MOLECULAR AND CELLULAR BIOLOGY, Nov. 1994, p. 7592–7603
0270-7306/94/$04.00+0

Vol. 14, No. 11

Selection of New HSF1 and HSF2 DNA-Binding Sites Reveals Differences in Trimer Cooperativity

PAUL E. KROEGER AND RICHARD I. MORIMOTO*

Department of Biochemistry, Molecular Biology and Cell Biology, Northwestern University, Evanston, Illinois 60208

Received 7 July 1994/Returned for modification 18 August 1994/Accepted 24 August 1994

Multiple heat shock transcription factors (HSFs) have been discovered in several higher eukaryotes, raising questions about their respective functions in the cellular stress response. Previously, we had demonstrated that the two mouse HSFs (mHSF1 and mHSF2) interacted differently with the HSP70 heat shock element (HSE). To further address the issues of cooperativity and the interaction of multiple HSFs with the HSE, we selected new mHSF1 and mHSF2 DNA-binding sites through protein binding and PCR amplification. The selected sequences, isolated from a random population, were composed primarily of alternating inverted arrays of the pentameric consensus 5′-nGAAn-3′, and the nucleotides flanking the core GAA motif were nonrandom. The average number of pentamers selected in each binding site was four to five for mHSF1 and two to three for mHSF2, suggesting differences in the potential for cooperative interactions between adjacent trimers. Our comparison of mHSF1 and mHSF2 binding to selected sequences further substantiated these differences in cooperativity as mHSF1, unlike mHSF2, was able to bind to extended HSE sequences, confirming previous observations on the HSP70 HSE. Certain selected sequences that exhibited preferential binding of mHSF1 or mHSF2 were mutagenized, and these studies demonstrated that the affinity of an HSE for a particular HSF and the extent of HSF interaction could be altered by single base substitutions. The domain of mHSF1 utilized for cooperative interactions was transferable, as chimeric mHSF1/mHSF2 proteins demonstrated that sequences within or adjacent to the mHSF1 DNA-binding domain were responsible. We have demonstrated that HSEs can have a greater affinity for a specific HSF and that in mice, mHSF1 utilizes a higher degree of cooperativity in DNA binding. This suggests two ways in which cells have developed to regulate the activity of closely related transcription factors: developing the ability to fully occupy the target binding site and alteration of the target site to favor interaction with a specific factor.

Heat shock transcription factor (HSF) is known to be the transcriptional activator responsible for the inducible expression of genes such as HSP70 (1, 16, 21, 23, 33, 36, 37, 40). However, multiple distinct HSFs have been isolated from the human, mouse, chicken, and tomato genomes, demonstrating that HSF is a family of factors (24, 27, 30–32). Subsequent experiments have demonstrated that the HSF family members respond to different stimuli (29, 35). HSF1 is a monomeric protein that is latent in the cytosolic and nuclear compartments. In response to environmental stress such as heat or heavy-metal treatment, HSF1 is rapidly activated within minutes—a process that involves oligomerization to a trimeric form, phosphorylation, complete nuclear translocation, and binding to the heat shock element (HSE) of stress-responsive genes such as HSP70 (1, 4, 5, 16, 23, 29). HSF1 has been unambiguously shown to be the major factor induced during heat stress (29). In contrast, HSF2 is relatively unaffected by heat and although many treatments have been examined, only hemin treatment of K562 erythroleukemia cells has been shown to activate HSF2 (35). The activation of HSF2 is much slower, requiring hours, and proceeds through oligomerization and nuclear translocation; however, HSF2 is not phosphorylated. Recent biochemical studies of human HSF2 have strongly suggested that it is a dimer in the control or latent form (22). As yet there is no evidence regarding the composition of this dimer; however, this apparent difference in the control forms of HSF1 and HSF2 suggests distinct regulatory mechanisms.

The HSE, with which HSF interacts, has been previously defined as an array of adjacent inverted pentamers with the consensus sequence 5′-nGAAn-3′ (2, 41). The HSE in the promoter of the human HSP70 gene is composed of five pentameric binding sites arranged as adjacent inverted arrays. Three of these repeats match the current consensus, 5′-nGAAn-3′, and two repeats deviate from the consensus although they have the essential G at the second position. In vivo footprinting analysis revealed that heat-induced HSF1 was bound to all five repeats of the HSP70 HSE, whereas hemin-induced HSF2 failed to contact the first repeat (1, 35). In vitro footprinting with recombinant mouse HSFs (mHSFs) substantiated this result (18). Studies with *Drosophila* HSF (dHSF) demonstrated the basic nature of the HSE and that HSF utilizes cooperative interactions to achieve specificity and high-affinity interactions with its target site (25, 42). Additionally, studies with yeast HSF and dHSF have demonstrated that the cooperative interactions between HSF trimers are necessary for HSF activation of transcription (3, 8). These studies have been very informative; however, since there are multiple HSFs in higher eukaryotes and yeasts and *Drosophila melanogaster* have only a single HSF, we wondered if each HSF would have similar or distinct properties.

To address this issue, we determined the DNA-binding specificities of mHSF1 and mHSF2 by random oligonucleotide selection (26). In contrast to mutagenesis, this is an unbiased method for the analysis of binding sites, as the protein selects the preferred binding sequences from a random pool of possible sites. This protocol has been successfully used to

* Corresponding author. Mailing address: Dept. of Biochemistry, Molecular Biology and Cell Biology, Northwestern University, 2153 Sheridan Rd., Evanston, IL 60208. Phone: (708) 491-3340. Fax: (708) 491-4461.

examine the DNA-binding characteristics of several transcription factors including c-fos, myoD, SRF, and the GATA family of factors (7, 17, 20, 26). In some instances, information to guide the preparation of binding-site probes that were partially random and anchored at known nucleotides was available (7). In other instances, the binding site was unknown and totally randomized probes were utilized (26).

Utilizing a completely random pool of sequences, we have isolated new mHSF1 and mHSF2 binding sites. Our analysis has defined the sequence of the HSE for both mHSFs and was consistent with the definition of the HSE developed in previous studies, 5′-nGAAn-3′. Our studies have further demonstrated that there are differences in the abilities of mHSF1 and mHSF2 to bind to certain arrangements of repeats and that mHSF1 has a greater capacity to bind DNA cooperatively.

MATERIALS AND METHODS

Purification of proteins. mHSF1 and mHSF2 were purified to near homogeneity by utilizing a T7 expression system in BL21-DE3 bacteria essentially as described previously, except that recombinant mHSF1 was chromatographed through an S-Sepharose column after the heparin-Sepharose step (18).

Plasmid DNAs. pBluescript KS− was used for all clonings of the selected oligonucleotides (Stratagene). p89XL-CAT contains 1,800 bp of the human HSP90 promoter fused to chloramphenicol acetyltransferase and was a gift of Eileen Hickey and Lee Weber (14).

Random oligonucleotide selection. We utilized the protocol of Pollack and Treisman and synthesized a 77-bp oligonucleotide, 5′-CAGGTCAGTTCAGCGGATCCTGTCG-$(N)_{27}$-GAGGCGAATTCAGTGCAACTGCAGC-3′ (where N is any nucleotide), and primers complementary to each end (26). Primer F was 5′-GCTGCAGTTGCACTGAATTCGCCTG-3′, and primer R was 5′-CAGGTCAGTTCAGCGGATCCTGTCG-3′. The 77-mer was purified on an 8% denaturing acrylamide gel and used to prepare a probe for gel shift analysis. The 77-mer oligonucleotide was labeled and made double stranded by annealing primer F and extending the bottom strand with Klenow fragment in the presence of [α-^{32}P]dCTP as described previously (26). Approximately 5 ng of labeled probe and 1 μg of poly(dI-dC) · poly(dI-dC) were mixed with 10 nM mHSF1 or mHSF2 and incubated at 25°C for 30 min as described previously. The extended binding reaction permitted the proteins to cycle through several association and dissociation events, leading to the isolation of higher-affinity selected sequences. The binding reaction mixture was then subjected to electrophoresis on a 4% (40:1) acrylamide gel in 0.25× Tris-borate-EDTA buffer for 2 h at 150 V. The gel was dried and exposed to XAR-5 film at −70°C overnight. The mHSF1 shift was composed of two complexes, A and B, with B the more slowly migrating of the two, as noted previously (18). The mHSF2 complex was composed primarily of a single band designated U. For the subsequent rounds of selection, the mHSF1 B and mHSF2 U shifted DNA complexes were excised from the dried gel and incubated in 200 μl of 10 mM Tris-HCl, pH 8.0, for 3 h at 37°C. Ten microliters of the eluted DNA was used in a PCR to make probe for the next round of selection. PCR conditions were 10 mM Tris-HCl, pH 8.8; 50 mM KCl; 6 mM $MgCl_2$; 1 mM dithiothreitol; 0.18 μM primers F and R; 10 μCi of [α-^{32}P]dCTP; 50 μM each dATP, dTTP, and dGTP; and 20 μM dCTP. The final reaction volume was 100 μl, and the parameters were 20 cycles at 94°C for 1 min, 62°C for 1 min, and 72°C for 1 min. In subsequent rounds of selection, 1.5 nM protein was used in the binding reaction mixture. After five rounds of selection by mHSF1 and mHSF2, both pools of amplified oligonucleotides were digested with *Bam*HI and *Eco*RI and cloned into Bluescript KS− (Stratagene). The blue and white colony selection method was used to identify possible recombinants, and the composition of the insert was determined by dideoxy sequencing of denatured double-stranded templates (28). After the third round of selection, we sequenced some clones to check our selection process and found that the random region varied from 23 to 27 nucleotides in length. This was likely due to sequence heterogeneity of the original 77-base oligonucleotide that was not detected during gel purification; however, it did not affect the isolation or interpretation of selected sequences.

Preparation of labeled probes. To determine the relative affinities of mHSF1 and mHSF2 for selected and natural sequences, PCR was used to amplify the selected oligonucleotide probes (15). The T7 and T3 promoter primers complementary to Bluescript KS− were used to prepare single end-labeled probes. This was accomplished by labeling the T7 primer to high-level specific activity with T4 DNA kinase for use in the PCR (15). PCR buffer conditions were as described above. The T7 and T3 primers were added at 0.25 μM, and the parameters were 30 cycles at 94°C for 1 min, 50°C for 1 min, and 72°C for 1 min followed by 5 min at 72°C. The labeled probe was purified of free primer and nucleotide by addition of 0.25 volumes of 8 M ammonium acetate, 5 μg of glycogen, and 1 volume of isopropanol. After 10 min at 25°C, the labeled probe was recovered by centrifugation (12,000 × *g*, 10 min). The integrity and purity were checked by electrophoresis on a 10% (19:1) polyacrylamide gel. The concentration of the labeled DNA was estimated by a direct spectrophotometric analysis of the entire sample (400 μl) at 260 nm.

Active protein determination and equilibrium DNase I footprinting. Footprinting reaction mixtures were established and treated with DNase I as described previously, except that the template concentration for each reaction was 10^{-10} M except when indicated otherwise and the binding reaction was performed in a 100-μl volume (11, 18). The binding reaction was determined to be at equilibrium by ≈1 to 2 min as judged by DNase I protection. To ensure equilibrium, all reaction mixtures were incubated for 20 min prior to DNase I digestion. The specific activity of the mHSF1 and mHSF2 protein preparations was determined by titration of DNase I protection with a known amount of unlabeled DNA. Binding reaction mixtures with 1.6×10^{-10} M labeled 1B5-34 DNA and $\approx 6 \times 10^{-9}$ M HSF trimer were adjusted with various amounts of unlabeled 1B5-34 so that the final DNA concentrations were 1.6×10^{-10} to 2.02×10^{-8} M (see legend to Fig. 3). DNase I footprinting was performed, and the extent of protection was determined by direct quantitation with a PhosphorImager (Molecular Dynamics, Sunnyvale, Calif.). We determined the point at which the level of protection from DNase I cleavage was reduced to 50%. The total concentration of active protein $[P_t]$ at this point was estimated by the equation $[P_t] = K_d + 1/2[D_t]$, where K_d is the apparent equilibrium constant and $[D_t]$ is the total concentration of DNA in the reaction mixture. By utilizing a high-affinity binding site (1B5-34) at a low DNA concentration (0.16 nM) and an excess of HSF (≈6 nM), we were able to calculate the amount of active protein in each HSF preparation with little error and without the necessity of reiteration. The 1B5-34 K_d values (≈0.2 nM) for mHSF1 and mHSF2 would contribute only minimal error to the calculation, and it is likely, as stated below, that the apparent K_d values for HSF binding to this site may actually be lower.

The apparent K_d values for mHSF1 and mHSF2 binding to selected and natural sequences were determined by direct quantitation of the dried gels with the PhosphorImager. The

DNA concentration in each reaction mixture was held constant at 10^{-10} M, and the amount of mHSF1 or mHSF2 was varied from below to above this fixed DNA concentration (see figure legends for specific concentrations). The results were expressed as the percentage bound versus the log of the active HSF concentration, and the data were fitted to a sigmoidal binding equation with the Igor program (Wavemetrics, Inc., Lake Oswego, Wash.). This program calculated the curve that best fit the data by reiterative passes and estimated the apparent K_d from the midpoint of this fitted curve.

Chimeric HSF construction and footprinting. Chimeric HSFs were prepared by digestion of the PETmHSF1 and PETmHSF2 clones with *Sph*I. The factors share an *Sph*I site at amino acids 179 of mHSF1 and 167 of mHSF2. Additionally, there is an *Sph*I site in the PET3a vector at nucleotide position 654, upstream of the T7 promoter. The two resulting fragments from each *Sph*I digestion were gel purified. The large *Sph*I fragment that contains most of the PET3a and mHSF1 C-terminal sequences was ligated to the N-terminal *Sph*I fragment from the PETmHSF2 vector. The resulting construct, HSF2DBD/HSF1, has amino acids 1 to 167 of mHSF2 and amino acids 180 to 503 of mHSF1. The fusion of the fragments at the *Sph*I site preserved the reading frame, and a chimeric protein was produced. An identical protocol was used to create the chimera HSF1DBD/HSF2, which has amino acids 1 to 179 of mHSF1 and 168 to 517 of mHSF2. These chimeric proteins were expressed in BL21-DE3 bacteria as soluble proteins with isopropyl-β-D-thiogalactopyranoside induction, and the DNase I footprinting was done as described above with lysates of induced bacteria.

RESULTS

Analysis of recovered HSF binding sequences. We utilized the purified mHSF1 and mHSF2 proteins to select new binding sequences. We prepared a labeled 77-bp oligonucleotide, randomized in the middle 27 bp; established binding reactions with mHSF1 and mHSF2; and amplified the bound DNA as described in Materials and Methods. This procedure was repeated a total of five times, and then the recovered oligonucleotides were cloned. All of the mHSF1- and mHSF2-selected sequences that were characterized are presented in Fig. 1A and B, respectively. The cloned sequences were analyzed for the presence of potential HSF binding sites through nucleotide alignment. Our analysis was guided by the knowledge of known HSF binding sites such as those found in the mammalian HSP70, HSP90, and small HSP promoters, as well as the current consensus site (2, 41). We examined each selected sequence for appropriately spaced guanine residues and then examined the flanking nucleotides to determine the number of potential pentameric repeats present in each oligonucleotide. On the basis of this alignment, we were able to confirm that 5′-nGAAn-3′ was the consensus sequence for an HSE pentamer selected by mHSF1 or mHSF2. After aligning the various selected sequences, we noted a difference in the number of potential pentamers selected in each binding site by mHSF1 and mHSF2, and this is graphically demonstrated in Fig. 1C. mHSF1 and mHSF2 selected an average of four to five and two to three pentamers per binding site, respectively. Notably, the number of pentamers per binding site selected by mHSF2 was skewed toward two to three. This was of interest, since we had examined some clones after three rounds of selection and noted that the populations selected by both mHSF1 and mHSF2 were predominantly composed of two to three pentameric repeats (data not shown). Thus, with repeated binding and amplification, mHSF1 selected oligonucleotides from the population that had an average of four to five pentamers per binding site. In contrast, the average number of pentamers per binding site selected by mHSF2 did not increase with further selection. One interpretation is that, in contrast to that of mHSF1, the stability of mHSF2 binding was not enhanced by the binding of adjacent trimers on a single oligonucleotide. This point is further substantiated below.

The initial alignment of all mHSF1 and mHSF2 monomeric selected sequences (5′-nGAAn-3′) suggested that there was little preference for nucleotides in the first and fifth positions. However, a subsequent alignment of dimeric selected sequences (5′-nGAAnnTTCn-3′) demonstrated that there was a preference for nucleotides in these positions. The nucleotide frequencies at each position were calculated and used to derive consensuses for mHSF1 and mHSF2 binding: 5′-aGAA(c/t)gTTCg-3′ and 5′-(a/g)GAAnnTTC(g/t)-3′ (capital letters indicate the core nucleotides of each pentamer), respectively (Fig. 2). This analysis demonstrated that there was a preference for particular nucleotides adjacent to the conserved guanine (2 and 2′) residues and that these were different depending on the position of the pentamer within the binding site. In the mHSF1-selected sequences, there was a distinct alternating appearance of adenine and cytosine in the first position of the pentamer (positions 1 and 1′). There was also strong selection against thymine in the first position (1 and 1′) and adenine and guanine at the fifth position (5 and 5′), which is consistent with previous studies that utilized partially purified human HSF1 (10). The consensus for mHSF2 binding also demonstrated some preference in the "n" positions, as adenine or guanine was favored in position 1 and adenine or cytosine was favored at the 1′ position. This demonstrated that mHSF1 and mHSF2 recognized similar, but not identical, consensus sequences.

Equilibrium DNase I footprinting: determination of active protein concentration. To reveal possible differences in mHSF1 and mHSF2 affinity for selected HSEs, we used quantitative DNase I footprinting (19). All probes were prepared by PCR amplification of selected binding sequences with the flanking T7 and T3 primers, as described in Materials and Methods. Initially, we measured the amount of active mHSF1 and mHSF2 protein in our purified preparations with the mHSF1-selected template 1B5-34 (Fig. 1A), a site that both HSFs bound with high affinity. Equilibrium binding reaction mixtures with 6 nM trimer of either mHSF1 or mHSF2 and 0.16 nM 1B5-34 DNA were established in the presence of various amounts of identical unlabeled 1B5-34 and treated with DNase I (Fig. 3A). As the total concentration of 1B5-34 in the binding reaction mixture was increased from 0.16 to 20.2 nM (Fig. 3A, lanes B to K and M to U), both mHSF1 and mHSF2 DNase I footprints were titrated or inhibited by competition by the unlabeled 1B5-34. We quantitated the level of DNase I protection in each lane and determined the amount of unlabeled 1B5-34 that resulted in a 50% decrease in the level of protection (Fig. 3B). The active protein concentration was calculated as described in Materials and Methods, and values of 2.25 and 5.9 nM for mHSF1 and mHSF2 were obtained, corresponding to 43 and 98% active protein, respectively. In subsequent footprinting experiments, we evaluated the level of protection from DNase I cleavage and correlated this with the amount of active protein.

Binding of mHSF1 and mHSF2 to certain selected sequences mimics interactions on the HSP70 promoter. Our previous analysis of mHSF1 and mHSF2 binding to the HSP70 HSE demonstrated that mHSF1 protected all five sites and that mHSF2 protected only sites 2 through 5 (18). To reevaluate our original result in the context of the present experiments, we cloned an oligonucleotide containing the HSP70

A

HSF1 SELECTED SITES

```
1B5-4                    cgCGCagGAAggttgtaaatgctgaccgaggcg
1B5-15                   cgATCtgGTGgttagcggttggatgctttcgaggcg
1B5-18                cgatgtttggcGAAtgTTCtaatcttgaggcg
1B5-21                   cgGACctGTTtatcgtcggagggaacgaagaggcg
1B5-24.1  cggaatacaaatttcagTTCcaGAAaggagcg
1B5-32       cgagaatattggagaagcGCAgtGTCagaggcg
1B5-36         tGTCgaGAActttacgGCAcaTTCggttccgagaggcg
1B5-39      cgacacggctttaagGCCatGGAacatgcgaggcg

1B5-24                 cgggTTCtgGAAagAACcagattgagaggcg
1B5-24.2                 cgGACtaGAAttTTCgcagctttagaggcg
1B5-26       cgtagtagggGATtgTTCgaGAAttcgata....
1B5-33          cgatgtaGAAcgCTCtcGAAttgtggaggcg
1B5-37            cgaccGTTacCACatGACtaagactcgaggcg
1B5-38            cgactGAAtgGTCgaGAAgatttggagaggcg

1B5-1        cgtaattatatacgcTTCgaGGAgtTTCgaGGCg
1B5-6                 cgtagTTCgaGAAtgTGCgaGAAgaagaggcg
1B5-7              tGTCgaTTCtaGAAcaCTCtctcctgctacgaggcg
1B5-8              tGTCgcGACcaGAAcgTTCtgcgttccagtgaggcg
1B5-9             cgGTAggTACgtGAAcaTTCtaattacatgaggcg
1B5-16       cgggggagGAAgaGTCgaGAAgcAACgaggcg
1B5-17                cGAAggGTCgaGAActTTCgaaccagctggaggcg
1B5-20                cGATacTTCgaGATatTTCcgaggcgaggcg
1B5-22          cgaacctattaTTCgaGAAcaTGCatGTGaggcg
1B5-23                cGAGaaTGCtcGAAatTTCcttagtcggaggcg
1B5-25          cgagattagctgTTCtaGCAtgTACcaGAGgcg
1B5-28         cggagaatgcagtaTTCatGAAtgCGCgaGGCg
1B5-31                cGAAtaATCgaGAGttATCtatgtcactagaggcg

1B5-2          cgagattGAAagTTCctGAAagATCacGAGgcg
1B5-3        cgggtATCggGAAtaCACggGAAacTTCtgaggcg
1B5-5         cgagctatgGTGatTACcaGAAggTTCgcGAGgcg
1B5-11      cgaccggggagGATcaTACgaGTActTTCgaGGCg
1B5-13              cGAGtgTTCtaGATcaGCCgaGGAtattgaggcg
1B5-14           tGTCgaGATtgTGCttGAAtgGTCcactcctgaggcg
1B5-27              cGAAatGCCgaGTAtaTTCgaGTAtgcgaggcg
1B5-29         cgcgagggGAAcaTACcaGAAtgATCgaGGCg
1B5-34      cgttcTTCtgGATggTTCtgGAAggTCCcgaggcg
1B5-35         cggactaaGAAgaTTCgcGGAggTTCgaGGCg
1B5-40              cGATgcTACtaGAAggTTCacGAGtagaggcg

1B5-10          cgTACtcGAAtaGCCtaGAAcgTGCttGAGg

1B5-12           tGTCgaGAAtgTTCggGACcgATCagGGTgaGGCg
1B5-30      cGCTtgGACagGTTatTGCtaGAActGTCgaGGCg
```

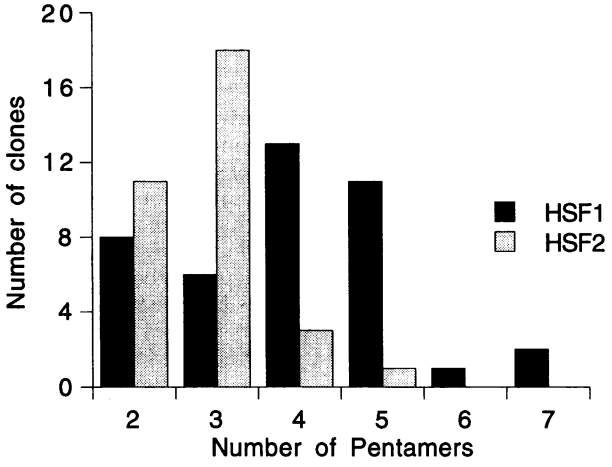
B

HSF2 SELECTED SITES

```
2U5-1                 cgaagATCtgGAAcactgaccatccgaggcg
2U5-3                    cgcgGAAggTTCtatacgaccttgaggcg
2U5-6                 cggtAGCacGAAtctggaaaacctagaggcg
2U5-10              cgagtgccagaGAAtaTTCggcacgggaggcg
2U5-24             cgtagttATCtcGAAgcaTACggGAGgaggcg
2U5-25          cgtgGAAcaCTCggtGAAcgTTCcgtccggaggcg
2U5-26             cgacaagtagtaGAAtcTTCggacgcgaggcg
2U5-34             cGAAggAGCggaGAAgtATCagaagaggcg
2U5-36          cgattaggaacaaagGATtaTTCcataaggaggcg

2U5-2        cggggtagtatgcttttcGAAacTTCgaGAGagaggcg
2U5-4                       cGAAtgCCCgaGAAcccaaacactgatagaggcg
2U5-8          cgtaataacgtagacGATtgTTCgtGTTcgaggcg
2U5-12                   tGTCggGAActTTCtaccgctttgacgtctgaggcg
2U5-19                    cgacaGAAtcTTCacGTAagttagggcgaggcg
2U5-21                   cgatggGAGcaAACgtGTTctagatcgaggcg
2U5-27.1                   cggaGAAcaATCtcGCGtccaaagtgaggcg
2U5-27.2                 tGTCgaGAAtgTTCcgcgaatcaacttacccgaggcg
2U5-28                   tGTCgaGAAtaATCtacggagacccctctgttgaggcg
2U5-29     cgcgttccttaagcttatgGAAacCACgaGGCg
2U5-30.1                 tGTCgaGAAacATCtatagcgcttctgtgaggcg
2U5-31              cgacgggtagGATgcTTCtaGAAaagagaggcg
2U5-32           cgaggcgatcgaagcGAAgtTTCgaGAGaggcg
2U5-32.1                    cGAAtaTTCtaGTTctaaaactaaggccgaggcg
2U5-32.2                 tGTCgaGAAggATCgacccaaaagtggccttagaggcg
2U5-33   cgtgtatgcttaacgtgcagcGATgcTTCgaGGCg
2U5-35   cgtggctatgcgatatgttgaGAAttAACgaGGCg
2U5-37     cgaagggtgcgaaggatatGATtgTTCgaGGCg
2U5-38                      cGAAttTACgcGATcgaaagaaacatgaggcg
2U5-40             cgcagagggtatGACtaTTCcaGAAtaagaggcg

2U5-27   cgccaggacggagggaACCccGAAtaATCgaGGCg
2U5-30                   cgattaGAAtaATCccGAGcgGGCctttggaggcg
2U5-39        cggcgagGGAttGGCtcGAAgaTTCtaaggaggcg

2U5-22           cggGCAggTACaaGGAagATCtgGAAagaggccg
```

FIG. 1. Compilation and analysis of selected oligonucleotides. (A and B) The sequences of the oligonucleotides that were selected by mHSF1 and mHSF2 as described in Materials and Methods are shown in 5′-to-3′ orientation. Clones are identified by a number and a letter (1B or 2U) to indicate which protein selected the site (mHSF1 or mHSF2) and the protein-DNA complex isolated (B or U); the number 5 indicates that the clones were isolated after five rounds of selection, and the final number identifies the specific clone. In some instances, more than one oligonucleotide was cloned into a single vector, and these clones have an additional number to designate their order and relationship (e.g., 1B5-24, 1B5-24.1, and 1B5-24.2). The random sequence that was selected lies between 5′-cg-3′ at the 5′ side and 5′-gaggcg-3′ on the 3′ side. Even though the flanking sequences have no HSF binding sites alone, in some instances the binding site that was selected included sequences adjacent to the randomized region, and this has been included (e.g., 1B5-7 and 1B5-14). The core motifs of the binding sequences within each oligonucleotide are capitalized and underlined. The selected oligonucleotides for both mHSF1 and mHSF2 have been arranged according to increasing numbers of potential pentameric binding sites. (C) Histogram that demonstrates the relationship between the number of pentamers per binding site and the number of clones of each type isolated by mHSF1 and mHSF2.

C



HSE into Bluescript and used a PCR-amplified template for DNase I footprinting. We confirmed our previous results—specifically, that mHSF1 contacted all five pentamers and that mHSF2 failed to contact the first pentamer of the HSP70 HSE—with this new clone (Fig. 4A). The footprints are of the bottom strand and have DNase I protection boundaries identical to those reported previously (18). The apparent K_d values for the binding of mHSF1 and mHSF2 to the HSP70 HSE were calculated to be 0.83 and 2.9 nM, respectively (Table 1), and this agreed well with our previous estimate of binding affinity that utilized saturation binding (18).

In the course of our footprinting analysis, we also found selected sequences to which mHSF1 and mHSF2 bound differently. An mHSF1-selected site, 1B5-40, which had five potential pentameric sites, was bound by both mHSF1 and mHSF2 with apparent K_d values of 0.42 and 1.0 nM, respectively (Table 1). Interestingly, mHSF1 protected all five sites (footprint size = 34 nucleotides) in 1B5-40, whereas mHSF2 bound to only sites 1 through 3 (footprint size = 24 nucleotides) (Fig. 4B, compare lanes K and U). The 3′ boundary of mHSF2 bound to 1B5-40 was in pentamer 4 (5′-gTTCa-3′), and this suggested that even though this was a consensus pentamer, it was not contacted. This differential binding was also observed on 1B5-10, where mHSF1 protected all six sites

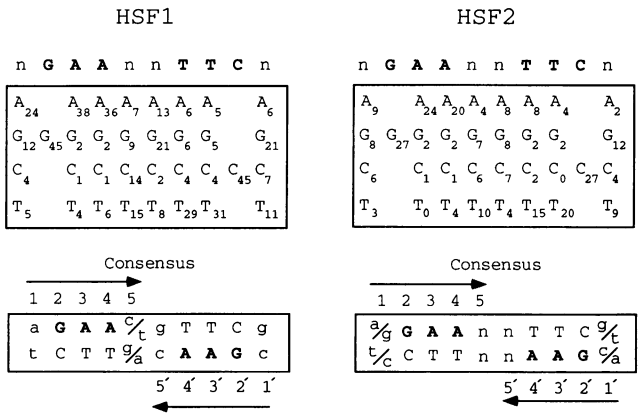


FIG. 2. Derivation of the consensus binding sites for mHSF1 and mHSF2. The potential dimeric arrays (5′-nGAAnnTTCn-3′) contained in each oligonucleotide selected by mHSF1 and mHSF2 were aligned, and the number of nucleotides (A, G, C, or T) at each position was calculated. In order not to bias the analysis, only those arrays that were selected entirely from random sequence were included in the analysis. The dimeric consensus sequences for mHSF1 and mHSF2 derived from this analysis are shown at the bottom as double-stranded DNA to allow easier interpretation. The nucleotide positions in the inverted pentameric repeats (indicated by arrows) are numbered 1 to 5 and 1′ to 5′ to permit identification of specific nucleotide positions as discussed in the text.

(footprint size = 37 nucleotides) from DNase I digestion and mHSF2 protected only sites 1 through 5 (footprint size = 31 nucleotides) (data not shown). The simplest explanation for this observation was that different numbers of mHSF1 and mHSF2 trimers were bound to these selected sequences.

To examine the basis for this difference between mHSF1 and mHSF2 binding, we first measured the dissociation rate for mHSF1 and mHSF2 binding on 1B5-40 by DNase I footprinting. Both factors dissociated with half life of ≈10 min, suggesting that it was not the rate of dissociation that affected the differential appearance of the footprints (data not shown). We considered that there might be some specific sequence requirements for mHSF2 binding that were not obvious from our consensus. It was possible that the final pentamers in 1B5-40 (5′-gTTCacGAGt-3′) and 1B5-10 (5′-tGAGg-3′) were low-affinity sites for mHSF2 interaction. Notably, there was a thymine in the 1′ position on the bottom strand of 1B5-40 pentamer 4 and the first position on the top strand of 1B5-10 pentamer 6, and this was not favored by either mHSF1 or mHSF2 (Fig. 2). We noted other oligonucleotides, 1B5-5 and 1B5-25, that had terminal 5′-nGAGn-3′ pentamers that were bound equally well by mHSF1 and mHSF2, suggesting that the core sequence 5′-GAG-3′ was not responsible for the differential interaction of mHSF2. This led us to conclude that it was more likely the thymine residues present in the first position of pentamers 4 and 6 of 1B5-40 and 1B5-10, respectively, that hindered mHSF2 binding. Why should mHSF1 completely protect these sites when it had an equally strong selection against thymine in the first position of the pentameric repeat? One interpretation was that there were stronger cooperative interactions between adjacent mHSF1 trimers and that it was this interaction that resulted in the complete protection of 1B5-40 and 1B5-10 in spite of the unfavored T residues.

Comparison of mHSF1 and mHSF2 binding on selected sequences. We compared the relative affinities of mHSF1 and mHSF2 for a variety of selected sequences, focusing primarily on the sequences selected by mHSF1, as they contained more pentameric repeats on average and were bound with higher affinity. The mHSF1-selected sequences offered the opportunity to compare the binding affinities of mHSF1 and mHSF2 on a variety of sequences with characteristics similar to those of known HSEs (see Fig. 4A and B for an example). A summary of the apparent K_d values obtained is shown in Table 1. Many of the sequences that were selected by mHSF1 were bound by both HSFs with an affinity comparable to that of the HSP70 and HSP90 HSEs. mHSF1-selected sequences had a range of affinities for mHSF1 interaction from 0.17 to 4 nM. mHSF2 bound to mHSF1-selected sequences; however, the affinity was generally somewhat weaker and ranged from 0.22 to >15 nM. This observation suggested that mHSF1 and mHSF2 had similar sequence requirements for binding, which our alignment analysis supported, but that there were in some instances significant differences in affinity. We note that since the apparent K_d values for mHSF1 and mHSF2 binding to sequences such as 1B5-34 and HSP90 were near the concentration of DNA in the reaction mixture, it should be considered that the K_d might be lower than 0.1 nM. Additionally, our estimation of mHSF1 and mHSF2 K_d values has not been corrected for the number of trimers bound to individual sequences, as we have not yet determined the stoichiometry of binding. However, for comparative purposes, our current analysis serves to address the issue of differences between mHSF1 and mHSF2 rather than the absolute affinities of each site.

The sequences selected by mHSF2 contained fewer pentameric repeats than did those selected by mHSF1 and in general were of lower affinity (Table 1). The apparent K_d values obtained for mHSF2 binding to mHSF2-selected sequences ranged from 1 to >15 nM, and mHSF1 also bound with similar affinity (3.3 to >12 nM). One mHSF2-selected site, 2U5-19 (Fig. 5), was bound with higher affinity by mHSF2 (apparent K_d = 1 nM) than by mHSF1 (apparent K_d > 7 nM). This trimeric site, 5′-aGAAtcTTCacGTAa-3′, had no distinguishing characteristics that would suggest weaker mHSF1 binding, except for the thymine at the 1′ position in the second pentamer. As noted above, thymines in the first position of pentamers were not favored in the selection process by either mHSF1 or mHSF2 and apparently interfered with mHSF2 binding in particular (e.g., 1B5-40, 1B5-10, and 1B5-2). The addition of increasing amounts of mHSF1 protein to 2U5-19 revealed that this protein was not able to fully protect the site, and the level of DNase I protection did not reach saturation, suggesting that the apparent K_d might actually be higher than 7 nM (Fig. 5, lanes C through F). In contrast, increasing the amount of mHSF2 protein resulted in complete protection of 2U5-19 from DNase I digestion (Fig. 5, lanes H through K). This result was surprising, as we did not suspect that mHSF1 would be so significantly affected by the thymine residue. In order to determine if the thymine in the 1′ position of the central pentamer was responsible for the poor mHSF1 binding, we synthesized a variant (2U5-19mtp) in which this thymine was changed to an adenine, resulting in a consensus pentamer (Fig. 5). We tested if this new site (2U5-19mtp) was bound by mHSF1 with higher affinity than that seen with 2U5-19, and we also compared this with the binding of mHSF2. Titration of this new site revealed that both HSFs bound 2U5-19mtp with higher affinity and at similar levels (0.3 nM), suggesting that the thymine at this position did have deleterious effects on binding. There was a greater-than-10-fold increase in the affinity of mHSF1 for the 2U5-19mtp site, and the affinity of mHSF2 binding also increased severalfold. This experiment and those mentioned above demonstrate that the composition

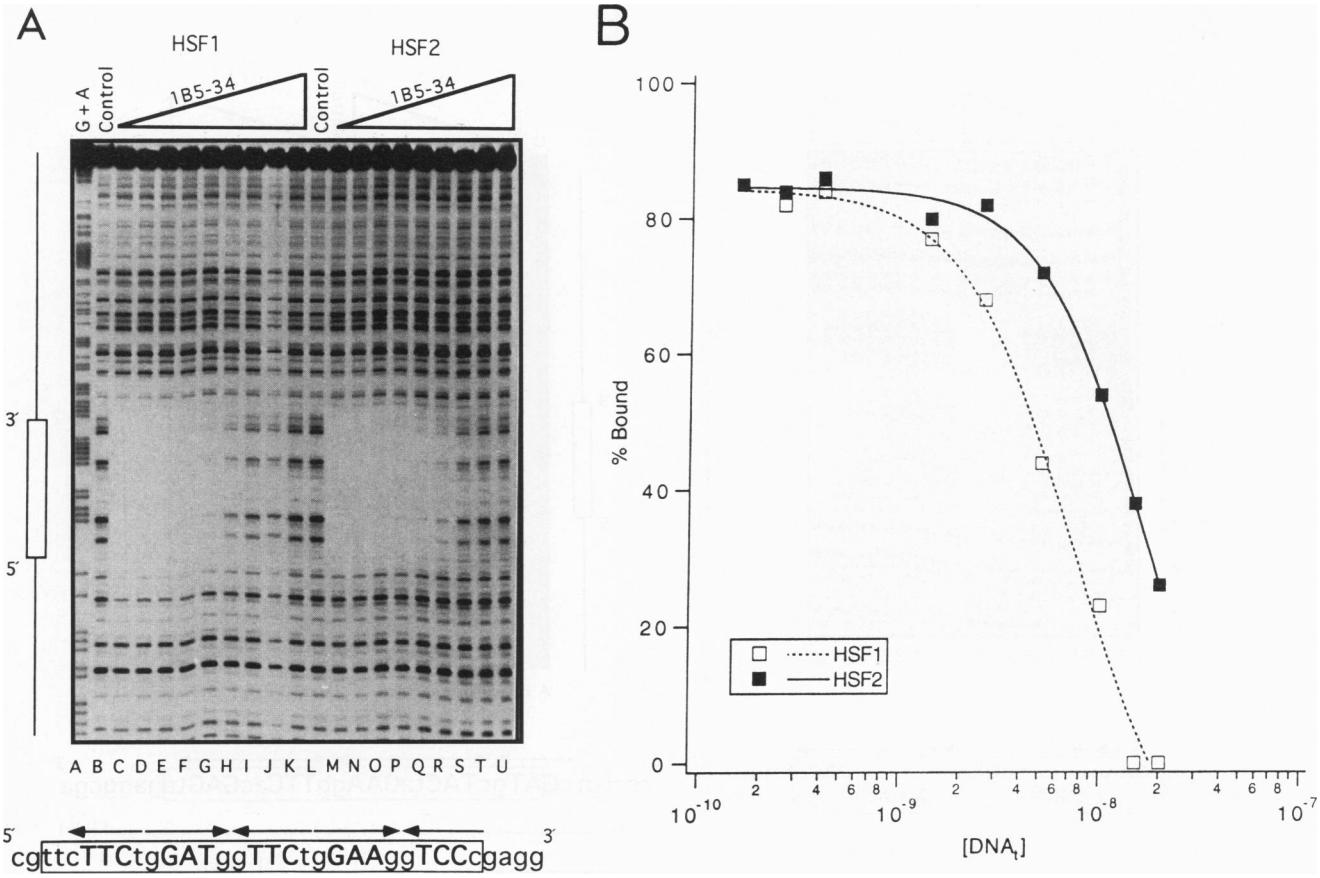


FIG. 3. Determination of active protein concentrations for mHSF1 and mHSF2 by DNase I footprinting. (A) Binding reaction mixtures that contained approximately 6 nM mHSF1 or mHSF2 and 0.16 nM labeled 1B5-34 DNA plus increasing amounts of unlabeled 1B5-34 DNA were established. The reaction mixtures were treated with DNase I and subjected to electrophoresis as described in Materials and Methods. The total concentrations of DNA in lanes B through K and L through U are 0.16, 0.16, 0.26, 0.41, 1.41, 2.66, 5.16, 10.2, 15.2, and 20.2 nM, respectively. Lane A contains a G+A sequencing ladder of 1B5-34, and lanes B and L show control reactions done in the absence of mHSF1 and mHSF2. The box at the left of the gel is marked with the orientation of the sequence and is also overlaid on the sequence below the gel. (B) The extent of DNase I protection by mHSF1 or mHSF2 in each lane of the gel shown in panel A was determined by direct measurement with the PhosphorImager and expressed as the percentage bound versus the log concentration of total DNA. The amount of active mHSF1 or mHSF2 protein was determined by calculation of the concentration of DNA required to reduce the amount of mHSF1 or mHSF2 bound to the labeled probe by 50% by using the formula $[P_t] = K_d + 1/2[D_t]$ as described in Materials and Methods. $[DNA_t]$, total concentration of DNA.

of the pentamer, particularly the nucleotide adjacent to the essential guanine, can have differential effects on mHSF1 and mHSF2 binding depending on the sequence context.

mHSF1 exhibits a higher degree of cooperativity in DNA binding than does mHSF2. In our studies, we encountered certain sequences on which only mHSF1 bound across extended regions containing groups of pentamers. The binding site titration of mHSF1 and mHSF2 on 1B5-13 substantiated this observation (Fig. 6A). As the amount of mHSF1 or mHSF2 protein was increased in the binding reaction mixture, the protection from DNase I cleavage in the region of the selected oligonucleotide increased (Fig. 6A, lanes H through K and R through U). We noted that binding of mHSF1 and mHSF2 occurred initially in the region of the selected oligonucleotide that contained the five pentameric repeats (compare lanes H and I and lanes R and S in Fig. 6A) and that the two proteins had the same 5′ boundary. However, when the concentration of mHSF1 protein was raised from 1.6 to 6.6 nM, there was a striking increase in the size of the mHSF1 footprint at the 3′ boundary (Fig. 6A, compare lanes I and J). This increase in the extent of the mHSF1 footprint resulted from protection of an additional 15 bp corresponding to three additional pentameric repeats. mHSF2 showed no protection of this region, even as the concentration of protein increased (Fig. 6A, lanes S through U).

Why was mHSF1 able to protect this adjacent region whereas mHSF2 could not? We examined the sequences flanking the 3′ boundary of the selected oligonucleotide for potential pentameric sites. The intervening region immediately adjacent to the selected region was

5′-attga-3′
3′-taact-5′

which did not have the conserved guanine residue at the 2′ position on the bottom strand; yet, it was appropriately spaced and retained the consensus adenine residues in positions 3′ and 4′ (striped box above sequence schematic in Fig. 6A). Adjacent to this site were two pentameric repeats that could provide a basis for mHSF1 interaction (Fig. 6A, bottom).

We considered that the region protected in the extended mHSF1 footprint on 1B5-13 simply represented a low-affinity

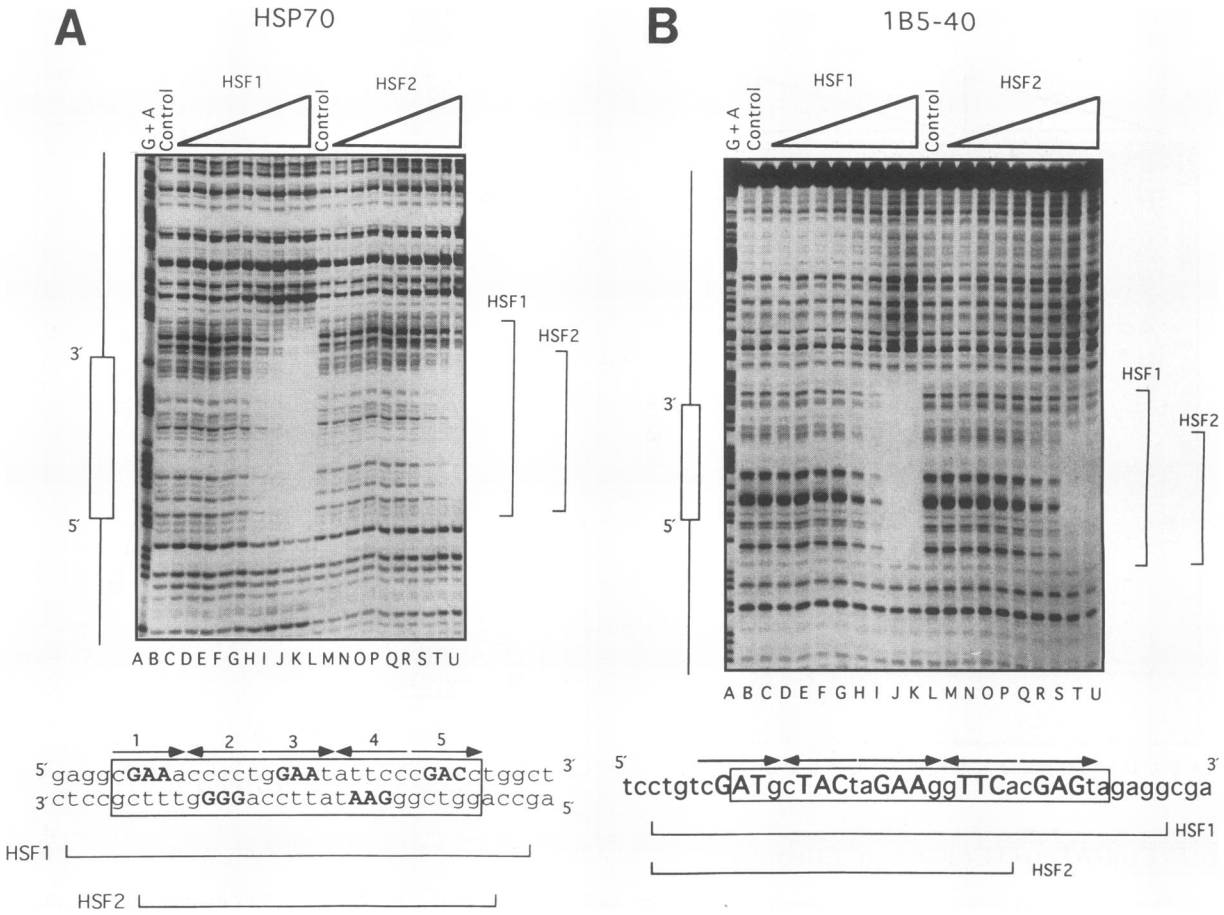


FIG. 4. Differential interaction of mHSF1 and mHSF2 on 1B5-40 mimics binding to the HSP70 HSE. (A) Equilibrium DNase I footprinting of mHSF1 (lanes C through K) and mHSF2 (lanes M through U) binding to the HSP70 HSE. The concentration of HSP70 HSE probe in all reaction mixtures was 0.1 nM, and the probe was labeled at the 5′ end on the bottom strand. The concentrations of mHSF1 in lanes C through K were 0.026, 0.052, 0.103, 0.206, 0.413, 0.826, 1.65, 3.3, and 6.6 nM, respectively. The concentrations of mHSF2 in lanes M through U were 0.03, 0.06, 0.12, 0.24, 0.48, 0.96, 1.92, 3.84, and 7.68 nM, respectively. Lane A contains the G+A sequencing ladder, and lanes B and L show control DNase I footprinting reactions. The extents of mHSF1 and mHSF2 binding are indicated at the right with brackets. At the bottom, the sequence of the HSP70 HSE is shown schematically and the boundaries of mHSF1 and mHSF2 interaction are indicated with brackets. The box at the left of the gel is marked with the orientation of the sequence and is also overlaid on the sequence below for additional assistance. The orientations of sites 1 through 5 are indicated with arrows, and the core motifs are boldfaced. (B) Equilibrium DNase I footprinting of mHSF1 and mHSF2 bound to 1B5-40. Lanes are as described for panel A. The concentrations of mHSF1 protein in lanes C through K were 0.012, 0.026, 0.052, 0.103, 0.206, 0.413, 0.826, 3.3, and 6.6 nM, respectively. The concentrations of mHSF2 protein in lanes M through U were 0.03, 0.06, 0.12, 0.24, 0.48, 0.96, 1.92, 7.68, and 15.4 nM, respectively. The extents of mHSF1 and mHSF2 binding are indicated at the right, and below the sequence, with brackets. The core of the pentameric repeats and their orientations are indicated with arrows. The boxed sequence corresponds to the random region of the oligonucleotide that was selected by mHSF1.

site of interaction. To address this issue and the apparent differences in cooperativity between mHSF1 and HSF2 directly, we prepared two variants of the 1B5-13 site (Fig. 6B). In one instance, the fifth pentameric site in 1B5-13 was mutated by changing the consensus guanine to a cytosine (1B5-13mtp). This created a 10-bp gap between the two regions tested, showing the necessity for proximity in mHSF1 binding at the selected region to stabilize mHSF1 interaction in the adjacent sequences. To test if mHSF2 was also capable of binding to the adjacent region, we made 1B5-13th2 in which the 5-base intervening region, noted above, was mutated to create an HSF pentamer binding site. We suspected that if a consensus pentamer was present at this position, both mHSF1 and mHSF2 could bind to the selected and adjacent regions of 1B5-13 equally. Utilizing DNase I footprinting, we compared the interactions of mHSF1 and mHSF2 on these substrates. When the fifth pentamer of the selected region was mutated (Fig. 6B, lanes G and H), there was protection of the remaining selected region (sites 1 through 4) but no binding of mHSF1 to site 5 or the adjacent region as seen for 1B5-13 (lanes C and D). mHSF2 bound to 1B5-13mtp with equal affinity and had boundaries of interaction similar to those of mHSF1. We concluded that when the fifth pentamer of the 1B5-13 selected region was mutated, mHSF1 trimers bound to the selected region were not in sufficient proximity to stabilize binding in the adjacent sequences. In contrast, when the 5-base sequence between the selected and adjacent regions was changed to a consensus pentamer, both mHSF1 and mHSF2 were able to bind across both the selected and adjacent regions equally (Fig. 6B, lanes K and L). This demonstrated that mHSF2 was capable of interaction with the adjacent region under appropriate circumstances and that lack of mHSF2

TABLE 1. Compilation of K_d values for certain selected binding sequences[a]

Binding site	Apparent K_d (10^{-9} M) ± SD		Pentameric repeats	
	HSF1	HSF2	No. of repeats	No. of consensus repeats
1B5-34	0.17 ± 0.005	0.22 ± 0.014	5	3
HSP90	0.24 ± 0.025	0.31 ± 0.012	6	5
1B5-10	0.29 ± 0.009	1.42 ± 0.021	6	2
1B5-17	0.35 ± 0.038	0.42 ± 0.025	4	3
1B5-40	0.42 ± 0.032	1.00 ± 0.11	5	2
1B5-12	0.59 ± 0.11	2.79 ± 0.66	7	2
HSP70	0.83 ± 0.11	2.90 ± 1.1	5	3
1B5-25	0.95 ± 0.03	3.71 ± 1.1	4	1
1B5-11	0.99 ± 0.071	2.64 ± 0.41	5	1
1B5-5	1.02 ± 0.09	1.95 ± 0.33	5	2
1B5-2	1.14 ± 0.15	>15.0	5	3
1B5-22	1.39 ± 0.19	2.99 ± 0.79	4	2
1B5-13	1.55 ± 0.09	5.16 ± 1.09	5	1
1B5-23	1.55 ± 0.04	2.40 ± 0.23	4	2
1B5-18	3.02 ± 0.27	3.48 ± 0.52	2	2
2U5-30.1	3.3	3.8	3	1
1B5-30	3.91 ± 0.4	>10.0	7	1
2U5-12	4.0	3.0	3	2
2U5-19	>7.0	1.0 ± 0.8	3	2
2U5-3	>10.0	10.0	2	2
2U5-10	>10.0	>10.0	2	2
2U5-29	>10.0	>10.0	3	1
2U5-22	>12.0	>15.0	5	1
2U5-30	>12.0	>15.0	4	1

[a] The apparent K_d values ± standard deviations for all sequences are listed and were determined from quantitative DNase I footprinting as shown in Fig. 4 and 7. In some instances, there were insufficient points examined for calculation of the standard deviation. The total number of pentamers and number of consensus repeats (5′-nGAAn-3′) present in each binding site are indicated. The table is arranged according to increasing mHSF1 K_d values.

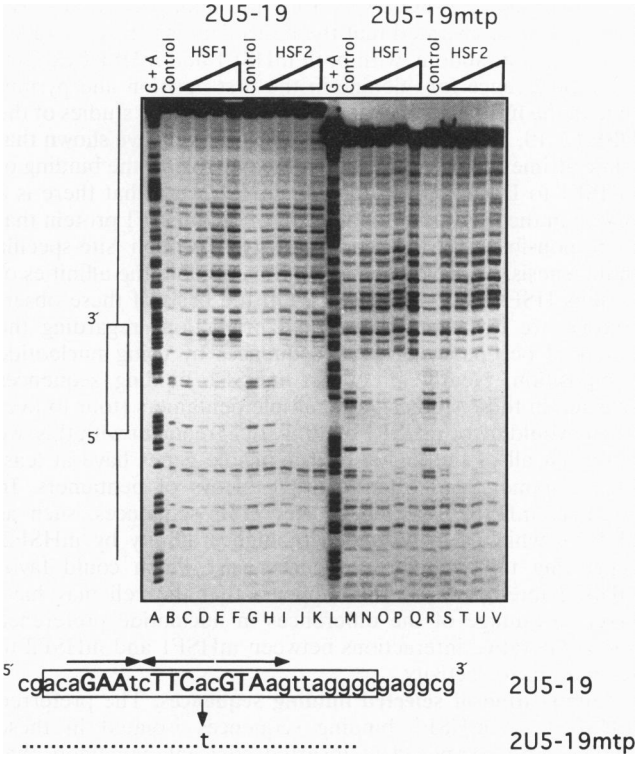


FIG. 5. Binding of mHSF1 and mHSF2 to 2U5-19 and 2U5-19mtp. DNase I footprinting was performed as described in the legend to Fig. 4. Lanes: A, G+A sequencing ladder of 2U5-19 probe; B and G, control DNase I reactions in the absence of protein; C through F, increasing concentrations of mHSF1 protein (0.41, 1.65, 6.6, and 24 nM, respectively); H through K, increasing concentrations of mHSF2 protein (0.48, 1.9, 7.8, and 31 nM, respectively). Lanes L through V are the same as A through K except that 2U5-19mtp DNA was used. Below the gel, the sequence of the top strand of 2U5-19 is shown with the pentameric repeats marked. The mutation that was introduced to make 2U5-19mtp is shown (discussed in the text).

binding in 1B5-13 was not because the sequence itself precluded mHSF2 interaction but instead was due to a lack of cooperative interactions between mHSF2 trimers. These results demonstrate that only mHSF1 can bind to the adjacent region of 1B5-13 through cooperative interactions with an mHSF1 trimer bound in the selected site. These results are also supported by earlier observations of the number of pentameric repeats selected by each protein and demonstrate that mHSF1 utilizes stronger cooperative interactions between adjacent trimers to stabilize binding across extended regions.

Chimeric HSFs reveal a new domain that regulates trimer-trimer cooperativity. Was there a specific domain of mHSF1 that could confer positive interactions between adjacent trimers? To address this question, chimeric HSFs were constructed, as shown in Fig. 7A. The chimeras contain the DNA-binding domain (DBD) and part of the oligomerization domain of one factor and the C terminus of the other factor. The DNA-binding properties of the chimeras were compared with those of the authentic mHSFs in a DNase I footprinting assay with 1B5-13 as the target site (Fig. 7B). As the concentration of full-length mHSF1 protein increased, the primary and adjacent binding sites were protected (Fig. 7B, lanes C through G), whereas full-length mHSF2 bound well only to the primary site and failed to bind to the adjacent sequences as described above (lanes N through R). In comparison, the DNA footprint of HSF1DBD/HSF2 was extended, as was observed with mHSF1 alone (Fig. 7B, lanes H through L), whereas HSF2DBD/HSF1 exhibited the same footprint as that observed for mHSF2 (lanes S through W). The HSF1DBD/HSF2 chimera also produced the characteristic DNase I hypersensitivity (indicated in Fig. 7B with an arrow at the left) observed with the intact mHSF1 protein (compare lanes G and L). These results suggest that the enhanced cooperative interactions of mHSF1 were due to amino acid sequences in the DBD that increased the stability of adjacent bound trimers. Attempts to further delineate the position of the cooperativity domain in mHSF1 have been inconclusive, as additional chimeras that contain hybrid DBDs either were inactive or exhibited no cooperativity. Thus, it will be necessary to construct specific mutants of the mHSF1 DBD to address this question further. The results obtained with 1B5-13 binding were further corroborated by footprinting 1B5-40, another site that exhibited differential mHSF1 and mHSF2 binding (data not shown). These results further support our contention that the differences in binding by mHSF1 and mHSF2 were related to differences in the potential for cooperative interactions between adjacent trimers.

DISCUSSION

In the studies presented here, we have compared two related transcription factors, mHSF1 and mHSF2, with respect to their DNA recognition properties. Our experiments have established in an unbiased manner the consensus sequence of the

HSE pentamer (5′-nGAAn-3′) for the multiple mHSFs. We have also demonstrated that the bases flanking the core GAA motif are nonrandom, with both mHSF1 and mHSF2 exhibiting a preference for adenine in the first position and pyrimidine in the fifth position, consistent with previous studies of the HSE (2, 10, 12, 25, 41). Most significantly, we have shown that trimer-trimer cooperativity has a greater role in the binding of mHSF1 to DNA than it does for mHSF2 and that there is a region in the N-terminal sequences of the mHSF1 protein that is responsible for this cooperativity. Through site-specific mutagenesis, we have been able to manipulate the affinities of various HSEs for certain HSFs. On the basis of these observations, we can then make some predictions regarding the potential occupancy of HSE sequences by using nucleotide composition. Nearly all of the mHSF1 binding sequences selected in these studies had multiple pentamers (four to five) which would favor mHSF1 binding. In agreement with this, we note that all of the highly heat-inducible genes have at least five pentamers and often multiple arrays of pentamers. In contrast, mHSF2 selected shorter HSE sequences, such as 2U5-19, which were bound with higher affinity by mHSF2, suggesting that cellular sequences exist which could favor mHSF2 interaction. It thus appears that the cell may have taken advantage of the differences in nucleotide preference and cooperative interactions between mHSF1 and mHSF2 to regulate factor activity.

Composition of selected binding sequences. The preferred mHSF1 and mHSF2 binding sequences isolated in these studies were composed of inverted adjacent pentamers that contained the primary sequence 5′-nGAAn-3′. This was consistent with previous analyses of the HSEs found in genes regulated by HSF (2, 41). Through alignment of dimeric repeats, we demonstrated that there was a preference for nucleotides in the first and fifth positions of the pentamer. Most notably, there was an alternating preference for A or C in the first position (1 and 1′) of adjacent pentamers selected by mHSF1. Our analysis demonstrated, as predicted from previous studies, that the spacing of the pentamers was critical for high-affinity interactions (2, 41). We tested several mHSF2-selected sequences that contained two dimeric repeats spaced inappropriately 3 bp apart (2U5-24 and 2U5-34) and observed no DNase I protection with an excess of either mHSF1 or mHSF2 (data not shown). This observation reinforced the model in which all pentameric sites must be on the same face of the DNA helix for efficient HSF interaction to occur (2, 25, 41). We can conclude that it is the overall composition of the binding site, primarily the number and arrangement of consensus pentamers, that dictates the affinity of interaction. The oligonucleotides we have selected contain a variety of binding sequences, which we utilized to establish the consensus. The diversity of sequences isolated indicates that HSF is a flexible protein tolerant of sequence changes.

Differential affinity of HSFs for specific sequences. The affinity of protein binding to certain sequences can be modulated by subtle changes in the consensus sequence, as exemplified by studies of Sp1 and the GATA family of factors (17, 19, 20). Thus, one way for the cell to modulate transcriptional activity of a gene regulated by a family of factors is to utilize binding sites that have a specificity or preference for one of the factors. We found that while mHSF1 and mHSF2 recognized nearly identical consensus sequences, there were certain sequences bound by only one HSF with high affinity. For example, the mHSF1-selected sequences, 1B5-2, 1B5-30, and 1B5-12, were bound with higher affinity by mHSF1 than by mHSF2. Conversely, the mHSF2-selected site, 2U5-19, was bound more avidly by mHSF2 than by mHSF1 and mutagenesis of this site could predictably change the site into one that bound mHSF1 and mHSF2 equally. These results suggest that along with the other mechanisms involved in the regulation of HSF activity (oligomerization, nuclear translocation, and phosphorylation), the composition of the HSE may dictate whether a particular HSF will bind. This is relevant, since recent studies have demonstrated that there can be more than one active HSF present in a cell (34). There is the potential for mHSF1 or mHSF2 to have positive or negative effects on gene expression when bound, and this might be dependent to some extent on the nature of the binding site. Studies with *D. melanogaster* have identified numerous HSF binding sites along the polytene

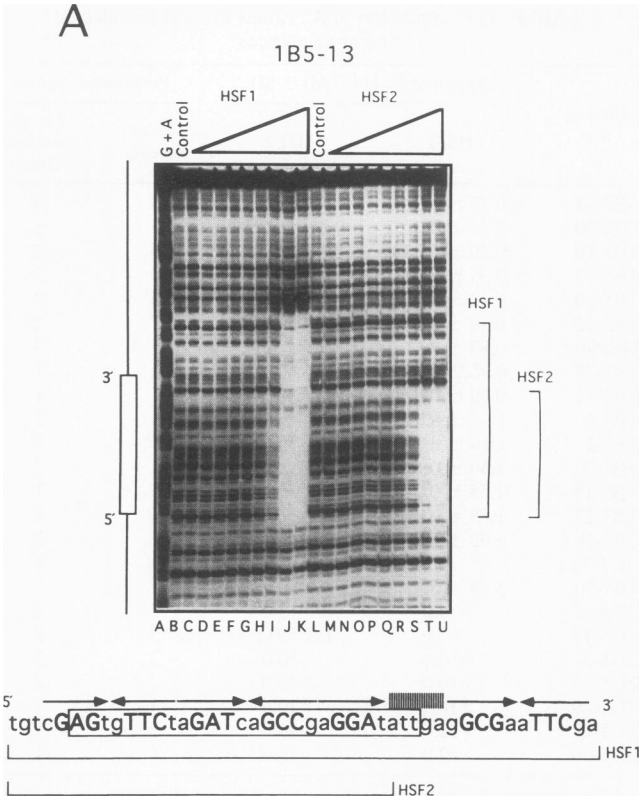


FIG. 6. DNase I footprinting of mHSF1 and mHSF2 reveals differences in trimer cooperativity. The binding and footprinting reactions were performed as described in the legend to Fig. 4. (A) Binding of mHSF1 and mHSF2 to 1B5-13. The concentrations of mHSF1 in lanes C through K were 0.026, 0.052, 0.103, 0.206, 0.413, 0.826, 1.65, 6.6, and 12 nM, respectively. The concentrations of mHSF2 in lanes M through U were 0.03, 0.06, 0.12, 0.24, 0.48, 0.96, 1.92, 7.68, and 15.4 nM, respectively. Lanes A, B, and L are as described in the legend to Fig. 4. The labeling is also as described in the legend to Fig. 4, and the striped box above the 1B5-13 sequence schematic indicates the pentameric sequence 5′-attga-3′ discussed in the text. (B) Comparison of mHSF1 and mHSF2 binding on 1B5-13, 1B5-13mtp, and 1B5-13th2. Reaction conditions were as described for panel A except that binding reaction mixtures contained 10 nM mHSF1 or 15 nM mHSF2. Lanes: A, E, and I, G+A sequencing ladders of the probes; B, F, and J, control DNase I reactions done in the absence of protein; C and D, mHSF1 and mHSF2 binding to 1B5-13; G and H, mHSF1 and mHSF2 binding to 1B5-13mtp; K and L, mHSF1 and mHSF2 binding to 1B5-13th2. At the right, the sequence of 1B5-13 is shown. The sizes of the DNase I-protected regions in each construct for mHSF1 and mHSF2 are delineated with arrows below the sequence. The mutations that were created are also shown below, and asterisks indicate which bases in the 1B5-13 sequence were mutated.

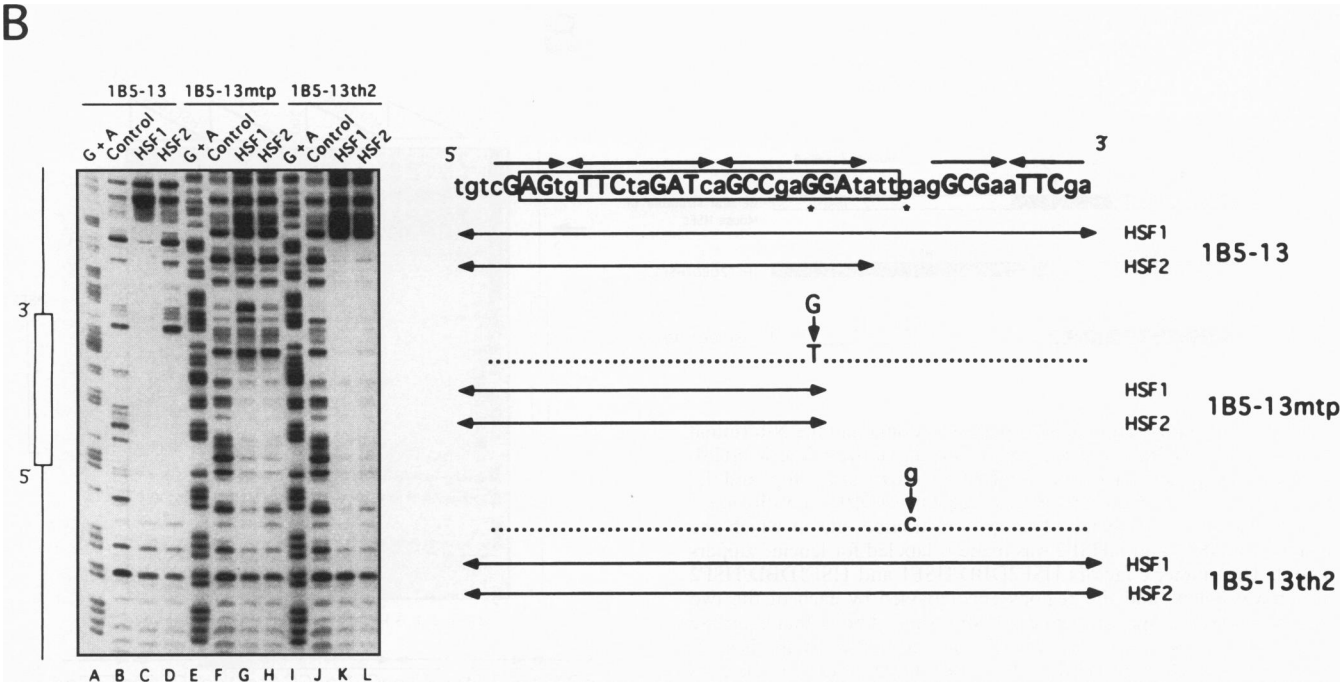


FIG. 6—*Continued.*

chromosomes, and it was suggested that some of these sites might be involved in the negative regulation of gene expression during heat stress (39). When more HSF-responsive genes are isolated, we will be better able to address the functional consequences of HSE specificity.

mHSF1 exhibits a higher level of cooperativity. Our experiments revealed that cooperative interaction between trimers could affect the binding of mHSF1 to certain sequences. To determine which region of the mHSF1 protein was responsible for increased cooperativity, we utilized chimeric HSF proteins. We found that the increased trimer-trimer cooperativity of mHSF1 could be transferred to mHSF2 when the DBD and part of the oligomerization domain of the two proteins were exchanged. Therefore, we suggest that the differences in mHSF1 and mHSF2 binding on most sequences are likely the function of differences in cooperative interactions between adjacent trimers and that the DBD or adjacent sequences of mHSF1 contain a region responsible for this increased cooperativity. Cooperativity was known to be important for *Drosophila* HSF binding, since a previous analysis demonstrated that the number of pentamers in a binding site correlated well with the stability of the interaction (42). Additionally, studies with yeast HSF have demonstrated that HSE composition can dictate the number of HSF trimers that are bound and the degree of heat inducibility possible (8). In support of our contention, a recent report has demonstrated a cooperativity domain in the hoxB5 protein immediately adjacent to the DNA-binding homeodomain (13).

Conclusions. Our experiments have demonstrated that both HSFs can bind to a variety of sequences. What are the implications of these results with respect to the regulation of known heat shock genes or other genes that contain HSEs within the promoter? The representation of a particular pentameric sequence within the genome can be calculated. For example, the human genome should contain approximately 10,000 to 12,000 copies of a perfect consensus trimeric array (5′-nGAAnnTTCnnGAAn-3′). Comparably, there would then be only three copies of a perfect 5-unit array (5′-nGAAnnT TCnnGAAnnTTCnnGAAn-3′) per genome. Thus, on the basis of our analysis, if the essential guanine residues were maintained and there were two or more consensus pentamers, preferably in a dyad array, the affinity of binding would be comparable to that of the HSE from the inducible HSP70 gene. Such an HSE as represented by 1B5-34 or HSP70, which both contain three consensus and two nonconsensus sites, would be found at 150 to 200 sites per genome. Likewise, the 5-unit arrays which conserve the guanine at the first position and some of the adenine residues in the third and fourth positions and have only two consensus pentamers would be represented on the order of 200 to 800 per genome. Are these estimations for the number of sites correct? Immunofluorescence studies of *Drosophila* HSF binding during heat shock have suggested that there are over 150 sites in addition to the heat shock genes with which HSF interacts; therefore, our estimation seems reasonable given the size of the human genome (39).

From our knowledge of the HSE and the requirements for binding, we can make some predictions about HSF occupancy. Since we know the number of molecules of mHSF1 per cell, we can calculate the molarity of mHSF1 trimers in the nucleus during heat shock, which for HeLa cells would be approximately 2.5×10^{-7} M (19, 29). For mHSF2 during hemin treatment, the concentration of trimeric protein in the nucleus would be approximately 1.3×10^{-7} M. Given this concentration of both HSFs, most high-affinity binding sequences should be occupied for a significant amount of the time, as the local concentration of HSF in the nucleus during activation exceeds the apparent K_d values for most of the binding sequences measured in our studies. However, if activation of transcription required full occupancy of the HSE by an HSF, then in some instances HSF2 would be unable to affect substantially the level of transcription, regardless of the protein concentration. Lower-affinity binding sequences, of which there would be upwards of 10^4 to 10^5, would be only partially occupied as the

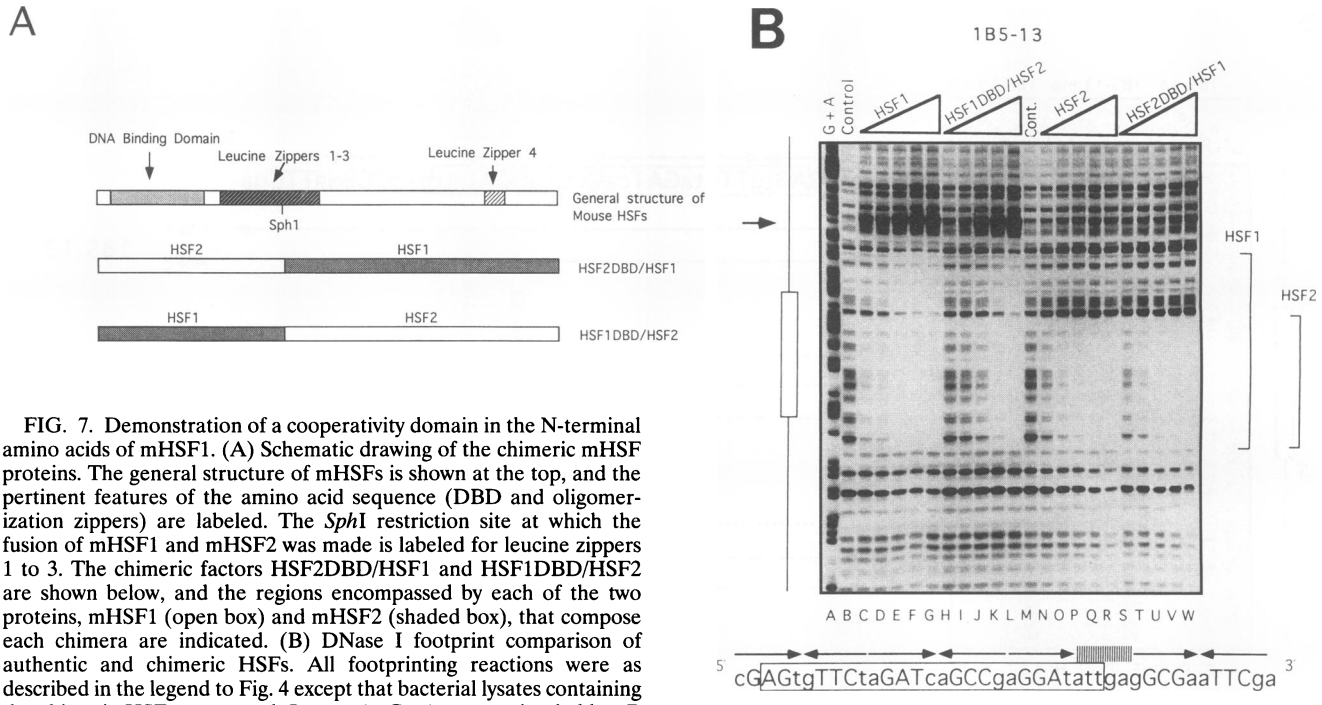


FIG. 7. Demonstration of a cooperativity domain in the N-terminal amino acids of mHSF1. (A) Schematic drawing of the chimeric mHSF proteins. The general structure of mHSFs is shown at the top, and the pertinent features of the amino acid sequence (DBD and oligomerization zippers) are labeled. The *Sph*I restriction site at which the fusion of mHSF1 and mHSF2 was made is labeled for leucine zippers 1 to 3. The chimeric factors HSF2DBD/HSF1 and HSF1DBD/HSF2 are shown below, and the regions encompassed by each of the two proteins, mHSF1 (open box) and mHSF2 (shaded box), that compose each chimera are indicated. (B) DNase I footprint comparison of authentic and chimeric HSFs. All footprinting reactions were as described in the legend to Fig. 4 except that bacterial lysates containing the chimeric HSFs were used. Lanes: A, G+A sequencing ladder; B and M, control DNase I footprinting reactions in the absence of added HSF protein; C through G, 0.83 to 12 nM mHSF1; H through L, 0.05, 0.1, 0.2, 0.5, and 1.5 μl of HSF1DBD/HSF2 lysate, respectively; N through R, 0.96 to 15.4 nM mHSF2; S through W, 0.05, 0.1, 0.2, 0.5, and 1.5 μl of HSF2DBD/HSF1 lysate, respectively. The extents of DNase I protection by mHSF1 and mHSF2 are indicated at the right, and all other labeling is as described in the legend to Fig. 4.

number of sequences exceeds the number of HSF molecules, although they might serve as a source of competition for the higher-affinity binding sites.

Of course, this analysis is in the absence of other factors, such as the effects of proteins that might interfere with HSF binding, the affinity of HSF for nonspecific sequences within the genome, or the location and accessibility of these sequences. With regard to the last point, previous studies have demonstrated that HSF is unable to bind in vitro to a chromatin template unless the chromatin has been disrupted by the binding of TFIID or the GAGA factor (6, 9, 38). Therefore, while there may be many potential binding sequences for HSF in the genome, if they are masked by chromosomal proteins then HSF would not be able to bind and would require an additional factor to gain access to the HSE. Whether all HSFs require an accessory factor to access the HSE in chromatin remains to be elucidated. One additional observation that may impact on our calculations is that, at least with HeLa cells, we have previously demonstrated that during heat shock HSF1 is concentrated in punctate structures within the nucleus (29). This suggests that the local concentration of HSF1 within these structures may be significantly higher than estimated from the number of molecules present in the cell. Further studies will elucidate whether these regions of high HSF1 concentration correspond to the sites of transcription for HSPs. In addition, there are undoubtedly other genes that can be activated or repressed by the interaction of HSF1 or HSF2, and the analysis of HSF binding sequences will aid in understanding the functionality of new HSEs as they are discovered.

ACKNOWLEDGMENTS

We thank the members of the Morimoto laboratory and Jonathan Widom for critical review of the manuscript. We are especially grateful to Jonathan Widom and Ed Feng for their time and help with the analysis of equilibrium binding and use of the Igor program. We also thank Mary Lin and Dave Kotzbauer for their help during the course of this work. We are indebted to Sue Fox for invaluable help with the photography.

These studies were supported by grants from the National Institutes of Health and the March of Dimes (R.I.M.). P.E.K. was supported by an individual National Research Service award.

REFERENCES

1. **Abravaya, K., B. Phillips, and R. I. Morimoto.** 1991. Heat shock-induced interactions of heat shock transcription factor and the human hsp70 promoter examined by in vivo footprinting. Mol. Cell. Biol. **11:**586–592.
2. **Amin, J., J. Ananthan, and R. Voellmy.** 1988. Key features of heat shock regulatory elements. Mol. Cell. Biol. **8:**3761–3769.
3. **Amin, J., M. Fernandez, J. Ananthan, J. T. Lis, and R. Voellmy.** 1994. Cooperative binding of heat shock transcription factor to the HSP70 promoter in vivo and in vitro. J. Biol. Chem. **269:**4804–4811.
4. **Ashburner, M.** 1970. Pattern of puffing activity in the salivary gland chromosomes of *Drosophila*. V. Response to environmental treatments. Chromosoma **31:**356–376.
5. **Baler, R., G. Dahl, and R. Voellmy.** 1993. Activation of human heat shock genes is accompanied by oligomerization, modification, and rapid translocation of heat shock transcription factor HSF1. Mol. Cell. Biol. **13:**2486–2496.
6. **Becker, P. B., S. K. Rabindran, and C. Wu.** 1991. Heat shock-regulated transcription in vitro from a reconstituted chromatin template. Proc. Natl. Acad. Sci. USA **88:**4109–4113.
7. **Blackwell, T. K., and H. Weintraub.** 1990. Differences and similarities in DNA-binding preferences of myoD and E2A protein complexes revealed by binding site selection. Science **250:**1104–1110.
8. **Bonner, J. J., C. Ballou, and D. L. Fackenthal.** 1994. Interactions between DNA-bound trimers of the yeast heat shock factor. Mol. Cell. Biol. **14:**501–508.
9. **Capdevila, M. D., and A. Garcia-Bellido.** 1974. Development and

genetic analysis of bithorax phenocopies in *Drosophila*. Nature (London) **250**:500–502.
10. **Cunniff, N. F. A., and W. D. Morgan.** 1993. Analysis of heat shock element recognition by saturation mutagenesis of the human HSP70.1 gene promoter. J. Biol. Chem. **268**:8317–8324.
11. **Dynan, W. S.** 1987. Dnase I footprinting as an assay for mammalian gene regulatory proteins, p. 75–87. *In* J. Setlow (ed.), Genetic engineering: principles and methods, vol. 9. Plenum Press, New York.
12. **Fernandes, M., H. Xiao, and J. T. Lis.** 1994. Fine structure analyses of the Drosophila and Saccharomyces heat shock factor—heat shock element interactions. Nucleic Acids Res. **22**:167–173.
13. **Galang, C. K., and C. A. Hauser.** 1993. Cooperative DNA binding of the human hoxB5 (hox-2.1) protein is under redox regulation in vitro. Mol. Cell. Biol. **13**:4609–4617.
14. **Hickey, E., S. E. Brandon, G. Smale, D. Lloyd, and L. A. Weber.** 1989. Sequence and regulation of a gene encoding a human 89-kilodalton heat shock protein. Mol. Cell. Biol. **9**:2615–2626.
15. **Higuchi, R., B. Krummel, and R. K. Saiki.** 1988. A general method of in vitro preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. Nucleic Acids Res. **16**:7351–7367.
16. **Kingston, R. E., T. J. Schuetz, and Z. Larin.** 1987. Heat-inducible human factor that binds to a human hsp70 promoter. Mol. Cell. Biol. **7**:1530–1534.
17. **Ko, L. J., and J. D. Engel.** 1993. DNA-binding specificities of the GATA transcription factor family. Mol. Cell. Biol. **13**:4011–4022.
18. **Kroeger, P. E., K. D. Sarge, and R. I. Morimoto.** 1993. Mouse heat shock transcription factors 1 and 2 prefer a trimeric binding site but interact differently with the HSP70 heat shock element. Mol. Cell. Biol. **13**:3370–3383.
19. **Letovsky, J., and W. S. Dynan.** 1989. Measurement of the binding of transcription factor Sp1 to a single GC box. Nucleic Acids Res. **17**:2639–2653.
20. **Merika, M., and S. H. Orkin.** 1993. DNA-binding specificity of GATA family transcription factors. Mol. Cell. Biol. **13**:3999–4010.
21. **Morimoto, R. I.** 1992. Transcriptional regulation of heat shock genes: a paradigm for inducible genomic responses. J. Biol. Chem. **267**:21987–21990.
22. **Morimoto, R. I.** 1993. Chaperoning the nascent polypeptide chain. Curr. Biol. **3**:101–102.
23. **Mosser, D. D., N. G. Theodorakis, and R. I. Morimoto.** 1988. Coordinate changes in heat shock element-binding activity and hsp70 gene transcription rates in human cells. Mol. Cell. Biol. **8**:4736–4744.
24. **Nakai, A., and R. I. Morimoto.** 1993. Characterization of a novel chicken heat shock transcription factor, HSF3, suggests a new regulatory pathway. Mol. Cell. Biol. **13**:1983–1997.
25. **Perisic, O., H. Xiao, and J. T. Lis.** 1989. Stable binding of Drosophila heat shock factor to head-to-head and tail-to-tail repeats of a conserved 5 bp recognition unit. Cell **59**:797–806.
26. **Pollack, R., and R. Treisman.** 1990. A sensitive method for the determination of protein-DNA binding specificities. Nucleic Acids Res. **18**:6197–6204.
27. **Rabindran, S. K., G. Giorgi, J. Clos, and C. Wu.** 1991. Molecular cloning and expression of a human heat shock factor, HSF1. Proc. Natl. Acad. Sci. USA **88**:6906–6910.
28. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
29. **Sarge, K. D., S. P. Murphy, and R. I. Morimoto.** 1993. Activation of heat shock gene transcription by HSF1 involves oligomerization, acquisition of DNA binding activity, and nuclear localization and can occur in the absence of stress. Mol. Cell. Biol. **13**:1392–1407.
30. **Sarge, K. D., V. Zimarino, K. Holm, C. Wu, and R. I. Morimoto.** 1991. Cloning and characterization of two mouse heat shock factors with distinct inducible and constitutive DNA-binding ability. Genes Dev. **5**:1902–1911.
31. **Scharf, K.-D., S. Rose, W. Zott, F. Schoff, and L. Nover.** 1990. Three tomato genes code for heat stress transcription factors with a remarkable degree of homology to the DNA-binding domain of the yeast HSF. EMBO J. **9**:4495–4501.
32. **Schuetz, T. J., G. J. Gallo, L. Sheldon, P. Tempst, and R. E. Kingston.** 1991. Isolation of a cDNA for HSF2: evidence for two heat shock factor genes in humans. Proc. Natl. Acad. Sci. USA **88**:6910–6915.
33. **Shuey, D. J., and C. S. Parker.** 1986. Binding of Drosophila heat-shock transcription factor to the hsp 70 promoter: evidence for symmetric and dynamic interactions. J. Biol. Chem. **261**:7934–7940.
34. **Sistonen, L., K. D. Sarge, and R. Morimoto.** 1994. Human heat shock factors 1 and 2 are differentially activated and can synergistically induce hsp70 gene transcription. Mol. Cell. Biol. **14**:2087–2099.
35. **Sistonen, L., K. D. Sarge, B. Phillips, K. Abravaya, and R. Morimoto.** 1992. Activation of heat shock factor 2 during hemin-induced differentiation of human erythroleukemia cells. Mol. Cell. Biol. **12**:4104–4111.
36. **Sorger, P. K.** 1990. Yeast heat shock factor contains separable transient and sustained response transcriptional activators. Cell **62**:793–805.
37. **Sorger, P. K., and H. R. B. Pelham.** 1988. Yeast heat shock factor is an essential DNA-binding protein that exhibits temperature-dependent phosphorylation. Cell **54**:855–864.
38. **Taylor, I. C. A., J. L. Workman, T. J. Schuetz, and R. E. Kingston.** 1991. Facilitated binding of GAL4 and heat shock factor to nucleosomal templates: differential function of DNA-binding domains. Genes Dev. **5**:1285–1298.
39. **Westwood, J. T., J. Clos, and C. Wu.** 1991. Stress-induced oligomerization and chromosomal relocalization of heat-shock factor. Nature (London) **353**:822–827.
40. **Wu, C., S. Wilson, B. Walker, I. Dawid, T. Paisley, V. Zimarino, and H. Ueda.** 1987. Purification and properties of *Drosophila* heat shock activator protein. Science **238**:1247–1253.
41. **Xiao, H., and J. T. Lis.** 1988. Germline transformation used to define key features of the heat shock response element. Science **239**:1139–1142.
42. **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.