value of the distances upon the oligomeric transition, which is obvious, as the protein becomes more compact in the dodecameric form. The interesting point to be noted here is that this flexibility in the overall protein conformation is lost, when the C-terminal tail engages itself in binding to the DNA in a non-specific manner. The conformational heterogeneity of the dodecameric species of Dps in the DNA-bound form seems to be interesting and may have significance in terms of its biological function. Absence of such phenomenon during conformational switch between the trimer and dodecamer, adds credence to this observation. Anisotropy saturation data with labeled DNA of varying length supports the model proposed by Ren et al. (*51*) satisfying the structural alterations that need to accompany the DNA–Dps complex formation.

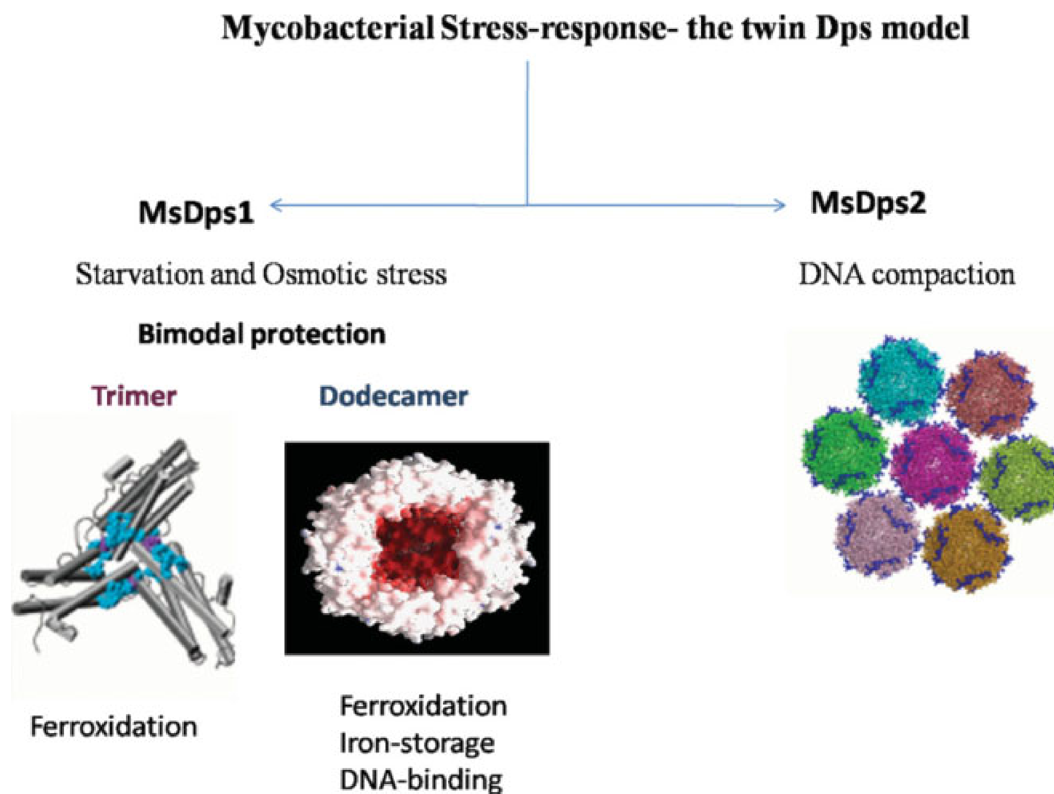

Figure 5. Model depicting the stress response of *Mycobacterium smegmatis* mediated by MsDps1 and MsDps2, the "twin-sister" Dps homologues. The trimer figure taken from (*62*), and the MSDps2 dodecamer figure, from (*38*). [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

Molecular Mechanism of Oligomerization

Symmetry plays an important role toward folding of a multisubunit protein *in vitro* (*58-61*). It reflects the multiplicity at the interface contacts and hence the number and position of required mutation(s) to disrupt the interfaces. In most of the cases it has been seen that the thermal folding kinetics follows an irreversible path and most probably the kinetic barrier for the reverse path is too high to overcome, which makes the multimeric assembly highly stable in solution (*59*). Hence, it becomes relevant to explore the structural determinants that drive the assembly of such identical monomeric subunits to form a highly symmetric stable oligomer. Non-bonded side-chain communications are sometimes essential for the structural stability of proteins which further manifests into its folding process as well function. We have adopted a method of identifying the interface residues from the protein structure networks, which can capture the local details of interactions in the global structural context. We have also identified clusters of amino acids at different interfaces of MsDps1 and have predicted key residues for mutation(s) that can probably disrupt the interfaces. Based on these predictions, single and double mutants are constructed through site directed mutagenesis at both tA and tB interfaces and named as tAm (E146A) and tBm (F47E) respectively. The two trimeric interfaces derived from the crystal structure of MsDps1 dodecamer are named as tA and tB, respectively. "t" stands for "trimer" and interfaces "A" and "B" indicate the "ferritin-like" and "Dps-like" interfaces respectively (*59*). The monomeric chains are oriented in different arrangemet in the two interface forms but the fold contains a C3 principal axis of symmetry. Though it is expected, that disruptions of one of the clusters in tA interface by mutating the buried glutamate residue (E) at 146th position will impair the trimer formation, to our surprise, we obtained a trimer in this case (*62*). Furthermore, it was also capable of forming a dodecamer (*62*). Hence, we propose that tAm is actually a prototype of the native Dps itself. By prototype we mean that the tAm (single mutation in the tA interface) mutant of the protein behaves similarly as the native protein with respect to oligomeric state as well as biological activities. The mutated protein (E146A) is able to fold into a stable trimer as shown by gel analysis (*62*). Similarly, tBm is also purified as a trimer. However, this mutant was found to be incapable of forming a dodecamer under optimum conditions. From the experimental results obtained in this study we would like to propose that due to the two trimeric interfaces, tA and tB, both trimers are present in solution and based on this, we have proposed a model depicting the trimer to dodecamer conversion in MsDps1 as shown in Figure 4. Hence, E146 from tA and F47 from tB interface clusters, are identified as crucial for stabilizing the MsDps1 trimers. We show here that F47 residue

is crucial for dodecamerization and hydrophobic interaction amongst structural building blocks makes major contribution. We think that MsDps1 in this context can be used further as a model system to explore oligomeric assembly of symmetric multimeric proteins.

CONCLUSIONS

The first mycobacterial Dps protein was discovered in *M. smegmatis,* when grown in carbon depleted media. Subsequent studies showed that the protein is quite different from the ones known in other organism and uniquely showed bimodal protection of the genomic material. Over the last decade several structural and functional studies have been carried out with Dps proteins. The conclusions of these studies have been documented elsewhere and summarized in this review. Discovery of the second Dps, its nucleoid formation property, opened up new routes to persue further the function of this protein. Furthermore, the second Dps transcription is controlled by vegetative sigma factors. Some unique properties of Dps warrant additional study, as listed below. It would be interesting to see whether new insights into the structure-function relationship of Dps can evolve from either of these observations.

a) One would like to see the evolutionary relationship among multiple Dps proteins in the same organisms and the functional significance of different Dps proteins.
b) One would look for any other function that Dps can perform other than DNA binding, ferroxidation activity or nucleoid formation. A new function for Dps may take us to explore whether oligomeric arrangement of the protein has any special significance.
c) Unique mutations that create a monomeric or trimeric Dps should be solved structurally, which will throw light into the pathway for oligomerization of a protein. This is an important question as we have shown two interfaces in the trimeric arrangements of Dps, out of which one is kinetically favored for oligomerization. This pathway is depicted in Figure 5. If we understand why one trimer is preferred over the other then we can address the next question.
d) Whether Dps oligomer can be used to trap DNA or any other material inside the core and then utilize the protein as a vehicle for delivery system.

Indeed, such possibilities have been proposed before as Dps can form a very stable nanostructure (*63*). Dps2 from *M. smegmatis,* with a rigid 12-mer structure appears to us as an excellent possibility.

REFERENCES

1. Ojha, A. K., Varma, S., and Chatterji, D. (2002) Synthesis of an unusual polar glycopeptidolipid in glucose-limited culture of *Mycobacterium smegmatis*. *Microbiology* **148**, 3039–3048.
2. Smeulders, M. J., Keer, J., Speight, R. A., and Williams, H. D. (1999) Adaptation of *Mycobacterium smegmatis* to stationary phase. *J. Bacteriol.* **181**, 270–283.
3. Finkel, S. E. and Kolter, R. (1999) Evolution of microbial diversity during prolonged starvation. *Proc. Natl. Acad. Sci. USA* **96**, 4023–4027.
4. Finkel, S. (2006) Long-term survival during stationary phase: evolution and the GASP phenotype. *Nat. Rev. Microbiol.* **4**, 113–120.
5. Blokpoel, M. C. J., Smeulders, M. J., Hubbard, J. A. M., Keer, J., and Williams, H. D. (2005) Global analysis of proteins synthesized by *Mycobacterium smegmatis* provides direct evidence for physiological heterogeneity in stationary-phase cultures. *J. Bacteriol.* **187**, 6691–6700.
6. Dubnau, E., Fontán, P., Manganelli, R., Soares-Appel, S., and Smith, I. (2002) *Mycobacterium tuberculosis* genes induced during infection of human macrophages. *Infect. Immun.* **70**, 2787–2795.
7. Bacon, J., Dover, L. G., Hatch, K. A., Zhang, Y., Gomes, J. M., Kendall, S., Wernisch, L., Stoker, N. G., Butcher, P. D., Besra, G. S., and Marsh, P. D. (2007) Lipid composition and transcriptional response of *Mycobacterium tuberculosis* grown under iron-limitation in continuous culture: identification of a novel wax ester. *Microbiol. (Reading, Engl.)* **153**, 1435–1444.
8. Kendall, S. L., Rison, S. C., Movahedzadeh, F., Frita, R., and Stoker, N. G. (2004) What do microarrays really tell us about *M. tuberculosis*? *Trends. Microbiol* **12**, 537–544.
9. Motiwala, A. S., Janagama, H. K., Paustian, M. L., Zhu, X., Bannantine, J. P., Kapur, V., and Sreevatsan, S. (2006) Comparative transcriptional analysis of human macrophages exposed to animal and human isolates of *Mycobacterium avium* subspecies paratuberculosis with diverse genotypes. *Infect. Immun.* **74**, 6046–6056.
10. Grassi, M., Bocchino, M., Marruchella, A., Volpe, E., Saltini, C., Colizzi, V., and Mariani, F. (2006) Transcriptional profile of the immune response in the lungs of patients with active tuberculosis. *Clin. Immunol.* **121**, 100–107.
11. Beste, D. J., Laing, E., Bonde, B., Avignone-Rossa, C., Bushell, M. E., and Mcfadden, J. J. (2007) Transcriptomic analysis identifies growth rate modulation as a component of the adaptation of mycobacteria to survival inside the macrophage. *J. Bacteriol.* **189**, 3969–3976.
12. Golby, P., Hatch, K. A., Bacon, J., Cooney, R., Riley, P., Allnutt, J., Hinds, J., Nunez, J., Marsh, P. D., Hewinson, R. G., and Gordon, S. V. (2007) Comparative transcriptomics reveals key gene expression differences between the human and bovine pathogens of the *Mycobacterium tuberculosis* complex. *Microbiol. (Reading, Engl.)* **153**, 3323–3336.
13. Sidders, B., Withers, M., Kendall, S. L., Bacon, J., Waddell, S. J., Hinds, J., Golby, P., Movahedzadeh, F., Cox, R. A., Frita, R., Tenbokum, A. M., Wernisch, L., and Stoker, N. G. (2007) Quantification of global transcription patterns in prokaryotes using spotted microarrays. *Genome Biol.* **8**, R265.
14. Zhu, X., Tu, Z. J., Coussens, P. M., Kapur, V., Janagama, H., Naser, S., and Sreevatsan, S. (2008) Transcriptional analysis of diverse strains *Mycobacterium avium* subspecies paratuberculosis in primary bovine monocyte derived macrophages. *Microbes Infect* **10**, 1274–1282.
15. Talaat, A. M., Lyons, R., Howard, S. T., and Johnston, S. A. (2004) The temporal expression profile of *Mycobacterium tuberculosis* infection in mice. *Proc. Natl. Acad. Sci. USA* **101**, 4602–4607.
16. Rachman, H., Strong, M., Ulrichs, T., Grode, L., Schuchhardt, J., Mollenkopf, H., Kosmiadi, G. A., Eisenberg, D., and Kaufmann, S. H. (2006) Unique transcriptome signature of *Mycobacterium tuberculosis* in pulmonary tuberculosis. *Infect. Immun.* **74**, 1233–1242.
17. Chowdhury, R., Gupta, S., and Chatterji, D. (2007) Identification and characterization of the *dps* promoter of *Mycobacterium smegmatis*: promoter recognition by stress-specific extracytoplasmic function sigma factors σ^H and σ^F. *J. Bacteriol.* **189**, 8973–8981.
18. Martinez, A. and Kolter, R. (1997) Protection of DNA during oxidative stress by the nonspecific DNA-binding protein Dps. *J. Bacteriol.* **179**, 5188–5194.

19. Facey, P. D., Hitchings, M. D., Saavedra-Garcia, P., Fernandez-Martinez, L., Dyson, P. J., and Del Sol, R. (2009) *Streptomyces coelicolor* Dps-like proteins: differential dual roles in response to stress during vegetative growth and in nucleoid condensation during reproductive cell division. *Mol. Microbiol.* **73**, 1186–1202.
20. Zhang, J., Xing, X., Herr, A. B., and Bell, C. E. (2009) Crystal structure of *E. coli* RecE protein reveals a toroidal tetramer for processing double-stranded DNA breaks. *Structure* **17**, 690–702.
21. Olsson, M., Tang, K. W., Persson, C., Wilhelmsson, L. M., Billeter, M., and Elias, P. (2009) Stepwise evolution of the herpes simplex virus origin binding protein and origin of replication. *J. Biol. Chem.* **284**, 16246–16255.
22. Nimonkar, A. V., Sica, R. A., and Kowalczykowski, S. C. (2009) Rad52 promotes second-end DNA capture in double-stranded break repair to form complement-stabilized joint molecules. *Proc. Natl. Acad. Sci. USA* **106**, 3077–3082.
23. Almirón, M., Link, A. J., Furlong, D., and Kolter, R. (1992) A novel DNA-binding protein with regulatory and protective roles in starved *Escherichia coli*. *Genes Dev* **6**, 2646–2654.
24. Gupta, S. and Chatterji, D. (2003) Bimodal protection of DNA by *Mycobacterium smegmatis* DNA-binding protein from stationary phase cells. *J. Biol. Chem.* **278**, 5235–5241.
25. Nicodème, M., Perrin, C., Hols, P., Bracquart, P., and Gaillard, J. L. (2004) Identification of an iron-binding protein of the Dps family expressed by *Streptococcus thermophilus*. *Curr. Microbiol.* **48**, 51–56.
26. Dundon, W., Nishioka, H., Polenghi, A., Papinutto, E., Zanotti, G., Montemurro, P., Del, G., Rappuoli, R., and Montecucco, C. (2002) The neutrophil-activating protein of *Helicobacter pylori*. *Int. J. Med. Microbiol.* **291**, 545–550.
27. Lacqua, A., Wanner, O., Colangelo, T., Martinotti, M. G., and Landini, P. (2006) Emergence of biofilm-forming subpopulations upon exposure of *Escherichia coli* to environmental bacteriophages. *Appl. Environ. Microbiol.* **72**, 956–959.
28. Li, X., Pal, U., Ramamoorthi, N., Liu, X., Desrosiers, D. C., Eggers, C. H., Anderson, J. F., Radolf, J. D., and Fikrig, E. (2007) The Lyme disease agent *Borrelia burgdorferi* requires BB0690, a Dps homologue, to persist within ticks. *Mol. Microbiol.* **63**, 694–710.
29. Hanna, E. S., Roque-Barreira, M. C., Bernardes, E. S., Panunto-Castelo, A., Sousa, M. V., Almeida, I. C., and Brocchi, M. (2007) Evidence for glycosylation on a DNA-binding protein of *Salmonella enterica*. *Microb. Cell. Fact.* **6**, 11.
30. Wayne, L. and Sohaskey, C. (2001) Nonreplicating persistence of *Mycobacterium tuberculosis*. *Annu. Rev. Microbiol.* **55**, 139–163.
31. Nyka, W. (1974) Studies on the effect of starvation on mycobacteria. *Infect. Immun.* **9**, 843–850.
32. Stillman, T. J., Upadhyay, M., Norte, V. A., Sedelnikova, S. E., Carradus, M., Tzokov, S., Bullough, P. A., Shearman, C. A., Gasson, M. J., Williams, C. H., Artymiuk, P. J., and Green, J. (2005) The crystal structures of *Lactococcus lactis* MG1363 Dps proteins reveal the presence of an N-terminal helix that is required for DNA binding. *Mol. Microbiol.* **57**, 1101–1112.
33. Kim, S. G., Bhattacharyya, G., Grove, A., and Lee, Y. H. (2006) Crystal structure of Dps-1, a functionally distinct Dps protein from *Deinococcus radiodurans*. *J. Mol. Biol.* **361**, 105–114.
34. Cuypers, M. G., Mitchell, E. P., Romão, C. V., and Mcsweeney, S. M. (2007) The crystal structure of the Dps2 from *Deinococcus radiodurans* reveals an unusual pore profile with a non-specific metal binding site. *J. Mol. Biol.* **371**, 787–799.
35. Antelmann, H., Engelmann, S., Schmid, R., Sorokin, A., Lapidus, A., and Hecker, M. (1997) Expression of a stress- and starvation-induced *dps/pexB*-homologous gene is controlled by the alternative Sigma Factor sigmaB in *Bacillus subtilis*. *J. Bacteriol.* **179**, 7251–7256.
36. Papinutto, E., Dundon, W. G., Pitulis, N., Battistutta, R., Montecucco, C., and Zanotti, G. (2002) Structure of two iron-binding proteins from *Bacillus anthracis*. *J. Biol. Chem.* **277**, 15093–15098.
37. Roy, S., Gupta, S., Das, S., Sekar, K., Chatterji, D., and Vijayan, M. (2004) X-ray analysis of *Mycobacterium smegmatis* Dps and a comparative study involving other Dps and Dps-like molecules. *J. Mol. Biol.* **339**, 1103–1113.
38. Roy, S., Saraswathi, R., Chatterji, D., and Vijayan, M. (2008) Structural studies on the second *Mycobacterium smegmatis* Dps: invariant and variable features of structure, assembly and function. *J. Mol. Biol.* **375**, 948–959.
39. Liu, X., Kim, K., Leighton, T., and Theil, E. C. (2006) Paired *Bacillus anthracis* Dps (mini-ferritin) have different reactivities with peroxide. *J. Biol. Chem.* **281**, 27827–27835.
40. Roy, S., Saraswathi, R., Gupta, S., Sekar, K., Chatterji, D., and Vijayan, M. (2007) Role of N and C-terminal tails in DNA binding and assembly in Dps: structural studies of *Mycobacterium smegmatis* Dps deletion mutants. *J. Mol. Biol.* **370**, 752–767,
41. Lomovskaya, O. L., Kidwell, J. P., and Matin, A. (1994) Characterization of the sigma 38-dependent expression of a core *Escherichia coli* starvation gene, *pexB*. *J. Bacteriol.* **176**, 3928–3935,
42. Grainger, D. C., Goldberg, M. D., Lee, D. J., and Busby, S. J. W. (2008) Selective repression by Fis and H-NS at the *Escherichia colidps* promoter. *Mol. Microbiol.* **68**, 1366–1377,
43. Schnetz, K. (2008) Fine-tuned growth phase control of *dps*, encoding a DNA protection protein, by FIS and H-NS. *Mol. Microbiol.* **68**, 1345–1347,
44. Stephani, K., Weichart, D., and Hengge, R. (2003) Dynamic control of Dps protein levels by ClpXP and ClpAP proteases in *Escherichia coli*. *Mol. Microbiol.* **49**, 1605–1614,
45. Gupta, S., Pandit, S. B., Srinivasan, N., and Chatterji, D. (2002) Proteomics analysis of carbon-starved *Mycobacterium smegmatis*: induction of Dps-like protein. *Protein. Eng.* **15**, 503–512,
46. Heil, A. and Zillig, W. (1970) Reconstitution of bacterial DNA-dependent RNA-polymerase from isolated subunits as a tool for the elucidation of the role of the subunits in transcription. *FEBS Lett.* **11**, 165–168.
47. Ishihama, A. and Ito, K. (1972) Subunits of RNA polymerase in function and structure. II. Reconstitution of *Escherichia coli* RNA polymerase from isolated subunits. *J. Mol. Biol.* **72**, 111–123.
48. Fujita, N. and Ishihama, A. (1996) Reconstitution of RNA polymerase. *Meth. Enzymol.* **273**, 121–130,
49. Dainese, E., Rodrigue, S., Delogu, G., Provvedi, R., Laflamme, L., Brzezinski, R., Fadda, G., Smith, I., Gaudreau, L., Palù, G., and Manganelli, R. (2006) Posttranslational regulation of *Mycobacterium tuberculosis* extracytoplasmic-function sigma factor sigma L and roles in virulence and in global regulation of gene expression. *Infect. Immun.* **74**, 2457–2461,
50. Homerova, D., Halgasova, L., and Kormanec, J. (2008) Cascade of extracytoplasmic function sigma factors in *Mycobacterium tuberculosis*: identification of a sigmaJ-dependent promoter upstream of sigI FEMS. *Microbiol. Lett.* **280**, 120–126,
51. Ren, B., Tibbelin, G., Kajino, T., Asami, O., and Ladenstein, R. (2003) The multi-layered structure of Dps with a novel di-nuclear ferroxidase center. *J. Mol. Biol.* **329**, 467–477.
52. Ilari, A., Stefanini, S., Chiancone, E., and Tsernoglou, D. (2000) The dodecameric ferritin from *Listeria innocua* contains a novel intersubunit iron-binding site. *Nat. Struct. Biol.* **7**, 38–43.
53. Zhao, G., Ceci, P., Ilari, A., Giangiacomo, L., Laue, T. M., Chiancone, E., and Chasteen, N. D. (2002) Iron and hydrogen peroxide detoxification properties of DNA-binding protein from starved cells. A ferritin-like DNA-binding protein of *Escherichia coli*. *J. Biol. Chem.* **277**, 27689–27696.
54. Ceci, P., Cellai, S., Falvo, E., Rivetti, C., Rossi, G. L., and Chiancone, E. (2004) DNA condensation and self-aggregation of *Escherichia coli* Dps are coupled phenomena related to the properties of the N-terminus. *Nucl. Acids. Res.* **32**, 5935–5944.

55. Bhattacharyya, G. and Grove, A. (2007) The N-terminal extensions of *Deinococcus radiodurans* Dps-1 mediate DNA major groove interactions as well as assembly of the dodecamer. *J. Biol. Chem.* **282**, 11921–11930.
56. Romão, C., Mitchell, E., and Mcsweeney, S. (2006) The crystal structure of *Deinococcus radiodurans* Dps protein (DR2263) reveals the presence of a novel metal centre in the N terminus. *J. Biol. Inorg. Chem.* **11**, 891–902.
57. Chowdhury, R. P. and Chatterji, D. (2007) Estimation of Förster's distance between two ends of Dps protein from mycobacteria: distance heterogeneity as a function of oligomerization and DNA binding. *Biophys. Chem.* **128**, 19–29.
58. Grueninger, D., Treiber, N., Ziegler, M. O., Koetter, J. W., Schulze, M. S., and Schulz, G. E. (2008) *Designed protein-protein Association. Science* **319**, 206–209.
59. Bellapadrona, G., Chiaraluce, R., Consalvi, V., Ilari, A., Stefanini, S., and Chiancone, E. (2007) Asn disrupt an intersubunit salt bridge and convert *Listeria innocua* Dps into its natural mutant *Listeria monocytogenes* Dps *Effects on Protein Stability at Low pH. Proteins* **66**, 975–983.
60. Frederick, K. K., Marlow, M. S., Valentine, K. G., and Wand, A. J. (2007) Conformational entropy in molecular recognition by proteins. *Nature* **448**, 325–329.
61. Cole, C. and Warwicker, J. (2002) Side-chain conformational entropy at protein-protein interfaces. *Protein Sci.* **11**, 2860–2870.
62. Chowdhury, R., Vijayabaskar, M., Vishveshwara, S., and Chatterji, D. (2008) Molecular mechanism of *in vitro* oligomerization of Dps from *Mycobacterium smegmatis*: mutations of the residues identified by "interface cluster" analysis. *Biochemistry* **47**, 11110–11117.
63. Heddle, J. G. (2008) Protein cages, rings and tubes: useful components of future nanodevices? *Nanotechnol. Sci. Appl.* **1**, 67–78.