Structure of human heat-shock transcription factor 1 in complex with DNA

Tobias Neudegger, Jacob Vergheze, Manajit Hayer-Hartl, F. Ulrich Hartl & Andreas Bracher

Heat-shock transcription factor 1 (HSF1) has a central role in mediating the protective response to protein conformational stresses in eukaryotes. HSF1 consists of an N-terminal DNA-binding domain (DBD), a coiled-coil oligomerization domain, a regulatory domain and a transactivation domain. Upon stress, HSF1 trimerizes via its coiled-coil domain and binds to the promoters of heat shock protein–encoding genes. Here, we present cocrystal structures of the human HSF1 DBD in complex with cognate DNA. A comparative analysis of the HSF1 paralog Skn7 from Chaetomium thermophilum showed that single amino acid changes in the DBD can switch DNA binding specificity, thus revealing the structural basis for the interaction of HSF1 with cognate DNA. We used a crystal structure of the coiled-coil domain of C. thermophilum Skn7 to develop a model of the active human HSF1 trimer in which HSF1 embraces the heat-shock-element DNA.
motifs—a tail-to-tail HSE (ggTTTcaGAAcc, in which capital letters denote the HSF1-recognition motifs) or a head-to-tail SatIII repeat (cgGATAgGAAAg)38—and solved the structures of the corecysts by molecular replacement at 2.9-Å and 2.55-Å resolution, respectively (Table 1 and Fig. 1b,c). Full-length HsHSF1 failed to crystallize, presumably because of high structural dynamics in the RegD and TAD.

The asymmetric unit of the HsDBD–HSE complex contained one DBD chain and one DNA molecule. Crystal symmetry created the second DBD chain bound to the inverted GAA repeat (Fig. 1b). The asymmetric unit of the HsDBD–SatIII complex contained two independent copies of the DBD, each contacting a GAA motif (Fig. 1c). The three crystallographically independent copies of the DBD in the two complexes were essentially identical in structure (Cα r.m.s. deviation (r.m.s.d.) of 0.570–0.658 Å) (Supplementary Fig. 1a) and closely resembled the previously reported NMR structure of the DBD without bound DNA14 (r.m.s.d. of 1.488–1.952 Å) (Supplementary Fig. 1b) and the structure of KIDBD bound to HSE15 (r.m.s.d. of 1.223 Å) (Supplementary Fig. 1c). Notably, in both HsDBD–DNA complexes, the C-terminal residues 101–120 of the DBD, including a short α10 helix (helix 32, residues 109–114) were structured (Supplementary Fig. 1d). These residues were absent in the DBD structure but have been shown to be essential for function in yeast HSF1 (ref. 43). The conformation of the C-terminal residues of HsDBD was stabilized by hydrophobic interactions of Phe104, Leu112 and Ile115 with the DNA double helix (helix 32, residues 109–114) were structured (Supplementary Fig. 1d). These residues were absent in the DBD structure but have been shown to be essential for function in yeast HSF1 (ref. 43).

Table 1 Data collection and refinement statistics

<table>
<thead>
<tr>
<th>Data collection</th>
<th>HsDBD–HSE</th>
<th>HsDBD–SatIII</th>
<th>CIDD–HSE</th>
<th>CIDD–SSRE</th>
<th>S-SAD</th>
<th>CIskn7(160–209)–I</th>
<th>CIskn7(160–209)–II</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Space group</strong></td>
<td>P22_1</td>
<td>P22_1</td>
<td>P22_2</td>
<td>P3</td>
<td>P22_2</td>
<td>P22_1</td>
<td>P22_2</td>
</tr>
<tr>
<td><strong>Cell dimensions</strong></td>
<td>a, b, c (Å)</td>
<td>90, 90, 90</td>
<td>90, 92.56, 90</td>
<td>90, 90, 90</td>
<td>90, 92.78, 90</td>
<td>90, 90, 90</td>
<td>90, 90, 90</td>
</tr>
<tr>
<td><strong>Wavelength (Å)</strong></td>
<td>1.70074</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Resolution (Å)</strong></td>
<td>47.19–2.91</td>
<td>47.19–2.91</td>
<td>47.53–1.7</td>
<td>47.19–2.91</td>
<td>47.53–1.7</td>
<td>47.19–2.91</td>
<td>47.53–1.7</td>
</tr>
<tr>
<td><strong>Rmerge</strong></td>
<td>2.91(3.09–2.91)</td>
<td>2.91(3.09–2.91)</td>
<td>2.91(3.09–2.91)</td>
<td>2.91(3.09–2.91)</td>
<td>2.91(3.09–2.91)</td>
<td>2.91(3.09–2.91)</td>
<td>2.91(3.09–2.91)</td>
</tr>
<tr>
<td><strong>Completeness (%)</strong></td>
<td>94.38</td>
<td>94.38</td>
<td>94.38</td>
<td>94.38</td>
<td>94.38</td>
<td>94.38</td>
<td>94.38</td>
</tr>
<tr>
<td><strong>Redundancy</strong></td>
<td>90.3</td>
<td>90.3</td>
<td>90.3</td>
<td>90.3</td>
<td>90.3</td>
<td>90.3</td>
<td>90.3</td>
</tr>
<tr>
<td><strong>Refinement</strong></td>
<td>1.70074</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Resolution (Å)</strong></td>
<td>30.29</td>
<td>30.29</td>
<td>30.29</td>
<td>30.29</td>
<td>30.29</td>
<td>30.29</td>
<td>30.29</td>
</tr>
<tr>
<td><strong>No. reflections</strong></td>
<td>3,137</td>
<td>3,137</td>
<td>3,137</td>
<td>3,137</td>
<td>3,137</td>
<td>3,137</td>
<td>3,137</td>
</tr>
<tr>
<td><strong>Rwork / Rfree</strong></td>
<td>0.223 / 0.267</td>
<td>0.223 / 0.267</td>
<td>0.223 / 0.267</td>
<td>0.223 / 0.267</td>
<td>0.223 / 0.267</td>
<td>0.223 / 0.267</td>
<td>0.223 / 0.267</td>
</tr>
<tr>
<td><strong>No. atoms</strong></td>
<td>810</td>
<td>810</td>
<td>810</td>
<td>810</td>
<td>810</td>
<td>810</td>
<td>810</td>
</tr>
<tr>
<td><strong>B factors</strong></td>
<td>1.669</td>
<td>1.669</td>
<td>1.669</td>
<td>1.669</td>
<td>1.669</td>
<td>1.669</td>
<td>1.669</td>
</tr>
<tr>
<td><strong>No. water molecules</strong></td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td><strong>r.m.s. deviations</strong></td>
<td>0.005</td>
<td>0.005</td>
<td>0.005</td>
<td>0.005</td>
<td>0.005</td>
<td>0.005</td>
<td>0.005</td>
</tr>
</tbody>
</table>

Table 2 Data collection and refinement statistics

© 2016 Nature America, Inc. All rights reserved.

ARTICLES
Figure 1 Interaction of HsHSF1 with cognate DNA. (a) Schematic of the domain structure of HsHSF1. The DNA-binding domain (DBD; blue), coiled-coil regions with hydrophobic heptad repeats (HR; red) and the transactivation domain (TAD; yellow) are indicated. Disordered segments are shown in white. RegD, regulatory domain. (b,c) Crystal structures of the DBD of HsHSF1 (HsDBD) with a tail-to-tail HSE (b) or SatIII repeats (c). Protein is shown in ribbon representation; the copies in the HsDBD–HSE complex are symmetry related. DNA is shown in stick representation with the phosphate backbone and the GAA motifs highlighted. The DNA sequence is shown below the models. Chain termini are indicated. (d) Specific interactions of HsDBD with the GAA motif. The area boxed in c is shown. Side chains of critical residues are shown in stick representation. Hydrogen bonds are represented by dotted lines. The C-terminal tail of the DBD is omitted for clarity. (e) Polar DNA-backbone interactions of HsDBD, viewed as in d.

contacted the phosphate backbone flanking the GAA binding motif (Fig. 1e). Hydrogen-bond contacts from the peptide-backbone amides of Lys62 and Lys80 to phosphate groups rigidified this interaction. The conserved residues Arg117 and Lys118 at the C-terminal end of the HsDBD formed additional contacts across the major groove; the conserved Lys116 was solvent exposed (Fig. 1e). This cluster of basic residues is strongly conserved in metazoan but not fungal HSF1 sequences. Importantly, the accessible lysine residues Lys80, Lys116 and Lys118 are subject to modification by acetylation36,37. Acetylation removes the positive charge from lysine residues, which is required for ionic interactions with the DNA backbone, and thus lowers overall affinity to DNA, as supported by functional binding data36,37. The cysteine residues Cys36 and Cys103, which have been implicated in the response to oxidative stress45,46, were distant from each other in space and pointed toward the solvent, and thus in principle were accessible for intermolecular disulphide-bond formation.

Role of DBD–DBD interactions in DNA binding

Adjacent DBD molecules bound to tail-to-tail HSE repeats were oriented for a symmetrical protein-protein interaction involving helix α2 (Fig. 2a). The closest intermolecular contact, corresponding to a distance of 4.95 Å, occurred between the Gly50 residues located at the N-terminal ends of the α2 helices. Gly50 is conserved in HSF1 sequences from chordata. The residues Gln49 and Gln51, surrounding Gly50, are polar but uncharged and may engage in van der Waals contacts. Peripheral electrostatic interactions of Lys54 from one DBD molecule with Asp48 or Asp95 from the other supported these contacts (Fig. 2b). The structure of KnDBD in complex with HSE exhibited distinct intermolecular interactions involving β-sheet contacts between the wing loop of one DBD molecule and the α2-α3 linker of the other15 (Supplementary Fig. 2a,b). These contacts were absent in the HsDBD–HSE complex. Synergistic effects of DBD–DNA and DBD–DBD interactions are consistent with the strong preference of tail-to-tail HSE motifs over head-to-head and head-to-tail motifs47.

The contacts between DBDs along the stacked DNA motifs in the lattice of HsDBD–SatIII crystals may serve as a model for head-to-tail interactions within HSF1 oligomers bound to HSE motifs with more than two inverted NGAAN repeats. Although adjacent DBDs did not interact, we observed contacts between every other DBD. These contacts were exclusively mediated by the extended wing loop. In the HsDBD–SatIII crystal lattice, the wing loop in one DBD molecule exhibited an extended conformation, reaching out along the DNA to a symmetry-related DBD molecule (Supplementary Fig. 2c). However, crystal-packing forces turned the stacked DNA double helices out of register in the crystal lattice, and thus the interface cannot fully reflect the natural situation. Along a continuous double helix, the hydrophobic tip of the extended wing loop (sequence QGLGV) of one DBD would be likely to extend further, reaching a conspicuous hydrophobic groove between helices α1 and α2 on the adjacent DBD (Supplementary Fig. 2d,e). The interaction might be subject to regulation by acetylation or SUMOylation at wing-loop residue Lys91 (refs. 35,37).

Characterization of a thermostable HSF protein

Structural and functional studies of HsHSF1 have been hampered by the inherent structural dynamics of the protein and its tendency to form various oligomeric states in vitro. To circumvent these problems, we sought to structurally characterize a presumably less dynamic HSF homolog from the thermophilic fungus C. thermophilum. The genome of this organism contains two HSF homologs, CTHT_0005280 and CTHT_0048700, representing orthologs of Saccharomyces cerevisiae HSF1 and SKN7, respectively48,49. Skn7 regulates the response to cell-wall and oxidative stress50. Whereas the DBD and coiled-coil regions of HSF1 and Skn7 are highly similar, the regulatory segment

Figure 2 Symmetrical interactions between the HsDBDs in the complex with the tail-to-tail HSE. (a) Overview along the two-fold symmetry axis. (b) Zoomed-in view of the area boxed in a, showing molecular contacts at the interface. Key residues are shown in stick representation, and DNA is shown in surface representation.
of Skn7 contains a ‘receiver’ domain that is phosphorylated upon cell-wall stress\(^{11}\) (Fig. 3a). Skn7 interacts not only with HSEs but also with DNA motifs (Sln1 star response elements (SSREs)) that are responsive to activation of the osmosensor Sln1 (refs. 52,53).

CtSkn7 exhibited higher sequence homology to HsHSF1 than did ChHSF1 (Supplementary Fig. 3). Specifically, the linker between the DBD and HR-A/B coiled-coil domain is extended by 31 residues in CtHSF1. We therefore analyzed CtSkn7 and recombinantly expressed a nearly full-length CtSkn7 (CtSkn7\(^{ΔQA}\); residues 35–765, excluding residues 1–34 and 653–684, which contain polyglutamine and polyalanine low-complexity sequences, respectively) (Fig. 3a). We also analyzed a truncated form comprising the DBD, linker and HR-A/B regions (residues 40–220; conserved in HSF1 homologs) and the DBD alone (residues 40–143, CtDBD). CtSkn7\(^{ΔQA}\) and CtSkn7\(^{ΔQA}\) bound to SSRE or tail-to-tail HSE DNA duplexes with similar efficiencies, as monitored by electrophoretic mobility shift assays (EMSA) (Fig. 3b).

The isolated DBD did not bind detectably to DNA under these conditions \(\text{in vitro}\), thus suggesting that oligomer formation via the HR-A/B domain was necessary to achieve sufficient avidity. Indeed, CtDBD behaved as a monomer in solution (~13.7 kDa; theoretical mass 12.6 kDa),

![Figure 3](image)

**Figure 3** Characterization of the HSF homolog Skn7 of *C. thermophilum*. (a) Comparison of the domain structures of HsHSF1 and CtSkn7. The additional Receiver domain of CtSkn7 is indicated in purple. The construct CtSkn7\(^{ΔQA}\) is also shown. (b) EMSA analysis of CtSkn7 DNA binding activity, showing a band shift of fluorescently labeled (Fl) HSE and SSRE, but not of a noncognate control DNA, upon binding of CtSkn7\(^{(40–220)}\) and CtSkn7\(^{ΔQA}\) (DNA sequences in Supplementary Table 1). A representative gel of 3 independent experiments is shown. Ctrl, control. (c) SEC-MALS analysis of CtDBD and CtSkn7\(^{(40–220)}\). The calculated molar masses are indicated. RI, refractive index.

![Figure 4](image)

**Figure 4** Interaction of CtSkn7 with DNA. (a,b) Crystal structures of CtDBD with a tail-to-tail HSE (a) or SSRE (b). Protein is shown in ribbon representation; the CtDBD–HSE complex has perfect C\(_2\) symmetry. DNA is shown in stick representation with the phosphate backbone and the GAA, GCC and GCC motifs highlighted. The DNA sequence is shown below the models. Chain termini are indicated. (c) Specific interactions of CtDBD with the GCC motif. The area boxed in b is shown. Side chains of highly conserved residues are shown in stick representation. Hydrogen bonds are represented by dotted lines. The C-terminal tail of the DBD is omitted for clarity. (d) Polar DNA-backbone interactions of CtDBD. (e) Reversal of DNA binding selectivity of CtSkn7\(^{(40–220)}\) by the K100M mutation, as shown by EMSA analysis. A representative gel of 3 independent experiments is shown. WT, wild type; ctrl, control.
as determined by size-exclusion chromatography coupled to multangle light scattering (SEC-MALS). In contrast, the mass of CtSnk7(40–220) (theoretical mass ~21.3 kDa) was determined as ~62.7 kDa, a value indicative of a constitutive trimer in solution (Fig. 3c).

Switching DNA binding specificities between HSF homologs

The high protein concentrations used for crystallization allowed us to obtain crystal structures of the CtDBD in complex with the tail-to-tail tandem HSE (gtTTCtaGAAcc) and the asymmetric SSRE motif (GGCCCCAGCCAAAT) (Table 1). The complex with the HSE motif closely resembled the HsDBD–HSE structure (r.m.s.d. of 0.798 Å), a result consistent with the high degree of sequence conservation between the DBDs (Fig. 4a and Supplementary Fig. 4a). Surprisingly, the complex of the CtDBD with the asymmetric SSRE motif exhibited the same topology as that of the HsDBD–HSE structure (r.m.s.d. of 0.684 Å) (Supplementary Fig. 4b), although the SSRE contained linear repeats of GCC motifs (Fig. 4b). This was explained by one CtDBD molecule interacting in reverse orientation with the first GCC motif, i.e., GCC in regular 5′–3′ notation, and the other interacting in forward orientation with the second GCC triplet. The interactions with the DNA were largely conserved (Supplementary Fig. 4g), with the notable exception of Lys100 in CtDBD, which is conserved in Skn7 sequences and corresponds to Met75 in helix α3 of the HsDBD (Supplementary Fig. 4c). In the CtDBD–SSRE complex, the Lys100 residues were engaged in hydrogen-bond interactions with guanine carbonyl groups, apparently compensating for less favorable van der Waals contacts of the SSRE DNA with helix α3 (Fig. 4c and Supplementary Fig. 4d–f). In contrast, in the CtDBD–HSE complex, Lys100 acted as a steric support for Arg96, performing a role similar to that of Met75 in the HsDBD (Supplementary Fig. 4g). The high side chain plasticity of lysine thus relaxed the DNA sequence specificity of Skn7 compared to that of HSF1, thereby allowing recognition of both GCC and GGC motifs in addition to GAA. Consistently with this result, the substitution K100M abolished the interaction of CtSnk7(40–220) with SSRE but preserved binding to HSE motifs, switching the DNA binding specificity of the mutant to that of HSF1 (Fig. 4e). Thus, a single amino acid change can markedly alter the DNA binding specificity of HSF homologs.

Structural basis for HSF oligomerization

Oligomer formation, a critical step in the activation of HSF proteins, is mediated by the HR-A/B coiled-coil region, which is connected to the C terminus of the DBD by a linker segment of 8–64 residues (refs. 24,54). The HR-A/B coiled-coil domains of CtSnk7 and HsHSF1 are remarkably similar (Supplementary Fig. 3). The constitutive formation of trimeric complexes by Skn7 (Fig. 3c) thus offered the opportunity to obtain insight into the structural mechanism of HSF1 oligomer formation. In CtSnk7(40–220), the 13-residue linker between the DBD and the HR-A/B was sensitive to Proteinase K, both in the presence and absence of cognate DNA, thus suggesting that it is flexible (Supplementary Fig. 5a,b). We obtained crystal structures of the

Figure 5 Oligomerization of CtSnk7 via the coiled-coil domain. (a) Crystal structure of a CtSnk7 (160–209) homotrimer. Two perpendicular views are shown. Side chains of hydrophobic-layer residues are shown in stick representation. The position of the two Cys185 disulfide bonds is indicated by an arrow. Cross-section of the coiled coil at the position of Cys185. The two-fold crystallographic-symmetry axis is indicated. (b) Crystal structure of a CtSnk7 (160–209) homotrimer. (c) Schematic of the domain structures of HsHSF1 and the HsHSF1–CtSnk7 chimera, HsHSF1(CHR-A/B), (d) Induction of firefly luciferase (Fluc) expression under the Hsp70 promoter relative to constitutive expression of Renilla luciferase under the CMV promoter in HeLa cells transferred with Venus (V) fusion proteins HsHSF1-V, HsHSF1(CHR-A/B)-V or Venus alone. Luciferase activities were monitored with and without heat shock (+HS). Induction levels were normalized to Venus alone without heat shock. Error bars, s.d. (n = 3 independent experiments).
CtSkn7 coiled–coil construct CtSkn7 (160–209) (Fig. 5a,b). We solved the structure of a crystal of space group C222₁ by sulfur single-wavelength anomalous diffraction (S-SAD) at 1.95-Å resolution and refined the final model against native data at 1.03-Å resolution (Table 1). The asymmetric unit of this crystal contained two protomers forming one half of a crystallographic two-fold-symmetric tetramer (Fig. 5a). The α-helical protomers in the asymmetric unit aligned in parallel, engaging in typical coiled-coil interactions, and were linked by a disulfide bridge involving Cys185 in each chain (Fig. 5a). A second crystal form of CtSkn7(160–209) in space group P2₁2₁2₁ contained a trimeric coiled coil in addition to a disulfide-bonded half-tetramer, which was similar to the half-tetramer observed in space group C222₁ (r.m.s.d. of 0.753 Å) (Table 1 and Fig. 5b). This result was consistent with our observation that CtSkn7(160–209) (~14.5 kDa; theoretical mass ~5.7 kDa) behaved as a mixture of dimers and trimers in solution (Supplementary Fig. 5c), in contrast to CtSkn7(40–220), which behaved as a trimer (Fig. 3c). The protomers in the trimer showed greater variation in backbone curvature (r.m.s.d. of 1.386–1.499 Å) than the protomers in the dimer units (r.m.s.d. of 0.475–0.979 Å). Of note, Cys185 is not conserved in the Skn7 proteins from closely related fungi such as Chactomium globosum and Neurospora crassa, thus suggesting that the observed disulfide-bond formation was a crystallographic artifact.

A functional HSF1-Skn7 chimera

To test whether the HR-A/B coiled-coil domain of CtSkn7 could functionally substitute for the HR-A/B domain of HsHSF1, we generated a chimeric construct in which residues 133–193 of HsHSF1 were replaced with residues 160–220 of CtSkn7 (HsHSF1(CHR-A/B); Fig. 5c). SEC-MALS indicated that HsHSF1(CHR-A/B) formed a constitutive trimer in vitro (Supplementary Fig. 5d). For functional analysis, we used HsHSF1(CHR-A/B)-C-terminally tagged with Venus fluorescent protein (V) and used HsHSF1-V as a control. Upon expression in HeLa cells under the control of the strong CMV promoter, both HsHSF1(CHR-A/B)-V and HsHSF1-V predominantly localized in the nucleus (Supplementary Fig. 5e,f). Using a reporter plasmid containing the firefly luciferase (Fluc) gene under control of the Hsp70 promoter, we found that expression of HsHSF1(CHR-A/B)-V resulted in an approximately six-fold induction of Fluc in the absence of heat stress. This constitutive activity suggested that the HR-C element of HsHSF1 in the chimera did not inhibit trimer formation of the CHR-A/B coiled-coil24–26 (Fig. 5c). We observed a similar induction only upon heat stress in control cells expressing Venus protein alone, thus reflecting activity of endogenous HSF1, or in cells expressing HsHSF1-V (Fig. 5d). Cells expressing the constitutively active chimera showed only a moderate additional Fluc induction upon heat stress, thus suggesting that HsHSF1(CHR-A/B)-V effectively occupied endogenous HSEs. Interestingly, binding of HsHSF1(CHR-A/B)-V to SatIII repeats, as indicated by the formation of nuclear stress granules34, was not constitutive but was observed only upon heat stress (Supplementary Fig. 5e,f). Together, these results indicate that the trimeric coiled-coil structure of Skn7 is functional in the context of HsHSF1.

DISCUSSION

The crystal structures solved in this study allowed us to assemble a tentative model for the trimer of HsHSF1(13–182) bound to a canonical HSE motif composed of three inverted GAA repeats (Fig. 6). We used the CHR-A/B coiled-coil trimer to homology-model the HR-A/B of HsHSF1 and found that the three DBDs are bound to one face of a rather straight DNA duplex, and the coiled-coil helical bundle is positioned on the opposite face. As a result, the DBDs and the linker regions connecting to the helical bundle embrace the DNA strand. This arrangement is supported by our finding that the C-terminal ends of the DBDs point toward the DNA, in contrast to an earlier proposal suggesting that DBDs and the coiled coil are positioned on the same side of the DNA strand55. The contacts between the DBDs would synergize with the avidity effect by coiled-coil trimerization, thus rendering the HSF1–DNA complex less dynamic than expected from the interactions of a single DBD with DNA. The interdomain linker of 13 residues is long enough to allow for irregularities in the distance between the GAA repeats that are arranged head to head. The linker is probably also of sufficient length to span pairs of SatIII repeats (Supplementary Fig. 6). Here, the DBDs would point alternately in opposite directions along these 33.8-Å-spaced pairs of head-to-tail HSE repeats, and only every second subunit would contribute to the same coiled-coil trimmer. Two of the linkers would have to adopt fully extended conformations unless relaxed by bending of the DNA or partial unzipping of the coiled coils.

Only the two DBDs bound to 3′ tail-to-tail HSE motifs would be in tight contact with each other, whereas the DBD bound to the 5′ GAA repeat would be flexibly linked to the 3′ DBD via its extended wing loop (Fig. 6). This 5′ DBD would form the most dynamic contacts to the DNA and thus may be preferentially accessible for acetylation at residues Lys80, Lys116 or Lys118. Acetylation and hence dissociation from DNA might therefore occur in a stepwise manner starting with the DBD linked via the wing loop.

HsHSF1 activation and inhibition are possible strategies for the treatment of protein-misfolding diseases and cancer, respectively56–58. However, the surface of the DBD does not exhibit conspicuous deep pockets as potential docking sites for small-molecule effectors. A rather shallow groove is located between helices α1 and α2, the hypothetical interaction site for the extended wing loop in the head-to-tail DBD–DBD interface. A compound directed to this site might thus reduce HSF1 activity. Structural information on the regulatory and transactivation domains will be important in developing drugs that can modulate HSF1 function.

The HSF1 paralog Skn7 has a broader DNA sequence specificity, recognizing both inverted-repeat HSE and SSRE motifs, enabled by the systematic substitution of Met75 in the DNA-recognition helix α3 of HsHSF1 by lysine (Lys100 in CtSkn7). Thus, a single amino acid substitution is sufficient for this expansion in DNA selectivity. Similarly, the substitution of Gln50 to lysine in the recognition helix of homeodomain transcription factors switches DNA target binding specificity between paralogs59,60. As in the case of HSFs, the change in specificity involves hydrogen-bond interactions between lysine and the C-6 carbonyl group of guanine.

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. Coordinates and structure factors have been deposited in the Protein Data Bank under accession codes PDB 5DSU (HsDBD–HSE complex), PDB 5DV (HsDBD–SatIII complex), PDB 5DW (CrDBD–HSE complex), PDB 5DX (CrDBD–SSRE complex), PDB 5DY (CtSkn7(160–209) crystal form I) and PDB 5DSZ (CtSkn7(160–209) crystal form II).

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank R. Körner for MS analysis, R. Lange and A. Jungclaus for assistance with protein purification, and the staff at the Core and Crystalization Facilities at the Max Planck Institute of Biochemistry and at the European Synchrotron Radiation Facility (ESRF) in Grenoble, France, for their excellent services. J.V. is supported by a Rudolf Haas Fellowship from the Jung Foundation for Science and Research.
AUTHOR CONTRIBUTIONS. T.N. performed the biochemical and functional analysis and obtained the crystals. A.B. and T.N. solved the crystal structures. J.V. performed the cell culture experiments, and M.H.-H. collected SEC-MALS measurements. A.B. and E.U.H. supervised the experimental design and data interpretation. All authors contributed to experimental design, data analyses and manuscript preparation.

COMPETING FINANCIAL INTERESTS. The authors declare no competing financial interests.

Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.


Molecular cloning. The HsHSF1 DBD construct, residues 1–120 (HsDBD), was PCR-amplified from pCMV-HsHSF1-Flag (plasmid 1932, Addgene) and cloned into pProEx-HtB (Life Technologies) between the EcoRI and HindIII restriction sites for expression as a TEV protease-cleavable, N-terminally His6-tagged fusion protein in E. coli. As a consequence of the cloning strategy, HsDBD contained the additional N-terminal sequence GAMGSSLRLGG after TEV protease treatment.

The expression constructs CSkn7(40–143) (CSDK), CSkn7(40–220) and CSkn7(160–209) were PCR-amplified from synthetic DNA and cloned into the pHEU plasmid for expression as N-terminally His6-ubiquitin (His6-Ub)-tagged fusion proteins by using SacII and HindIII restriction sites. Point mutants of CSkn7 and HsHSF1 were produced by PCR-based mutagenesis.

Constructs HSPA1A-Fluc and CMV-Rluc were as previously described. HsHSF1-Venus was constructed by subcloning HsHSF1 from pCMV-HsHSF1-Flag and Venus from mVen1 N1 (plasmid 27793, Addgene) into pRS413 with the Xhol-EcoRI and EcoRI-XbaI restriction sites, respectively. HsHSF1-Venus was then cloned into pCDNA3.1(Hyg+) with the Xhol and XbaI restriction sites. Venus and HsHSF1(CHR-A/B)-Venus were individually cloned into pCDNA3.1(Hyg+) with Xhol and XbaI restriction enzymes, respectively. All plasmid inserts were verified by DNA sequencing.

Protein expression and purification. All protein-purification steps were performed at 4 °C unless otherwise noted. Protein concentrations in the final preparations were determined by absorbance at 280 nm. His6-HsDBD was expressed in E. coli BL21-CodonPlus-RIL cells. Cells were grown to an OD600 of 0.6 at 37 °C in 1 l Luria-Bertani medium, respectively; this was followed by addition of 20 mM DTT and incubation on ice for 30 min. Initial crystals were obtained by the sitting-drop vapor-diffusion method (24 ml 0.1 M HEPES-NaOH, pH 7.4, 300 mM NaCl, 10% poly(dI-dC) (Sigma) and 1 mM DTT, the samples were incubated at RT for 20 min. Subsequently, protein–DNA complexes and unbound DNA were separated by gel electrophoresis on 3% agarose in Tris-acetate-EDTA buffer (2 h at 0 V), in an ice-cooled chamber. The fluorescence was recorded with an FLA-2000 image reader (Fujifilm) and analyzed with the Aida Image Analyzer software package (Raytest).

Size-exclusion chromatography coupled to multiangle light scattering (SEC-MALS). Protein samples (~60 to 120 μg) were analyzed with static and dynamic light scattering by autoinjection of the sample onto an SEC column (5 μm, 7.8 × 300 mm column, Wyatt Technology, WTC-030NS) at a flow rate of 0.2 ml min−1 in buffer F supplemented with 250 μg ml−1 poly(dl-dc) (Sigma) and 1 mM DTT; the samples were incubated at RT for 20 min. Subsequently, protein–DNA complexes and unbound DNA were separated by gel electrophoresis on 3% agarose in Tris-acetate-EDTA buffer (2 h at 0 V), in an ice-cooled chamber. The fluorescence was recorded with an FLA-2000 image reader (Fujifilm) and analyzed with the Aida Image Analyzer software package (Raytest).
mixtures of 100 nl complex with 100 nl precipitant containing 0.09 M MgAc and 31% PEG 4000. For cryoprotection, crystals were dipped into 34% PEG 4000 containing 5% glycerol before being flash cooled in liquid nitrogen.

The CIBDB–HSE complex was formed by mixture of protein solution at ~20 mg ml\(^{-1}\) with duplex DNA in a volume ratio of 2:60:1, resulting in a final concentration of ~1.2 mM CIBDB and ~0.7 mM HSE duplex; this was followed by incubation on ice for 30 min. Crystals were obtained by the hanging-drop vapor-diffusion method at 20 °C with a precipitant containing 0.18 M triaminomium citrate and 11% PEG 3350. The crystals were transferred stepwise into cryosolution containing 17% PEG 3350 and 15% glycerol.

The CIBDB–SSRE complex was formed with the same protocol as for CIBDB–HSE. The crystals were obtained by the hanging-drop vapor-diffusion method at 20 °C with a precipitant containing 0.1 M HEPES–NaOH, pH 7.5, and 25% PEG 2000 MME. For cryoprotection, crystals were soaked in 30% PEG 2000 MME containing 5% glycerol.

Crystals of CrSnk(160–209) in space groups C222\(_1\) and P2\(_1\)2\(_1\)2\(_1\)2 were obtained by the hanging-drop method at a protein concentration of ~37 mg ml\(^{-1}\) with precipitants containing 0.1 M HEPES–NaOH, pH 7.0, and 20% PEG 6000 or 21% PEG 6000, respectively. For cryoprotection, the crystals were transferred stepwise into cryosolution containing 25% PEG 6000 and 10% glycerol.

Structure solution and refinement. Diffraction data were collected at beamlines ID2-3, ID29 and ID30A at ESRF, Grenoble, France. The data were integrated and scaled with XDS\(^{69}\), Pointless\(^{66}\), Scale\(^{69}\), Aimless\(^{68}\) and Truncate\(^{68}\), as implemented in the CCP4i interface\(^{65}\), were used to convert the data to CCP4 format.

The crystal form of CIBSnk(160–209) in space group P2\(_1\)2\(_1\)2\(_1\)2 was solved by sulfur single-wavelength anomalous diffraction (SAD) at a wavelength of 1.70 Å with SHELX-CDE as implemented in HKL2MAP\(^{70,71}\). Phases were calculated for a 0.93–Å native data set collected at 0.9786–Å wavelength from the same crystal.

The model was improved by autobuilding with ARP-wARP\(^{72}\). The final model contains a half-tetramer of CrSnk(160–209).

The crystal structure of CIBSnk(160–209) in space group P2\(_1\)2\(_1\)2\(_1\)2 was solved in two steps. First, a CIBSnk(160–209) half-tetramer was localized by molecular replacement with Molerep\(^{53}\). In the resulting map of this partial structure solution, spurious electron density for an additional CIBSnk(160–209) trimmer was detected, which was then built manually with Coot\(^{54}\).

The CIBDB–HSE crystallographic model was solved by molecular replacement with Molerep with the K. lactis HSF1–HSE crystal structure (PDB 3HTS\(^{53}\)) as a search template. The refined model was used as a search template for the other DBD–DNA complex structures. The asymmetric units of HSE–HSE and CIBDB–HSE contain one DBD and a single strand of DNA, whereas the unit cells of CIBDB–SSRE and HSE–HSE–SatIII contain two protein chains and the complete DNA duplex.

All models were refined with Refmac5 (ref. 75). Coot was used for manual model building\(^{56}\). Residues with disordered side chains facing solvent channels were modeled as alamines.

Structure analysis and modeling. Coordinates were aligned with only Cα positions with Lsqkab and Lsqman\(^{76}\). Homology models were created with Modeller\(^{77}\). Linker regions were modeled with Modeller with the ModLoop server\(^{79,78}\) or the Chimera interface to Modeller\(^{80}\). Figures were generated with PyMOL (http://www.pymol.org/) and ESPript\(^{81}\).

Cell culture and transfection. HeLa cells (ATCC) were maintained in Dulbecco’s Modified Eagle’s medium (DMEM, Biologchrom AG) supplemented with 10% FBS (Gibco), 1% penicillin-streptomycin (Gibco) and 1% l-glutamine (Gibco) at 37 °C in an atmosphere of 5% CO\(_2\). Transfections were performed with Lipofectamine and Plus reagents according to the manufacturer’s instructions (Invitrogen). A final plasmid DNA amount of 100 ng was used for transfection per well in a 96-well plate format, containing ~20,000 cells. 2 µg of plasmid DNA was used for transfection of a 100-mm-diameter dish. The HeLa cells as obtained from ATCC were not tested for mycoplasma contamination. They displayed typical cell shape and showed no sign of endoparasite infection after DAPI staining.

Fluc/Rluc activity ratios in HeLa cells. Cells were transfected with heat stress (HS)-inducible firefly luciferase (Fluc) under control of the HSP7A promoter (HSP7A–Fluc) and constitutively expressed Renilla luciferase (Rluc) under control of the CMV promoter (CMV–Rluc)\(^{77}\) along with Venus, HsHsf1–Venus or HsHsf1–Venus (Hsf–A/B)–Venus. Rluc expression was used as a control for transfection efficiency. After 48 h, cells were exposed to HS for 2 h at 43 °C and 2 h recovery at 37 °C (+HS) or were maintained at 37 °C (–HS). Luciferase activities from each well were measured on a Lumat LB 9507 luminometer (Berthold Technologies) with the Dual-Glo Luciferase Assay System (Promega). The ratios of Fluc/Rluc activities were calculated, averaged and normalized to the activity ratio measured in cells expressing Venus-only control in absence of HS. Error (s.d.) values were calculated from three independent experiments.

Fluorescence microscopy. HeLa cells were cultured on poly-l-lysine–coated coverslips (BD Biocoat Cellware), fixed with 4% formaldehyde and washed with PBS, stained with DAPI (Invitrogen) and mounted on slides for microscopy. Fluorescence imaging was performed on a confocal laser scanning microscope (Leica TCS SP8) with a 63× oil-immersion objective with 405-nm (for DAPI) and 488-nm (for Venus) laser wavelengths.