

# Genome-Wide Analysis of the Biology of Stress Responses through Heat Shock Transcription Factor

Ji-Sook Hahn,<sup>1</sup>†‡ Zhanzhi Hu,<sup>2</sup>† Dennis J. Thiele,<sup>1\*</sup> and Vishwanath R. Iyer<sup>2\*</sup>

Department of Biological Chemistry, University of Michigan Medical School, Ann Arbor, Michigan 48109-0606,<sup>1</sup> and Institute for Cellular and Molecular Biology and Section of Molecular Genetics and Microbiology, University of Texas at Austin, Austin, Texas 78712-0159<sup>2</sup>

Received 7 November 2003/Returned for modification 26 January 2004/Accepted 15 March 2004

**Heat shock transcription factor (HSF) and the promoter heat shock element (HSE) are among the most highly conserved transcriptional regulatory elements in nature. HSF mediates the transcriptional response of eukaryotic cells to heat, infection and inflammation, pharmacological agents, and other stresses. While HSF is essential for cell viability in *Saccharomyces cerevisiae*, oogenesis and early development in *Drosophila melanogaster*, extended life span in *Caenorhabditis elegans*, and extraembryonic development and stress resistance in mammals, little is known about its full range of biological target genes. We used whole-genome analyses to identify virtually all of the direct transcriptional targets of yeast HSF, representing nearly 3% of the genomic loci. The majority of the identified loci are heat-inducibly bound by yeast HSF, and the target genes encode proteins that have a broad range of biological functions including protein folding and degradation, energy generation, protein trafficking, maintenance of cell integrity, small molecule transport, cell signaling, and transcription. This genome-wide identification of HSF target genes provides novel insights into the role of HSF in growth, development, disease, and aging and in the complex metabolic reprogramming that occurs in all cells in response to stress.**

Heat shock factors (HSFs) are transcription factors that activate the expression of genes in response to stress, thereby playing a central role in cellular homeostatic control mechanisms (33). HSF is highly conserved in its overall fundamental structure from yeasts to humans, harboring a winged helix-turn-helix DNA binding domain, one or more hydrophobic coiled-coil regions essential for the stability and regulation of homotrimer formation, and a carboxyl-terminal *trans*-activation domain (27, 51). HSFs bind to the major groove of *cis*-acting DNA promoter elements called heat shock elements (HSEs), which are also functionally conserved from yeasts to humans, consisting of tandem inverted repeats of the short consensus sequence 5'-nGAAn-3' (28, 51, 52). A number of posttranslational HSF modifications and protein-protein interactions have been described that contribute to both the activation and the repression of HSF activity *in vivo* (18, 32).

While the fundamental structure and function of HSF and the HSEs are conserved from yeasts to humans, there is wide variability in the numbers of HSF genes in nature. The baker's yeast, *Saccharomyces cerevisiae*, harbors a single HSF essential for cell survival (44, 50). Yeast HSF has both amino-terminal (NTA) and carboxyl-terminal (CTA) *trans*-activation domains,

which are thought to differentially contribute to the activation of specific target genes as well as in the regulation of HSF activity (7, 42). While initial studies suggested that yeast HSF is constitutively trimerized and bound to HSEs (22, 43), subsequent reports have demonstrated that yeast HSF binds constitutively to specific HSEs in the *HSP82* gene promoter and in a stress-inducible manner to other HSEs within this promoter (10, 13, 40). In contrast, although *Drosophila melanogaster* encodes a single HSF, this gene has been shown to be dispensable for cell growth and viability but essential for oogenesis, early larval development, and survival in response to acute stress (23).

Plants and mammals harbor multiple genes encoding HSF isoforms, with *Arabidopsis thaliana* possessing 21 distinct HSF genes and mammals possessing three genes encoding the HSF isoforms HSF1, HSF2, and HSF4 (34, 36). Targeted-deletion studies of the stress-responsive mouse HSF1 gene demonstrate that, while HSF1 is not essential for viability, it is essential for both basal and stress activation of heat shock protein (Hsp) gene expression, normal growth, extraembryonic development, inflammatory responses, fertility, and resistance to stress-induced apoptosis (9, 31, 53). Recent studies of *Caenorhabditis elegans* have demonstrated that a reduction in HSF activity shortens life span while HSF overexpression extends life span (11, 19). These and other studies strongly suggest that specific HSF target genes promote longevity (49).

Given the essential nature of yeast HSF and the range of phenotypes of HSF-knockout flies, worms, and mice affecting normal growth and development, stress resistance, and aging, the constellation of HSF target genes is likely to encompass additional, nonclassical Hsp genes. Furthermore, given the molecular events that accompany cellular responses to both chronic and acute stress, these adaptive responses are likely to

\* Corresponding author. Mailing address for Dennis J. Thiele: Department of Biological Chemistry, University of Michigan Medical School, Ann Arbor, MI 48109-0606. Phone: (919) 684-5776. Fax: (919) 668-6044. E-mail: dennis.thiele@duke.edu. Mailing address for Vishwanath R. Iyer: Institute for Cellular and Molecular Biology and Section of Molecular Genetics and Microbiology, University of Texas at Austin, Austin, TX 78712-0159. Phone: (512) 232-7833. Fax: (512) 232-3432. E-mail: vishy@mail.utexas.edu.

† J.-S.H. and Z.H. contributed equally to the work.

‡ Present address: Department of Pharmacology and Cancer Biology and the Sarah W. Stedman Nutrition and Metabolism Center, Duke University Medical Center, Durham, NC 27710.

invoke changes in the expression of a number of genes. However, only a small group of direct HSF target genes have been identified, which largely comprise the Hsp protein chaperones such as Hsp70, Hsp90, and the small Hsps (4, 35, 47). To begin to understand the precise molecular responses to stress and the central role played by HSF, we have used chromatin immunoprecipitation (ChIP) combined with DNA microarray approaches to identify virtually all of the direct target genes bound by *S. cerevisiae* HSF in vivo. Contrary to previous assumptions, these studies revealed that HSF is stress-inducibly bound to the majority of its targets. Newly identified HSF targets represent nearly 3% of the genome, and the diversity of their functions supports a central and broad role for HSF in orchestrating the multitude of cellular reprogramming events that occur in response to the stress of normal cell growth and as a consequence of acute stress. Furthermore, the identification of direct HSF target genes in yeast provides a simple model system to identify HSF and stress-responsive genes that may function in aging and disease.

#### MATERIALS AND METHODS

**Yeast strains.** Yeast strain W303-1A (*MATa ade2-1 can1-100 his3-11,15 leu2-3,112 trp1-1 ura3-1*) was used for ChIP and RNA blotting experiments.

**ChIP, microarray hybridization, and data analysis.** Yeast cells were grown in yeast extract-peptone-dextrose medium to an  $A_{600}$  of 0.7 at 30°C, and proteins were cross-linked to their DNA binding sites by adding formaldehyde to the cultures before heat shock for the time zero point or at indicated time intervals after heat shock at 39 or 42°C. For the time course experiment shown in Fig. 2B, formaldehyde was added to aliquots of the cultures at the indicated time intervals after heat shock. HSF-cross-linked DNA was isolated, sheared, and immunoprecipitated using anti-HSF antibody (a gift of David Gross) and purified after reversal of the cross-linking as described previously (21). The immunoprecipitated DNA and control DNA from the whole-cell extracts were amplified by PCR and fluorescently labeled with the Cy5 and Cy3 fluorophores, respectively, according to published protocols (20). Fluorescently labeled probes were cohybridized to a DNA microarray containing all of the intergenic and predicted coding regions of the yeast genome manufactured as described previously (21). The resulting microarrays were scanned with the GenePix 4000B scanner (Axon Instruments) and quantitated with GenePix 4.0 software. Data were uploaded into a relational database and filtered to pass minimum quality control thresholds before subsequent analysis of genomic enrichment (25).

**Determination of genomic enrichment.** For each genomic locus we defined the IP enrichment score  $E$  as  $E = \text{median}[R(\text{Cy5}/\text{Cy3})_1, R(\text{Cy5}/\text{Cy3})_2, \dots, R(\text{Cy5}/\text{Cy3})_n]$ , where  $R(\text{Cy5}/\text{Cy3})_i$  is the percentile rank of the Cy5/Cy3 ratio of the microarray element corresponding to that locus among all genomic loci in the  $i$ th independent ChIP experiment.  $E$  is expressed as a percent with values between 0 and 100 and is thus a measure of consistent in vivo association of HSF with a given genomic locus. If enrichment in each independent experiment were to be completely uncorrelated with other experiments, we assume that  $E$  would be normally distributed. We plotted IP enrichment ( $E$ ) against two different measures of heat shock induction (Fig. 1), and based on this graph, loci with  $E$  values greater than a threshold of 97.7 in heat shock experiments were considered to be direct binding targets of HSF. The graphs in Fig. 1 were generated using a sliding window of size 100. This window size yielded a better balance between reducing noise and maintaining sensitivity than did window sizes that were larger or smaller. To choose the cutoff value of 97.7, we examined the list of target genes sorted in descending order of their enrichment values and chose the threshold that would ensure maximal coverage of all genes that were significantly induced by heat shock and yet minimize the number of false positives. The average enrichment value for the top 210 genes was significantly greater ( $P = 5.5 \times 10^{-289}$ ) than if the ChIP experiments had been uncorrelated with each other.

**RNA blots and PCR assays.** Total RNA was isolated from yeast cells grown in yeast extract-peptone-dextrose medium and fractionated on 1.5% agarose-formaldehyde gels, and RNA blots were probed with  $^{32}\text{P}$ -labeled DNA fragments derived from PCR-amplified DNA fragments encompassing specified open reading frames. For PCR analysis of immunoprecipitated DNA, 2% of each IP sample was used to amplify specific promoter regions. PCR was performed for 27 cycles of 30 s at 94°C, 30 s at 50°C, and 1 min at 72°C.

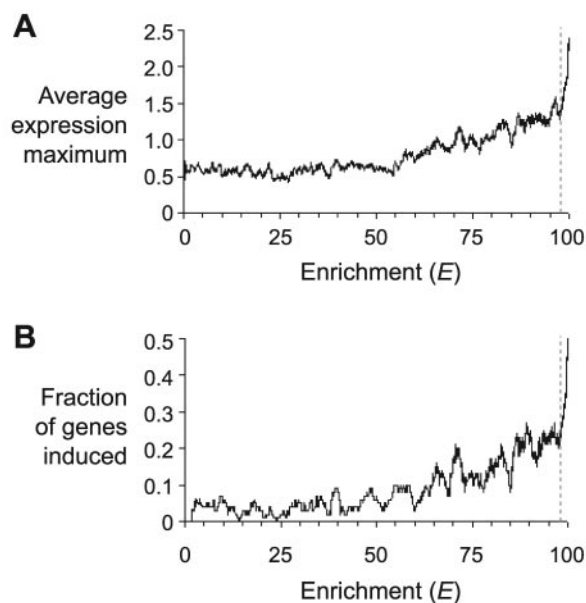


FIG. 1. Determination of threshold for HSF targets. Moving-window average analysis shows the relationship of HSF binding to downstream gene expression in heat shock experiments. Expression data are derived from previously published work (12). Genomic enrichment  $E$  values of loci indicating potential association with HSF based on 20 independent heat shock experiments are plotted on the x axis. The corresponding gene expression value of genes in a sliding window of size 100 is plotted on the y axis with the use of two different measures. (A) Average peak expression value (with  $\log_2$  scale) after heat shock relative to time zero in two independent data sets. (B) Fraction of genes within the window that are induced by at least threefold in at least two time points in the expression data sets. The threshold  $E$  value of 97.7 for designating a locus as an HSF target is indicated by the gray dashed line.

#### RESULTS

**Identifying the genome-wide targets of *S. cerevisiae* HSF.** We determined the in vivo association of *S. cerevisiae* HSF with each genomic locus based on 14 independent cross-linking experiments under unstressed conditions (30°C) and 20 independent cross-linking experiments under heat shock conditions (20 min at 39 or 42°C). There were no significant differences between the binding distribution of HSF to its targets at 39°C and that at 42°C, and we therefore combined the results from 10 experiments from each heat shock condition to determine chromosomal binding site enrichment in response to heat shock. We defined a score,  $E$ , for consistent chromosomal enrichment across multiple independent ChIP experiments (Materials and Methods). Since *S. cerevisiae* HSF is known to function as an activator upon exposure of cells to heat shock, loci with high  $E$  values at high temperature are taken as the physiological binding targets of HSF. However, to set an appropriately stringent threshold for defining HSF targets, we used the fact that the physiological targets of yeast HSF are expected to be transcriptionally induced by heat shock. Based on the relationship between enrichment ( $E$ ) of a genomic locus and the expression level of the gene downstream of it as measured by a moving-window average analysis (Fig. 1), we set our operational threshold for defining targets at 97.7. Loci with ( $E$ ) values greater than this threshold under heat shock conditions

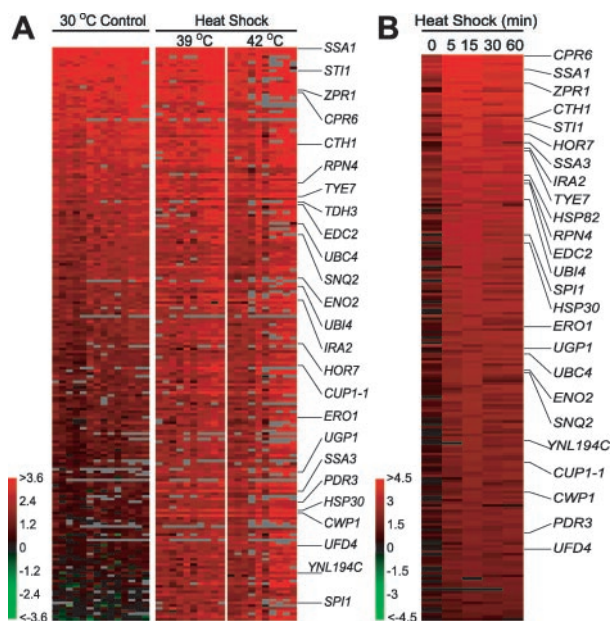


FIG. 2. Genome-wide binding distribution of HSF. Cy5/Cy3 ratios of genomic fragments with enrichment values above the threshold are displayed using a red and green color scale. Gray cells indicate missing data that were filtered out. (A) Each column represents an IP performed after independent HSF cross-linking carried out under 30°C unstressed conditions or after a 20-min heat shock at the indicated temperature. Genes downstream of selected target loci are shown on the right. The complete list of target loci and their ratio values are included in supplemental data. (B) Binding of HSF to all targets during the heat shock time course. HSF ChIP samples prepared at the indicated times during the 39°C heat shock were analyzed by DNA microarrays.

tend to be strongly induced by heat shock and are thus defined as the putative physiological binding targets of HSF *in vivo*.

Of all loci evaluated, 210 loci representing sequences upstream of 165 distinct open reading frames had enrichment values above this threshold (Fig. 2A) (supplemental data; see <http://www.iyerlab.org/hsf> for a complete list of targets). When a binding locus occurred between two divergently transcribed genes, either one or both of the downstream targets were designated as a putative target, depending on their expression profile during heat shock. Of the 210 target loci 146 were intergenic regions and an overwhelming majority of these (140, or 96%) were upstream of open reading frames rather than intergenic regions that are not upstream of any open reading frame. The enrichment of heat shock-induced genes in this group of loci is significantly higher than would be expected if enrichment were uncorrelated with heat shock expression ( $P = 1.3 \times 10^{-33}$ ).

The conclusion that this analysis has identified direct *in vivo* targets of *S. cerevisiae* HSF is supported by the following observations, which will be further discussed in detail below. First, the putative *in vivo* binding sites that we identified by ChIP microarray analysis are highly enriched for sequences matching or closely related to the defined HSF DNA binding site HSE that is conserved from yeasts to humans (51). Second, the majority of genes downstream of the putative HSF binding sites displayed heat-inducible expression according to published microarray data sets (12). Third, putative target genes

with known functions in the stress response, such as protein chaperones, are highly enriched in this analysis. However, it is possible that we have misidentified some genes as putative HSF targets and failed to identify some true HSF direct targets. For example, HSF may activate some genes only under a specific set of conditions; may repress the expression of other genes, as has been proposed for mammalian HSF1 (54); or may bind to specific chromosomal HSE sites without engaging in the regulation of transcription. It has also been reported that HSF binding is influenced by the yeast cell cycle (48), and our experiments were performed on asynchronous cell cultures.

**Regulation of HSF DNA binding activity.** The stress-responsive HSF isoforms of flies and mammals exist in unstressed cells as a monomer which in response to stress is converted to homotrimers that bind HSEs with high affinity and accumulate in the nucleus (32). Early experiments with *S. cerevisiae* HSF suggested that HSF binds constitutively to target gene promoters, even in the absence of heat shock (22, 43). However, *in vivo* footprinting and ChIP studies with the *HSP82* promoter demonstrated heat-inducible and cooperative binding of *S. cerevisiae* HSF to weaker HSEs, as well as constitutive binding to another, stronger HSE (10, 13). Our genomic data from 34 independent ChIP experiments showed a marked increase in the binding of *S. cerevisiae* HSF to a large subset of its target promoters after heat shock *in vivo* compared to that for cells grown at 30°C (Fig. 2A). Although HSF appears to bind many of its strongest target promoters equally well at 30°C and after heat shock, other prominent target loci such as the promoters of the *SSA3* and *HSP30* genes showed clear heat-inducible binding. This is the first evidence demonstrating that, similarly to higher eukaryotic cells, *S. cerevisiae* HSF binds to many of its targets in a heat-regulated manner in a much broader way than previously appreciated. Putative target loci that showed larger binding differences (between 30°C and heat shock temperatures) tended to show slightly lower enrichment than did the loci which showed smaller binding differences; however, our independent IP experiments indicate that the heat-inducible binding is remarkably consistent, even for targets with low enrichment (Fig. 2A).

To further investigate the broad utilization of stress-induced HSF target binding, we analyzed the kinetics of HSF binding to the 210 target loci *in vivo* through a time course of ChIP-DNA microarray experiments. While at 30°C HSF was bound to several targets constitutively, many targets were bound at a low level; binding increased rapidly after heat shock, with maximal binding occurring approximately 5 to 15 min after heat shock depending on the target sites and diminishing at later time points (Fig. 2B) (supplemental data). The difference in the binding of HSF to its targets at low and high temperatures was more readily apparent in this time course experiment than was observed in Fig. 2A. This is likely due to the fact that maximal binding occurs between 5 and 15 min after heat shock while all the other 34 experiments analyzed binding 20 min after heat shock, when binding levels are reduced relative to the peak. These results suggest that there is a significant increase in HSF binding to even the most consistently strong target loci upon heat shock. The apparent increase in DNA binding affinity could be due to both stress activation of HSF binding and a stress-inducible increase in the stability of prebound HSF.

This transient binding profile was independently verified by

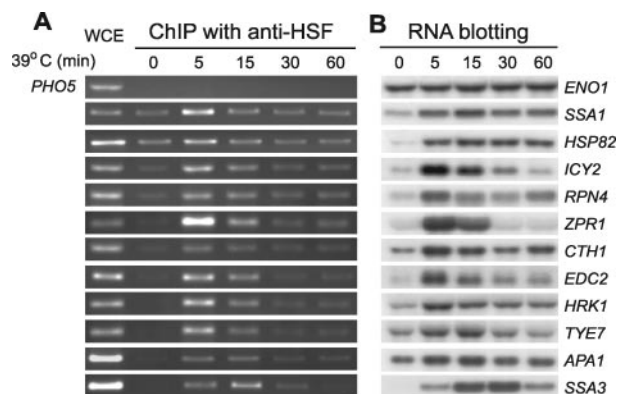


FIG. 3. Correlation between heat-inducible binding of HSF and mRNA expression. (A) Verification of HSF binding to selected targets by PCR. The same IP DNA used for microarrays as shown in Fig. 2B was used to verify binding of HSF to selected promoters by PCR. *SSA1*, *HSP82*, and *SSA3* were previously known targets of HSF, and the remainder were targets newly identified in this study. The *PHO5* promoter was used as a negative control. WCE, whole-cell extract. (B) Kinetics of HSF target gene expression by heat shock. Expression levels for the indicated genes were detected by RNA blot analysis. *ENO1* was used as a negative control.

PCR with primer pairs specific for newly identified HSF target gene promoters as well as previously known HSF targets with the same immunoprecipitated samples (Fig. 3A). Although the microarray data in Fig. 2B were generated from one representative set of experiments, several repeats of the time course IP experiments showed reproducible trends of heat-inducible binding of HSF to specific targets when HSF binding was detected by PCR. These data demonstrated that specific genes such as *SSA1* and *HSP82* exhibited constitutive HSF binding that was slightly but transiently elevated in response to heat shock over the 60-min period of the experiment. Other HSF targets (for example, *SSA3* and *ZPR1*) displayed robust and transient induction of HSF binding in response to heat shock. These binding profiles strongly parallel the transient mRNA expression profiles of the HSF target genes upon heat shock (Fig. 3B). For example, the mRNA levels of some chaperone genes such as *SSA1* and *HSP82* were maintained at high levels until 60 min after heat shock, consistent with the continued binding of HSF to their promoters over this time course, while *SSA3* mRNA induction upon heat shock was delayed compared with that of other targets, reflecting the observed delayed binding of HSF to the *SSA3* promoter.

The RNA blot analysis also verified several cases of the regulation of divergently transcribed genes by HSF bound in the intergenic regions; genes within the pairs *ZPR1-YGR210C*, *SIS1-LST8*, *SSC1-TAH11*, and *BUD7-MBF1* were all induced by heat shock, and this induction occurred in an HSF activation domain-dependent manner (data not shown). Taken together, these data demonstrate that *S. cerevisiae* HSF DNA binding is stress activated on a global scale and that the kinetics of this process suggests that it plays an important role in the activation of HSF target gene expression in response to heat shock.

**Expression profiles of HSF targets.** To further investigate the relationship between the DNA binding activity of the *S.*

*cerevisiae* HSF and transcriptional regulation of its corresponding binding targets, we examined the expression profiles of all HSF target genes within a previously published data set that measured changes in mRNA expression levels in response to heat shock and other environmental stresses (12). The data in Fig. 4 are from an S288C derivative. Although different strain backgrounds can potentially have different responses to stress, we note that in vivo binding of HSF and heat shock induction profiles are well correlated between W303 strains used in this study and an S288C derivative BY4741 strain (data not shown). The majority (77%) of the mRNAs representing putative HSF targets, identified by ChIP-DNA microarray experiments here, were rapidly and transiently induced by a factor of at least twofold during heat shock from 25 to 37°C, as evidenced by their expression profiles in two independent experiments (Fig. 4) (supplemental data). Ninety percent of the putative targets were induced by a factor of at least 1.5-fold. The set of target genes was also induced by variable temperature shocks. When cells were shifted from high to low temperature (a reversal of the heat shock), the putative targets were strongly down-regulated. This correlation suggests that the transcriptional response of the putative HSF target promoters to heat shock is mediated by HSF binding to these promoters. Interestingly,

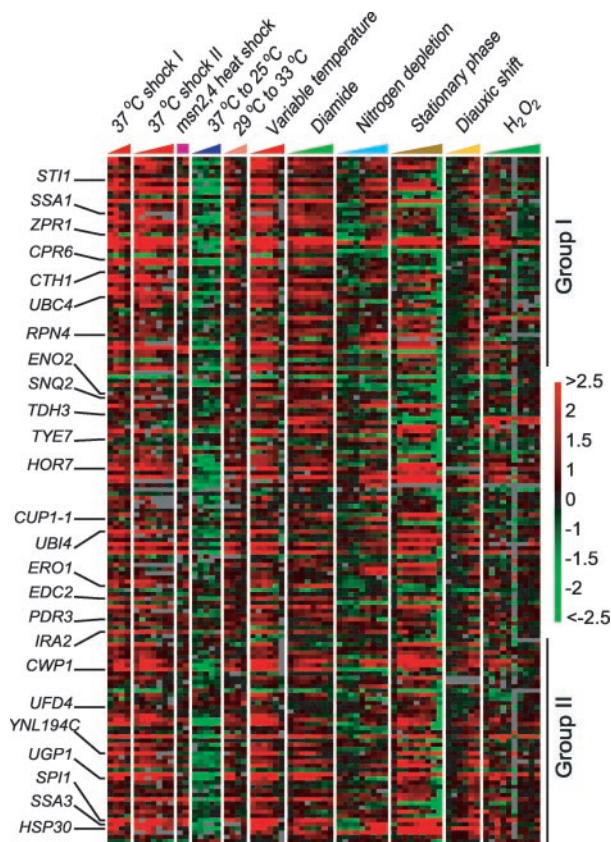


FIG. 4. Expression profiles of HSF targets. Expression during multiple distinct stress conditions of all HSF targets identified here is represented using a red and green scale. Expression data are taken from the work of Gasch et al. (12). Targets that showed the smallest difference in HSF binding between 30°C and heat shock conditions are listed on top. The subset of target promoters used for motif discovery is indicated on the right.

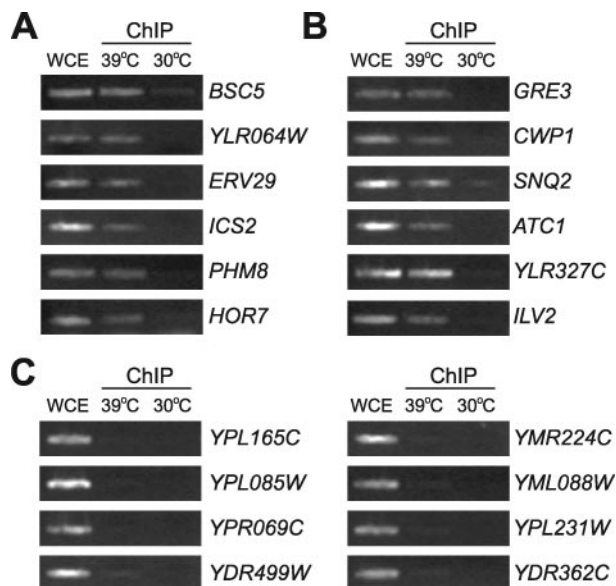


FIG. 5. Validation of microarray results. Promoter-specific PCR for the indicated genes was performed using ChIP samples that had earlier been analyzed on microarrays. (A) Genes predicted as targets by microarrays and containing a consensus HSE in their promoter. (B) Genes predicted as targets by microarrays and lacking a consensus HSE in their promoter. (C) Genes identified as nontargets by microarrays, containing a consensus HSE in their promoter. *ILV2*, *ATC1*, *ICS2*, and *BSC5* were not appreciably induced by heat shock. WCE, whole-cell extract.

most of the putative HSF targets also showed induction by other stresses in addition to heat shock. The targets were transcriptionally induced by treatment with the sulfhydryl oxidant diamide and during the transition to stationary phase. A smaller subset of the targets were induced, to slightly lower overall levels, by depletion of nitrogen sources and the diauxic shift. However, induction of the putative HSF targets does not reflect a generalized stress response, as indicated by the lack of appreciable induction of these targets following treatment with the oxidizing agent hydrogen peroxide.

A subset of the genes that were not induced at least twofold in the above data set appeared to be in fact moderately induced in other large-scale data sets measuring gene expression changes during heat shock (6). A few are potential false positives arising because of overlap between a bona fide target promoter enriched in the IP and another unrelated nearby promoter. Thus, only a small number of putative HSF targets (7% of the total) do not show any evidence of induction after heat shock. The enrichment values for the promoters of these genes indicated that they were consistently and strongly bound by HSF *in vivo*. This binding could be independently validated by promoter-specific PCR (Fig. 5). This anomalous expression profile is what would be expected of targets of HSF either that are repressed during stress or to which HSF might bind for another, as yet unknown function.

**Binding of HSF is directed by conserved motifs.** While the role of the highly conserved HSE in directing HSF binding to specific loci has been well established in yeasts and metazoans (33, 51, 52), a number of studies have demonstrated that there are distinct types of HSEs bound by *S. cerevisiae* HSF, as well

as distinct HSEs that are preferentially bound by the HSF1 and HSF2 isoforms in mammals (1). Indeed, the cooperative binding of *S. cerevisiae* HSF to multiple nonconsensus HSEs may be enhanced by heat shock or other stresses (10, 13, 45). To attempt to understand the determinants that govern the heat shock-dependent binding distribution of HSF within the *S. cerevisiae* genome, we applied the pattern-finding program MDscan (29) to distinct subsets of the HSF target loci identified here by ChIP-DNA microarray experiments. We first ranked the 210 HSF target loci by the difference in binding observed between cells exposed to 30°C and those exposed to heat shock conditions. We then designated sequences from the top 50 loci (43 unique sequences) that showed the smallest difference in binding between the two temperatures as group I and sequences from the bottom 50 loci (44 unique sequences) that showed the biggest difference in binding between the two temperatures as group II (Fig. 4). All of the top 10 motifs identified by MDscan from group I contained the canonical HSE, which is consistent with the notion that the canonical HSE motif directs constitutive binding of HSF. A consensus HSE is present upstream of about 46% of all the HSF target promoters that we identified here, compared with its presence upstream of approximately 27% of all the loci in the whole yeast genome ( $P = 8.1 \times 10^{-8}$ ). We could not identify any significant correlation between the extent of HSF binding and the composition of the consensus HSE in terms of the number of pentamer repeats or their orientation.

MDscan failed to identify the canonical HSE in any of the top 10 motifs overrepresented in group II. Instead, an HSE-like sequence containing a TTC motif as well as an STRE (stress-responsive element, 5'-AGGGG-3')-like motif (30, 39) was identified (supplemental data). While the consensus HSE has inverted repeats of the TTC/GAA motif with fixed spacing, the variant identified here has perfect repeats of TTC with variable spacing. The exact spacing of the 5'-nTTCn-3' inverted repeat in HSE has recently been found to be flexible for both the *CUP1* and *MDJ1* promoters (37, 45), and a TTC-rich sequence was also discovered in the *HSP82* and *HSP104* promoters (10, 14). These findings suggest that the TTC-rich sequence could be a less conserved variant of the canonical HSE which could be bound by HSF in a heat-inducible manner or also in response to other stresses. The STRE is the binding site for the transcriptional activators Msn2 and Msn4, which bind to the STRE following a variety of stresses such as heat shock and nutrient depletion and activate the transcription of downstream stress response genes (30). It is possible that the novel motif present in the group II subset of HSF target promoters is a variant of the STRE and contributes to the induction of these promoters through binding by Msn2-Msn4. Indeed, Fig. 4 shows that, in strains lacking Msn2 and Msn4, the heat shock induction of the group II targets is in general reduced, whereas the group I target genes are relatively unaffected ( $P = 2.65 \times 10^{-5}$ ).

**Validation of HSF binding data.** We performed additional promoter-specific PCR to independently validate our microarray results for the binding of HSF (or its lack thereof) to different classes of promoters. The results of these validation experiments are shown in Fig. 5. We could thus verify the binding of HSF in a heat shock-dependent manner to promoters that lacked a consensus HSE (*ICY2*, *CTH1*, and *EDC2* in

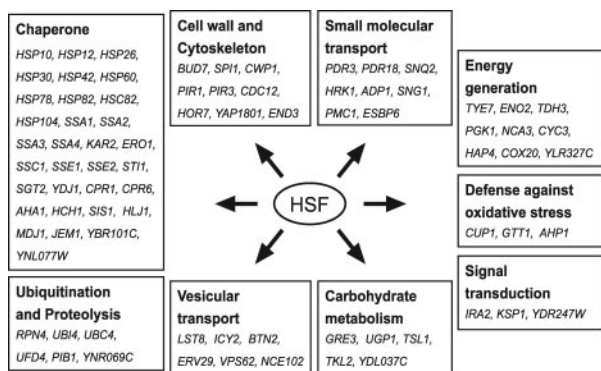


FIG. 6. The constellation of HSF target genes in the yeast genome. Representative HSF targets are categorized according to their known or hypothetical functions based upon sequence homology. See supplemental data at <http://www.iyerlab.org/hsf> for the entire set of HSF targets identified in this work.

Fig. 3A and all the genes in Fig. 5B). We also verified the binding of HSF to several target promoters that contained a consensus HSE (the remaining genes in Fig. 3A and all the genes in Fig. 5A), as well the lack of binding of HSF to promoters that were not identified as targets by microarray analysis, even when those promoters contained a consensus HSE (Fig. 5C). Although an HSE is known to be important for HSF binding, its presence is clearly not the sole determinant of whether a promoter is bound and activated by HSF. We also verified the binding of HSF to the promoters of the small number of genes that were not induced by heat shock. *ILV2* and *ATC1* lack a consensus HSE, while *ICS2* and *BSC5* contain a consensus HSE; all four promoters were identified as targets by microarrays and validated by promoter-specific PCR (Fig. 5A and B), and none of them were appreciably induced by heat shock (see supplemental data for expression levels).

**The constellation of yeast HSF target genes.** HSF proteins and their cognate HSEs are fundamentally conserved in their structure and function from baker's yeast to humans. Indeed, previous studies have demonstrated that both mammalian HSF1 and HSF2 can rescue the viability defect of *S. cerevisiae* cells lacking endogenous yeast HSF (28). Given the tools available in *S. cerevisiae*, this is an ideal system to begin to decipher the biological role of HSF through the identification of virtually all of the direct HSF target genes. Our results revealed approximately 165 genes directly bound by HSF, approaching nearly 3% of the total genes in yeast. The products of many of these HSF target genes are known or predicted to function in a wide range of cellular processes. These include protein chaperone functions, ubiquitination and proteolysis, vesicular transport, maintenance of the cell wall and cytoskeleton, small molecule transport, carbohydrate metabolism, energy generation, signal transduction, and oxidative stress defense mechanisms (Fig. 6) (see Discussion and supplemental data).

While the broad functional groups of HSF target genes identified in this analysis might be unanticipated, one can predict important roles for these genes under stressful conditions. For example, cells can maintain cellular integrity under stress conditions by appropriate expression of cell wall proteins. Although it has been known that the majority of cell wall proteins

and enzymes involved in cell wall biosynthesis are induced by Slt2 mitogen-activated protein kinase pathways in response to cell wall perturbations including heat shock (24), HSF might also contribute to the induction of cell wall proteins. It is also known that some glycolytic enzymes such as phosphoglycerate kinase and enolase are induced by heat shock, which might reflect the higher energy demands during stress (17). Newly identified HSF targets include genes encoding enzymes in the glycolysis pathway such as *PGK1*, *ENO2*, and *TDH3* and the gene for transcription factor *Tye7*, which is involved in regulation of glycolytic genes (38), suggesting a role for HSF in the expression of glycolytic genes during heat shock. Dozens of target genes that do not belong to the category of well-characterized Hsps or chaperones have consensus HSEs in their promoters (supplemental data). The requirement for yeast HSF even at low temperatures strongly suggests that HSF may activate the basal expression of specific sets of genes or have other roles that are essential for yeast cell growth and viability.

## DISCUSSION

### The genome-wide binding distribution of *S. cerevisiae* HSF.

One of the cellular responses to cope with stressful conditions is regulating gene expression to protect macromolecules, recycle irreversibly damaged cellular components, and facilitate appropriate adaptive responses until more favorable conditions are encountered. One such category of stress-responsive proteins is the Hsps, which play diverse roles in biology that include protein folding, prevention of protein aggregation, and protein trafficking and degradation and roles in the maturation and activation of signal transduction proteins and transcription factors. While Hsps are well-characterized targets of yeast and mammalian HSFs, the essential nature of yeast HSF and the function of the stress-responsive mammalian and *Drosophila* HSF in normal growth, development, and other processes suggest that HSF proteins activate the expression of many, as yet unidentified genes.

The *in vivo* genome-wide binding distribution of HSF in baker's yeast described here has revealed novel aspects of HSF function. First, we have identified dozens of new direct downstream transcriptional targets of this important and highly conserved transcriptional regulator, suggesting a much broader and central role for HSF in the physiological response to heat shock and other stresses. Second, we found that the binding of HSF to a large subset of its target promoters in yeast is markedly induced by heat shock. This study thus permits a dissection of the precise role and contribution of a single transcriptional regulator to the global transcriptional and metabolic reprogramming that is common to all cells in response to stress. The transcript levels for a small number of HSF target genes that we have identified here (approximately 7%) appear to be either unchanged or lower after heat shock. It is therefore possible that, in addition to its predominant role in activating genes after heat shock, HSF binding to some promoters is transcriptionally neutral or even mediates repression during heat shock. The physiological significance of this is not clear at present, but a repressor function for HSF has been described earlier in the case of mammalian HSF (5, 41). Alternatively, HSF binding at these promoters could affect transcription under other physiological conditions not examined here.

**Heat shock-induced binding of the *S. cerevisiae* HSF.** It is currently thought that, although yeast HSF is similar in its fundamental structure to metazoan HSF and binds to a similar *cis*-acting HSE site, it is distinguished by the fact that it binds to its target promoters equally well under control or heat shock conditions. The transcriptional activation function of yeast HSF is believed to be predominantly at the level of a post-DNA binding event, perhaps phosphorylation or another, as yet uncharacterized modification or interaction. Our results indicate that yeast HSF binds with higher affinity to a significant fraction of its target promoters following heat shock, similarly to its metazoan counterparts, underscoring the evolutionary and functional conservation of this stress-responsive transcription factor. While the heat-dependent increase in binding of mammalian HSF1 to its cognate targets is the result of a stress-inducible monomer to homotrimer transition and subsequent nuclear accumulation, currently there is no evidence that a similar mechanism may be responsible for the heat-inducible binding of HSF in *S. cerevisiae*.

While this work was in progress, an independent large-scale study used a similar ChIP-microarray approach to identify the binding targets of a large number of epitope-tagged transcription factors in yeast, including HSF, but only in cells grown under non-heat shock conditions (26). Of the HSF binding loci identified in that study, 31 of 35 were also selected as targets in our work; however, our studies, using ChIP-microarray analysis to determine binding targets of the native HSF protein in control as well as heat-shocked cells, have identified approximately three times the number of HSF target genes (supplemental data). Therefore, it is likely that the heat shock induction of HSF binding has revealed HSF targets both that are constitutively bound and for whom HSF is strongly heat-inducibly bound. Although only 46% of the targets that we identified have a consensus HSE in their upstream regions, there is a significant enrichment for HSE-containing promoters in the target set. At the promoters lacking a consensus HSE, it is possible that HSF binds to nonconsensus sites, perhaps in conjunction with other factors.

**Regulation of HSF targets by other environmental stresses.** Many of the HSF direct targets belong to the class of general stress response genes that are activated in response to multiple stresses, implying that a subset of HSF targets are coregulated by Msn2-Msn4 as previously observed for *HSP104* and *HSP26* (2, 14). Most of the HSF targets show induction by diamide, transition to the stationary phase, and depletion of nitrogen or carbon sources, as well as by heat shock. This might reflect coregulation of HSF targets with other regulators specific for these stresses and/or activation of HSF by these multiple stresses. Stresses other than heat shock can activate HSF either by a mechanism similar to heat shock or by distinct signal transduction pathways. In the case of diamide, which can induce protein misfolding by disulfide bond formation, this stress could trigger effects similar to those caused by heat shock to activate HSF.

It has been shown that *CUP1* is activated by HSF under glucose starvation conditions in an Snf1 kinase-dependent manner, whereas the heat shock induction of *CUP1* is independent of Snf1 (46). In addition, *HSP26* induction by carbon source starvation has been shown to be dependent on HSF as well as Msn2-Msn4 (2). We have recently shown that Snf1 is

involved in the phosphorylation of HSF under glucose starvation conditions, followed by activation of a subset of HSF targets (16). Therefore, it is possible that different stresses may activate HSF through signal transduction pathways distinct from heat shock activation.

**Roles of HSF targets.** Identification of novel targets of HSF revealed more diverse roles for HSF than previously established. Based on known or predicted functions of homologous proteins, HSF targets are involved in many aspects of cellular function during the stress of normal cell proliferation, in protecting cells from thermal damage and other stresses, and in metabolic and regulatory reprogramming to allow cells to adapt to new environmental conditions. The broad and important functions associated with HSF direct target genes underscore the essential roles for HSF both under normal cell growth conditions and in response to acute stressful conditions. Given the roles of mammalian HSF1 in normal cell growth and development and in the stress response (8, 31, 53, 55), it will be important to ascertain if mammalian HSF1 has a similar broad range of targets that have been identified here for yeast HSF. Furthermore, given the established role of *C. elegans* HSF as an important determinant of life span (11, 19), the genome-wide identification of yeast HSF targets may provide a framework for the identification of genes directly involved in aging. For example, one HSF target identified here, *PNC1*, encodes an enzyme that deaminates nicotinamide, is induced by heat stress and glucose starvation, and has been shown to be necessary and sufficient for yeast life span extension by caloric restriction and low-intensity stress (3). *ZK1290.5*, a *C. elegans* gene known to be induced upon heat shock, has been identified as an aldo-keto reductase (15), as is *GRE3*, an HSF target identified here. It is possible that other yeast HSF targets identified in this study will provide important insights into genes that function in stress-related disease states and in aging.

#### ACKNOWLEDGMENTS

We thank David Gross for providing polyclonal anti-HSF antiserum, Chen Kuang for excellent technical assistance, Jonathan Davies for all aspects of microarray production, Jonghwan Kim for access to unpublished data, and Patrick Killion for the microarray database.

This work was supported by a grant from the National Institutes of Health (GM59911) to D.J.T. Work in the lab of V.R.I. is supported in part by grants from the NIH (AA13518) and from the Texas Higher Education Coordinating Board.

#### REFERENCES

- Ahn, S. G., P. C. Liu, K. Klyachko, R. I. Morimoto, and D. J. Thiele. 2001. The loop domain of heat shock transcription factor 1 dictates DNA-binding specificity and responses to heat stress. *Genes Dev.* **15**:2134–2145.
- Amoros, M., and F. Estruch. 2001. Hsf1p and Msn2/4p cooperate in the expression of *Saccharomyces cerevisiae* genes *HSP26* and *HSP104* in a gene- and stress type-dependent manner. *Mol. Microbiol.* **39**:1523–1532.
- Anderson, R. M., K. J. Bitterman, J. G. Wood, O. Medvedik, and D. A. Sinclair. 2003. Nicotinamide and *PNC1* govern lifespan extension by caloric restriction in *Saccharomyces cerevisiae*. *Nature* **423**:181–185.
- Boy-Marcotte, E., G. Lagniel, M. Perrot, F. Bussereau, A. Boudsocq, M. Jacquet, and J. Labarre. 1999. The heat shock response in yeast: differential regulations and contributions of the Msn2p/Msn4p and Hsf1p regulons. *Mol. Microbiol.* **33**:274–283.
- Cahill, C. M., W. R. Waterman, Y. Xie, P. E. Auron, and S. K. Calderwood. 1996. Transcriptional repression of the prointerleukin 1 $\beta$  gene by heat shock factor 1. *J. Biol. Chem.* **271**:24874–24879.
- Causton, H. C., B. Ren, S. S. Koh, C. T. Harbison, E. Kanin, E. G. Jennings, T. I. Lee, H. L. True, E. S. Lander, and R. A. Young. 2001. Remodeling of yeast genome expression in response to environmental changes. *Mol. Biol. Cell* **12**:323–337.

7. **Chen, T., and C. S. Parker.** 2002. Dynamic association of transcriptional activation domains and regulatory regions in *Saccharomyces cerevisiae* heat shock factor. *Proc. Natl. Acad. Sci. USA* **99**:1200–1205.
8. **Christians, E., A. A. Davis, S. D. Thomas, and I. J. Benjamin.** 2000. Maternal effect of Hsf1 on reproductive success. *Nature* **407**:693–694.
9. **Christians, E. S., L. J. Yan, and I. J. Benjamin.** 2002. Heat shock factor 1 and heat shock proteins: critical partners in protection against acute cell injury. *Crit. Care Med.* **30**:S43–S50.
10. **Erkine, A. M., S. F. Magrogan, E. A. Sekinger, and D. S. Gross.** 1999. Cooperative binding of heat shock factor to the yeast *HSP82* promoter in vivo and in vitro. *Mol. Cell. Biol.* **19**:1627–1639.
11. **Garigan, D., A. L. Hsu, A. G. Fraser, R. S. Kamath, J. Ahringer, and C. Kenyon.** 2002. Genetic analysis of tissue aging in *Caenorhabditis elegans*. A role for heat-shock factor and bacterial proliferation. *Genetics* **161**:1101–1112.
12. **Gasch, A. P., P. T. Spellman, C. M. Kao, O. Carmel-Harel, M. B. Eisen, G. Storz, D. Botstein, and P. O. Brown.** 2000. Genomic expression programs in the response of yeast cells to environmental changes. *Mol. Biol. Cell* **11**:4241–4257.
13. **Giardina, C., and J. T. Lis.** 1995. Dynamic protein-DNA architecture of a yeast heat shock promoter. *Mol. Cell. Biol.* **15**:2737–2744.
14. **Grably, M. R., A. Stanhill, O. Tell, and D. Engelberg.** 2002. HSF and Msn2/4p can exclusively or cooperatively activate the yeast *HSP104* gene. *Mol. Microbiol.* **44**:21–35.
15. **GuhaThakurta, D., L. Palomar, G. D. Stormo, P. Tedesco, T. E. Johnson, D. W. Walker, G. Lithgow, S. Kim, and C. D. Link.** 2002. Identification of a novel cis-regulatory element involved in the heat shock response in *Caenorhabditis elegans* using microarray gene expression and computational methods. *Genome Res.* **12**:701–712.
16. **Hahn, J. S., and D. J. Thiele.** 2004. Activation of the *Saccharomyces cerevisiae* heat shock transcription factor under glucose starvation conditions by Snf1 protein kinase. *J. Biol. Chem.* **279**:5169–5176.
17. **Hohmann, S., and W. H. Mager.** 1997. Yeast stress responses. R. G. Landes Company, Georgetown, Tex.
18. **Holmberg, C. I., S. E. Tran, J. E. Eriksson, and L. Sistonen.** 2002. Multisite phosphorylation provides sophisticated regulation of transcription factors. *Trends Biochem. Sci.* **27**:619–627.
19. **Hsu, A. L., C. T. Murphy, and C. Kenyon.** 2003. Regulation of aging and age-related disease by DAF-16 and heat-shock factor. *Science* **300**:1142–1145.
20. **Iyer, V. R.** 2003. Microarray-based detection of DNA protein interactions: chromatin immunoprecipitation on microarrays, p. 453–463. In D. Bowtell and J. Sambrook (ed.), *DNA microarrays: a molecular cloning manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
21. **Iyer, V. R., C. E. Horak, C. S. Scafe, D. Botstein, M. Snyder, and P. O. Brown.** 2001. Genomic binding sites of the yeast cell-cycle transcription factors SBF and MBF. *Nature* **409**:533–538.
22. **Jakobsen, B. K., and H. R. Pelham.** 1988. Constitutive binding of yeast heat shock factor to DNA in vivo. *Mol. Cell. Biol.* **8**:5040–5042.
23. **Jedlicka, P., M. A. Mortin, and C. Wu.** 1997. Multiple functions of *Drosophila* heat shock transcription factor in vivo. *EMBO J.* **16**:2452–2462.
24. **Jung, U. S., and D. E. Levin.** 1999. Genome-wide analysis of gene expression regulated by the yeast cell wall integrity signalling pathway. *Mol. Microbiol.* **34**:1049–1057.
25. **Killion, P. J., G. Sherlock, and V. R. Iyer.** 2003. The Longhorn Array Database (LAD): an open-source, MIAME compliant implementation of the Stanford Microarray Database (SMD). *BMC Bioinformatics* **4**:32.
26. **Lee, T. I., N. J. Rinaldi, F. Robert, D. T. Odom, Z. Bar-Joseph, G. K. Gerber, N. M. Hannett, C. T. Harbison, C. M. Thompson, I. Simon, J. Zeitlinger, E. G. Jennings, H. L. Murray, D. B. Gordon, B. Ren, J. J. Wyrick, J. B. Tagne, T. L. Volkert, E. Fraenkel, D. K. Gifford, and R. A. Young.** 2002. Transcriptional regulatory networks in *Saccharomyces cerevisiae*. *Science* **298**:799–804.
27. **Littlefield, O., and H. C. Nelson.** 1999. A new use for the 'wing' of the 'winged' helix-turn-helix motif in the HSF-DNA cocrystal. *Nat. Struct. Biol.* **6**:464–470.
28. **Liu, X. D., P. C. Liu, N. Santoro, and D. J. Thiele.** 1997. Conservation of a stress response: human heat shock transcription factors functionally substitute for yeast HSF. *EMBO J.* **16**:6466–6477.
29. **Liu, X. S., D. L. Brutlag, and J. S. Liu.** 2002. An algorithm for finding protein DNA binding sites with applications to chromatin-immunoprecipitation microarray experiments. *Nat. Biotechnol.* **20**:835–839.
30. **Martinez-Pastor, M. T., G. Marchler, C. Schuller, A. Marchler-Bauer, H. Ruis, and F. Estruch.** 1996. The *Saccharomyces cerevisiae* zinc finger proteins Msn2p and Msn4p are required for transcriptional induction through the stress response element (STRE). *EMBO J.* **15**:2227–2235.
31. **McMillan, D. R., X. Xiao, L. Shao, K. Graves, and I. J. Benjamin.** 1998. Targeted disruption of heat shock transcription factor 1 abolishes thermo- tolerance and protection against heat-inducible apoptosis. *J. Biol. Chem.* **273**:7523–7528.
32. **Morimoto, R. I.** 1998. Regulation of the heat shock transcriptional response: cross talk between a family of heat shock factors, molecular chaperones, and negative regulators. *Genes Dev.* **12**:3788–3796.
33. **Morimoto, R. I., A. Tissières, and C. Georgopoulos.** 1994. The biology of heat shock proteins and molecular chaperones. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
34. **Nover, L., K. Bharti, P. Doring, S. K. Mishra, A. Ganguli, and K. D. Scharf.** 2001. *Arabidopsis* and the heat stress transcription factor world: how many heat stress transcription factors do we need? *Cell Stress Chaperones* **6**:177–189.
35. **Parsell, D. A., and S. Lindquist.** 1993. The function of heat-shock proteins in stress tolerance: degradation and reactivation of damaged proteins. *Annu. Rev. Genet.* **27**:437–496.
36. **Pirkkala, L., P. Nykanen, and L. Sistonen.** 2001. Roles of the heat shock transcription factors in regulation of the heat shock response and beyond. *FASEB J.* **15**:1118–1131.
37. **Santorino, N., N. Johansson, and D. J. Thiele.** 1998. Heat shock element architecture is an important determinant in the temperature and transactivation domain requirements for heat shock transcription factor. *Mol. Cell. Biol.* **18**:6340–6352.
38. **Sato, T., M. C. Lopez, S. Sugioka, Y. Jigami, H. V. Baker, and H. Uemura.** 1999. The E-box DNA binding protein Sgc1p suppresses the *gcr2* mutation, which is involved in transcriptional activation of glycolytic genes in *Saccharomyces cerevisiae*. *FEBS Lett.* **463**:307–311.
39. **Schmitt, A. P., and K. McEntee.** 1996. Msn2p, a zinc finger DNA-binding protein, is the transcriptional activator of the multistress response in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* **93**:5777–5782.
40. **Sekinger, E. A., and D. S. Gross.** 2001. Silenced chromatin is permissive to activator binding and PIC recruitment. *Cell* **105**:403–414.
41. **Singh, I. S., R. M. Viscardi, I. Kalvakolanu, S. Calderwood, and J. D. Hasday.** 2000. Inhibition of tumor necrosis factor- $\alpha$  transcription in macrophages exposed to febrile range temperature. A possible role for heat shock factor-1 as a negative transcriptional regulator. *J. Biol. Chem.* **275**:9841–9848.
42. **Sorger, P. K.** 1990. Yeast heat shock factor contains separable transient and sustained response transcriptional activators. *Cell* **62**:793–805.
43. **Sorger, P. K., M. J. Lewis, and H. R. Pelham.** 1987. Heat shock factor is regulated differently in yeast and HeLa cells. *Nature* **329**:81–84.
44. **Sorger, P. K., and H. R. Pelham.** 1988. Yeast heat shock factor is an essential DNA-binding protein that exhibits temperature-dependent phosphorylation. *Cell* **54**:855–864.
45. **Tachibana, T., S. Astumi, R. Shioda, M. Ueno, M. Uritani, and T. Ushimaru.** 2002. A novel non-conventional heat shock element regulates expression of *MDJ1* encoding a DnaJ homolog in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **277**:22140–22146.
46. **Tamai, K. T., X. Liu, P. Silar, T. Sosinowski, and D. J. Thiele.** 1994. Heat shock transcription factor activates yeast metallothionein gene expression in response to heat and glucose starvation via distinct signalling pathways. *Mol. Cell. Biol.* **14**:8155–8165.
47. **Treger, J. M., A. P. Schmitt, J. R. Simon, and K. McEntee.** 1998. Transcriptional factor mutations reveal regulatory complexities of heat shock and newly identified stress genes in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **273**:26875–26879.
48. **Venturi, C. B., A. M. Erkine, and D. S. Gross.** 2000. Cell cycle-dependent binding of yeast heat shock factor to nucleosomes. *Mol. Cell. Biol.* **20**:6435–6448.
49. **Verbeke, P., J. Fonager, B. F. Clark, and S. I. Rattan.** 2001. Heat shock response and ageing: mechanisms and applications. *Cell. Biol. Int.* **25**:845–857.
50. **Wiederrecht, G., D. Seto, and C. S. Parker.** 1988. Isolation of the gene encoding the *S. cerevisiae* heat shock transcription factor. *Cell* **54**:841–853.
51. **Wu, C.** 1995. Heat shock transcription factors: structure and regulation. *Annu. Rev. Cell Dev. Biol.* **11**:441–469.
52. **Xiao, H., O. Perisic, and J. T. Lis.** 1991. Cooperative binding of *Drosophila* heat shock factor to arrays of a conserved 5 bp unit. *Cell* **64**:585–593.
53. **Xiao, X., X. Zuo, A. A. Davis, D. R. McMillan, B. B. Curry, J. A. Richardson, and I. J. Benjamin.** 1999. HSF1 is required for extra-embryonic development, postnatal growth and protection during inflammatory responses in mice. *EMBO J.* **18**:5943–5952.
54. **Xie, Y., C. Chen, M. A. Stevenson, P. E. Auron, and S. K. Calderwood.** 2002. Heat shock factor 1 represses transcription of the IL-1 $\beta$  gene through physical interaction with the nuclear factor of interleukin 6. *J. Biol. Chem.* **277**:11802–11810.
55. **Yan, L. J., E. S. Christians, L. Liu, X. Xiao, R. S. Sohal, and I. J. Benjamin.** 2002. Mouse heat shock transcription factor 1 deficiency alters cardiac redox homeostasis and increases mitochondrial oxidative damage. *EMBO J.* **21**:5164–5172.