**Structure and Regulation of the SSA4 HSP70 Gene of Saccharomyces cerevisiae**

William R. Boorstein§ and Elizabeth A. Craig†§*

(Received for publication, June 4, 1990)

From the ‡Molecular and Cellular Biology Program and the †Department of Physiological Chemistry, University of Wisconsin, Madison, Wisconsin 53706

SSA4 is the only one of five heat-inducible HSP70 genes in Saccharomyces cerevisiae whose expression is restricted to conditions of stress. Comparison of the nucleotide sequences of the SSA4 gene with other HSP70 genes indicates that it diverged from its most closely related yeast homologues hundreds of millions of years ago. However, a high degree of identity has been maintained between SSA4 and other yeast 70-kDa heat-shock proteins at the amino acid level suggesting, in light of its distinct pattern of regulation, that it performs an important function. A 44-base pair region of the SSA4 promoter containing an extended match to the conserved eukaryotic heat-shock element (HSE) is necessary and sufficient to mediate heat-inducible regulation. HSESSA4 is capable of promoting only a low level of transcription under nonstress conditions. We present evidence in support of a revised definition of the functional HSE in S. cerevisiae, similar to the recently proposed modular Drosophila HSE.

Elevated expression of several heat-shock proteins in anssa1issa2 double-mutant strain has previously been reported. The SSA4 promoter is activated in this strain. The increase in expression of SSA4 caused by deletion of these closely related genes is mediated via the same upstream activating sequences that activate transcription in response to heat shock. Activation of HSE-mediated transcription by disruption of constitutively expressed HSP70 genes supports an autoregulatory model of control of the heat-shock response.

---

* This work was funded by United States Public Health Service grants from the National Institutes of Health (to E. A. C.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† Supported by a United States Public Health Service training grant in molecular and cellular biology. Present address: Howard Hughes Medical Institute, Division of Biology 156-29, California Institute of Technology, Pasadena, CA 91125.

§ The abbreviations used are: HSP, heat-shock protein; hsp70, 70-kDa heat-shock protein; HSE, heat-shock element; HSF, heat-shock transcription factor; UAS, upstream activation sequence; ORF, open reading frame; bp, base pair; kb, kilobase pair; ATF, activating transcription factor.

Craig, 1985). Stress induction of HSPEs is regulated primarily at the level of transcription. The mechanism by which heat-shock genes are activated has also been highly conserved (reviewed in Bienz and Pelham, 1987). Identification of related heat-shock elements (HSEs) in many eukaryotic species, including yeast, plants, and mammals, has led to the derivation of the HSE consensus sequence, CNNGANNTTCCNG, where N is any nucleotide (Pelham, 1985). Recently, the definition of HSEs in Drosophila has been revised, based on the effects of site-directed mutations (Amin et al., 1988; Xiao and Lis, 1989). Functionally defined Drosophila HSEs can be viewed as modular, consisting of at least three repeats of the sequence GAA in alternating orientation and separated from each other by two nucleotides. HSEs are specific binding sites for the heat-shock transcription factor, HSF (Topol et al., 1985; Perisic et al., 1989; Wiederrecht et al., 1987; Wu et al., 1987). Transcriptional activation of heat-shock genes occurs by a posttranslational mechanism believed to involve phosphorylation of HSF (Sorger et al., 1987; Sorger and Pelham, 1988; Zimarino and Wu, 1987). Studies of Drosophila suggest that HSF/HSE-mediated control acts at an early step in transcriptional elongation (Rougvie and Lis, 1988).

The 70-kDa heat-shock proteins (hsp70s) are among the most highly conserved of all proteins. Species as distantly related as archeobacter and human contain related genes (Craig, 1985). In S. cerevisiae, a large multigene family encodes eight hsp70s that are both expressed and localized differentially within the cell. The hsp70 genes have been divided into four subfamilies based on conservation at the amino acid level (reviewed in Lindquist and Craig, 1988). Genetic analyses suggest that members of the subfamilies are functionally related as well (reviewed in Craig, 1989). SSA4, the subject of this study, is one of four members of the most complex group, the SSA subfamily. While none of the individual SSA genes are required for viability, this subfamily is essential (Werner-Washburne et al., 1987). SSA1 and SSA2 are the only two members of the SSA subfamily that are expressed at detectable levels in cells under nonstressed growth conditions. ssa1issa2 double mutants are temperature-sensitive for growth but viable at 30 °C. SSA1SSA2SSA4 cells are inviable, indicating that the SSA4 gene can partially compensate for the ssa1issa2 defects (Werner-Washburne et al., 1987).

SSA4 is the only classic HSP70 gene in yeast; it is expressed at extremely low levels under normal conditions and is rapidly and dramatically induced following heat-shock treatment. The previously analyzed SSA1 and SSA3 genes are also heat-inducible; however, they differ from SSA4 in that SSA1 is expressed at high basal levels and SSA3 is induced under conditions of lowered intracellular cAMP (such as starvation) (Werner-Washburne et al., 1989). Heat-shock activation of
the SSA1 and SSA3 promoters is dependent upon upstream activating sequences (UAS) similar to the HSEs of higher eukaryote (Boorstein and Craig, 1990b; Slater and Craig, 1987). HSEs are also involved in mediating basal expression of SSA1 and enhancing the level of expression of SSA3 under conditions of lowered CAMP. The role of HSEs in promoting transcription under nonstress conditions is consistent with evidence demonstrating that yeast HSF is bound to HSEs under normal, as well as stressed conditions (Jakobsen and Pelham, 1988; Sorger and Pelham, 1987).

Here we present the initial characterization of the SSA4 gene and its transcriptional regulation. We define the promoter element that mediates heat-inducible regulation of SSA4 and compare its activity with that of HSEs from SSA1 and SSA3, which contribute both to heat induction and to high expression levels under non-stressed conditions. We have also analyzed SSA4 promoter activity in mutant strains lacking constitutively expressed, closely related homologues, SSA1 and SSA2. We evaluate the modular ISE concept, defined by experiments in Drosophila, in light of the data presented here and recent analyses of heat-inducible S. cerevisiae promoters.

MATERIALS AND METHODS

S. cerevisiae Strains—DS10 (MAT a, his3-11, his3-15, leu2-3, leu2-112, lys1, lys2, Δtry1, ura3-52) was utilized in all experiments presented here unless specifically noted. DS16 is isogenic to DS10 and contains disrupted alleles of SSA1 and SSA2 (ssal::HIS3 and ssaz::LEU21) (Craig and Jacobsen, 1984; Werner-Washburne et al., 1987).

Plasmid Constructions—An SSA4/lacZ translational fusion (Fig. 1A) was constructed in a derivative of the promoter cloning vector phT102 (Tü and Casadaban, 1990). The EcoRI to SacI fragment of phT102 was replaced with the EcoRI to Smol fragment of the E. coli nick uroc operon from pAP55 (Brusilow et al., 1983) ligated to the Smal to SacI fragment of pSK105 (Casadaban et al., 1983) to generate the translational lacZ fusion vector pWB202. pWB201 includes an extended polylinker (5'-AGCATTCCCGGGATCCGT-3') immediately upstream of lacZ, as well as uroc operon sequences, to facilitate construction of promoter deletions by using an upstream buffer of non-essential vector sequences and an XhoI site to provide constant vector sequences upstream of the SSA4 deletion end points. An SpI linker (5'-pGGAATTCGAC-3') was cloned into the SacI site of the polylinker of pWB201. The 856-bp SSA4H NlIII fragment (Fig. 2B) was subsequently inserted into this SpeI site. The central four base pairs in this polylinker site was one-base-pair in length to allow for subsequent insertion of a DNA restriction fragment of pLG669-Z (Guarente and Ptashne, 1981). SSA4 sequences in pWB220 extend from 200 bp upstream to 200 bp downstream (Fig. 2B). The SSA4/lacZ fusion junction was verified by DNA sequence analysis. The remaining promoter end points were mapped to within ±3 nucleotides by polycarylamide gel electrophoresis of small restriction fragments.

A centromeric UAS cloning vector, pWB220 (Fig. 1B), was constructed by replacing the SSA4 sequences of pWB213 with CYCl promoter sequences. Specifically, the small XhoI to SacI region of pWB213 (containing the entire SSA4 region and the amino-terminus portion of the lacZ-coding region) was replaced with the corresponding restriction fragment of pLG669-Z (Guarente and Ptashne, 1981). SSA4 sequences in pWB220 extend from 175 bp upstream to 175 bp downstream of the primary initiation site, through the amino-terminus coding region. pWB220 also contains different linker and lacZ fusion sites from those in pWB213. pWB227 and pWB227-R were constructed by inserting a synthetic double-stranded oligonucleotide matching the 44 bp indicated in Fig. 2A, flanked by sequences necessary to generate an XhoI site and an XhoI complementary overhang, into the XhoI site of pWB220. The pWB227 and pWB227-R inserts were verified to be identical with the native SSA4 sequence. All plasmid constructions utilized standard methods (Ausbel et al., 1987).

DNA Sequence Determination and Analysis—Appropriate regions of the SSA4H plasmid were subcloned into MI3mp vectors and sequenced by the dyeode chain termination method (as described in Ausbel et al., 1987). Sequence was obtained across all restriction cies used for cloning and on both DNA strands. DNA and protein sequence analyses were performed with University of Wisconsin Genetics Computer Group programs (Devereux et al., 1984).

mRNA Analysis—RNA was isolated by rigorous vortexing of cells with a solution of Tris-HCl, 1 M NaCl, 0.5 M dithiothreitol, in the presence of glass beads, 0.5% sodium dodecyl sulfate, phenol, and chloroform. Lysis was followed by multiple phenol/chloroform extractions and ethanol precipitations as described (Ellwood and Craig, 1984). Yields were determined by A260nm and normalized by comparison of ethidium bromide-stained rRNA in polyacrylamide gel electrophoresis.

A Sau3A to AulA (+147 to -402) fragment was used as a probe to generate the translational lacZ fusion vector pWB201. pWB201 in-

FIG. 1. Plasmids utilized to identify sequences controlling SSA4 expression. A, pWB213, the plasmid containing the SSA4/lacZ hybrid gene (used to construct the 5' deletion series plasmids). B, the centromeric UAS cloning vector pWB220, a derivative of pWB213. pWB220 and pWB213 differ only in the region between lacZ and ARS/1 sequences. Genes and other functional plasmid sequences are labeled in bold type; the location and size, and orientation of genes are indicated by curved arrows. The two naphthared arcs in each plasmid represent sequences of pBR322 origin. Lightly stippled arcs represent sequences from the E. coli uroc operon. SSA4 sequences are represented by solid black arcs. Restriction enzyme recognition sites used in plasmid constructions are labeled. The entire nucleotide sequence of these vectors is known with the exception of approximately 400 nucleotides at each end of the CEN11 sequence. RI, EcoRI; Sa, SacI; Sp, SpHl; X, XhoI.
5 days of growth at 23 °C. Cells were collected in siliconized 1.5-ml microcentrifuge tubes; cell pellets were rapidly frozen in a -70 °C bath and were stored at this temperature. Optimization of growth conditions by using rich medium resulted in lower basal and higher heat-shocked levels of expression than observed in defined, plasmid selective, medium (data not shown). Plasmid stability was routinely assessed as follows: an aliquot of cells was plated on rich medium and the fraction of resulting colonies that contained the T7 promoter was determined. The plasmid was carried in bacteria that were grown in E. coli DH5α (E. coli strain 5287) with 10 μg/ml ampicillin and 0.1% arabinose. The T7 promoter is functional in E. coli DH5α and the plasmid is stably maintained in E. coli DH5α. More than 80% of the cells routinely retained plasmid at the time of sample collection.

β-Galactosidase activity was determined by microfluorometric assay utilizing the substrate 4-methylumbelliferyl-β-D-galactoside (to be described in detail elsewhere). Briefly, cells were suspended and permeabilized in 1 ml of Z buffer (Miller, 1972) with 0.05% sodium dodecyl sulfate, 50 mM β-mercaptoethanol, 1 mM phenylmethylsulfonyl fluoride, vortexed vigorously, diluted to yield fluorogenic product in the linear detection range of the fluorimeter, and incubated with 1 mM 4-methylumbelliferyl-β-D-galactoside for 15-120 min at room temperature. The reactions were performed in a final volume of 200 μl in 96-well microtiter plates. Fluorescence of samples with extremely low β-galactosidase levels was enhanced by the addition of 50 μl of 1 M Na2CO3 after 2 h of incubation. Fluorescence was determined using a Titertek Fluoroskan I fluorimeter. β-Galactosidase activities are presented as (μM/(μM x s)) x 10^3, where μM is the increase in fluorescence between times 1 and 2, t is the time in minutes, between the two fluorescence readings, μM is a measure of cell density at the time of cell collection, s is the volume of the culture assayed, in milliliters, and is a correction factor for plasmid stability that is equal to the fraction of cells retaining the plasmid at the time of cell collection. The arbitrary fluorescence units were converted to Miller units by the empirically derived constant, C = 5.7 x 10^3. N is a normalization factor determined by assay of a known standard quantity of β-galactosidase to control for differences in temperature and pH which affect β-galactosidase activity and fluorescence of methylumbelliferone, respectively. Selected samples were permeabilized as described above and were assayed in parallel with the substrate o-nitrophenyl-β-D-galactopyranoside, as described (Slater and Craig, 1987); results of the two assays were similar, with the exception of the low activity samples that were below the accurate detection limits of the O-nitrophenyl-β-D-galactopyranoside assay. Data presented are averages from two to six determinations of at least two independent transformants. Individual determinations were almost always within 20% of the averages shown here. Occasionally, high or low values were obtained. In these cases, samples from all treatments of a single culture were similarly affected (i.e. basal, heat-shock, and stationary phase); thus, patterns of induction were highly reproducible.

RESULTS

Characterization of the SSA4 Gene and Encoded Transcripts—The SSA4 gene was isolated on a 9-kb genomic clone.3 We constructed a restriction map of the pSSA4H genomic insert and localized the SSA4 coding region by DNA hybridization analysis utilizing fragments of the related SSA1 gene as probes (Fig. 2A and data not shown). The nucleotide sequence of the coding region, as well as 1.2 kb of flanking DNA upstream of the gene, was determined (Fig. 2B). A single large open reading frame with similarity to SSA1 was identified. This SSA4 ORF has the potential of encoding a protein of 642 amino acids with a predicted molecular mass of 69,650 Da and an isoelectric point of 4.86. These values are consistent with previous estimates based on polyacrylamide gel electrophoresis analysis of the SSA4 protein, Ssa4p (Werner-Washburne et al., 1987).

While Ssa4p is closely related to hsp70s and noninducible cognates of distantly related species (68-75% identity to Drosophila and human homologues), it is most closely related to the yeast hsp70 Ssa1p and Ssa2p, which are 97% identical. Ssa4p and Ssa1p are each 642 amino acids in length, although optimal alignment involves the introduction of at least three

3 J. C. A. Bardwell and E. A. Craig, unpublished data.

FIG. 2. Restriction map and nucleotide sequence of SSA4 and flanking genomic DNA. A, restriction map of the 9.0-kb SSA4 HindIII fragment, SSA4-H. The SSA4 protein coding region and the upstream open reading frame are represented by arrows. The filled region of the map indicates the part of the clone whose DNA sequence

Downloaded from http://www.jbc.org/ by guest on November 13, 2017
Fig. 3. Comparison of predicted complete SSA4 and SSA1 hsp70 protein sequences. Solid bars, colons, or dots between corresponding amino acids of the two proteins indicate identity, or high or low levels of similarity, respectively. This alignment was obtained by optimizing matches while introducing the smallest number of gaps (represented by dots in the sequence lines) according to the algorithm of Needleman and Wunsch (1970). Matches between nonidentical amino acids were weighted as described by Gribskov and Burgess (1986). Alternative alignments can be obtained with a larger number of gaps, particularly in the C-terminal region, but these do not significantly increase the similarity. SSA1p is from Slater and Craig (1989).

Relative to the region of the hsp70 genes in yeast, there is an abrupt decrease in similarity near the carboxyl termini of the proteins. The amino-terminal 525 amino acids of Ssa4p are 89% identical with SSA1p, while the carboxy-terminal 117 amino acids share only 50% identity. However, two short highly conserved domains are present within the C-terminal region, the eighth and the second domain amino acids of Ssa4p are identical to those of Drosophila hsp70A, as well as to those of SSA1p. 85% of the amino acids between these two conserved domains (610 through 636) are glycine, alanine, and proline, including six glycine-alanine repeats; therefore, this region is likely to be highly flexible.

The nucleotide sequence similarity is strictly limited to the coding region (Fig. 4), in which the genes share 67% identity. The high degree of divergence between SSA4 and SSA1 based on silent (synonymous) substitutions (182%, calculated and corrected for multiple substitution events by the method of Perler et al. (1980)) indicates that these genes are likely to have arisen from a duplication event a minimum of 100 million years ago. The relatively low divergence, 13%, based on replacement (nonsynonymous) substitutions indicates that a high level of selective pressure has maintained the similarity in the coding potential of these two genes.

An additional open reading frame of at least 494 nucleotides was identified on the same strand as SSA4 (Fig. 2, and data not shown). This ORF begins 5' to the region included in Fig. 2B and terminates 889 bp upstream of the SSA4 coding domain. The upstream ORF is predicted to encode a protein by the TESTCODE algorithm, an empirically derived method based primarily on the periodic compositional bias of coding sequences (Fickett, 1982). The ORF sequence is not closely correct for multiple substitution events by the method of Needleman and Wunsch (1970). Matches between nonidentical amino acids were weighted as described by Gribskov and Burgess (1986). Alternative alignments can be obtained with a larger number of gaps, particularly in the C-terminal region, but these do not significantly increase the similarity. SSA1p is from Slater and Craig (1989).
related to entries in GenBank, EMBL, and NBRF data bases.

Mapping the 5' Termini of SSA4 mRNA and Identification of Putative SSA4 Transcriptional Regulatory Sequences—The 5' ends of SSA4 transcripts were mapped to the gene by primer extension and S1 nuclease techniques (Fig. 5). Heat-induced transcripts from the native gene mapped to a five-nucleotide region. The HSP70-related open reading frame begins at the first AUG downstream of this putative initiation site. The region flanking this AUG matches the S. cerevisiae translational initiation consensus sequence in 8 of 10 positions (Cigan and Donahue, 1987). The length, 53 nucleotides, and the adenine-rich content of the predicted 5' untranslated SSA4 mRNA leader are typical of yeast genes transcribed by RNA polymerase II.

The sequence TATAAAA, which matches the functionally defined TATA element (Chen and Struhl, 1988), is present 75 bp upstream of the putative transcriptional initiation site(s). Interestingly, several occurrences of sequences that have been shown to act as preferred initiation sites in CYCI, PHO5, and other yeast promoters, TC(G/A)A and RRYYR (where R and Y are purine and pyrimidine nucleotides, respectively) (Hahn et al., 1985; Rudolph and Hin nen, 1987), occur near the actual initiation sites; however, these do not appear to be active sites of transcriptional initiation.

Six matches to the canonical dyad heat-shock element CNNGAANNNCTCNG, where N is any nucleotide (Pelham, 1985), occur upstream of the transcribed region of SSA4. Three of these, including the two closest matches (seven of eight) to the consensus, are present in an overlapping arrangement centered at position −188. The −188 region contains four perfect plus three imperfect GAA blocks with relative spacing and orientation consistent with the extended modular HSE definition from Drosophila (Amin et al., 1988). In addition, an exact match to the ATF consensus binding site, (G/T)(A/T)CGTCA, occurs at position −100, 17 bp upstream of the TATA sequence (reviewed in Jones et al., 1988). Very similar sequences have been shown to bind yeast ATF and exhibit UAS function (Kornuc et al., 1988; Jones and Jones, 1989; Lin and Green, 1989). It is intriguing that an HSP70 gene from both yeast and human genomes contain putative ATF-binding sites, although the function of these sequences in heat-shock regulation is not known (Greene et al., 1987).

Expression of an SSA4/lacZ Fusion Gene—To facilitate identification of SSA4 transcriptional regulatory sequences, SSA4 DNA from −801 to the initiation codon was fused to the lacZ gene of E. coli. The fusion gene included three HSE-like regions, the putative TATA element, the transcriptional initiation region, and the entire untranslated leader from SSA4. β-Galactosidase activity from cells transformed with this construct was very low under optimal growth conditions at 23 °C. Activity increased 70-fold following 39 °C heat-shock treatment (Fig. 6, top line). Basal activity was only slightly (0.8 Miller units) higher in stationary phase than in exponentially growing cells at 23 °C.

Localization of Regulatory Sequences—A progressive series of deletions of SSA4 upstream sequence was constructed to delineate regions involved in transcriptional control. The deleted fusion constructs retaining sequences from at least −234 to +59 were regulated in a heat-inducible manner (Fig. 6). The largest and smallest deletion derivatives of the SSA4/lacZ hybrid gene that gave rise to heat-inducible β-galactosidase activity were analyzed in greater detail. Both constructs exhibited a high degree of heat-shock regulation at the RNA level (Fig. 5B). Furthermore, the 5' termini of the fusion mRNAs mapped to the same nucleotides in the SSA4 sequence as did the native SSA4 transcript termini (Fig. 5B). β-Galactosidase activity from both constructs increased dramatically within 30 min of temperature upshift, reached a maximum after 60 min, and then gradually declined (Fig. 7). Therefore, expression of the fusion genes corresponds to that of the native SSA4 gene.

Removal of an additional 60 bp downstream of the −234 deletion endpoint essentially abolished heat-inducible regulation, indicating that the −188 HSE-like region is critical to SSA4 heat induction. This region contains the only match to the newly extended HSE definition within 1.2 kb upstream of the SSA4 gene, as well as the closest matches to the canonical HSE 14-bp palindromic sequence. In addition, β-galactosidase activity under basal (23 °C) conditions is 3-fold lower from constructs lacking the −188 HSE region than from slightly larger constructs that retain these HSE-like sequences, suggesting that the HSE sequences may be capable of mediating a low level of expression under optimal conditions, as well as increased expression under conditions of stress. The deletion analysis also demonstrates that the HSE match at position −549 is not essential for heat-shock regulation and that both the ATF- and HSE-like sequences at −100 and −14, respectively, are not sufficient to drive heat-
We have localized the heat-inducible UAS of SSA4 to a 
44-bp region containing extensive similarity to sequences that
inducible expression, at least from their native locations. An
increase in both basal and induced levels of expression upon
removal of sequences between -366 and -240 suggests ele-
ments that negatively modulate SSA4 expression may be
present in this region. However, no match to the upstream
repression sequence that negatively modulates SSA1 tran-
scription was identified in the SSA4 promoter (Park and
Craig, 1989).

**Definition of UASHs**

To determine whether the sequence shown to be required for heat-shock activity of the SSA4 promoter by 5' deletion analysis were sufficient to mediate heat-inducible transcription in the absence of flanking SSA4 sequences, a short region containing the HSEs centered at 
-188 was tested for upstream activation sequence activity. The vector employed for these experiments, pWB220, is a derivative of the SSA4/lacZ deletion plasmids in which the SSA4 sequences were replaced with DNA from the down-
stream portion of the CYC1 promoter that includes TATA
elements and initiation sites, but lacks upstream activation
sequences. Thus, expression of this lacZ fusion is dependent
upon UAS activity from DNA inserted into a cloning site
adjacent to the upstream boundary of CYC1 sequences. An
oligonucleotide identical with 44 bp encompassing the HSE-like
sequences at -188 was synthesized and cloned into the UAS test vector (Fig. 8). The insertion of this HSESSA4
oligonucleotide upstream of the CYC1/lacZ fusion resulted in a 4-
fold increase in basal levels of activity of the CYC1/lacZ
hybrid protein. Following heat shock, the heterologous pro-
moter mediated a further 12-fold increase in activity. The
kinetics of the change in β-galactosidase activity following
temperature upshift mediated by the hybrid SSA4-oligonucle-
otide/CYC1 promoter was similar to that from the “complete”
SSA4 promoter (Fig. 7). Since the 44-bp SSA4 sequence was
able to mediate heat-inducible transcription in either orien-
tation (Fig. 8B), we designate it UASHs.

**UASHs Mediates High Levels of Transcription in ssa1ssa2 Mutants**—SSA1 and SSA2 are the only SSA genes actively
transcribed in cells under normal growth conditions. Viability
of ssa1ssa2 double-mutant strains at 23 °C depends on the
presence of an intact SSA4 gene, even though SSA4 is not
expressed in wild-type cells growing at this temperature. The
SSA4 protein has been shown to be abundant in ssa1ssa2 cells
(Werner-Washburne et al., 1987). To determine if the SSA4 promoter exhibits altered regulation in ssa1ssa2 cells at 23 °C,
expression of the Δ-594 construct, which is regulated like the
native SSA4 gene in wild-type cells, was assayed in a double-
mutant background. β-Galactosidase activity from the SSA4/
lacZ fusion in an ssa1essa2 double mutant strain was 600 fold
greater than levels in wild-type isogenic strains during growth
at 23 °C (Table I). To determine whether UASHs was alone
sufficient to drive this normally high basal activity, expres-
sion driven by the UASHs/CYC1 hybrid promoter was assayed in
cells with disrupted SSA1 and SSA2 genes. The UAS
cloning vector, pWB220, yielded an extremely low basal level
of β-galactosidase activity, 0.3 units, in the double-mutant
strain. Insertion of the 44-bp HSESSA4 oligonucleotide into the
UAS cloning site of this vector resulted in 185 units of β-
galactosidase activity under the same growth conditions
(Table I). This confirms that the constitutive activation in
the mutant strains is at the level of transcription and dem-
onstrates that a common upstream activating sequence, UASHs,
mediates heat-inducible expression in wild-type strains and high basal levels of transcription in cells lacking
SSA1 and SSA2.

Neither the SSA4/lacZ fusion nor the UASHs/CYC1 hybrid
promoter were activated to higher levels by a 39 °C heat shock
in the ssa1ssa2 background, although both constructs exhibited
induction in wild-type cells following heat shock. Heat-
shock treatment may have failed to activate the SSA4 pro-
moter over the high basal expression levels in ssa1ssa2 mut-
tants because the promoter was already maximally induced.
Alternatively, the prolonged activation of the stress response
may have induced a steady state that was unable to readily
respond to increased stress.

**DISCUSSION**

We have localized the heat-inducible UAS of SSA4 to a 44-
bp region containing extensive similarity to sequences that

![Figure 6](http://www.jbc.org/content/189/17/3947/F6.large.jpg)

**FIG. 6. β-Galactosidase activities from the SSA4/lacZ fusion gene and the promoter deletion deriv-
atives.** The solid black lines represent SSA4 sequences. The positions of the initiation codon, the transcription
initiation site, and the deletion end points are indicated; pWB213A–294 contains 1 bp beyond that indicated
identical with that in the native gene because of the addition of linker sequences. Diamonds represent partial
matches to the HSE consensus, as overlined in Fig. 2B. The putative TATA element is denoted by T. The plasmid
on the top line, pWB213, is the parental plasmid from which the deletions were derived as described under
“Materials and Methods.” Deletion derivative plasmids, from top to bottom, are named pWB213A–594 to
pWB213A–61. pWB213 contains an 8-bp linker inserted at position -594 of the SSA4 DNA as described under
“Materials and Methods.” pWB213 also contains 665 bp of vector sequence immediately upstream of SSA4 that
are not present in the deletion derivatives. All other plasmids represented here differ from each other only in the
extent of 5' noncoding SSA4 sequence included (and consequently the relative spacing of DNA sequences flanking
the deletion end points and the actual junction sequences), β-Galactosidase activities from transformants growing
exponentially at 23 °C and following a 1-h 39 °C heat-shock treatment are presented in Miller units.

<table>
<thead>
<tr>
<th>Temperature</th>
<th>23°C Galactosidase Activity</th>
<th>39°C Galactosidase Activity</th>
<th>39°C/23°C Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2</td>
<td>14</td>
<td>70</td>
<td>0.2</td>
</tr>
<tr>
<td>0.2</td>
<td>12</td>
<td>60</td>
<td>0.2</td>
</tr>
<tr>
<td>0.1</td>
<td>10</td>
<td>100</td>
<td>0.1</td>
</tr>
<tr>
<td>0.1</td>
<td>10</td>
<td>100</td>
<td>0.1</td>
</tr>
<tr>
<td>0.2</td>
<td>11</td>
<td>55</td>
<td>0.2</td>
</tr>
<tr>
<td>0.7</td>
<td>17</td>
<td>24</td>
<td>0.7</td>
</tr>
<tr>
<td>0.6</td>
<td>15</td>
<td>25</td>
<td>0.6</td>
</tr>
<tr>
<td>0.2</td>
<td>0.4</td>
<td>2.0</td>
<td>0.2</td>
</tr>
<tr>
<td>0.2</td>
<td>0.6</td>
<td>3.0</td>
<td>0.2</td>
</tr>
<tr>
<td>0.3</td>
<td>0.3</td>
<td>1.0</td>
<td>0.3</td>
</tr>
</tbody>
</table>
Structure and Regulation of a Yeast Heat-shock Gene

The difference in regulation between the SSA subfamily genes, resulting from different promoter composition, may provide an explanation for the maintenance of similar genes over a long period of evolutionary time; the gene products may be functionally equivalent, multiple genes being required for an appropriate pattern of expression and/or rapidity of induction under certain circumstances. However, it is possible that the amino acid differences between the proteins, perhaps in the divergent C termini, may confer distinct functions, or at least specialized abilities to perform a common function under different environmental conditions.

HSEs in Yeast—The functional Drosophila HSE is comprised of at least three inverted GAA modules separated from each other by 2 bp (Amin et al., 1988; Xiao and Lis, 1988). The three modules need not be contiguous, a gap of precisely one module width (5 bp) is tolerated, provided the modules flanking the gap are direct repeats (i.e. oriented as if the gap contained a GAA sequence). All seven heat-shock elements that have been functionally defined from heat-inducible yeast by the isolated HSEs of SSA3 and SSA1 (Boorstein and Craig, 1990b; Slater and Craig, 1987). Therefore, the role of HSEs in contributing to the high expression of the SSA1 and SSA3 promoters under nonstress conditions (in which the HSE-containing SSA4 promoter is not active) appears to result from differences in regions of the promoters distinct from the HSEs (Slater and Craig, 1987; Boorstein and Craig, 1990a). For example, expression of SSA3 in stationary phase is dependent upon a distinct promoter element, UASPDS, but is further enhanced by HSESSA1 (Boorstein and Craig, 1990a).

The difference in regulation between the SSA subfamily genes, resulting from different promoter composition, may provide an explanation for the maintenance of similar genes over a long period of evolutionary time; the gene products may be functionally equivalent, multiple genes being required for an appropriate pattern of expression and/or rapidity of induction under certain circumstances. However, it is possible that the amino acid differences between the proteins, perhaps in the divergent C termini, may confer distinct functions, or at least specialized abilities to perform a common function under different environmental conditions.

HSEs in Yeast—The functional Drosophila HSE is comprised of at least three inverted GAA modules separated from each other by 2 bp (Amin et al., 1988; Xiao and Lis, 1988). The three modules need not be contiguous, a gap of precisely one module width (5 bp) is tolerated, provided the modules flanking the gap are direct repeats (i.e. oriented as if the gap contained a GAA sequence). All seven heat-shock elements that have been functionally defined from heat-inducible yeast
genes match the extended modular HSE definition derived from analysis of a Drosophila HSP70 promoter. The yeast HSEs are comprised of at least three and as many as five appropriately positioned GAA modules spanning up to four turns of the DNA helix (Table II). The functionally defined HSE of SSA4 meets the criteria of the modular HSE definition; sequences upstream of the transcribed portion of the SSA4 gene that match the canonical HSE but not the modular HSE appear to be nonfunctional. Strong matches to the modular HSE are also found upstream of heat-inducible yeast genes whose promoters have not yet been analyzed (Table II: SSA4 and ST7). Yeast HSE regions also contain as many as four appropriately positioned sequences with a single mismatch to the core consensus GAA (TTC) in addition to the perfect matches; one such imperfect core sequence, TTA, is present with particularly high frequency (for example SSA3; see Table II). These occurrences of imperfect GAA blocks suggest that they function as components of HSF-binding sites. In fact, in vitro binding experiments indicated that HSF contacts at least some imperfect GAA blocks (Shuey and Parker, 1986).

Direct evidence supporting an extended HSE definition in yeast comes from reevaluation of previous analysis of HSE2. An oligonucleotide containing 22 bp of the HSE2 region contains three GAA blocks and is sufficient to mediate heat-inducible transcription (Table II, line 3). However, a shorter (15 bp) region that includes the identical 14 bp match with the canonical HSE, but only two of the three GAA modules, was unable to function as a UAS (Table II, line 11). Interestingly, an even shorter HSE2 oligonu-

### Table II

<table>
<thead>
<tr>
<th>Gene</th>
<th>Function</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSA4</td>
<td></td>
<td>TTTTTAGTTTCGCAGAACTTTTT=TT;CTTTTTTTCTAGAACGCCGTGGAAGAA --- -- - - -- =</td>
</tr>
<tr>
<td>SSA3</td>
<td>+</td>
<td>TTTT&amp;ZAGA&amp;ZGECATCGGC</td>
</tr>
<tr>
<td>SSA1-HSE2 (22-mer)</td>
<td>+, B</td>
<td>TGTAAACTTTCCAGAAATCTTCTAGAAAAG --- -- - - --</td>
</tr>
<tr>
<td>SSA1-HSE3</td>
<td>B</td>
<td>ATCGAAGGTTCTGAAGGCGGGAAAGGGTTTAGTACCA</td>
</tr>
<tr>
<td>HSP92</td>
<td>B, M</td>
<td>ATCGAAGGTTCTGAAGGCGGGAAAGGGTTTAGTACCA</td>
</tr>
<tr>
<td>PGK</td>
<td>M</td>
<td>TGTCTAGAAAGCTTCTGAAATATCTTGGAAGAA</td>
</tr>
<tr>
<td>UBI4</td>
<td>M</td>
<td>TGTCTAGAAAGCTTCTGAAATATCTTGGAAGAA</td>
</tr>
<tr>
<td>8. Consensus</td>
<td>Frequent:</td>
<td>TA cg TA</td>
</tr>
<tr>
<td>Core:</td>
<td>Examples:</td>
<td></td>
</tr>
<tr>
<td>Rare:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. KAR2</td>
<td>ND</td>
<td>ATAGAAACCTTTGCAAGAATTTCCAGG</td>
</tr>
<tr>
<td>10. ST7</td>
<td>ND</td>
<td>ATCTTCGTGAAAGTTCTGAGTATGAGATAGAAGACCTTCCAGGAAAA</td>
</tr>
</tbody>
</table>
| 11. SSA1-HSE2 (15-mer) | - | tttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt...
celotide was able to function efficiently as a heat-inducible UAS, the vector nucleotides flanking the short SSA1 insert created appropriately oriented and positioned matches (2 of 3 bp) to the GAA module on either side of the SSA1 oligonucleotide (Table II, line 12). This result further supports the idea that partial matches to the core consensus, GAA, can contribute to HSE function in yeast.

The base pairs flanking highly conserved core GAA modules of an HSE are also conserved, but to a lesser extent. The bias of nucleotides between GAA blocks is indicated in line 8 of Table II for two model HSEs that each consist of three adjacent GAA repeats. The bias is strongest upstream of 5'-GAA-3' sequences on the plus strand, particularly between two exact consensus modules (TTC--GAA). The sequence bias we observed between GAA blocks differs somewhat from that in HSE-like regions from heat-shock genes of higher eukaryotic species (Amin et al., 1988). Most notably, G was not present in the position two nucleotides upstream of GAA in any of the 16 perfect or imperfect core GAA modules in the region of functional HSEs (Table II; Amin et al., 1988) observed that G occurs frequently in this position of HSEs from Drosophila and other species and demonstrated its ability to contribute to HSE function in Drosophila cells. It is clear from examination of the sequences in Table II and the analysis of HSE-containing constructs that HSEs of different yeast genes share many common features. However, the flexibility of the functional HSE sequence is also striking (Wei et al., 1986). The required degree of specificity for HSF binding appears to be achieved by a long but degenerate recognition site, rather than a short strictly conserved UAS; the vector nucleotides flanking the short UAS elements allows a single transcription factor, HSF, to stimulate transcription of many genes at different relative levels in response to a common signal.

**Autoregulation of the Heat-shock Response**—Transcription of heat-shock genes has been proposed to be negatively regulated by heat-shock proteins. This autoregulatory hypothesis is based on studies in which blocked translation caused abnormally high levels of heat-shock gene transcription following stress exposure (McAlister and Finkelstein, 1980; DiDomenico et al., 1982; Plesset et al., 1982). We have shown that disruption of two constitutively expressed HSP70 genes causes increased transcription of the heat-inducible SSA4 gene. Because the same short UAS5 elements is necessary and sufficient to mediate basal expression in an ssa1ssa2 strain as mediates stress-activated transcription in wild-type strains, we surmise that the general stress response is constitutively activated in the double-mutant strain. Additional mutant phenotypes of ssa1ssa2 strains, including constitutive thermotolerance and active synthesis of Hsp82, support this conclusion (Craig and Jacobsen, 1984).

These results suggest that SSA proteins function in a negative feedback loop, regulating the transcription of heat-shock genes. Thus, hsp70 is critical to maintenance of the heat-shock response in the OFF (uninduced) state, and a decrease in hsp70 availability, either by disruption of the constitutively expressed HSP70 genes or the increased demand for hsp70s under conditions of stress, is sufficient to activate the response. ISF may be modulated by hsp70 directly or indirectly via proposed activities in protein folding and mediation of interactions between polypeptides (reviewed by Lindquist and Craig, 1988). Overexpression of the SSA1 HSP70 gene (from a heterologous promoter) does not appear to decrease either basal or heat-induced HSF/HSE-mediated transcription, although it does repress SSA1 promoter activity via sequences distinct from the HSEs (Stone and Craig, 1990). Therefore, it is likely that there are other components, possibly additional HSPs, in the control loop modulating the HSP activity state. In *E. coli*, mutations in any of three HSP genes, grpE, dnaJ, or the HSP70 homologue, dnaK, result in σ70-dependent increased transcription of heat-shock genes* (Tilly et al., 1983). Since DnaK can interact directly with DnaJ and GrpE (Sell, 1987; Johnson et al., 1989), analogues of the latter two proteins may interact with hsp70 in yeast to control the activation state of HSF. Perhaps the mechanism by which the primary intracellular stress stimulus is transduced, resulting in induction of the stress response, is conserved between prokaryotes and eukaryotes, although the mechanism of activating RNA polymerase on heat-shock promoters has diverged. Elucidation of the mechanisms by which HSP70 expression controls transcription of heat-shock genes will require further definition of the function of the HSP70 gene products and the means by which the stress pathway is activated.

**Acknowledgments**—We wish to thank Jeffrey Shilling and the University of Wisconsin Biotechnology Center for sequencing portions of the SSA4 gene and Diane Gambill, Lynn Manseau, and Carolyn Norris for comments on the manuscript. We also thank Helen Tu for pHT102 and Dave Stone for the yeast strains used in this study.

**REFERENCES**


* D. Straus, W. Walter, and C. Gross, personal communication.
Structure and Regulation of a Yeast Heat-shock Gene


Miller, J. H. (1972) Experiments in Molecular Genetics, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY


Perisic, O., Xiao, H., and Lis, J. T. (1989) Cell 50, 797-800


Structure and regulation of the SSA4 HSP70 gene of Saccharomyces cerevisiae.
W R Boorstein and E A Craig


Access the most updated version of this article at http://www.jbc.org/content/265/31/18912

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC’s e-mail alerts

This article cites 0 references, 0 of which can be accessed free at http://www.jbc.org/content/265/31/18912.full.html#ref-list-1