In recent years, it has become clear that, in addition to the regulation of the expression of specific genes, there are global regulatory systems that control the simultaneous expression of a large number of genes in response to a variety of environmental stresses. The first of these global control systems, and of substantial importance, is the heat-shock response. The heat-shock response is characterized by the induction of a large set of genes in response to a variety of environmental stresses. The heat-shock proteins (heat-shock proteins—HSPs) upon shifts to higher temperatures are denatured (i.e., alcohols, heavy metals). The heat-shock response is universal and many of the heat-shock proteins are highly conserved among species. In bacteria, the heat-shock response has been studied extensively in several Gram-positive bacteria (Bacillus subtilis) and in the Gram-negative bacteria (i.e., Escherichia coli, Agrobacterium tumefaciens). The first recognition of the molecular abundance of the bacterial heat-shock proteins took place with the introduction of high-resolution two-dimensional polyacrylamide gels (2D gels) to analyze complex mixtures of cellular proteins. Two-dimensional gels, followed by mass spectrometry, were used to define the heat-shock stimulons in several bacteria, and to study the regulatory elements that control the heat-shock response. Here, we review the heat-shock response and its regulation in bacteria. The review will emphasize the use of proteome analysis in the study of this response, and will point out those open
I. INTRODUCTION

In recent years, it has become clear that, in addition to the regulation of the expression of specific genes, there are global regulatory systems that control the simultaneous expression of a large number of genes. These systems are called Global Regulatory Systems, and they respond to a variety of environmental stresses, such as change of temperature, pH, nutrients, and oxidation. First discovered, and one of the most important of the global regulatory systems, is the heat-shock response. This response is general, found in all living cells examined (Craig, 1985), and is a protective and homeostatic cellular process that increases thermotolerance. The heat-shock response is characterized by the induction of a large set of proteins (heat-shock proteins—HSPs) as a result of a rapid increase in the environmental temperature. Many of the HSPs are molecular chaperons (e.g., GroEL, GroES, DnaK, and DnaJ) and ATP-dependent proteases [e.g., ClpP, Lon (La) and HslVU] that play a critical role in the repair of protein folding and in protein degradation under normal and stress conditions. Proteins such as GroEL (the bacterial homolog of Hsp 60) and DnaK (the bacterial homolog of Hsp 70) are highly conserved in evolution between Escherichia coli and humans (Boorstein, Ziegelhoffer, & Craig, 1994; Gupta, 1995). Although the major proteins in the heat-shock response are highly conserved, the regulation of the response varies between different bacterial species. Several regulatory systems evolved in bacteria and will be discussed here.

The global regulatory systems were studied, using transcriptomics and proteomics. The latter studies enable one to determine not only the level of the protein but also post-translational modifications, which appear to constitute an integral part of the stress response.

Here, we review the heat-shock response and its regulation in bacteria. The review will emphasize the use of proteome analysis in the study of this response, and will point out open questions that can be investigated with proteomics, including mass spectrometry techniques.

II. GLOBAL REGULATORY NETWORKS

Bacteria live in environments that constantly undergo physical and chemical changes. In order to survive the changes in the environment, as well as with the physiological changes (such as these related to the cell cycle), various global regulatory networks have evolved. The first attempts to study the extent of such regulatory networks were started by comparing proteomes, using the O’Farrell two-dimensional gels, and resulted in the identification of the large group of E. coli heat-shock proteins (O’Farrell, 1975; Neidhardt et al., 1981). Later, proteomic-based experiments revealed the size and composition of the various-stress induced stimulons of E. coli (VanBogelen, Kelley, & Neidhardt, 1987a). This induction of large groups of proteins in a response to a specific environment suggested the existence of global regulatory systems that control the expression of large regulons.

Gene expression can be regulated either at the level of transcription or post-transcription. The level of transcription can be regulated by positive control elements—activators or by negative control elements—repressors. Some of these control elements are specific for one gene, whereas others control a large group of genes, thus creating a regulon. In addition to transcriptional regulation, many post-transcriptional regulatory systems evolved on the way from the gene to the active protein. The post-transcriptional regulatory systems control the stability of the mRNA and the rate of translation initiation. In addition, they can determine the stability of the protein and its activity by post-translational modifications. The existence of all of the control elements described here was demonstrated in the global regulatory systems that control the response to heat-shock and other environmental and physiological conditions.

Transcriptional regulation is the primary level that regulates gene expression. The process of RNA synthesis and its control was extensively studied in bacteria, especially in E. coli and Bacillus subtilis (Burgess & Anthony, 2001). The E. coli DNA-dependent RNA polymerase is the enzyme responsible for all cellular RNA synthesis. This enzyme consists of a core (subunits 7, 54, 54, 245) that is capable of elongation and termination of transcription, and an additional subunit (σ), which is bound to the RNA polymerase to form the holoenzyme, increases the efficiency of transcription initiation, and determines specific promoter recognition (Burgess et al., 1969). In E. coli, there are seven known sigma factors: σ 70; the vegetative sigma factors, σ 54, σ 32, σ E, and σ 43; and σ 54 (Helmann & Chamberlin, 1988; Lonetto, Grishkov, & Gross, 1992; Burgess & Anthony, 2001). The sigma factors serve as master regulators mainly by competition for the core RNA polymerase, which is the limiting component of the transcription machinery (Ishihama, 2000). Additional regulation of transcription is exerted by repressors, transcriptional activators, sigma-binding anti-sigma factors, and even by small RNAs (Hughes & Mathee, 1998; Helmann, 1999; Vicente, Chater,
De Lorenzo, 1999; Ishihama, 2000; Severinov, 2000; Wassarman & Storz, 2000).

These various control elements regulate the expression of genes during various conditions such as starvation, sporulation, and various stress conditions. For example, the E. coli stationary phase is regulated by the master regulator $\sigma^S$ (Lange & Hengge-Aronis, 1991). The levels of $\sigma^S$ itself are affected by cis- and trans-elements—small molecules such as ppGpp and homoserine lactone, and proteins that react to them, such as CRP-cAMP (Hengge-Aronis, 2000). Sigma factor S regulates the induction of more than 50 genes (Hengge-Aronis, 2000). All of these elements create a complex regulatory network that enables the bacterial cell to adapt to the changing environment.

### III. HEAT-SHOCK RESPONSE

The heat-shock response is one of the fundamental responses of living cells. It is characterized by the induction of a set of proteins called heat-shock proteins, many of which are highly conserved in evolution from bacteria to human (Boorstein, Ziegelhoffer, & Craig, 1994; Gupta, 1995). This response has been studied in many cellular systems such as bacteria, yeast, insects (Drosophila melanogaster) (Michaud, Marin, & Tanguay, 1997), worms (Caenorhabditis elegans) (Rose & Rankin, 2001), and mammals (Christians, Yan, & Benjamin, 2002; Li, Menoret, & Srivastava, 2002). The heat-shock proteins are important for protection against environmental stress, and they produce tolerance against high temperature, high salt, and heavy metals (VanBogelen, Acton, & Neidhardt, 1987b; Inbar & Ron, 1993; Hecker & Volker, 1998). Heat-shock proteins also play critical roles in protective systems such as the human immune system (Christians, Yan, & Benjamin, 2002; Li, Menoret, & Srivastava, 2002).

Heat-shock—a rapid up-shift in the environmental temperature—results in various physical and chemical changes in bacterial proteins and membranes. It is presumed that these changes, such as protein unfolding, are detected by cellular systems, which induce the large set of heat-shock proteins to cope with the changes and the potential damage. This heat-shock response is regulated by several control elements, thus dividing the major stimulus of heat-shock proteins into several regulatory groups (regulons). One of the most efficient methods to study these regulons is by global analysis of the transcriptome and the proteome. The contribution of these methods to the field is discussed below.

Several heat-shock proteins were found to protect against damage induced by temperature shift-ups. Among the characterized proteins are the main cellular chaperone machineries GroE and DnaK, the ATP-dependent proteases Lon (La), HslVU, ClpP, DegP, and FtsH (FhIB), and other proteins involved in protein folding, refolding, quality control, and degradation. GroE and DnaK are both multimeric complexes that have ATP-dependent activity (Sherman & Goldberg, 1992, 1996; Kandror et al., 1994). The GroE catalytic complex involves GroEL and GroES in a ratio of 1:2, creating a football-shape molecular structure (Sparrer, Rutkat, & Buchner, 1997). This complex catalyzes protein refolding, and is involved in protein degradation by the ATP-dependent proteases (Sherman & Goldberg, 1992, 1996; Kandror et al., 1994). These ATP-dependent proteases degrade abnormal proteins under stress and non-stress conditions, and in addition play major regulatory functions by controlling the degradation of specific proteins (Goldberg, 1972; Maurizi, 1992; Gottesman, 1996; Deuerling et al., 1997; Zhou et al., 2001). The role of these and other E. coli heat-shock proteins in protection against temperature-induced damage is summarized in Table 1.

The heat-shock proteins are highly conserved, whereas the control of their expression is highly variable among organisms and even among various bacteria. One of the control elements found in Gram-negative bacteria is a heat-shock sigma factor that regulates the transcription of the major heat-shock proteins. The Gram-negative E. coli is a good example for this system because the synthesis of the heat-shock proteins is regulated by alternative sigma factor called $\sigma^{32}$. In addition, there is a group of proteins induced under conditions of elevated temperature that is regulated by another heat-shock sigma factor, $\sigma^{31}$ (encoded by rpoE). In other Gram-negative bacteria, such as the Agrobacterium tumefaciens, of the proteobacteria, the control systems are more complicated. For example, the transcription of GroESL is stimulated during heat-shock by a $\sigma^{32}$-like activator; but, in non-heat-shock conditions, transcription is repressed by the HrcA protein that binds to the CIRCE (Control Inverted Repeat of Chaperone Expression) sequence in the upstream of the promoter region (Segal & Ron, 1993; Nakahigashi et al., 1999). The control system of HrcA-CIRCE was first described in the Gram-positive Bacillus subtilis (Zuber & Schumann, 1994).

The following sections will describe the specific control mechanisms in various bacterial groups. In short, the heat-shock response in bacteria is controlled by one of the following control systems, or by a combination of a few of them:

- Alternative sigma factors that act as transcriptional activators by recognizing specific heat-shock promoters upstream to heat-shock genes. Among these factors are $\sigma^{12}$ and $\sigma^{32}$ of the Gram-negative bacteria, and $\sigma^{B}$ of the Gram-positive bacteria.
- Transcriptional repressors. The most conserved and the most ubiquitous among these repressors is HrcA (Heat
Regulation at CIRCE, which binds to a conserved inverted repeat (CIRCE) that is present upstream to heat-shock operons. Heat-shock operons controlled by HrcA-CIRCE are transcribed by the vegetative sigma factor \(\sigma^70\) in Gram-positive bacteria and by the heat-shock sigma factor \(\sigma^32\) in Gram-negative bacteria.

Figure 1 contains a graphical illustration of several general heat-shock control mechanisms of bacteria, the consensus sequences of the vegetative and heat-shock promoters are shown in Figure 2, and the inverted repeat CIRCE is shown in Figure 3.

### A. Control Elements in Gram-Positive Bacteria

The model organism for studying the heat-shock response in Gram-positive bacteria is *B. subtilis*. The heat-shock response of this bacterium includes the induction of proteins from several regulatory groups: general stress proteins (GSPs)—whose induction is regulated by the alternative sigma factor \(\sigma^B\), the major chaperones—whose induction is regulated by the HrcA-CIRCE control elements (Zuber & Schumann, 1994; Hecker & Volker, 1998), and several groups of proteins regulated by specific control elements, all of which are discussed below.

#### 1. Sigma B-Controlled Genes

Sigma B was found to control a stress/starvation regulon that comprises a very large set of general stress genes (for review, see Hecker, Schumann, & Volker, 1996; Hecker & Volker, 1998). These \(\sigma^B\)-dependent genes are strongly induced by heat-, ethanol-, acid-, or salt-stress, as well as by starvation for a carbon source, phosphate, and oxygen.

**TABLE 1. Heat-shock proteins of *Escherichia coli***

<table>
<thead>
<tr>
<th>Protein (Protein)</th>
<th>Function</th>
<th>Molecular weight (kDa)</th>
<th>Theoretical pI</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>ClpB</td>
<td>Chaperone</td>
<td>96</td>
<td>5.37</td>
<td>Kitagawa et al. (1991)</td>
</tr>
<tr>
<td>ClpP</td>
<td>Protease</td>
<td>24</td>
<td>5.52</td>
<td>Maurizi et al. (1990)</td>
</tr>
<tr>
<td>ClpX</td>
<td>Protease</td>
<td>46</td>
<td>5.24</td>
<td>Gottesman et al. (1993)</td>
</tr>
<tr>
<td>DegP (HtrA)</td>
<td>Protease</td>
<td>50</td>
<td>8.65</td>
<td>Lipinska et al. (1988)</td>
</tr>
<tr>
<td>DnaJ</td>
<td>Chaperone</td>
<td>39</td>
<td>7.98</td>
<td>Bardwell et al. (1986)</td>
</tr>
<tr>
<td>DnaK</td>
<td>Chaperone</td>
<td>69</td>
<td>4.83</td>
<td>Bardwell &amp; Craig (1984)</td>
</tr>
<tr>
<td>FkpA</td>
<td>Peptidyl-prolyl isomerase</td>
<td>29</td>
<td>8.39</td>
<td>Danese &amp; Silhavy (1997)</td>
</tr>
<tr>
<td>FtsH (Hf1B)</td>
<td>Protease</td>
<td>71</td>
<td>8.91</td>
<td>Herman et al. (1995)</td>
</tr>
<tr>
<td>FtsJ</td>
<td>Unknown</td>
<td>23</td>
<td>9.44</td>
<td>Herman et al. (1995)</td>
</tr>
<tr>
<td>GapA</td>
<td>Dehydrogenase</td>
<td>36</td>
<td>6.61</td>
<td>Charpentier &amp; Branlant (1994)</td>
</tr>
<tr>
<td>GroEL</td>
<td>Chaperone</td>
<td>57</td>
<td>4.85</td>
<td>Neidhardt et al. (1981)</td>
</tr>
<tr>
<td>GroES</td>
<td>Chaperone</td>
<td>10</td>
<td>5.15</td>
<td>Tilly et al. (1983a)</td>
</tr>
<tr>
<td>GrpE</td>
<td>Nucleotide exchange factor</td>
<td>22</td>
<td>4.68</td>
<td>Lipinska et al. (1988)</td>
</tr>
<tr>
<td>HslI (HipH)</td>
<td>(\alpha)-lactate dehydrogenase</td>
<td>36</td>
<td>5.29</td>
<td>Chuang &amp; Blattner (1993)</td>
</tr>
<tr>
<td>HslU (ClpY, HtpI)</td>
<td>Protease</td>
<td>49</td>
<td>5.24</td>
<td>Chuang et al. (1993)</td>
</tr>
<tr>
<td>HslV (ClpQ, HtpO)</td>
<td>Protease</td>
<td>19</td>
<td>5.96</td>
<td>Chuang et al. (1993)</td>
</tr>
<tr>
<td>Hsp33 (HslO)</td>
<td>Chaperone</td>
<td>33</td>
<td>4.65</td>
<td>Chuang &amp; Blattner (1993)</td>
</tr>
<tr>
<td>HtpG</td>
<td>Chaperone</td>
<td>71</td>
<td>5.09</td>
<td>Bardwell &amp; Craig (1987)</td>
</tr>
<tr>
<td>HtpX</td>
<td>Unknown</td>
<td>32</td>
<td>6.60</td>
<td>Kornitzer et al. (1991)</td>
</tr>
<tr>
<td>HtpY (HtgA)</td>
<td>Unknown</td>
<td>21</td>
<td>9.44</td>
<td>Missiakas, Georgopoulos, &amp; Raina (1993)</td>
</tr>
<tr>
<td>HtrC</td>
<td>Unknown</td>
<td>21</td>
<td>9.33</td>
<td>Raina &amp; Georgopoulos (1990)</td>
</tr>
<tr>
<td>HtrM (RafD)</td>
<td>Epimerase</td>
<td>35</td>
<td>4.80</td>
<td>Raina &amp; Georgopoulos (1991)</td>
</tr>
<tr>
<td>IbpA (HtpN, HslT)</td>
<td>Chaperone</td>
<td>16</td>
<td>5.57</td>
<td>Allen et al. (1992)</td>
</tr>
<tr>
<td>IbpB (HtpE, HslS)</td>
<td>Chaperone</td>
<td>16</td>
<td>5.19</td>
<td>Allen et al. (1992)</td>
</tr>
<tr>
<td>Lon (La)</td>
<td>Protease</td>
<td>87</td>
<td>6.01</td>
<td>Gayda et al. (1985)</td>
</tr>
<tr>
<td>LysU</td>
<td>Lysyl-tRNA synthetase</td>
<td>58</td>
<td>5.10</td>
<td>Leveque et al. (1990)</td>
</tr>
<tr>
<td>MetA (HTS)</td>
<td>Homoserine transsuccinylase</td>
<td>36</td>
<td>5.06</td>
<td>Biran et al. (1995)</td>
</tr>
<tr>
<td>PpiD</td>
<td>Peptidyl-prolyl isomerase</td>
<td>68</td>
<td>4.94</td>
<td>Dartigalongue &amp; Raina (1998)</td>
</tr>
<tr>
<td>(\sigma^B) (RpoH, HtpR, Hin, Fam)</td>
<td>Sigma factor</td>
<td>32</td>
<td>5.64</td>
<td>Landick et al. (1984)</td>
</tr>
<tr>
<td>(\sigma^{10}) (RpoD, Alt)</td>
<td>Sigma factor</td>
<td>70</td>
<td>4.69</td>
<td>Burton et al. (1981)</td>
</tr>
<tr>
<td>(\sigma^{24}) (\sigma^{24}, RpoE)</td>
<td>Sigma factor</td>
<td>22</td>
<td>5.38</td>
<td>Raina, Missiakas, &amp; Georgopoulos (1995)</td>
</tr>
</tbody>
</table>
Two groups of signals were found to trigger the induction of \( \text{sigB} \), the gene that codes for \( \sigma^B \). The first group contains extracellular signals that result in a drop of the ATP level—starvation to glucose, oxygen, or phosphate, but not to amino acids, which triggers the induction of ppGpp and keeps the ATP pool constant (Maul et al., 1995). The second group of stimuli includes physical stress-factors such as heat-, salt-, and acid-stress, but not oxidative-stress (Hecker & Volker, 1998). This group of stimuli induces the synthesis of \( \sigma^B \) via a two-component system of RsbS and RsbT that changes the balance of a complex network of antisigma (RsbW) factor and its agonist (non-phosphorylated RsbV) to activate \( \sigma^B \) (Akbar & Price, 1996; Yang et al., 1996). For the expression of some genes, the involvement of \( \sigma^B \) is essential, whereas for others it seems to be only involved because it can be replaced by alternative stress-induction mechanisms (Hecker & Volker, 1998).

The group of GSPs consists of 125 genes (Petersohn et al., 2001), and not much is known about their physiological role in the complex general stress response.
identified GSPs can be assigned to five main groups (Hecker & Volker, 1998):

1. σB-dependent genes that encode subunits of stress-inducible proteases. ClpP, ClpC, and ClpX are probably essential for the renaturation or degradation of misfolded or denatured proteins that accumulate in the cell upon exposure to stress conditions (Gottesman, 1996; Gerth et al., 1998). Null mutants of clpC, clpP, and clpX are extremely sensitive to heat-, salt-, or ethanol-stress, much more than mutants of sigB (Kruger, Volker, & Hecker, 1994; Msadek, Kunst, & Rapoport, 1994; Gerth et al., 1998).

2. σB-dependent genes that encode general oxidative stress-protective proteins such as katE that encode catalase (Engelmann, Lindner, & Hecker, 1995) and the DNA-protecting protein Dps (Antelmann et al., 1997a). Other σB-dependent proteins such as thiordoxin ClpC, ClpP, and the fifth and sixth gene products of the clpC operon [sms, yacK; (Kaan, Jurgen, & Schweder, 1999)] may also be involved in adaptation to oxidative stress (Hecker & Volker, 1998).

3. Proteins with a putative role in the adaptation to salt- or water-stress. A proline-uptake system encoded by a functional copy of opuE is required by B. subtilis for the use of external proline as an osmoprotectant (Hecker & Volker, 1998). However, the physiological role of σB in the expression of opuE is still unclear because exogenously provided proline was used as an osmoprotectant in a sigB mutant (von Blohn et al., 1997). YtxH and GsiB are homologous to plant-desiccation proteins, which are involved in water-stress protection, and YkzA is a homolog of the E. coli Osmc, which is involved in osmo-adaptation (Mueller, Bukusoglu, & Sonenshein, 1992; Volker et al., 1994; Maul et al., 1995; Varon, Brody, & Price, 1996).

4. A heterogeneous group of proteins: their role in adaptation to stress is yet to be determined. One of these proteins is GspA (Antelmann et al., 1995), which is also induced upon starvation to amino acids (Eymann & Hecker, 2001), seems to be involved in the expression of hag, that encodes flagelin, or UDP-glucose pyrophosphorylase with a function in cell-wall metabolism (Varon et al., 1993). Some proteins seem to participate in NAD synthesis (e.g., nadC and nadE gene products) or might catalyze NADP(H)-dependent reactions (Antelmann et al., 1997a,b; Hecker & Volker, 1998; Scharf et al., 1998).

5. The fifth group consists of a large number of proteins that, so far, show no significant similarity to known proteins (Petersohn et al., 2001).

This very large stress regulon seems to give a basal level of protection against a large variety of stress conditions.

2. HrcA-CIRCE Controlled Genes

The second important group of heat-shock-induced proteins includes the chaperones that are controlled by the HrcA-CIRCE repression system. That system is comprised of an inverted repeat cis element and a trans protein-repressor encoded by the hrcA gene. The first reported inverted repeat upstream to the groE operon was found in Mycobacterium tuberculosis in 1989 (Baird, Hall, & Coates, 1989). It took several years until this element was recognized as a widespread heat-shock control element for the groE and dnaK operons. Several lines of direct evidence for the role of CIRCE as a negative cis element were obtained (Narberhaus, 1999): (1) deletion of the inverted repeat relieved the repression of a reporter gene fusion (amyS) (van Asseldonk et al., 1993);
(2) placement of CIRCE behind a foreign promoter reduced the expression of the downstream gene (Zuber & Schumann, 1994); and (3) site-directed mutation, or the removal of three or four nucleotides in one arm of the inverted repeat, resulted in an elevated transcription of the downstream genes at normal growth temperature (Zuber & Schumann, 1994; Babst, Hennecke, & Fischer, 1996). Transcription remained derepressed when the inverted repeat was restored by compensating mutations in the second arm of the inverted repeat. Therefore, the CIRCE is not only a potential stem-and-loop structure because its sequence by itself is required for repression. These data suggest that the CIRCE is a binding site for a sequence-specific repressor protein that binds to the CIRCE.

The elucidation of CIRCE as a potential repressor binding site initiated a search for the counterpart repressor. Major steps towards tracking the repressor were accomplished by two observations (Narberhaus, 1999): (1) a deletion of orf39—the first gene of the dnaK operon of B. subtilis resulted in an elevated levels of groE transcript (Schulz, Tschaschel, & Schumann, 1995); and (2) B. subtilis mutants affected in the regulation of groE and dnaK operons were mapped orf39 (Yuan & Wong, 1995a). Moreover, production of Orf39 from a plasmid that carries a functional copy of orf39 restored the repression activity in one of the mutants (Yuan & Wong, 1995a). The binding of Orf39 to CIRCE was shown by gel retardation (Narberhaus, 1999), and the name HrcA-CIRCE was given to this protein after disruption of the equivalent gene (Narberhaus, 1999). The binding of Orf39 to CIRCE is not only of dnaK but also of clpC, clpP, and trxA operons are under the control of the vegetative sigma factor $\sigma^A$ and the stress sigma factor $\sigma^B$ (Kruger, Msadek, & Hecker, 1996; Gerth et al., 1998; Scharf et al., 1998). Although both promoters were used under a number of stress conditions, the induction pattern of the genes varied for the different genes and for the particular stress condition. A cis element that contains a heptameric tandem consensus sequence was found upstream to the clpC, clpE, and clpP B. subtilis operons, and was shown to be the binding site of the CtsR repressor (Kruger & Hecker, 1998; Derre, Rapoport, & Msadek, 1999a; Derre et al., 1999b). CstR was lately found also in Listeria monocytogenes (Nair et al., 2000).

In Streptomyces coelicolor and S. albus, the groESL$_1$ operon and the groEL$_2$ gene are regulated by tandem CIRCE elements, whereas the dnaK operon encodes its own autoregulatory repressor (Bucca et al., 1995; Bucca, Hindle, & Smith, 1997). Heat-inducible transcription of the dnaK operon (dnaK, grpE, dnaJ, hspR) initiates from the vegetative promoter. Disruption of hspR led to high and constitutive transcription levels of the dnaK operon, but had no effect on the groE expression level (Bucca, Hindle, & Smith, 1997). Similar to the GroE modulation of HrcA activity, it was shown that the DnaK protein forms a specific ATP-independent complex with the Streptomyces HspR repressor, and that this interaction is necessary for HspR to bind a dnaKp fragment in gel-shift assays (Bucca et al., 2000). The dnaK heat-induction model suggested by Bucca et al. (2000) suggests a decrease in the availability of DnaK due to the accumulation of heat-damaged proteins. This model has many similarities to the heat induction of the $\sigma^{32}$-dependent transcription in E. coli, a model that will be discussed in details below.

Another heat-shock control element found in S. albus is the RheA, which represses the transcription of hsp18 (encoding a small heat-shock protein) by binding specifically to the hsp18 promoter (Servant & Mazodier, 1996; Servant, Rapoport, & Mazodier, 1999). Transcription analysis of rheA in the S. albus wild-type and in rheA mutant strains suggested that RheA represses transcription not only of hsp18 but also of rheA itself (Servant, Rapoport, & Mazodier, 1999).

### B. Control Elements in Gram-Negative Bacteria

The first model organism for studying the heat-shock response in Gram-negative bacteria was E. coli. Most of the heat-shock genes of this bacterium are regulated by alternative sigma factors $\sigma^{32}$ or $\sigma^E$. However, in recent years it became clear that other Gram-negative bacteria differ from E. coli in the control of the heat-shock response. Thus, most other Gram-negative bacteria contain the HrcA-CIRCE control elements, usually regulating the groESL operon, and the only group that does not have this control element is the group of $\gamma$-proteobacteria, which includes E. coli (Segal & Ron, 1996b; Ron et al., 1999; Rosen et al., 2002b).

#### 1. Sigma 32-Controlled Genes

The heat-shock response of Gram-negative bacteria is regulated mainly by the alternative sigma factors $\sigma^{32}$ and $\sigma^E$ (Yura, Kanemori, & Morita, 2000). Sigma-32 is a master regulator encoded by the rpoH (htpR, hin) gene that was the first of a group of minor sigma factors discovered in
heat-shock proteins are required. Upon a rapid shift to repression and protein instability, whereas low amounts of these low levels are maintained due to transcriptional regulation at several levels. At low temperature (30°C), the synthesis rates of individual proteins, or proteomics and genomics approaches. This regulon includes all the major cytoplasmic heat-shock proteins of E. coli.

The σ^32 regulon includes most of the proteins involved in protein folding, repair, and degradation. Such proteins are the heat-shock-induced molecular chaperons ClpB, DnaK, DnaJ, GroEL, and GroES, which are involved in protein folding and prevention of protein aggregation (Neidhardt et al., 1981; Tilly et al., 1983a; Bardwell & Craig, 1984; Bardwell et al., 1986; Kitagawa et al., 1991; Tomoyasu et al., 2001). The regulon comprises also all of the important cytosolic proteases Lon (La), ClpP, ClpX, HslV (ClpY), HslU, and (ClpQ) (Goldberg, 1972; Gayda et al., 1985; Maurizi et al., 1990; Chuang et al., 1993), and the membranal metalloprotease FtsH (HflB) (Herman et al., 1995; Tomoyasu et al., 1995). Other important σ^32-regulated proteins are HTS (Homoserine transsucyclinylase), which is a key enzyme in methionine biosynthesis (Biran et al., 1995), proteins involved in protein isomerization [PpiD (Dartigalongue & Raina, 1998) and HtrM (Raina & Georgopoulos, 1991)], and the vegetative sigma factor (σ^70) (Burton et al., 1981).

Homologs of rpoH were identified in more than twenty species of eubacteria from α, β, and γ subgroups of proteobacteria (Narberhaus et al., 1997; Sahu, Chowdhury, & Das, 1997; Andersson et al., 1998; Emetz & Klug, 1998; Huang, Tseng, & Yang, 1998; Karls et al., 1998; Nakahigashi, Yanagi, & Yura, 1998, 2001; Nakahigashi et al., 1999). In some of these bacteria, the rpoH homologs demonstrates translational induction and stabilization upon heat-shock that are very similar to those found in E. coli (Nakahigashi, Yanagi, & Yura, 1998).

The general function of the σ^32 regulon was studied in several bacterial species by analysis of rpoH mutants. These mutants were usually found to be temperature-sensitive (Huang, Tseng, & Yang, 1998; Zhou et al., 1988; Nakahigashi et al., 1999). As expected from their temperature-sensitive phenotype, some of the heat-shock proteins are essential at elevated temperature.

The levels of σ^32 and its activity are temperature-regulated at several levels. At low temperature (30°C), the intracellular levels of σ^32 is fewer than 50 molecules per cell (Straus, Walter, & Gross, 1987; Craig & Gross, 1991). These low levels are maintained due to transcriptional repression and protein instability, whereas low amounts of heat-shock proteins are required. Upon a rapid shift to 42°C, the level increases 15–20-fold within 5 min, and declines to a new steady state level, 2–3-fold higher than the pre-shift level (Straus, Walter, & Gross, 1987). The levels and the time-course of σ^32 induction are sufficient for the necessary induction of heat-shock-gene expression upon temperature upshift. A relatively modest heat-shock activates the translation of rpoH transcripts, and transiently stabilizes σ^32 (Straus, Walter, & Gross, 1987; Nagai, Yuzawa, & Yura, 1991), whereas a severe heat-shock (a rapid shift from 30 to 50°C) can also activate rpoH transcription (Yura, Kanemori, & Morita, 2000). The decrease in the synthesis of heat-shock proteins upon temperature downshift is primarily a result of the decrease in σ^32 activity (rather than its levels) caused mainly by an excess of the DnaK-chaperone machinery (Straus, Walter, & Gross, 1989; Taura et al., 1989).

The transcriptional regulation of the rpoH gene is very complex. It can be transcribed from at least four promoters, three of them (P1, P4, and P5) are recognized by the vegetative σ^70, and the fourth (P3) is recognized by σ^32 (Erickson et al., 1987; Nagai et al., 1990). P3- and P4-transcription of rpoH is negatively regulated by DnaA (Wang & Kaguni, 1989), and P4-P5 transcription controlled by an additional negative control system—the cAMP-CRP/CytR nucleoprotein complex (Kallipolitis & Valentin-Hansen, 1998).

Several findings indicate that the heat-shock induced σ^32 levels are also regulated at the translational level. Expression of rpoH-lacZ translational fusion, but not transcriptional fusion, can be induced. Furthermore, heat induction of the fusion protein occurs even when RNA synthesis is inhibited (Nagai, Yuzawa, & Yura, 1991). Recent results based on extensive in vivo and in vitro experiments related to the secondary RNA structure have shown that the translation regulation of RpoH is mediated by the rpoH mRNA’s secondary structure (Morita et al., 1999; Morita et al., 2000; Yura, Kanemori, & Morita, 2000).

Sigma-32 level is regulated not only by its expression level but also by the turnover of the protein. Although this protein is unstable during normal growth at 30°C (or even at 42°C), significant stabilization occurs immediately upon temperature upshift from 30°C to 42°C and continues for 4–5 min (Straus, Walter, & Gross, 1987). The protein instability involves the DnaK chaperone machinery. Mutants in DnaK, DnaJ, or GrpE markedly stabilize σ^32 under non-stress conditions (Tilly et al., 1983b; Tilly, Spence, & Georgopoulos, 1989; Straus, Walter, & Gross, 1990), indicating this involvement of these proteins in σ^32 turnover. The initial studies suggested that the membrane-associated metalloprotease FtsH (HflB) is responsible for σ^32 degradation (Herman et al., 1995; Tomoyasu et al., 1995). However, later studies were able to show that the cytosolic proteases Lon (La), HslVU, and ClpP are also
involved in $\sigma^{32}$ degradation (Wawrzynow et al., 1995; Kanemori et al., 1997; Kanemori, Yanagi, & Yura, 1999; Yura, Kanemori, & Morita, 2000). Although it is difficult to determine the relative significance of each protease in $\sigma^{32}$ degradation, it seems that the latter three proteases take a significant part of the degradation, possibly even equivalent to that of FtsH (Kanemori et al., 1997). It is assumed that during heat-shock the DnaK machinery and the proteases become occupied by the misfolded and unfolded proteins that accumulate because of the denaturing effect of the temperature increase. Consequently, there are insufficient levels of the proteolytic machinery to bring about $\sigma^{32}$ degradation, and it accumulates and activates the transcription of the heat-shock genes. Because the DnaK chaperones and the proteases have $\sigma^{32}$ promoters, their synthesis is increased, and, therefore, a few minutes after the temperature upshift, the level of the proteases and chaperones is high enough to destabilize $\sigma^{32}$, bringing the level of the heat-shock proteins to a new steady state.

The final level of $\sigma^{32}$ regulation is activity regulation (Yura, Kanemori, & Morita, 2000). This regulation operates mainly by creating ternary complexes of (DnaK-ADP)-DnaJ-cyclophilin (Yura, Kanemori, & Morita, 2000). Although it is difficult to determine the relative significance of each protease in $\sigma^{32}$ degradation, it seems that the latter three proteases take a significant part of the degradation, possibly even equivalent to that of FtsH (Kanemori et al., 1997). It is assumed that during heat-shock the DnaK machinery and the proteases become occupied by the misfolded and unfolded proteins that accumulate because of the denaturing effect of the temperature increase. Consequently, there are insufficient levels of the proteolytic machinery to bring about $\sigma^{32}$ degradation, and it accumulates and activates the transcription of the heat-shock genes. Because the DnaK chaperones and the proteases have $\sigma^{32}$ promoters, their synthesis is increased, and, therefore, a few minutes after the temperature upshift, the level of the proteases and chaperones is high enough to destabilize $\sigma^{32}$, bringing the level of the heat-shock proteins to a new steady state.

The final level of $\sigma^{32}$ regulation is activity regulation (Yura, Kanemori, & Morita, 2000). This regulation operates mainly by creating ternary complexes of (DnaK-ADP)-DnaJ-cyclophilin (Yura, Kanemori, & Morita, 2000). GrpE binds to the ternary complex and stimulates ADP release and complex dissociation by ATP binding. This cycle of binding and release appears to play an important role in $\sigma^{32}$ activity (and possibly stability) in vivo (Gamer, Bujard, & Bukau, 1992; Yura, Kanemori, & Morita, 2000).

2. Genes Controlled by Other Minor Sigma Factors

Another alternative sigma factor involved in the heat-shock response is $\sigma^E$ ($\sigma^{32}$), which was found to be an essential gene in E. coli at all temperatures (De Las Penas, Connolly, & Gross, 1997). As mentioned above, $\sigma^E$ activates transcription of rpoH under conditions of severe heat-shock, and because it has a $\sigma^E$ promoter, it also regulates itself. The $\sigma^E$ regulon protects cells against extracytoplasmic stress-derived damage. Other members of this regulon are periplasmic proteases that are involved in protein metabolism: the protease DegP (HtrA) and the periplasmic peptidyl prolyl isomerase FkpA (Erickson & Gross, 1989; Strauch, Johnson, & Beckwith, 1989; Dartigalongue & Raina, 1998). In addition, there are proteins in this regulon that are also regulated by the Cpx two-component system that mediates the synthesis of periplasmic-protein-folding factors such as DsbA (disulfide bond-forming enzyme); these reactions indicate that there are at least two overlapping regulatory systems involved in the periplasmic response to stress (Connolly et al., 1997; Danese & Silhavy, 1997; Pogliano et al., 1997).

The phage shock-protein operon (pspABCE) is strongly induced in response to a variety of stress conditions, including severe heat-shock. The transcription of this operon is regulated by a third minor sigma factor-$\sigma^{54}$ and by its positive regulator called PspE, a member of the enhancer-binding family (Jovanovic, Weiner, & Model, 1996). This activator binds specifically to the cis-acting sequence upstream of the pspA-E operon via a helix-turn-helix motif in its C-terminal. The cis-acting sequence contains two binding sites: UAS I (upstream activating sequence I) and UAS II, which together constitute the psp enhancer (Jovanovic, Weiner, & Model, 1996; Jovanovic, Rakonjac, & Model, 1999).

3. HrcA-CIRCE-Controlled Genes

For several years, the HrcA-CIRCE system was found only in Gram-positive bacteria, and was considered as a Gram-positive heat-shock control element. However, since the first discovery of the CIRCE element in the Gram-negative $\alpha$-proteobacterium A. tumefaciens (Segal & Ron, 1993), many CIRCE elements were identified in other Gram-negative bacteria. The inverted repeat was detected in a large number of phylogenetically distant bacteria, including Gram-negative bacteria of the $\alpha$, $\beta$, and $\gamma$-purple proteobacteria. The only groups where it is probably not present at all are the $\gamma$2 and $\gamma$3 purple bacteria, which also include the Gram-negative model organism E. coli (Segal & Ron, 1996b; Ron et al., 1999). The inverted repeat (TTAGCACTC-N9-GAGTGCTAA) is highly conserved in all of the studied genes (Segal & Ron, 1996b, 1998). In contrast to Gram-positive bacteria, where CIRCE-regulated genes are transcribed with the vegetative sigma factor ($\sigma^V$), in A. tumefaciens the groEL operon is HrcA-CIRCE controlled, but is transcribed mainly by $\sigma^{32}$ (Nakahigashi et al., 1999). In A. tumefaciens, it was possible to show, using 2D gels, that GroE proteins are the only proteins whose synthesis is repressed by the HrcA-CIRCE system (Rosen et al., 2002b). In Bradyrhizobium japonicum, two groESL operons were found: groESL1 is $\sigma^{32}$ regulated whereas groESL2 is CIRCE-HrcA-$\sigma^{54}$ dependent (Ron et al., 1999). The control of chaperone expression by the HrcA-CIRCE system seems to be more ancient than the $\sigma^{32}$-dependent transcription of heat-shock genes because it is found in all the bacteria except two small groups that lost it during evolution, whereas $\sigma^{32}$-dependent transcription is found only in Gram-negative bacteria (Ron et al., 1999).

4. Minor Regulatory Elements

Expression of at least ten genes in B. japonicum, seven of which code for small heat-shock proteins, is under the
control of ROSE (repression of heat-shock gene expression) (Narberhaus et al., 1998; Munchbach, Nocker, & Narberhaus, 1999a). This negatively cis-acting DNA element confers temperature control to a σ70-type promoter. ROSE elements are not restricted to B. japonicum, but are also present in Bradyrhizobium sp. (Parasponia), Rhizobium sp. strain NGR234, and Mesorhizobium loti (Nocker et al., 2001). The latest model for ROSE activity suggests that ROSE controls heat-shock protein expression by a temperature-dependent secondary structure of ROSE mRNA that controls the access of the ribosome to the ribosome-binding site (Nocker et al., 2001).

Proteome analysis of A. tumefaciens in various stress conditions, including heat-shock, indicated that most of these stress-induced proteins (80/97) were specific to one stress stimulon (Rosen et al., 2001). Analysis of mutants in rpoH, hrcA, or in both showed that the heat-shock induction of 32 (out of 56) heat-shock proteins is independent of RpoH and HrcA. These results indicate the existence of additional regulatory factors in the A. tumefaciens heat-shock response (Rosen et al., 2002b). These uncharacterized regulatory elements may also involve ROSE because A. tumefaciens belongs to the Rhizobiaceae group.

An additional unique post-transcriptional control mechanism was demonstrated in A. tumefaciens involved a specific cleavage of the groESL operon transcript. The resulting groES transcript is rapidly degraded, whereas the groEL transcript is stable, leading to a differential expression of the two genes of the operon—as could be detected by quantitative analysis of the protein expression, using 2D-gels (Segal & Ron, 1995, Rosen et al., 2002b). This mRNA processing is temperature-dependent, and constitutes the first example of a controlled processing of transcripts in bacteria.

C. Networks of Positive and Negative Regulators

Various positive- and negative-control elements are involved in the regulation of the bacterial heat-shock response. Some genes/operons are regulated by several control elements [e.g., E. coli pspABCE (Jovanovic, Weiner, & Model, 1996), A. tumefaciens groESL (Segal & Ron, 1995, 1996a; Nakahigashi et al., 1999), and B. subtilis clpC (Kruger, Msadek, & Hecker, 1996; Gerth et al., 1998; Scharf et al., 1998)]. Other heat-shock proteins are directly regulated by only one control element. However, the heat-shock response is always a complex response that regulates itself—it is induced by damaged proteins, but because it contains the genes for proteases and chaperones, which are induced, the level of the damage proteins is reduced, thus reducing the level of induction of the heat-shock response.

Because the regulatory elements of these complex heat-shock response networks are associated with each other, any impairment of the cellular steady state at one point may affect directly or indirectly the whole network. The study of these global regulatory networks requires global analysis methods. Methods such as transcriptome analysis were recently implemented to this field (Petersohn et al., 2001). However, as demonstrated before, analysis of mRNA levels is required but insufficient, because higher regulatory levels control the expression and activity of the heat-shock genes. Thus, it is necessary to also perform global analysis at the protein level—proteomics studies. The contribution of proteomics studies to the study of the heat-shock response in bacteria and the potential of these methods for future studies will be discussed in the next three sections. It should be noted that this review is limited to the study of the heat-shock response in bacteria; however, similar technologies will solve even more complex problems in eukaryotic organisms.

IV. PROTEOME STUDIES OF HEAT-SHOCK REGULONS

The first recognition of the molecular richness of the bacterial heat-shock response took place with the introduction, in 1975, of high-resolution two-dimensional polyacrylamide gels (2D gels) to resolve complex mixtures of cellular proteins (O’Farrell, 1975). This technology introduced a phase shift in the study of bacterial physiology by setting up a new basic experimental approach (Neidhardt & VanBogelen, 2000). Previously, one or more proteins were chosen for the study of physiological phenomena; using the proteomic approach, it is possible to obtain information that will indicate which specific proteins should be studied (Neidhardt & VanBogelen, 2000). In its early days, the method suffered from many technical problems, mainly due to the low reproducibility of the carrier ampholine-based electrophoresis. However, since the introduction of IPG (immobilized pH gradient) gels (Gorg et al., 1988) for the separation of proteins in the first dimension, the 2D gels became a highly reproducible protein-separation method, which is usually the method of choice for a proteome study. Furthermore, in the last few years a large number of image analysis methods have been developed for the quantitative comparison of 2D-gel spot volumes. These types of software enable the rapid comparative analysis of large sets of 2D gels. In addition, visualization methods such as dual-channel imaging (Bernhardt et al., 1999) allows the rapid detection of heat-shock (stress)-induced proteins, by combining autoradiography and a silver-stained electropherogram.

Two-dimensional gel electrophoresis is a very powerful tool to resolve proteins. However, this method is incomplete because the routine analysis does not enable the analysis of several groups of proteins: membrane proteins,
highly alkaline proteins, and rare proteins. Many efforts to find a standard procedure to solubilize membrane proteins before separation by 2D gel electrophoresis have not been successfully put into practice. In the few successful cases, the differential extraction of membrane proteins with specific detergents was the method of choice (Santoni et al., 2000). The second group of proteins with low compatibility for studies in 2D gels is the group of alkaline proteins. The solutions for this group are easier than for the membrane proteins, because it is possible to separate them (with lower resolution) on wide pH-gradient gels (Ohlmeier, Scharf, & Hecker, 2000; Buttner et al., 2001). Because the number of highly alkaline non-membrane proteins is relatively low, this compromise in the resolving power of the gels does not affect the ability to identify the proteins with mass spectrometry techniques. The third group of proteins that is not easily studied on 2D gel is the rare proteins. Because the first-dimension gels have a limited capacity, the rare proteins are present in the final gels in concentrations that may be below the detection limits, of even the very sensitive radiolabeling methods. Solutions for the problem of rare proteins are usually based on a pre-fractionation of the proteome, either by chromatography, by free-flow electrophoresis, or by obtaining sub-cellular proteomes of defined compartments. Because some of the known heat-shock proteins are membrane-bound (e.g., *E. coli* PtsH), alkaline (e.g., *B. subtilis* DnaJ with a theoretical pI of 8.24), or rare (e.g., *E. coli* σ32), one should keep in mind the limits of the 2D procedure when analyzing the proteome. Although alternative technologies exist, such as multi-dimensional-HPLC coupled mass spectrometry and protein-chip techniques (which will not be discussed in this review), to the best of our knowledge they were not used in the study of the bacterial heat-shock response.

The first studies of a bacterial heat-shock response with 2D gels compared the proteomes of *E. coli* grown under balanced growth conditions with proteomes obtained after exposure to elevated temperatures (Neidhardt et al., 1981; VanBogelen, Kelley, & Neidhardt, 1987a). These experiments led to an understanding of the extent of the heat-shock response. Later, similar studies were performed in other bacteria (e.g., *B. subtilis*, *B. Japonicum*, *Brucella melitensis*, *A. tumefaciens*, and *Vibrio* sp. strain S14) to determine which proteins are involved in the heat-shock response, and how many of them are induced under more than one stress condition (Richter & Hecker, 1986; Nystrom, Olsson, & Kjelleberg, 1992; Bernhardt et al., 1997; Munchbach et al., 1999b; Teixeira-Gomes, Cloeckaert, & Zygmunt, 2000; Buttner et al., 2001; Rosen et al., 2001). The stress proteins induced under the various stress conditions can overlap with each other to different extents—
as can be seen by the five models presented in Figure 4. There are stimulons that are completely overlapped (Fig. 4A), such as the *E. coli* heat-shock stimulon and its ethanol stimulon. The model presented in Figure 4B represents the case of the cadmium chloride-induced stress that comprises all of the heat-shock-induced proteins and additional CdCl2-specific proteins (VanBogelen, Kelley, & Neidhardt, 1987a). In *B. subtilis*, the σB-regulated group of general stress proteins is the core of the response to various stress conditions. In addition to this group of proteins, stress-specific proteins are induced by heat-shock, salt-stress, acidic-stress, and other stress conditions (Richter & Hecker, 1986; Bernhardt et al., 1997; Buttner et al., 2001); this model is represented by Figure 4C. The induction of heat-shock proteins, acidic-stress proteins, and oxidative-stress proteins in *A. tumefaciens* (Rosen et al., 2001) and *B. melitensis* (Teixeira-Gomes, Cloeckaert, & Zygmunt, 2000) resembles the model presented in Figure 4D. Finally, the model found in Figure 4E, which represent non-overlapping stress stimulons, is found between the oxidative stress response of *E. coli* and its SOS response (response to damage in the DNA) (VanBogelen, Kelley, & Neidhardt, 1987a). So far, only a small number of bacterial heat-shock proteomes were defined with 2D gels. However, the conditions for protein separation on 2D gels are now being determined for an increasing number of bacterial species. A by-product of this functional proteome analysis is the identification of many heat-shock proteins that are induced in response to various stimulons and during virulence bacterial host interactions (Chen et al., 2000; Monahan et al., 2001; Singh, Jayaswal, & Wilkinson, 2001; Giard et al., 2002).
The stimulons represent groups of proteins whose expression is induced coordinately following an environmental change. At the molecular level, each stimulon may be composed of more than one regulon, each controlled by a different molecular factor. The dissection of stimulons into regulons is based on the comparison of the induction pattern of wild-type bacteria with those of mutants in known regulatory elements. This method was used to define the $\sigma^B$-dependent proteins of *B. subtilis*, induced upon heat-shock and other stress conditions (Antelmann et al., 1997b; Petersohn et al., 2001). The same approach was used to define the *A. tumefaciens* RpoH- and HrcA-dependent regulons (Rosen et al., 2002b). The results of these experiments indicated that the known regulators controlled less than half of the induced proteins in the heat-shock stimulon. These results provided evidence for the existence of additional unknown regulators in the control of the heat-shock response of *A. tumefaciens* (Fig. 5). Such experiments will eventually be performed for additional systems, will help to define the known regulons, and will indicate ways to look for new regulons and regulatory elements. A complementary approach involves the study of transcriptions, indicating levels of gene-specific transcripts. This method also enables the comparison of two transcriptomes to identify genes induced at the transcriptional level. An overall comparison of the mRNA expression profiles with the protein expression pattern (determined on the 2D gels) can also be used to obtain differences between the two patterns: transcriptional and translational. Such differences will point out the existence of additional, higher regulatory levels.

One of the main inputs of this kind of global analysis of stress response in a large number of micro-organisms is the possibility to gain an understanding of evolutionary aspects of the control systems. Comparative analysis of stress proteomes and their organization into stimulons and regulons, in relation to phylogeny, will contribute to the comprehension of microbial physiology and its response to environmental conditions.

Additional information about the heat-shock response can be obtained from examining sub-proteomes. For example, proteins damaged or unfolded by elevated temperatures during heat-shock tend to aggregate. Proteome studies of the aggregates define the group of proteins with high thermal instability. In addition, the study of aggregates is important for the elucidation of quality control mechanisms, because damaged proteins can also be refolded with the aid of chaperons, or can be degraded by proteases (Gottesman, 1999). An example of one such study is the investigation of *E. coli* aggregates at various temperatures (Tomoyasu et al., 2001). These aggregates contain 350–400 protein species, which are substrates of the DnaK and ClpB chaperones. Similar studies indicated that the major cytosolic energy-dependent proteases are also involved in preventing aggregation, because in their absence the aggregates increase in concentration and contain more protein species (Rosen et al., 2002a). Following these experiments, it may be possible to use proteomes from aggregates to identify specific substrates for the various chaperones and proteases.

Another advantage of using 2D gels to analyze heat-shock proteomes is the ability to resolve post-translational modified proteins from their non-modified forms. As an

**FIGURE 5.** Heat-shock proteome of wild type (upper gel) and a $\Delta rpoH$ mutant (bottom gel) of *A. tumefaciens*. The circled spots in the upper gel are heat-shock proteins that are not induced under heat-shock conditions in the $\Delta rpoH$ mutant.
example— in many bacterial species the major chaperones, GroEL and DnaK, appear as multiple spots on a 2D gels. Moreover, the appearance of some of these spots is heat-shock-dependent; that dependency suggests heat-shock-dependent post-translational modification. Such is the case with A. tumefaciens, in which the DnaK protein appears in one spot during balanced growth (Rosen et al., 2001), but upon long exposure to heat-shock (70 min) two additional DnaK forms appear, which have a lower apparent molecular weight and a more acidic pI (Rosen R, Büttner K, Becher D, Hecker M, and Ron EZ, unpublished data). Those data suggest heat-shock-specific cleavage of DnaK.

Analysis of the B. subtilis proteome on 2D gels also revealed a large number of proteins in multiple spots. Some of these multiple spots represent heat-shock-induced proteins, whose level of induction differs between the various spots of the same protein species (Antelmann et al., 1997b; Bernhardt et al., 1997; Buttner et al., 2001). The characterization of the post-translational modification sites and the modification themselves will be discussed in the next section.

**V. MASS SPECTROMETRY IN THE ANALYSIS OF HEAT-SHOCK-INDUCED POST-TRANSLATIONAL MODIFICATIONS OF PROTEINS**

For the use of 2D gels as a global analysis method for protein expression, it is essential to be able to identify a large number of proteins, some of which are present in low quantities. Because bacterial genomes are typically simple, and the number of paralogs (multiple proteins that result from gene duplication within an organism) is rather low, MALDI-TOF-based peptide mass fingerprinting is usually sufficient for protein identification. A typical MALDI-TOF mass spectrum of A. tumefaciens trypsin-digested GroEL is presented in Figure 6. This MALDI-TOF-based peptide mass fingerprinting method was used for the identification of hundreds of bacterial proteins, including heat-shock proteins of E. coli, B. subtilis, B. japonicum, A. tumefaciens, H. pylori, and others (Munchbach et al., 1999b; Jungblut et al., 2000; Buttner et al., 2001; Tomoyasu et al., 2001; Rosen et al., 2002b).

Identification of protein spots in the 2D gel of E. coli, B. subtilis, B. japonicum, and A. tumefaciens (Buttner et al., 2001; Tomoyasu et al., 2001; Rosen et al., 2002b) revealed the existence of multiple spots of the same gene product that suggested post-translational protein modifications. It is interesting to note that some of the modifications appear to be regulated, because their concentration changes in response to heat-shock or other stress conditions. Several post-translational modifications of E. coli ribosomal proteins have been analyzed by MALDI-TOF peptide mass fingerprinting. The characterized modifications were: acetylation, methylation, β-methylthiolation, multiple methylation, and amino acid cleavage. It was possible to determine the nature of the modification and the modified peptide, but not the amino acid residue that was modified (Arnold & Reilly, 2002).

The exact modification site can be determined by using several different MS-MS techniques. Methods such as MALDI-TOF PSD (post-source decay), ESI/MALDI tandem hybrid QqTOF mass spectrometry, or ion trap mass spectrometry should allow peptide sequencing and specific characterization of modified amino acids. Although several post-translational modifications have been characterized on bacterial heat-shock proteins, many of them were characterized by traditional methods, rather than by MS methods. Here, we will review the post-translational modifications known to be involved in heat-shock-related processes, and will point out the one that was characterized by MS techniques.

**A. Phosphorylation**

In eukaryotic systems, protein phosphorylations are known to be involved in protein labeling and in many signal transduction pathways. In bacteria, the number of known phosphorylated proteins is much lower. However, several phosphorylated proteins are involved in the heat-shock response of various bacteria.

The heat-shock transcriptional activation of the lonD gene of Myxococcus xanthus is controlled by a two-component histidine-aspartate phosphorylation system (Ueki & Inouye, 2002). This finding is based on an in vitro transcription analysis of M. xanthus lonD, which showed stimulation in the presence of an extract from heat-shocked cells, for which the upstream promoter region of lonD was found to be essential. The promoter is bound by the lonD promoter-binding protein, HsfA, which is homologous to a response regulator of the NtrC family two-component system. The HsfA protein is phosphorylated by the histidine kinase HsfB prior to autophosphorylation of HsfB itself. In vitro, HsfA with HsfB activated the transcription of lonD in a phosphorylated-dependent manner. In that study, the phosphorylation was demonstrated by in vitro phosphorylation assays, using 32P, but there was no characterization of the phosphorylation site.

The general-stress sigma factor of B. subtilis (σB) is regulated by several regulatory kinases and phosphatases (the Rsb proteins), which catalyze the release of σB from an anti-σB factor (Akbar & Price, 1996; Yang et al., 1996; Akbar et al., 2001; Zhang, Scott, & Haldenwang, 2001). To determine phosphorylation sites in this system, potential target sites were directly mutagenized, and in vitro and in vivo assays of phosphorylation and σB activity were performed. In those experiments, the exact phosphorylation sites were determined (Akbar et al., 2001).
Another phosphorylation of a heat-shock protein was found in Hsp70 of *Mycobacterium leprae* at threonine-175 (Peake, Winter, & Britton, 1998). To study this phosphorylation, threonine-to-alanine mutants at positions 175 and 193, were generated by site-directed mutagenesis and phosphoamino acid analysis was performed. It was shown that Thr175 was the dominant threonine residue that was autophosphorylated in *M. leprae* Hsp70. The phosphorylation led to an increased affinity for a model polypeptide substrate-reduced and alkylated bovine albumin.

The RssB stationary phase response regulator targets σ^S for degradation by ClpXP (Zhou et al., 2001). σ^S is a highly unstable global regulatory protein in *E. coli*, whose degradation is inhibited by various stress signals, such as carbon starvation, high osmolarity, and heat-shock. Acetyl phosphate, which phosphorylates RssB, is required to deliver σ^S to ClpXP, as was demonstrated *in vitro*, with purified components (Bouche et al., 1998; Zhou et al., 2001).

These examples demonstrate the importance of protein phosphorylation in bacterial stress response. All the studies were performed by traditional methods, probably because the complex spectrometric analysis of small amounts of phosphorylated proteins needs specific isolation procedures before analysis, (Sickmann et al., 2001). In addition,
sensitive and accurate MS and MS/MS phosphoamino- and amino-acid sequence analysis methods are required.

B. Other Post-Translational Modifications

1. Protein Oxidation

Protein oxidation can take place whenever aberrant proteins are synthesized (Dukan et al., 2000). For protein-oxidation experiments, the level of aberrant proteins was elevated in *E. coli* cultures by decreasing transcriptional or translational fidelity, using specific mutations or drugs. Protein carbonylation, an oxidative modification, was found to be increased in parallel to the induction of the heat-shock chaperone GroEL. It appears that the carbonylation results from the increased susceptibility of the misfolded proteins. For the detection of the protein oxidation, an Oncor OxyBlot kit (Oncor Ltd.) was used to modified carbonyl groups on the proteins, and the detection was done by blotting. No determination of a specific carbonylation site was performed.

2. Sulfoxidation

The chloroplast-localized Hsp21 contains a conserved methionine-rich sequence, predicted to form an amphipatic-helix with the methionine residues situated along one of its sides. This methionine undergoes sulfoxidation as was detected by mass spectrometry in proteolytically cleaved peptides that were produced from recombinant Hsp21, which had been treated with varying concentrations of hydrogen peroxide (Gustavsson et al., 1999). Sulfoxidation of the methionine residues in the conserved amphipatic helix corresponds to a significant conformational change in the Hsp21 protein oligomer. In that study, mass spectrometry techniques were used to determine sulfoxidation, because it was possible to obtain large amounts of protein by an over-expression of recombinant protein.

3. Acetylation

Monoacetyl spermidine accumulation in *E. coli* is induced under heat-shock, cold shock, ethanol stress, and alkaline shift. Acetylation occurs with nearly equal frequency at the N1 and N8 positions of this ubiquitous polyamine. Spermidine acetylation does not appear to be associated with known stress regulons, such as htpR, oxyR, and SOS. *E. coli*, capable of acetylating spermidine, constitutively expresses a spermidine acetyltransferase activity during all phases of growth (Carper et al., 1991; Limsuwun & Jones, 2000). The acetylation analysis in these studies was performed by traditional chemical methods and not by mass spectrometry.

The experiments summarized here point out that traditional analytical methods for the determination of post-translational modifications are still preferred by researchers in the field of bacterial heat-shock response. It is, therefore, essential to demonstrate to those researchers the potential of the various available MS-MS techniques and to develop more “user-friendly” and accurate approaches for the study of post-translational modifications. These new-developed approaches should be sensitive enough to use on a single spot from a 2D gel—still a difficult problem—as described by Sickmann et al. (2001).

VI. POTENTIAL SIGNIFICANCE OF PROTEOME STUDIES FOR ANALYSIS OF REGULONS

In the previous sections, we described the state-of-the-art in the study of the heat-shock response, and the contribution of proteomics to this field. Essentially, proteomics tools were used to study the heat-shock response only in a few bacterial species, among which the full potential is being used only in two species, *B. subtilis* and *E. coli*. In this section, we will describe the potential and the drawbacks of proteomics studies of regulons, with an emphasis on the heat-shock stimulon.

Two-dimensional gel electrophoresis is a powerful tool to define stimulons and regulons. Using quantitative visualization methods, such as Coomassie Brilliant Blue stain, Sypro ruby stain, or radiolabeling combined with autoradiography, it is possible to determine rates of protein synthesis, protein stability, and relative amounts of individual proteins. Thus, it is possible in a single experiment to characterize the stimulon and to determine the decrease in the expression level or synthesis rate of other proteins. This definition of the proteome in specific physiological and environmental conditions is the foundation for the study of the role of individual proteins and regulatory networks involved in bacterial growth under these specific conditions. As mentioned before, partial data on the size and the composition of a stimulon can also be obtained by transcriptome analysis. However, transcriptome analysis is limited to transcriptional regulation and cannot predict the final induction level of the protein—the active level in the cellular machinery. Another advantage of proteome analysis is the ability to distinguish among sub-populations of the same gene product, because post-translational modified proteins will usually appear as separate spots from the non-modified forms (Bernhardt et al., 1997; Buttner et al., 2001; Rosen et al., 2002b). This separation of a population of a singular protein can be the basis to study the role of a specific modification in the protein activity. Proteome analysis can also be performed in other methods such as protein chips and multi-dimensional chromatography. With various methods of ICAT (Isotope-Coded Affinity Tags),
it is also possible to use mass spectrometry per se for quantitative proteome analysis (Griffin et al., 2001; Han et al., 2001; Turecek, 2002; Zhou et al., 2002). Thus, in our opinion the major unique contribution of 2D gel-based stress response studies will be in the area of post-translational modifications.

Post-translational modifications are currently detected mainly on purified proteins in high quantities that are suitable for traditional methods to detect post-translational modification. In our experience, it is possible to identify post-translational modifications on proteins obtained from 2D gels (Becher D, Rosen R, Büttner K, Hecker M, and Ron EZ, unpublished data) with ESI-QqTOF MS-MS, although it requires the collection of protein spots from a large number of parallel gels.

In addition, mass spectrometry has a unique role in the determination of disulfide bond formation, and in proving their involvement in the mechanism that activates molecular chaperones (Barbirz, Jakob, & Glocker, 2000). Using E. coli Hsp33, it was possible to demonstrate that the activation process is accompanied by the formation of two intramolecular disulfide bonds within the protein: Cys(232)-S-S-Cys(234) and Cys(265)-S-S-Cys(268). This finding indicates a significant conformational change during the activation process of Hsp33. In this case, mass spectrometry enabled the discovery of a novel molecular mechanism that alters the structure of the disulfide bond, and that activates a molecular chaperone.

The large number of regulated post-translational modifications that appear in response to various stress conditions suggests that these modifications have an important biological role, probably in regulating protein activity. These modifications will undoubtedly become an interesting research area, depending on the development of highly sensitive methods for the analysis.

ACKNOWLEDGMENTS

We thank Prof. Dr. Michael Hecker, Dr. Dörte Becher, and Knut Büttner for their stimulating input. We thank Mr. and Mrs. Colton for awarding a Colton Scholarship (R.R.). We thank the Israel Center for Emerging Diseases. This work was partly supported by the Manja and Morris Leigh Chair (E.Z.R.) for Biophysics and Biotechnology.

REFERENCES


Eymann C, Hecker M. 2001. Induction of sigma(B)-dependent general stress genes by amino acid starvation in a spo0H


Yuan G, Wong SL. 1995a. Isolation and characterization of *Bacillus subtilis* *groE* regulatory mutants: Evidence for *orf39* in the *dnaK* operon as a repressor gene in regulating the expression of both *groE* and *dnaK*. J Bacteriol 177:6462–6468.


