All heat shock transcription factors (HSFs) share two regions of homology, identified as the DNA binding and trimerization regions. The DNA binding region consists of two parts, an 89-amino-acid minimal DNA-binding domain and an additional 21 amino acids which are not necessary for specific DNA binding of a monomeric DNA-binding domain. These 21 amino acids may act as a flexible linker between the DNA-binding and trimerization domains. Saccharomyces cerevisiae HSF has an additional 32 amino acids between the proposed flexible linker and the trimerization domain. Deletion of this unique region has no effect on the structural integrity or essential in vivo functions of HSF. To investigate the role of the 21-amino-acid proposed linker, a series of internal deletions was created in fragments containing the DNA-binding and trimerization domains. The deletions have no effect on the structural integrity of the protein as assayed by circular dichroism spectroscopy. However, alterations of the linker do affect affinity of trimeric HSF binding to its target DNA. In addition, deletion of part or all of the proposed linker from full-length yeast HSF, an essential protein, disrupts growth of yeast.

Eukaryotic cells respond to stress through highly regulated expression of a small set of proteins known as heat shock proteins. All promoters for heat shock proteins contain a multiple number of inverted repeats of a 5-base-pair sequence, nGAAn (1, 2). Distribution of these binding sites in different species and different heat shock promoters vary from single arrays of three to eight inverted 5-base-pair units to several short arrays for a single heat shock promoter (1-4). Each array of nGAAn sequences is referred to as a heat shock element (HSE). Heat shock transcription factor (HSF) binds to the HSE and subsequently regulates the cellular response to stress (5, 6). Although only two nGAAn sequences are necessary for binding HSF in vitro, at least three inverted nGAAn units are necessary for binding in vivo (1, 2, 7). The alternating orientation of the five base pair nGAAn units positions each successive binding site on alternate sides of the DNA thus facilitating binding of the trimeric protein (7, 8).

HSF has been identified and cloned from many eukaryotic cells including those of human, mouse, chicken, fruit fly, tomato, and yeast (9-18). Although the sequences of the transcriptional activation regions of HSF are divergent, the DNA binding and trimerization regions are conserved among all identified HSF genes. All amino acids necessary and sufficient for trimerization are found between residues 333 and 424 of Saccharomyces cerevisiae HSF (19). The DNA binding region was originally defined as a 118-amino-acid fragment (residues 167-284) in S. cerevisiae (16). A 110-amino-acid fragment containing the DNA binding region, residues 171-280, is capable of sequence-specific DNA binding as a monomer, although the affinity of this region is significantly lower than that of a fragment consisting of both the DNA binding and trimerization regions (see "Results"). The NH2-terminal 89 amino acids of the DNA binding region are necessary and sufficient for the sequence-specific DNA binding, and these residues have been redefined as the minimal DNA-binding domain. In all HSF genes, except S. cerevisiae and tomato HSFs, the extra 21 amino acids, previously included in the DNA binding region, are located adjacent to the trimerization domain. In S. cerevisiae and tomato HSFs there are additional non-conserved sequences between these 21 conserved amino acids and the trimerization domain. S. cerevisiae HSF has 52 unique amino acids in this region and each of the three tomato HSFs has regions which vary in length and content (14-16).

In order to permit binding of a trimeric protein to DNA, it has been proposed that flexibility is needed between the DNA-binding and trimerization domains (7, 8). To investigate whether the 21 conserved amino acids or the 52 non-conserved amino acids are needed for flexibility, fragments of S. cerevisiae HSF were constructed with deletions in these regions. The resulting proteins were assayed for structural and functional perturbations. These results indicate that the 52 non-conserved amino acids, referred to as the Sc unique region, are unnecessary for proper structure or function of HSF under physiological conditions. The 21 conserved amino acids are not necessary for the structure of HSF but are important for the DNA binding activity of trimeric HSF. This short conserved region may act as a flexible linker between the DNA-binding and trimerization domains.

MATERIALS AND METHODS

Construction of HSF and HSE Derivatives—pHF103 was created by ligation of the EcoRI-KpnI of pBl5.04 (gift of P. Sorger), which encodes amino acids 1-259 of S. cerevisiae HSF, into pBluescript KS+ (Stratagene). pHS124 was constructed by oligo-mediated site-directed mu-
tagenesis to create an NdeI site, which changes amino acid 170 to a methionine. The mutation was confirmed by single-stranded DNA sequencing and restriction digestion. The NdeI-SphI fragment of pHN124 was then cloned into a derivative of the pET3-b expression vector to yield pHN136. The protein, Sc-D, when expressed from pHN136 containing 174 amino acids, RHA, at the carboxyl terminus of the protein.

A fragment containing the first 424 amino acids of S. cerevisiae HSF was inserted into pBluescript KS+(Stevanage) to create pHF158. pHN126 was constructed by oligo-mediated site-directed mutagenesis to create an NdeI site, which changes amino acid 170 to a methionine. The mutation was confirmed by single-stranded DNA sequencing, and restriction digestion. pHN423 was constructed by oligo-mediated site-directed mutagenesis to create a SphI site immediately after amino acid 280 of pHN126. The mutation was confirmed by single-stranded DNA sequencing and restriction digestion. The NdeI-SphI fragment of pHN423 was then cloned into a derivative of the pET3-b expression vector to make pHN420. The protein, Sc-D(21L), when expressed from pHN420 contains amino acids 171–280 plus five amino acids, AClIN, at the carboxyl terminus of the protein.

Oligo-mediated site-directed mutagenesis was used to make the deletions in the linker region of the plasmid pHN186 in order to create the following variations of this pBluescript KS+ derivative: pHN148, pHN401, and pHN402. The linker region of pHN126 was digested with BamHI and KpnI to yield pHN136. The amino acid sequence created by these deletions was determined by amino acid analysis of the protein, Sc-D, when expressed from pHN136 containing 174 amino acids, RHA, at the carboxyl terminus of the protein.

PhN136 was then digested with NdeI and SphI, and the fragment containing the first 424 amino acids of S. cerevisiae HSF was inserted into pBluescript KS+(Stratagene) to create pHF158. pHN126 was constructed by oligo-mediated site-directed mutagenesis to create an NdeI site, which changes amino acid 170 to a methionine. The mutation was confirmed by single-stranded DNA sequencing, and restriction digestion. The NdeI-SphI fragment of pHN126 was then cloned into a derivative of the pET3-b expression vector to yield pHN136. The same method was used to create pHN403, pHN404, and pHN412 from pHN148, pHN401, pHN402, and pHN411, respectively. All of the proteins expressed from plasmids constructed by this cloning method have an additional four amino acids, GLMN, at the carboxyl terminus of the protein.

The BarnHI site of pRS414 (Stratagene) was removed by Klenow fill-in and religation. The resulting plasmid was digested with EcoRI and ligated with the EcoRI fragment of pE3.2 (15) in order to construct pHN1002. Oligo-mediated site-directed mutagenesis was used to create deletions in the linker region of pHN126 to create the following variations of pHN126: pHN1015, pHN1016, pHN1017, and pHN1018, respectively.

The plasmid pHN153 was generated by cloning BarnHI fragments of the plasmid pHN136 in order to create the following variations of the pBluescript KS+ derivative: pHN148, pHN401, and pHN402. The same method was used to create pHN1002, pHN1003, pHN404, and pHN412 from pHN148, pHN401, pHN402, and pHN411, respectively. All of the proteins expressed from plasmids constructed by this cloning method have an additional four amino acids, GLMN, at the carboxyl terminus of the protein.

The BarnHI site of pRS414 (Stratagene) was removed by Klenow fill-in and religation. The resulting plasmid was digested with EcoRI and ligated with the EcoRI fragment of pE3.2 (15) in order to construct pHN1002. Oligo-mediated site-directed mutagenesis was used to create deletions in the linker region of pHN126 to create the following variations of pHN126: pHN1015, pHN1016, and pHN1017.

Circular Dichroism Spectroscopy—CD spectra were recorded on a Aviv 60 DS spectropolarimeter at 25 °C. Measurements were taken at 1-nm intervals with a 1-s time constant and 1.5 nm bandwidth, and were averaged over five scans. A path length of 2 mm was used with protein concentrations of approximately 1 mg/ml. The buffer conditions were 25 mM sodium phosphate, pH 8.3, and 100 mM NaCl. The mean residue ellipticity (θ(m)) of each protein was calculated using the following equation:

\[ \theta(m) = \theta(\text{obs}) \times \frac{M(W/d)(\text{c})}{\theta(\text{calc})} \]  

where \( \theta(\text{obs}) \) is observed ellipticity in millidegrees, MW is the molecular mass in daltons, d is pathlength in mm, c is concentration in mg/ml, and n is the number of amino acids (23). The mean residue ellipticity allows direct comparison of the CD spectra independent of the size of the protein. The molar ellipticity (θ(m)) was calculated using a variant of the former equation:

\[ \theta(m) = \theta(\text{calc}) \times \frac{M(W/d)(\text{c})}{\text{MW}} \]  

Cross-linking—Stock solutions of ethylene glycol bis-[succinimidylsuccinate] (EGS) were prepared fresh in dimethyl sulfoxide and added to a 20 μM solution of the protein in 25 mM sodium phosphate, pH 8.8, to give final EGS concentrations ranging from 0.5 to 200 μM. The reactions proceeded for 15 min at room temperature and were then quenched by addition of glycine to final concentration of 0.1 M. The products were analyzed by SDS-polyacrylamide gel electrophoresis. Bands were visualized with Coomassie Blue.

Gel Retardation Assay—DNA-protein complexes were formed in 50 mM HEPES, pH 6.0, 100 mM NaCl, 1 mM EDTA, 10% glycerol, 0.05 mg/ml poly(dI-dC), and 0.1 mg/ml bovine serum albumin for Sc-DT(21L), Sc-DT(10L), and Sc-DT(10L), and the Sc-DT(19L) or in 25 mM Tris, pH 6.7, 25 mM NaCl, 2.5 mM dithiothreitol, 5% glycerol, 1 mg/ml bovine serum albumin, and 25 μg/ml sheared calf thymus DNA for Sc-D and Sc-D(21L). The DNA fragment, TA-3, was radiolabeled with [32P]dCTP using the terminal transferase superfamily. The radiolabeled DNA was added on a 6% polyacrylamide gel and electrophoresed in Tris-borate buffer for 1.5 h at 300 V. The gels were dried and autoradiographed.
Footprinting—The fragment, TA-3, was radiolabeled at the XhoI with all four (α-32P)NTPs. DNA-protein complexes were formed in 25 mM Tris, pH 6.7, 100 mM NaCl, 5 mM MgCl₂, 5% glycerol, 1 mg/ml bovine serum albumin, 10 μg/ml sheared calf thymus DNA, and 2 mM dithiothreitol. The DNA-protein complexes were incubated with 3.5 pg/ml serum albumin, 10 pg/ml sheared calf thymus DNA, and 2 mM EDTA, and the DNA was precipitated with ethanol. The DNA pellets were resuspended in formamide and loaded onto an 8% denaturing polyacrylamide gel. The G/A ladder was prepared as described (24). After electrophoresis, the gel was dried and autoradiographed.

Yeast Strains—Plasmids encoding the mutant HSF proteins were transformed into a haploid strain derived from W303 (ade2-1 trpl can1-100 leu2-3-112 his3-11,15 ura3) (25), which carried the HIS3-LEU2 chromosomal disruption of HSF and a wild-type gene on the URA3-containing plasmid YCp50 (15). The yeast colonies were grown on galactose plates (1.7 mg/ml yeast nitrogen base, 2% glucose, and all amino acids except leucine, uracil, and tryptophan). Expression of wild-type HSF was suppressed by growing the colonies on glucose plates (1.7 mg/ml yeast nitrogen base, 2% glucose, and all amino acids). The yeast colonies were transferred into a haploid strain derived from W303 (ade2-1 trpl can1-100 leu2-3-112 his3-11,15 ura3) (25), which carried the HIS3-LEU2 chromosomal disruption of HSF and a wild-type gene on the URA3-containing plasmid YCp50 (15). The yeast colonies were grown on galactose plates (1.7 mg/ml yeast nitrogen base, 2% galactose, and all amino acids except leucine, uracil, and tryptophan). Expression of wild-type HSF was suppressed by growing the colonies on glucose plates (1.7 mg/ml yeast nitrogen base, 2% glucose, and all amino acids except leucine, uracil, and tryptophan).

RESULTS

The Minimal DNA-binding Domain Contains Only 89 Amino Acids—HSF contains two areas of homology, the trimerization domain and DNA binding regions. The conserved DNA binding region contains 110 amino acids (residues 171–280 in S. cerevisiae HSF). In order to explore the boundaries of this region, 21 amino acids were removed from the carboxyl-terminal end. This newly defined minimal DNA-binding domain was compared with a fragment containing all of the conserved amino acids of the DNA binding region in order to confirm that the minimal DNA-binding domain contains all of the sequences necessary for specific DNA binding.

The 59-amino-acid DNA-binding domain (Sc-D) and the 110-amino-acid DNA binding region (Sc-D(21L)) have nearly identical apparent DNA binding affinities. Sc-D and Sc-D(21L) were mixed with radioactively labeled DNA containing an HSE with three nGAAn boxes. The mixture was then electrophoresed through a non-denaturing polyacrylamide gel. As shown in Fig. 2, the presence of the additional 21 amino acids has no significant effect on the ability of an isolated DNA-binding domain to bind DNA. Although both Sc-D and Sc-D(21L) are capable of binding an HSE in the presence of competitor DNA, the gel mobility shift assays must be performed in very low (25 mM) NaCl concentrations and the protein-DNA complexes appear as smeared bands. DNase I footprinting was used to confirm that the DNA-binding domain is specifically binding the nGAAn units of the HSE. As seen in Fig. 3, the monomeric DNA-binding domain binds specifically to the three nGAAn units.

Further, the cleavage patterns of Sc-D and Sc-D(21L) are indistinguishable. Thus, the conserved 21 amino acids at the carboxyl-terminal of the DNA binding region have no effect on the ability of the isolated DNA-binding domain to interact with DNA.

Deletions in the Linker Region—In addition to the short (21-amino-acid) conserved region, S. cerevisiae HSF also contains a unique 52-amino-acid region between the minimal DNA-binding domain and the trimerization domain. To study the role of the 73-amino-acid region between the DNA-binding and trimerization domains of yeast heat shock transcription factor, a set of internal deletions in S. cerevisiae HSF were created (Fig. 1). The resulting series of proteins are named showing that they are derived from a fragment of the S. cerevisiae HSF gene containing the DNA-binding and trimerization domain. A series of deletions, with one end at the defined NH₂-terminal end of the trimerization domain, have sequentially removed the Sc unique region (Sc-DT(73L)), then part or all of the 21-amino-acid conserved sequence (Sc-DT(10L) and Sc-DT(0L), respectively). An S. cerevisiae HSF derivative (Sc-DT(19*L)), with
FIG. 3. The minimal DNA-binding domain specifically binds nGAAn units. Sc-D and Sc-D(21L) were bound to the fragment TA-3 and subjected to DNase I digestion as described under "Materials and Methods." GA ladder; no HSF, Sc-DT(73L) 5 × 10⁻³ M; Sc-D 2 × 10⁻⁴ M; Sc-D(21L) 2 × 10⁻⁴ M.

deletion of all but the 19 carboxyl-terminal amino acids of the unique sequence, was also assayed for activity and structural integrity.

Deletions in the Linker Region Do Not Disturb the Oligomerization Properties of the HSF Derivatives—The ability of the HSF derivatives containing various forms of the linker to form homotrimers was tested by chemical cross-linking with EGS. Other techniques, such as analytical ultracentrifugation, confirm cross-linking as a valid assay for trimerization activity in fragments of yeast HSF (8, 19). At similar concentrations of EGS, all of the proteins with internal deletions between the DNA-binding and trimerization domains, including Sc-DT(19L), produce a band with an apparent molecular weight three times that of the monomer on SDS-polyacrylamide gel electrophoresis (data not shown). Thus changes in the linker do not cause any detectable disruption of the function of the trimerization domain.

Deletions in the Proposed Linker Region Do Not Alter the Secondary Structure of the HSF Derivatives—To test the structural integrity of the mutants, their secondary structures were probed using CD spectroscopy. Fig. 4A shows that the proteins produce very similar CD spectra. Each HSF derivative has the minima at 208 and 222 nm characteristic of a mostly α-helical secondary structure. This agrees with the predominantly α-helical nature of the trimerization domain (19), in combination with the partial α-helical character of the DNA-binding domain (26). The highly similar CD spectra of the HSF derivatives imply that the overall secondary structure has not been strongly altered by any of the internal deletions. Thus, it is unlikely that the deletions within the proposed linker have caused substantial misfolding of the protein.

The Proposed Linker Region Has No Ordered Secondary Structure—To look directly at the secondary structure of the linker, the CD spectra of the DNA-binding domain (S. cerevisiae residues 171–159) and the trimerization domain (S. cerevisiae residues 333–424) were subtracted from each of the CD spectra of the HSF derivatives. As shown in Fig. 4B, it is apparent that the difference CD spectrum representing the linker region of Sc-DT(21L) is characteristic of random coil with a minimum at 230 nm and maximum at 220 nm (27). As expected, the remainder spectrum of the linker of Sc-DT(0L) is nearly flat, indicating that deletion of the entire linker is not disturbing the secondary structure of the DNA-binding or tri-
merization domain. When the CD spectrum of Sc-D is subtracted from the CD spectrum of Sc-D(21L), the difference spectrum is also characteristic of random coil. Therefore, the 21 conserved amino acids have no defined secondary structure either when free at the COOH-terminal end or when tethered to the trimerization domain.

Deletions in the Linker Region Decrease DNA Binding Affinity—Although the deletions do not appear to radically affect the structure of the DNA binding and trimerization region, HSF fragments with deletions in the conserved linker region are deficient in DNA binding activity as assayed by gel mobility shift. Several concentrations of each HSF derivative were mixed with a radioactively labeled DNA fragment containing three nGAAn boxes, then electrophoresed through a non-denaturing polyacrylamide gel and autoradiographed (Fig. 5). The apparent dissociation constants were defined as the concentration of protein required to bind 50% of the target DNA in an HSF trimer-DNA complex. Sc-D(73L) and Sc-D(10L) have apparent dissociation constants of approximately 2 \( \times 10^{-7} \) M, respectively. Both Sc-D(10L) and Sc-D(10L), which have deletions in the conserved linker region and a complete deletion of the Sc unique region, bind DNA with approximately 10-fold lower affinity than the wild-type protein. The extent of the deletion of the proposed linker does not affect the extent of binding, as the dissociation constants of both Sc-D(10L) and Sc-D(10L) are both approximately 2 \( \times 10^{-6} \) M. Finally, Sc-D(19L), with only a portion of the Sc unique region, binds less well than any of the other mutant proteins studied, including the mutant with no linker sequence at all (data not shown). The extent of secondary structure seen in the CD spectra (Fig. 4A), along with the ability of the HSF derivatives to trimerize, argue against the possibility that the loss of binding is due to misfolded or unfolded proteins.

All of the trimeric HSF derivatives bind DNA with higher affinity than the monomeric DNA-binding domain. The dissociation constants of Sc-D and Sc-D(21L) are 2 \( \times 10^{-4} \) M in 25 mM NaCl rather than 100 mM NaCl as was used in the reactions of the trimeric HSF fragments with DNA. The lower salt concentration allows stronger interaction between the protein and DNA and is necessary for the monomeric DNA-binding domain to bind DNA.

Deletions in the Conserved Linker Region Are Lethal in Yeast—Yeast HSF is required for viability and has constitutive transcriptional activity (15, 16). The internal deletions described above were duplicated in the full-length gene, and the mutant HSFs were expressed in S. cerevisiae from the native HSF promoter on a single-copy plasmid. The chromosomal HSF gene was disrupted by LEU2, and wild-type HSF was expressed solely from a 10GAL1 promoter on a plasmid containing a selectable marker for uracil production (15, 18). Yeasts containing both plasmids were grown on media with either galactose or glucose as the sole carbon source. On galactose media, the wild-type HSF was produced, and all cells grew. When placed on glucose media, however, only the mutant HSF will be expressed. Since HSF is essential, any mutant HSF which is unable to fully function will not be able to support growth on glucose media. Removal of the Sc unique region had no effect on the viability of the yeast cells at any of the temperatures examined. All proteins with partial or total removal of the conserved linker sequences were unable to support growth at any temperature assayed (see Table 1). Thus any deletion in the 21-amino-acid conserved linker region results in non-functional HSF.

DISCUSSION

All known HSFs have a highly conserved DNA binding region. Although the homologous region includes at least 110 amino acids, an 89-amino-acid minimal DNA-binding domain is sufficient for specific binding to HSFs. The additional 21 amino acids have no effect on the apparent DNA binding affinity nor on the cleavage pattern of the DNase I footprint of the DNA-binding domain. Further, difference CD spectra indicate that the last 21 amino acids of the conserved DNA binding region have no defined secondary structure in solution. Thus, the minimal DNA-binding domain consists of residues 171 to 259 of S. cerevisiae HSF.

The recent x-ray crystal structure of the DNA-binding domain of HSF from the milk yeast, Kluyveromyces lactis, shows that this domain is a variant of the helix-turn-helix family of DNA-binding proteins (26). The fragment which was crystallized is equivalent to residues 171–259 of S. cerevisiae HSF, defined in this paper as the minimal DNA-binding domain. The limits of the DNA-binding domain are also supported by NMR studies of a 130-amino-acid fragment of Drosophila melanogaster HSF including residues amino- and carboxyl-terminal to the DNA-binding domain (29). These results demonstrate no defined secondary structure either amino-terminal of amino acid 48, which is equivalent to amino acid 174 of S. cerevisiae HSF, or carboxyl-terminal of amino acid 332 of D. melanogaster HSF which is equivalent to amino acid 259 of S. cerevisiae HSF.

Finally, limited proteolysis of Sc-D(21L) with trypsin yields two major proteolytic fragments, amino acids 171–242 and amino acids 244–264, as determined by electrospray-ionization mass spectrometry. The first proteolytic product represents the amino-terminal portion of the protein up to a cleavage site in an extended loop, as seen in the crystal structure (26). The second fragment is created by cleavage in the extended loop and in the proposed linker. The remaining residues in the proposed linker were not recovered, supporting the conclusion that the 21 conserved amino acids are highly flexible and probably solvent exposed.

Unlike most HSFs, S. cerevisiae HSF contains 52 non-conserved amino acids between the proposed linker and the trimerization domain. Deletion of the Sc unique region has no effect on the structure or function of HSF. Analysis by CD spectroscopy shows that deletion of this region does not substantially change the overall secondary structure of a fragment including...
Flexible Linker of Yeast HSF

the DNA binding and trimerization regions. Sc-DT(21L) binds DNA with an apparent affinity similar to the wild-type fragment Sc-DT(73L) and can trimerize as well as the non-deleted fragment. Finally, deletion of the Sc unique region from full-length HSF does not alter the viability of yeast in which the mutant HSF has replaced the wild-type HSF. These results imply that the 52-amino-acid Sc unique region does not provide the flexibility that is required for a trimeric protein to bind linear DNA. Although the Sc unique region does not affect the binding of HSF to DNA under physiological conditions, it may play another, as yet unknown, role in transcriptional activation as a response to heat shock.

HSF binds to DNA as a trimer (8) and biochemical studies show that the trimerization domain is a triple-stranded α-helical coiled-coil (19). Such a trimeric coiled-coil would force the DNA-binding domains to be placed at 120° relative to one another. DNA, on the other hand, has approximately 180° symmetry between inverted binding sites. Therefore, some accommodation must be made by either the protein or the DNA in order to contact nGAAn units with each of the three DNA-binding domains within a trimer. Circular permutations assays show that the DNA is not significantly bent when bound by HSF, implying that the protein must make most of the adjustments (22, 30). Fig. 6 illustrates the juxtaposition of a trimeric structure onto a DNA molecule consisting of a series of inverted nGAAn sequences. The model depicts how a trimeric molecule can not fully occupy three consecutive nGAAn repeats. A flexible linker would allow the adjustment of an 3-fold symmetric HSF to one that can bind all three sites in the major groove, where HSF has been shown to bind (7, 30).

Many observations presented in this paper lead to the conclusion that the 21 conserved amino acids at the carboxyl-terminal end of the DNA-binding domain shares characteristics found in other defined linkers and must be the necessary flexible linker. First, linkers have been shown to have little defined secondary structure, allowing great flexibility in the movement of the connected domains. For example, circular dichroism and nuclear magnetic resonance spectroscopy on synthetic peptides of a linker from the dihydrolipoyl acetyltransferase component of the E. coli pyruvate dehydrogenase multienzyme complex shows that the linkers have little secondary structure (31). The CD difference spectra indicate that the 21 conserved amino acids of HSF's DNA binding region have no defined secondary structure when tethered by the trimerization domain or when free at the carboxyl-terminal end. Thus, the 21 conserved amino acids likely provide the flexibility necessary for a trimeric molecule to bind linear DNA.

Second, linkers provide more than simple flexibility. Although linkers have little or no ordered secondary structure,

| TABLE I |
| --- |

| Viability of HSF derivatives in vivo |
| --- |

| Glucose | Galactose | Glucose | Galactose | Glucose | Galactose | Glucose | Galactose |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 37 °C | 30 °C | 25 °C | 18 °C |
| Vector | + | ND | ND | ND | ND | ND | + |
| Sc HSF | + | + | + | + | + | + | + |
| Sc-(21L) | + | + | + | + | + | + | + |
| Sc-(10L) | ND | ND | ND | ND | ND | ND | ND |
| Sc-(6L) | ND | ND | ND | ND | ND | ND | ND |
| Sc-(19*L) | ND | ND | ND | ND | ND | ND | ND |

* ND, not determined.

Fig. 6. Two different views (at 90° relative to one another) of a model of HSF bound to DNA illustrate the symmetry problems of a trimer protein. Each ellipse represents a single DNA-binding domain with the arrows indicating the orientation of the binding site in the major groove. The three solid bold lines in each view depict the extension of 3-fold rotational symmetry from the base of the trimerization domains.
particular sequences may be necessary for proper function or activity. For example, mutations in the "hinge region" between the DNA-binding and dimerization domains of LexA negatively affect the ability of the protein to form stable complexes with DNA (32). For HSF, the DNA-binding and trimerization domains must be separated by the conserved amino acid linker for proper function. Although Sc-DT(21L) and Sc-DT(19*L) have linkers of similar length, the apparent DNA binding affinity of the fragment with the substituted linker (Sc-DT(19*L)) is significantly lower. The conserved linker region may have the ability to adopt a particular conformation or set of conformations which other sequences, such as those found in the Sc unique region, cannot.

Finally, linkers have also been implicated in regulating the orientation of the linked domains. Deletion of the linker in endoglucanase A changes the activity of both domains as well as the protease sensitivity of the mutant protein. Structural data support the hypothesis that these changes are due to altered orientation of the two domains caused by deletion of the linker (33). Further, linkers have been implicated in regulating the affinity of DNA-binding proteins for sites on the DNA. For example, deletions and additions in the linker regions of type IC restriction enzymes result in altered DNA sequence specificity (34). The linker of members of the C2 zinc cluster family of yeast transcriptional activators directs binding to DNA target sites with different spacings between the inverted binding sites (35). Deletions in the linker region of HSF affect the DNA binding affinity only when the trimerization domain constrains the rotational symmetry of the protein. Thus it is possible that the linker region of HSF assists the high affinity binding to DNA by allowing the proper orientation of the DNA-binding domains that are necessary for the juxtaposition of the trimeric protein with linear DNA.

Acknowledgments—We thank the following people for construction of plasmids: Pis Abola (pHN126, pHN124, pHN136, and pHN138); Eva Grotkopp (pHN148 and pHN420); Becky Drees (pBL3); and Susan Hubl (pHN1002). We also thank Susan Hubl for help with the yeast experiments, Prof. Susan Marqusee for help with circular dichroism, and Dr. David King for help with the electrospray-ionization mass spectrometry.

REFERENCES
1. Xiao, H., and Liu, J. T. (1988) Science 238, 1139–1142
2. Amin, J., Ananthan, J., and Voellmy, R. (1990) Mol. Cell. Biol. 8, 3761–3769
3. Biemes, M., and Pelham, H. R. (1987) Adv. Genet. 24, 31–72
4. Boorstein, W. R., and Craig, E. A. (1990) J. Biol. Chem. 265, 19912–19921
5. Morimoto, R. I., Sarge, K. D., and Abravaya, K. (1992) J. Biol. Chem. 267, 2187–21990
6. Liu, J., and W., C. (1993) Cell 74, 1–4
7. Periasir, O., Xiao, H., and Liu, J. T. (1989) Cell 58, 79–806
8. Sorger, P. K., and Nelson, H. C. M. (1992) Cell 70, 807–813
9. Rahindran, S. K., Giorgi, C., Clo, J., and Wu, C. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 6911–6915
10. Schott, T. J., Gallo, G. J., Sheldon, L., Tempel, P., and Kington, R. E. (1991) Proc. Natl. Acad. Sci. U. S. A. 88, 6911–6915
11. Sarge, K. D., Zimario, V., Holm, K., Wu, C., and Morimoto, R. I. (1991) Genes & Dev. 6, 1902–1911
12. Nakai, A., and Morimoto, R. B. (1993) Mol. Cell. Biol. 13, 1983–1997
13. Clo, J., Westwood, J. T., Becker, P. B., Wilson, S., Lambert, K., and Wu, C. (1993) Cell 73, 1085–1097
14. Scharf, K. D., Rose, S., Zett, W., Scho, F., Never, L., and Scho, F. (1990) EMBO J. 9, 4409–4501
15. Sorger, P. K., and Pelham, H. R. (1988) Cell 54, 555–564
16. Wiederrecht, R., Seto, D., and Parker, C. S. (1988) Cell 54, 841–853
17. Jakobsen, B. K., and Pelham, H. R. (1991) EMBO J. 10, 389–397
18. Gallo, G. J., Prentice, H., and Kington, R. E. (1993) Mol. Cell. Biol. 13, 749–761
19. Peteranderl, R., and Nelson, H. C. M. (1992) Biochemistry 31, 12272–12276
20. Studier, F. W., Rosenberg, A. H., Dunn, D. J., and Dubendorff, J. W. (1989) Methods Enzymol. 185, 60–89
21. Stark, M. J. (1987) Gene (Amst.) 51, 255–267
22. Rye, H. S., Drees, B. L., Nelson, H. C., and Glazer, A. N. (1993) J. Biol. Chem. 268, 25229–25234
23. Canter, C. R., and Schimmel, P. R. (1980) Biophysical Chemistry, pp. 409–480, W. H. Freeman and Company, San Francisco, CA
24. Maxam, A. M., and Gilbert, W. (1980) Methods Enzymol. 58, 459–560
25. Shore, D., and Nasmuth, K. (1987) Cell 51, 721–732
26. Harrison, C. J., Bohm, A. A., and Nelson, H. C. M. (1994) Science 263, 224–227
27. Saxena, V. P., and Waelafer, D. E. (1971) Proc. Natl. Acad. Sci. U. S. A. 68, 969–972
28. Sorger, P. K. (1980) Cell 22, 785–805
29. Vuister, G. W., Kim, S. J., Wu, C., and Bax, A. (1994) Biochemistry 33, 10–16
30. Gross, D. S., English, K. E., Collina, K. W., and Lee, S. W. (1990) J. Mol. Biol. 216, 611–631
31. Radem, F. S., and A, D., Parham, R. N., and Appella, E. (1988) J. Biol. Chem. 264, 767–775
32. Oertel-Buchheit, P., Schmidt-Dor, T., Granger-Schnarr, M., and Schnarr, M. (1993) J. Mol. Biol. 229, 1–7
33. Shen, H., Schmuck, M., Pilz, I., Gilkes, N. R., Kulburn, D. G., Miller, R. J., and Warren, R. A. (1991) J. Biol. Chem. 266, 11335–11340
34. Gubler, M., and Bickle, T. A. (1991) EMBO J. 10, 951–957
35. Reece, R. J., and Ptasznie, M. (1993) Science 261, 909–911