PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

Crystal structure of the hexamer of human heat shock factor binding protein 1

Xueqi Liu,[1] Lingfeng Xu,[1] Yiwei Liu,[2] Xiaohang Tong,[1] Guangyu Zhu,[3] Xuejun C. Zhang,[1,3]* Xuemei Li,[1]* and Zihe Rao[1,2]

[1] National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China
[2] Laboratory of Structural Biology, Tsinghua University, Beijing 100084, China
[3] Crystallography Research Program, Oklahoma Medical Research Foundation, Oklahoma City, Oklahoma 73104

ABSTRACT

Heat shock response (HSR) is a ubiquitous cellular mechanism that copes with a variety of stresses. This response is mediated by a family of transcriptional activators, heat shock factors (HSFs), which are under tight regulation. HSF binding protein 1 (HSBP1) is a negative regulator of HSR and is reported to bind specifically with the active trimeric form of HSF1, thus inhibiting its activity. HSBP1 contains heptad-repeats in the primary sequence and was believed to stay in a trimer form in solution. We report the crystal structure of the trimerization domain of the M30I/L55P mutant of human HSBP1 at 1.8 Å resolution. In this crystal form, the HSBP1 fragment of residues 6–53 forms a continuous, 11-turn long helix. The helix self-associates to form a parallel, symmetrical, triple coiled-coil helix bundle, which further assembles into a dimer of trimers in a head-to-head fashion. Solution study confirmed that the wild-type HSBP1 shares similar biophysical properties with the crystallized variant. Furthermore, we identified Ser31, which buried its polar side chain in the hydrophobic interior of the helix bundle, as a stability weak-spot. Substitution of this residue with Ile increases the melting temperature by 24°C, implicating that this conserved serine residue is maintained at position 31 for functional purposes.

Proteins 2009; 75:1–11.

Key words: HSBP1; HSF1; heat shock response; triple helix bundle; coiled-coil; heptad-repeat.

INTRODUCTION

Living cells react to external stresses such as elevated temperatures, chemical toxicants, and pathogen infections with the so-called heat shock response (HSR) by which transcription of a small set of genes, the heat-shock protein (*hsp*) genes, is activated.[1] Heat shock proteins (HSPs) encoded by the *hsp* genes include mainly molecular chaperones, proteases, and other proteins essential for protection and recovery from cellular damage resulting from misfolded proteins. This cellular phenomenon is widely observed in bacteria, plants, and animal worlds. In addition, some of the *hsp* genes function in development and cell differentiation.[2] Malfunction of the underlying cellular mechanism has been directly linked to tumor formation, cancers, and numerous neuronal degeneration diseases.[3–5]

HSR is mainly regulated by heat shock transcription factors (HSFs), of which HSF1 is a prototype.[2] HSF1 upregulates *hsp* gene expression by functioning both as a transcription factor and a polyadenylation stimulatory factor.[2] HSF activities are regulated at multiple levels: posttranslational modification, the oligomeric status of HSF, its DNA-binding ability, transcriptional competence, nuclear and subnuclear localization, as well as its interactions with regulatory cofactors or other transcription factors all appear to be fine tuned.[6] Particularly, inactive HSF1 assumes a monomer form (HSF1-M), and transcriptionally active HSF1 has a trimer form (HSF1-T).

To understand the regulation of HSF1 activity, intensive studies have been focused on identifying proteins that potentially interact with HSF1.[7–14] Using either the full length or trimerization domain of HSF1 as the bait, a protein termed heat shock factor binding protein 1 (HSBP1) was identified in a yeast-two-hybridization (Y2H) analysis and subsequently characterized.[7,12] This protein is highly conserved across species and has been found in every organism studied so far except yeast.[2] HSBP1 is ubiquitously

Abbreviations: AUC, analytical ultracentrifugation; HR, heptad-repeat; HSBP1, heat shock-factor binding protein 1; HSF1, heat shock-factor 1; HSP, heat shock protein; HSR, heat shock response; MAD, multiwavelength anomalous dispersion; MW, molecular weight; rmsd, root mean square deviation; SAS, solvent accessible surface; WT, wild type.

Grant sponsor: National Science Foundation, China (NSFC); Grant number: 30670494; Grant sponsor: Ministry of Science and Technology of China (the Project 973); Grant number: 2007CB914304; Grant sponsor: the Protein Science Project (CAS); Grant number: 2006CB911002.
*Correspondence to: Xuejun C. Zhang, Crystallography Research Program, Oklahoma Medical Research Foundation, 825 N.E. 13th Street, Oklahoma City, OK 73104. E-mail: zhangc@omrf.org; Xuemei Li, National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, 15 Datun Road, Beijing 100101, China. E-mail: lixm@sun5.ibp.ac.cn
Received 3 May 2008; Revised 26 June 2008; Accepted 29 June 2008
Published online 2 September 2008 in Wiley InterScience (www.interscience.wiley.com).
DOI: 10.1002/prot.22216

expressed and is usually localized in the nucleus.[15] Members of the HSBP family are small proteins (<10 kDa); for example, human HSBP1 consists of 76 amino-acid residues. Like HSF1, HSBP1 is predicted to have a coiled-coil structure and is found to self-assemble into homo-oligomers in solution. NMR analysis suggests the existence of a long α-helix in HSBP1 with unstructured amino (N)- and carboxyl (C)-terminal regions.[12] Coexpression of HSBP1 with HSF1 reduces the DNA-binding ability of HSF1. In mammalian cells, overexpression of HSBP1 inhibits the transcriptional activator activity of HSF1. Furthermore, the biological function of HSBP1 has been analyzed in *C. elegans* where overexpression of an HSBP1 homolog results in inhibition of HSR and lowers the survival rate after a heat shock stress.[7] On the basis of these observations, HSBP1 is currently considered as a negative regulator of HSR. Its functions under physiological condition may include preventing incidental activation of HSF1.

In spite of these advances, the detailed structural information on HSBP1 oligomerization remains elusive. Toward understanding the structural basis of HSR regulation by HSBP1, we carried out structural studies on this protein. Here we report the crystal structure of a degradation-resistant fragment of HSBP1. The structure reveals a continuous, long-helix conformation of HSBP1 which assembles into a coiled-coil three-helix bundle and subsequently into an elongated, symmetrical hexamer.

MATERIALS AND METHODS

Protein expression and purification

The cDNA of full length human HSBP1 (GenBank ID: NM_001537) was cloned into the pET-28a(+) vector (Novagen). DNA sequence analysis of the gene that we used indicated that the encoded HSBP1 protein contained two incidental point mutations of Met30 substituted by Ile and Leu55 by Pro (M30I/L55P). The resulting plasmid was transformed into *E. coli* BL21(DE3) cells. The recombinant His-tag fusion protein, His-HSBP1, was purified by Ni-NTA (Qiagen) affinity chromatography. The sample was then applied to a Resource Q ion-exchange column (GE Healthcare), from which the protein was eluted at 200 m*M* NaCl using a 0–1*M* NaCl gradient in 20 m*M* HEPES (pH 7.5). Selenomethionyl (Se-Met) derivative of HSBP1 was expressed with the pGEX-6p-1 vector in the methionine auxotrophic *E. coli* strain B834 (DE3) (Novagen). The derivative protein was fused with a GST (glutathione *S*-transferase) tag and purified with Glutathione-Sepharose 4B beads (GE Healthcare). After being cleaved from GST with Precision Protease (GE Healthcare) overnight at 4°C, the Se-Met derivative HSBP1 protein was eluted with PBS (phosphate buffered saline) and then applied to the Resource Q ion-exchange column using the same protocol as the native protein. The recombinant full length protein (both native and Se-Met derivative) consisted of 76 residues of HSBP1 (except the incidental M30I/L55P point mutations) and an N-terminal tag. In the case of "native" protein, a 34-residue peptide was attached at the N-terminus, and the intact fusion protein had a molecular weight (MW) of 12 kDa. In the case of Se-Met derivative, a peptide of sequence GPLGS was added at the N-terminus, and the protein had a predicted MW of 8.9 kDa. Because the protein samples did not have 280-nm absorption, their concentration was determined using the bicinchoninic acid (BCA) method[16]; for example, the extinction coefficient of full-length HSBP1 without any tag at 220-nm absorption was calibrated as 0.1 mg/mL/AU.

Crystallization

The His–HSBP1 recombinant protein sample was concentrated to 10 mg/mL in 20 m*M* HEPES (pH 7.5) and 200 m*M* NaCl, and was crystallized using the hanging drop method with a 1:1 (v/v) ratio to the reservoir solution. The initial crystallization condition was identified from a sparse-matrix screening using the Crystal Screen Kit I from Hampton Research and further optimized. Crystals typically had a spindle shape and grew to 0.2 mm in the longest dimension in a few weeks at 16°C. The crystallization was reproducible though with varied time. The crystal used for data collection was grown with a reservoir solution of 1.7*M* ammonium sulfate, 15% (v/v) glycerol, and 20 m*M* HEPES (pH 7.5). SDS polyacrylamide gel electrophoresis (SDS-PAGE) analysis of the crystal content revealed that the crystallized protein sample was a mixture of HSBP1 fragments of molecular weights about 6 kDa. Subsequent mass-spectroscopic analysis of the crystal confirmed that major components had twin peaks at about 6.2 kDa. Further N-terminal amino acid sequence analysis was inconclusive. Although different at the N-terminus, Se-Met derivative of HSBP1 was crystallized under a similar condition as the His-tagged native recombinant protein.

Data collection, structure determination, and refinement

The native crystal of HSBP1 was soaked in a saturated sodium citrate solution as the cryo-protectant and flash-frozen in 100 K cold nitrogen stream for data collection. It diffracted to 1.8-Å resolution at the BL-5A beamline of the Photon Factory Synchrotron Facility (Japan). Crystals of the Se-Met derivative were treated similarly for cryo-protection. A three-wavelength data set of the Se-Met derivative crystal of HSBP1 was collected at 2.5 Å resolution using a MAR CCD detector on the beamline 3W1A at Beijing Synchrotron Radiation Facility (China). Processing diffraction images and scaling integrated intensities were performed using the HKL2000 software package.[17]

The crystal structure of HSBP1 was solved using the multiwavelength anomalous dispersion (MAD) method[18] and data collected at peak, edge, and remote regions of

Table I
Data Collection, Phasing, and Refinement Statistics

Data collection				
Space group	R3			
Cell parameters *a, b, c* (Å)	35.2,35.2,233.3	35.1, 35.1, 234.2		
	Native	Inflection	Peak	Remote
Wavelength (Å)	1.0	0.9802	0.9799	0.9000
Resolution (Å)	50 (1.86)[a]−1.80	50 (2.59)−2.50	50 (2.59)−2.50	50 (2.59)−2.50
R_{merge}	0.068 (0.21)	0.06 (0.33)	0.07 (0.42)	0.08 (0.44)
$<I/\sigma>$	29.5 (4.7)	20.7 (2.6)	17.6 (2.0)	13.4 (2.0)
Completeness (%)	97.4	99.7	99.2	99.0
Redundancy	3.6	3.7	3.8	3.3
No. of observations	35,026	13,755	13,897	10,972
No. of unique reflections	9,762	3,682	3,678	3,317

Phasing statistics	Inflection versus peak	Inflection versus remote	Peak versus remote
Correlation of the anomalous data	0.54	0.45	0.39
F.O.M. solve	0.37		
F.O.M. resolve	0.64		

Refinement statistics	
Resolution (Å)	50−1.8
No. of reflections	9,441
Twin fraction	0.366
Twinned R_{work}/R_{free}[b] (1.86−1.80 Å)	19.5/23.4 (32.6/35.4)
No. of protein atoms	732
No. of solvent atoms	22
Average *B*-factor (Å^2)	
Protein	41.1
Solvent	34.7
R.m.s. deviations	
Bond lengths (Å)	0.007
Bond angles (°)	1.09
Ramachandran plot (%)[c]	96.4/3.6/0/0

[a]Numbers in parentheses are corresponding values in the highest resolution shell.
[b]Reflections of $|F_{obs}| > 0.0$.
[c]Calculated using PROCHECK.[23] Numbers reflect the percentage of residues in the core, allowed, generously allowed, and disallowed regions, respectively.

the selenium absorption of the Se-Met derivative crystal. Four potential selenium atoms were located and refined at 2.6-Å resolution using the program SOLVE.[19] Density modification in RESOLVE was used to improve and extent the MAD phases to 1.8 Å resolution of the native data set.[19] The resulting electron density map was automatically traced using RESOLVE,[20] resulting in an initial model of approximately 80% completeness of the asymmetric unit. Model building was further completed manually and refined using the programs O and CNS.[21,22] During the refinement, we found that the crystals that we used for diffraction data collection were twinned with hemihedral twin factors ranging between 0.1 and 0.4. Subsequently, we refined the crystal structure using detwinning options in CNS and a 1.8-Å resolution data set collected from a single native crystal of a 0.37 twin factor. Data collection and refinement statistics are summarized in Table I.

Mutagenesis constructs

Site directed mutagenesis was used to probe the structure of HSBP1. First, the true WT HSBP1 gene was reconstructed using the overlap extension PCR method and the cDNA of HSBP1-M30I/L55P as the template. A number of point mutations were then constructed with the true WT HSBP1 as the template. All these HSBP1 variants were cloned into the pGEX-6p-1 vector, and the recombinant proteins were purified using the same protocol as the Se-Met derivative HSBP1. All mutations were verified by full-length DNA sequencing.

Circular dichroism analysis

Circular dichroism (CD) was used to analyze the secondary structures and thermal stability of HSBP1 variants. The experiments were carried out with the PiStar-180 Stopped-flow Spectrometer (Applied Photophysics, UK) and a 0.1-cm light-path cuvette. The protein concentrations of all samples were adjusted to $OD_{220} = 2.0$ in a working buffer of 10 m*M* Na_2HPO_4 (pH 8.0) and 10 m*M* NaCl. Wavelength scans were performed between 200 and 260 nm with a 1-nm step-size at 20°C, and temperature scans were performed between 20 and 66°C (or

90°C for the S31I mutant) with a 2°C step-size at 222 nm. The thermal denaturation was reversible.

Gel filtration

Gel filtration experiments were used to analyze the oligomerization of HSBP1 variants in solution. The assays were performed with a Superdex 75 HR 10/300 (GE Healthcare) column on the Akta Purifier System (GE Healthcare) and a buffer of 25 m*M* Tris-HCl (pH 8.0) and 100 m*M* NaCl.

Analytical ultracentrifugation

Analytical ultracentrifugation (AUC) was used to analyze the molecular size distribution of HSBP1 variants in solution. Sedimentation velocity experiments were carried out with the Proteomelab™ XL-A/XL-I Protein Characterization System (Beckman Coulter). Experiments were performed at 20°C and 60,000 rpm. Protein samples prepared after gel filtration were diluted to $OD_{220} = 1.0$ in a buffer of 20 m*M* Tris-HCl (pH 8.0) and 100 m*M* NaCl. Velocity data were collected in a continuous scan mode at 220 nm. Sedimentation coefficients were calculated with the program Sedfit.[24]

Accession numbers

Coordinates and the experimental structural factors of the HSBP1 crystal structure have been deposited in the RCSB Protein Data Bank with accession number 3CI9.

RESULTS

Overall crystal structure of HSBP1

The recombinant protein of full length human HSBP1 (with an incidental M30I/L55P double point mutation) was expressed in *Escherichia coli* and purified. The protein sample was crystallized reproducibly in an R3 space group crystal form. The hexagonal cell parameters are $a = b = 35$ Å and $c = 233$ Å. Mass spectrometry analysis showed that our recombinant protein sample had been subjected to incidental proteolysis from 12 kDa to about 6.2 kDa during the crystallization process. However, our effort to determine the boundaries of the remaining peptide was inconclusive. The native crystal diffracted beyond 1.8-Å resolution. Phases of the structural factors were solved using the Se-Met based MAD method. There were two HSBP1 monomers (denoted as A and B) per asymmetric unit; the corresponding Matthews coefficient is 2.2 Å^3/Da (i.e. ~45% solvent content). Residues from the N-terminus to Asp5 and from Lys50 to the C-terminus in molecule A and residues from the N-terminus to Thr8 and from Asp54 to the C-terminus in molecule B could not be built into the final refined model because of lack of interpretable electron density. The combined, visible part of the HSBP1 molecule (i.e. residues 6–53) has a calculated MW of 5.5 kDa. Because the protein sample in the crystal has a molecular weight of 6.2 kDa, the invisible, mobile regions accounted for 0.7 kDa, equivalent to about six amino acid residues. In the crystal lattice, we observed an unused space which could be shared by both N- and C-terminal peptides.

Each of the two crystallographically independent HSBP1 molecules consists of a single, curved, 11-turn α-helix (see Fig. 1). The two HSBP1 molecules in an asymmetric unit can be superimposed well; the root mean square deviation (rmsd) of the Cα atoms of residues 10–49 is 0.66 Å between the two protein molecules. Each single-helix molecule forms parallel coiled-coil homotrimers with its own symmetry mates around the crystallography three-fold axis [Fig. 1(A)]. The region of the three-helix bundle formed by residues 16–49 packs more tightly and is slightly left-hand twisted, while the N-terminal region splits into three fingers [Fig. 2 (A)]. Each monomer of the crystal structure buries about 1300 Å^2 (i.e. 30% of the total) solvent accessible surface (SAS) during the trimer formation, 73% of which is contributed by carbon atoms.

Furthermore, the two HSBP1 trimers assembled from molecules A and B, respectively, are related by a nearly perfect two-fold axis (178.5° rotation with a 0.3 Å screw length) in a head-to-head fashion. The dyad axis is perpendicular to the crystallography three-fold axis but in an arbitrary direction otherwise (~20° from the *a* axis of the hexagonal cell). The two trimers can superimpose reasonably well with each other, resulting in a 1.3-Å rmsd for 120 Cα atoms (10–49 residues from each molecular) from each trimer (using a 3.0 Å cutoff). Each trimer contributes its three, N-terminal, helix fingers (e.g. Thr8–Val15 in chain A) to assemble a short, antiparallel, six-helix bundle, whereas the 50s regions form the two ends of the elongated, symmetrical, hexamer spindle (see Fig. 1). The longest dimension of the HSBP1 hexamer is about 110 Å. Each trimer buries about 2,200 Å^2 SAS in the hexamer interface, which is formed predominantly by carbon atoms (87%). The extent of the trimer–trimer interaction appears comparable with the interaction within a trimer (i.e. 2×2200 vs. 3×1300 Å^2 buried SAS).

Structural details of the HSBP1 trimer and hexamer

Our crystal structure illustrated a simple topology of the HSBP1 protein: the visible part of the protein is comprised of a single α-helix. This segment corresponds to the most conserved region in amino acid sequence comparison among members of the HSBP1 family. Particularly, nearly all interior residues in the trimerization region (i.e. residues 15–49) are absolutely conserved in the alignment shown in Figure 3. With an exception of Ser31, all of these residues are hydrophobic. Helices in a typical coiled-coil structure consist of heptad-repeats

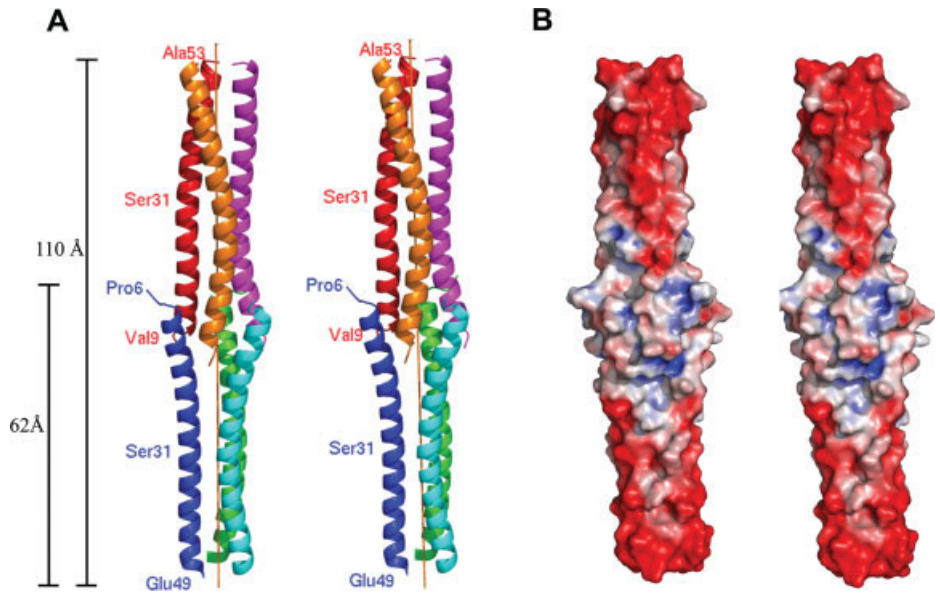

Figure 1

Crystal structure of HSBP1 hexamer. (**A**) Stereo view of the ribbon diagram of the hexamer structure. Each monomer is colored differently. The vertical line indicates the three-fold axis. The dimensions of the hexamer and trimer are labeled on the left. (**B**) Stereo view of HSBP1 hexamer surface electrostatic potential distribution. Electrostatic potential distribution was calculated with the program APBS.[23] Positively charged regions (>+10 kT/e) are colored blue, and negatively charged regions (<−10 kT/e) in red. The picture was drawn with the program PyMol (http://www.pymol.org).

(HRs). The first and fourth residues (i.e. *a* and *d*) in the sequence *a*–*g* of each HR occupy the interhelix interface and generally are hydrophobic. Residues *e* and *g*, which are usually bulky, polar, or charged, contribute toward specificity and establish proper chain registration between the coiled-coil helices.[27–29] In the three-helix bundle region of HSBP1, an anomaly occurs at the conserved Ser31. This serine residue occupies a position where Met30 (Ile in our case) should be the interior residue according to the heptad pattern prediction, and it separates the long helix into two HR sequences (see Fig. 3). Nevertheless, this anomaly appears smooth in the 3D structure [Fig. 1(A)]. The side chain and backbone amide group of Ser31 share the carbonyl oxygen of Phe27 from the previous helical turn for hydrogen bonding. We speculated that burying a cluster of three serine side chains in the generally hydrophobic interior of the helix bundle is energetically unfavorable[30]; thus the conservation of Ser31 likely bears some biological functions. It is likely that substitution of this Ser31 with isoleucine (or other hydrophobic) residue would enhance the stability of the HSBP1 trimer but sacrifice some of its functions.

The hexamerization region of each HSBP1 molecule (i.e. residues 6–16) simultaneously participates in two sets of antiparallel helix–helix interactions with neighboring helices. One is via Leu12 and Val15; and the other one is via the Leu12 and Val16 [Fig. 2(A)]. Each of the two residue pairs interacts with its dyad symmetric counterpart in a knobs-into-holes fashion. Because of their short side chain lengths, the interior residues of this six-helix bundle form a small hydrophobic cavity of ~10-Å diameter at the center of the hexamer region. The side chain hydroxyl group of conserved Thr13 has an energetically favorable gauche$^+$ (-60° χ_1) rotamer,[31] and participates in a 2.6-Å hydrogen bond with its symmetry counterpart on the hexamer surface, sealing the interface formed by Leu12 and Val16.

Gly36 is the only residue in the long, visible helix (residues 9–53) that has a low helix propensity. It is located on the ridge of the helix bundle and exhibits a normal, helical backbone φ-ϕ angle. In most species, this position in HSBP1 is conserved. An exception is the plant HSBP1, where the corresponding position is Ser, and mutation of this residue to Ala showed no effect on HSF-binding.[9]

Similar properties of the wild-type and M30I/L55P variant in solution

To rule out the possibility that the incidental M30I/L55P mutant contributes structural perturbation that results in the observed crystal structure, we constructed the wild type (WT) HSBP1 by a reverse mutation I30M/P55L from the HSBP1-M30/L55P template and compared the properties of this bona fide WT HSBP1 with those of the crystallized HSBP1-M30I/L55P.

The CD spectra showed that the WT and the M30I/L55P variant have near identical profiles including double

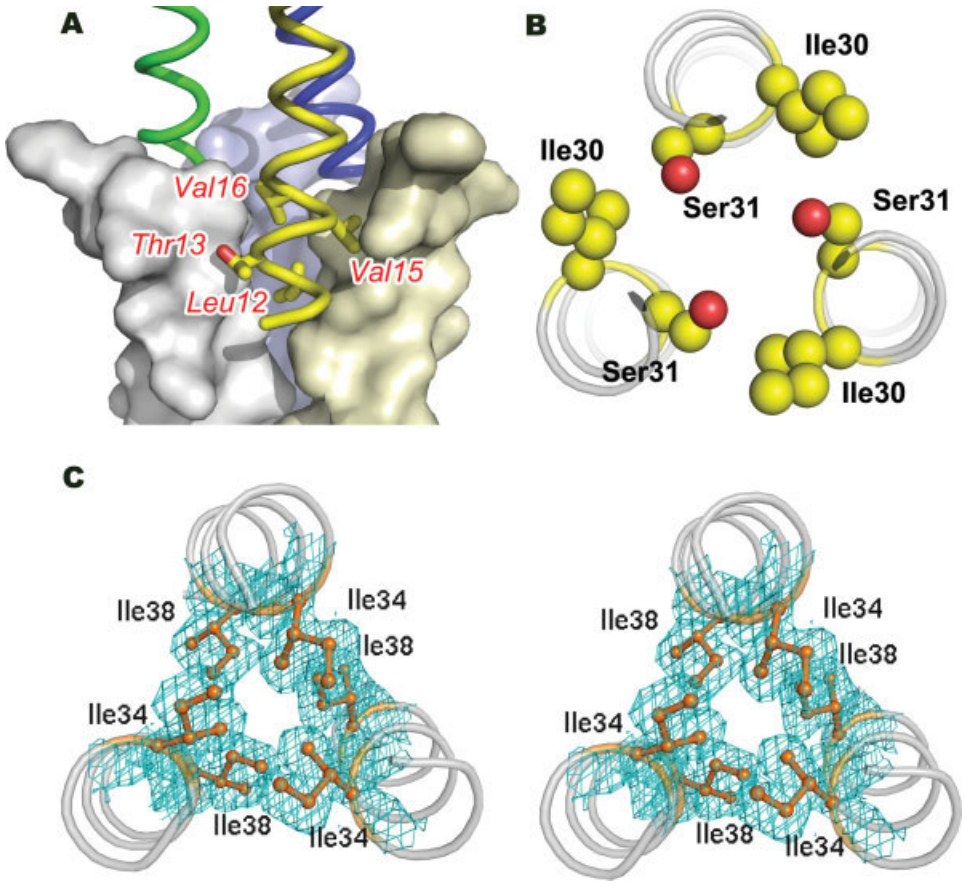

Figure 2

Interior packing of the helix bundle. The peptide backbone is shown in cartoon, and the residues are shown in stick-ball models. (**A**) Trimer–trimer interface. One trimer is shown in molecular surface model, and the other one is shown in cartoon. Residues Leu12, Thr13, Val15, and Val16 from one monomer are shown in stick models. (**B**) Packing of Ser31 in the three helix bundle. The Ser31 and the adjacent Ile30 are shown in CPK model. The protein complex is viewed along the crystallography three-fold axis. (**C**) Stereo view of the packing of three helix bundle around Ile34 and Ile38. The 2Fo-Fc map phased with the final refined model was contoured at 1.0 sigma level around the residues of interests. This picture was drawn with the program PyMol.

minima at 208 and 222 nm [Fig. 4(A)], suggesting that they share very similar secondary structural contents. Similar to a previous report,[12] the recombinant proteins of the WT as well as the M30I/L55P variant showed reversible temperature denaturation [Fig. 4(B)]. Nevertheless, the WT showed a steeper denaturing curve than the double mutant, indicating a more cooperative thermal unfolding in the WT. Furthermore, in a size-exclusion chromatography analysis, the WT and the double mutant showed identical elution profiles with a 220 nm absorption peak at the 9.5-mL elution volume [Fig. 4(C)]. Similarly, in a sedimentation velocity analytical ultracentrifugation (AUC) assay, the two recombinant samples showed identical sedimentation coefficients (1.96 $\pm$ 0.09 S) [Fig. 4(D)]. Therefore, the solution behaviors of the WT and the M30I/L55P variant are the same at both secondary and tertiary structure levels. These results allow us to transfer structural information obtained from the M30I/L55P crystal structure directly to the WT HSBP1. Thus, all subsequent mutagenesis studies were performed in the WT background. On the other hand, our repeated effort to crystallize the WT HSBP1 did not result in any usable crystals, and unlike the M30I/L55P variant the WT protein sample did not degrade during the crystallization trial. Met30 is solvent exposed in the HSBP1 hexamer but is not involved in crystal packing. Therefore, we believe that this point mutation was not a contributing factor for protein crystallization. On the other hand, the L55P mutation is located inside a predicted helix region. We speculate that this point mutation might cause a destabilization of this region thus making it more susceptible to proteolytic degradation. Therefore, the L55P point mutation might have resulted in a truncation of the C-terminal peptide which facilitated the crystallization. Such a "beneficial" degradation does not appear available for the WT protein.

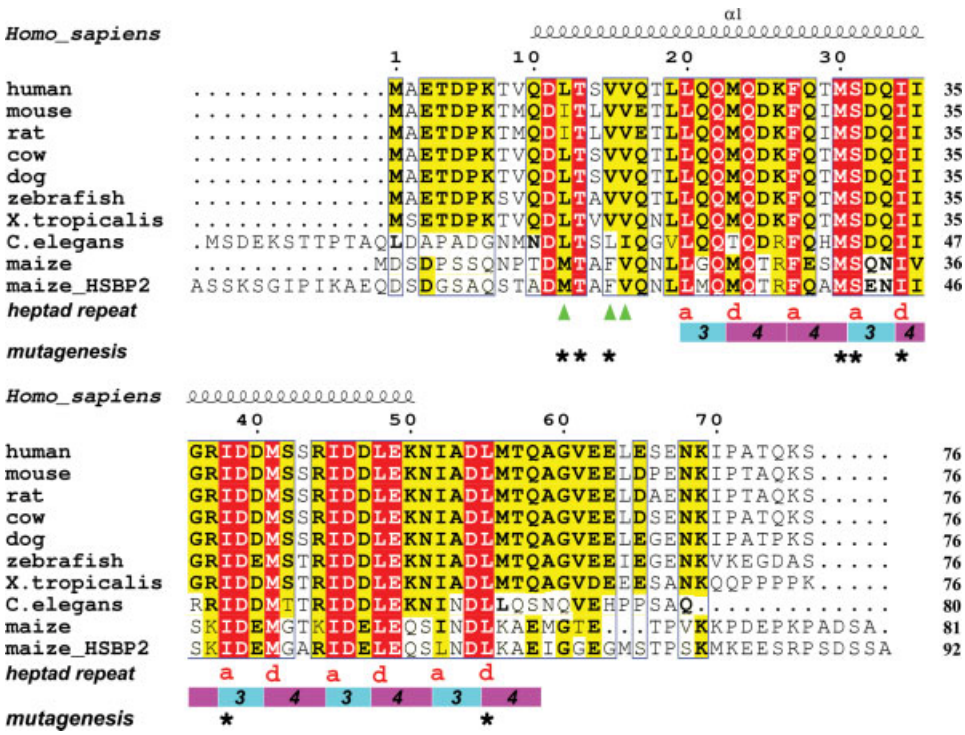

Figure 3

Sequence alignment of HSBP1 homologs from different species. The primary sequence of Human (*Human sapiens*) HSBP1 is from the GenBank file ***NP_001528***; mouse (*Mus musculus*) HSBP1, ***NP_077181***; rat (*Rattus norvegicus*) HSBP1, ***NP_775142***; cow (*Bos Taurus*) HSBP1, ***Q3ZC22***; dog (*Canis familiaris*) HSBP1, ***XP_852175***; zebrafish (*Danio_rerio*) HSBP1, ***AAH59566***; *Xenopus tropicalis* HSBP1, ***NP_001011422***; *Caenorhabditis elegans* HSBP1, ***NP_502406***; maize (*Zea mays*) HSBP1, ***AAM15929***; maize HSBP2, ***AAR18070***. Identical residues are highlighted in red, while the other conserved residues are highlighted in yellow and shown in black bold. The *a* or *d* positions of HRs are marked below the alignment. The green triangles indicate positions that are involved in trimer–trimer interaction. Mutation sites in this study are marked with asterisks. The position of the crystallographically observed α-helix is depicted as a helix, according as chain A of the refined structure. The alignment was calculated by CLUSTALW.[25] The figure was generated with ESPript 2.2.[26]

Breaking the oligomer with point mutations

To verify biological relevance of the observed hexamer helix bundle, we introduced mutations to disrupt the trimer and hexamer formation observed in the crystal. To break the HSBP1 trimer, a double mutation I34R/I38R was introduced in the interior of their helix bundle in middle of the long coiled-coil region [Fig. 2(C)]. Presumably, the six Arg residues (two from each HSBP1 molecule) would repel from each other and prevent association of the helices. Similarly, to break the hexamer, a double mutant L12D/V15D was introduced in the trimer–trimer interface [Fig. 2(A)].

As we expected, the I34R/I38R mutant protein lost the ability to oligomerize in a size exclusion chromatography assay [Fig. 4(C)] and sedimentation velocity ultracentrifugation analysis [Fig. 4(D)]. Meanwhile, CD spectroscopic analysis showed that the I34R/I38R variant had no regular α-helix or β-sheet secondary structure [Fig. 4(A)]. It indicates that in the absence of a helix bundle, HSBP1 does not possess a well-defined conformation and is likely to be unfolded in solution.

In the CD spectroscopic analysis of the L12D/V15D variant, we found that its secondary structure is essentially the same as the WT [Fig. 4(A)]. Consistent with our structure-based prediction, this double mutation did not cause a complete dissociation of the helix bundle and thus preserved significant α-helix content in the protein sample, in contrary to the I34R/I38R variant. Furthermore, the result of a sedimentation velocity AUC experiment on L12D/V15D was similar to the WT [Fig. 4(D)], and its melting temperature was even slightly higher than that of the WT [Fig. 4(B)]. However, we observed complete absence of the hexamer species of L12D/V15D from the size exclusion chromatography while the WT HSBP1 showed a recognizable hexamer population [Fig. 4(C)]. Together, these data suggest that the hexamer observed in the crystal structure also exits solution, although the trimer form appears the dominant species under the experimental condition.

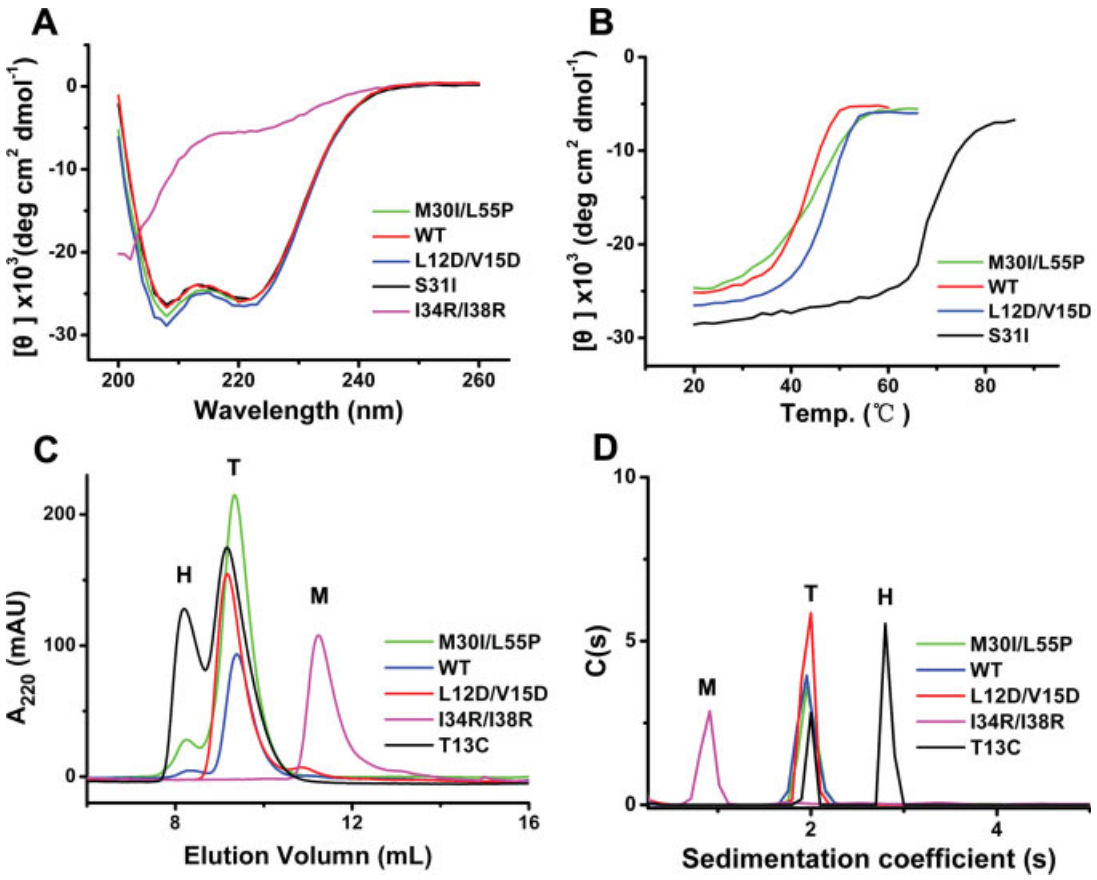

Figure 4

Solution analyses of HSBP1 variants. (**A**) Representative CD wavelength scans at 20°C. (**B**) Representative CD thermal scan at 220 nm. All the protein samples (with tags cleaved) were diluted to OD_{220} = 2.0 in a buffer of 10 m*M* Na_2HPO_4 and 10 m*M* NaCl. Both experiments were repeated multiple times, and the results were reproducible. (**C**) The size-exclusion chromatography experiments were performed on a Superdex 75 10/300 HR column. (**D**) The sedimentation velocity ultracentrifugation analysis was carried at 20°C, and the protein samples (with tag cleaved) were adjusted to OD_{220} = 1.0. The peaks at 0.95, 1.96, and 2.80 S correspond to an apparent MW of 9, 27, and 54 kDa and marked at monomer (M), trimer (T), and hexamer (H), respectively. Both experiments were monitored at 220 nm.

Analysis on hexamerization in solution

Because population of the WT HSBP1 hexamer was quite low in solution [Fig. 4(C,D)], we decided to use a mutagenesis approach to further confirm the structurally observed hexamer. We constructed a T13C point mutation based on the observation that this Thr13 formed a hydrogen bond with its symmetry counterpart in the hexamerization region. This T13C variant behaved similarly to the WT in both expression and purification. Importantly, in the size-exclusion chromatography experiment, we observed an increased elution peak at the position corresponding to the hexamer [Fig. 4(C)]. Consistent with the size-exclusion chromatography result, the sedimentation velocity AUC analysis showed a new peak at the hexamer position [Fig. 4(D)]. In both experiments, the ratio of hexamer to trimer population increased if the T13C sample was kept on ice for a few days compared with a freshly made sample (data not shown). These results confirmed that the hexamer of HSBP1 observed in the crystal structure also exists in solution and demonstrated that such a hexamer can be stabilized by engineered, symmetry related disulfide bonds.

A 24°C increase in melting temperature by a single point mutation S31I

As discussed earlier, Ser31 appears energetically unfavorable to reside in the hydrophobic interior of the helix bundle. To verify our structural prediction, we constructed a point mutation S31I and analyzed its solution property. Its CD wavelength scan at room temperature showed a profile similar to the WT [Fig. 4(A)]. More importantly, this single point mutation variant showed a 68°C melting temperature, 24°C higher than the WT [Fig. 4(B)]. This dramatic result strongly supports our hypothesis that burying a polar residue in the interior of the helix bundle has important functional implication.

DISCUSSION

Previously, it has been demonstrated that HSBP1 self-associates in a number of experiments including Y2H, sizing exclusion chromatography, native gel, and analytic ultracentrifugation.[7,12] The interaction strength between the full-length HSBP1 molecules is comparable with that between the full-length and a C-terminal truncation mutant, Δ65, which deletes residues from Ser66 to the C-terminus. In contrast, the interaction of full-length with another C-terminal truncation mutant, Δ45, was completely abolished. These data suggest that the region around residue 45, but not the C-terminal region, is critical for the homo-oligomerization of HSBP1. Furthermore, NMR analyses and limited proteolysis both establish that the C-terminus of HSBP1 is unstructured and susceptible to proteolytic cleavages.[12] In addition, it has been observed in an AUC experiment that HSBP1 equilibrates between a "monomer" of apparent molecular weight of 23 kDa and its dimer.[12] Our crystal structure of HSBP1 provides satisfactory explanations to these previous observations and further illustrates that the major part of HSBP1, that is residues 6–49, forms one continuous α-helix. This long helix assembles into a parallel three-helix bundle homotrimer, which further self-associates to form an elongated symmetrical hexamer (see Fig. 1). Our mutagenesis experiments of both hexamer-breaking (L12D/V15D) and hexamer-locking (T13C) demonstrate that the structurally observed hexamer also exits in solution [Fig. 4(C,D)]. The C-terminal region of our recombinant HSBP1 protein is missing from the crystal structure. They would be located beyond both observed ends of the elongated HSBP1 hexamer. Furthermore, our trimer breaking mutant I34R/I38R completely lost secondary structures [Fig. 4(A)], indicating that the tertiary structure is essential for the native conformation of HSBP1 and probably for its functions too.

On the basis of the sequence analysis, it was proposed that HSBP1 contains two HR regions: HR-N consists of residues 1–34, and HR-C consists of residues 35–59.[12] According to this prediction, there would be a "stutter" region at the junction of two HR sections, which might introduce a breaking, bending, or bulge in the helix. The crystal structure of HSBP1 revealed two features not expected from sequence analysis. First, the real anomaly of the long helix occurs at position 16 where the C-terminal three-helix bundle becomes split and the N-terminal region turns into a conformation that accommodates an antiparallel, symmetric, six-helix bundle formation. Second, the predicted "stutter" seems to be absorbed gradually into the coiled-coil conformation with only one polar residue, Ser31, located in the interior of the three-helix bundle.

According to the HR extrapolation shown in Figure 3, the residues C-terminal to the visible helix up to position 59 likely participate in a coiled-coil conformation; thus the coiled-coil structure might have two more helical turns in the full-length HSBP1 than what was observed in the crystal structure. The reason we did not see the helix extension at the C-terminal region is likely to be that a proteolytic cleavage by an unidentified protease occurred close to the visible C-terminus. HSBP1 homologs have highly conserved sequences across a wide range of species. Therefore, our crystal structure of the human HSBP1 likely provides reliable information for modeling other HSBP1 structures.

A 3-4-4-3 pattern is found in the middle of the long heptad repeats of HSBP1 to replace the usual 3-4-3-4 repetition. It has been observed that a 3-4-4-3 stutter is one of the two most frequently occurring irregularities in coiled-coils.[32] It often creates hot spots in a coiled-coil where low stability may be required.[33] The marginal stability of the WT HSBP1 may partially be contributed by the buried polar side chain of Ser31 in the helix-bundle interior [Fig. 2(B)] which creates the 3-4-4-3 stutter (see Fig. 3). The dramatic stability increase in the S31I mutant [Fig. 4(B)] strongly suggests that evolution favors functions of HSBP1 by sacrificing its stability. For example, one may speculate that Ser31 is required for HSBP1 to maintain a moderate stability. Such a low stability protein complex may serve as a sensor for the environmental changes and adjust its interactions with HSF1 in response to stresses.

Numerous proteins have been found to bind HSBP1 directly *in vivo*. The two well known examples are HSF1 and HSP70,[7] but their binding modes remain elusive. In principle, both HSBP1 and HSF1 could use either the monomer or trimer form to interact with each other. Using the monomer forms would require that such an interaction is energetically stronger than the homo-oligomerization of both HSF1 and HSBP1. Because the sequence homology within the HSBP1 family is much higher than that in the HSF1 family (see Fig. 3),[2] there appears no co-evolution between the two proteins. Therefore, the hetero-oligomerization interaction between HSF1 and HSBP1 is unlikely to be strong enough to over-ride that of homotrimerization. In fact, it has been demonstrated that HSBP1 specifically interacts only with the trimer form of HSF1.[7] If both HSF1 and HSBP1 used their trimer forms to interact with each other, would this interaction maintain their three-fold intrinsic symmetry? One possible scenario is that with a prior HSBP1 hexamer dissociation, a trimer of HSBP1 interacts with a trimer of HSF1 in a manner similar to the HSBP1 hexamer observed in the crystal structure. However, there is currently no evidence supporting such a model. In contrast, the involvement of the HSBP1 C-terminal region in HSF1 binding is in direct conflict with this model.[12] One alternative is that the HSF1 trimer interacts with the HSBP1 trimer in a nonsymmetrical fashion where the interaction is essentially between helix bundles. In this case, the requirement of trimerization on both

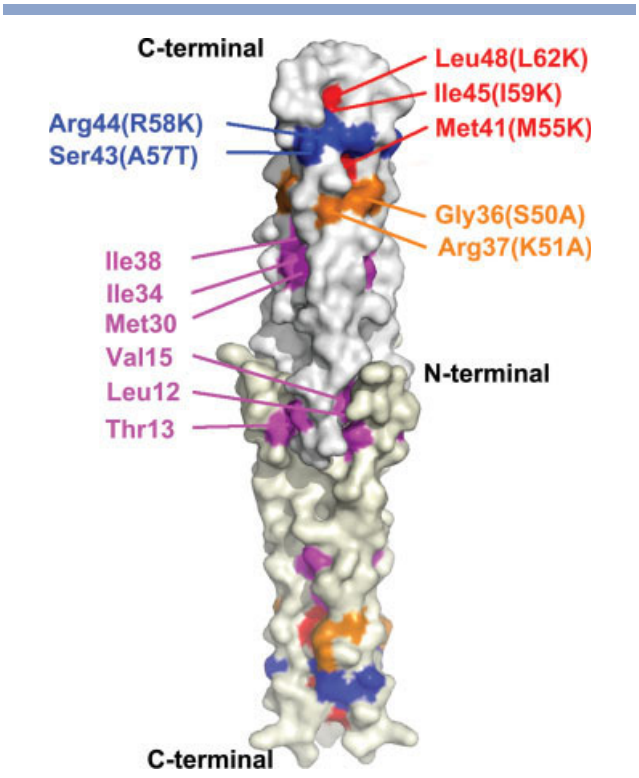

Figure 5

Summary of mutagenesis studies on HSBP1 homologs. The human HSBP1 hexamer is shown in a molecular surface representation. Residues investigated by the site-directed mutagenesis in this work are shown in magenta. Residues, mutations of which showed no effect on HSF1 binding in a previous study on plant HSBP2,[9] are shown in blue and specified in parentheses. Finally, residues, mutations of which showed detectable positive effects on HSF1 binding, are shown in red and similarly labeled.

HSF1 and HSBP1 would simply reflect the need for more extensive interaction to achieve binding specificity.

Both HSF1 and HSBP1 are acidic proteins with a pI of 5.0 and 4.2, respectively. Such a property appears to be unfavorable for their interaction. However, the trimer domain of HSF1 (residues 137–192), which is capable of directly binding with HSBP1, possesses a high concentration of basic residues with a calculated pI of 9.6. On the basis of HR prediction of this HSF domain, a number of Lys residues distribute on the surface of the predicted trimer helix bundle.[34] Complementarily, there is a long negatively charged ridge composed of acidic residues from Asp25 to the C-terminus in each HSBP1 monomer [Fig. 1(B)]. Electrostatic interaction between the two oppositely charged regions may facilitate their binding. Furthermore, conformational change associated with HSF1 monomer–trimer transition may serve as a switch to regulate this interaction.

There might be two possible ways for an HSBP1 mutation to disrupt the interaction between HSF1 and HSBP1: First, assuming that homo-oligomerization of HSBP1 is essential for the interaction with HSF1, any mutation that destabilizes the tertiary structure (e.g. our I34R/I38R double mutation) would decrease its binding with HSF1. Second, mutations at the direct interface would disrupt their binding. In a Y2H experiment, the plant HSBP2 mutations I52K/M55K and I59K/L62K (equivalent to interior residues 38/41 and 45/48 in human HSBP1 (see Fig. 5)) cannot interact with the HR domains of HSFA4a (a cognate HSF1 homolog for HSBP2). These mutations likely belong to the first category. On the other hand, a surface mutation of HSBP2, R58K (Arg44 in HSBP1), disrupts the binding with HSF, and A57T (Ser43 in human HSBP1) has a marginal effect. These mutations likely belong to the second group. In contrast, a surface mutant of HSBP2, S50A/K51A (Gly36/Arg37 in human HSBP1), shows no effect on HSF-binding. Moreover, Δ1–12 and Δ77–79 mutants of HSBP2 (no correspondence and VEE61–63, respectively, in human HSBP1) reduce the HSFA4a binding. Replacing both the N- and C-terminal domains of HSBP2 into EMP2 (HSBP1 orthorlog in plant) is sufficient to enable interaction of this EMP2 variant with HSFA4a.[9] Together, these results suggest that both N- and C-termini of HSBP are involved in HSF-binding. Considering the closer distance between the N- and C-termini from opposite trimers in a hexamer than the distance between termini from the same trimer (Figs. 1 and 5), the HSBP1 hexamer seems to have advantages over the trimer during an interaction with an HSF1 trimer. Therefore, a shift of the equilibration between the trimer and hexamer of HSBP1 may serve as a regulation mechanism of the interaction between HSF1 and HSBP1.

ACKNOWLEDGMENTS

The authors thank the staff of the National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences for technical assistance.

REFERENCES

1. Westerheide SD, Morimoto RI. Heat shock response modulators as therapeutic tools for diseases of protein conformation. J Biol Chem 2005;280:33097–33100.
2. Pirkkala L, Nykanen P, Sistonen L. Roles of the heat shock transcription factors in regulation of the heat shock response and beyond. FASEB J 2001;15:1118–1131.
3. Luparello C, Sirchia R, Pupello D. PTHrP [67–86] regulates the expression of stress proteins in breast cancer cells inducing modifications in urokinase-plasminogen activator and MMP-1 expression. J Cell Sci 2003;116(Pt 12):2421–2430.
4. Tonkiss J, Calderwood SK. Regulation of heat shock gene transcription in neuronal cells. Int J Hyperthermia 2005;21:433–444.
5. Whitesell L, Bagatell R, Falsey R. The stress response: implications for the clinical development of hsp90 inhibitors. Curr Cancer Drug Targets 2003;3:349–358.
6. Voellmy R. On mechanisms that control heat shock transcription factor activity in metazoan cells. Cell Stress Chaperones 2004;9:122–133.

7. Satyal SH, Chen D, Fox SG, Kramer JM, Morimoto RI. Negative regulation of the heat shock transcriptional response by HSBP1. Genes Dev 1998;12:1962–1974.
8. Xing H, Mayhew CN, Cullen KE, Park-Sarge OK, Sarge KD. HSF1 modulation of Hsp70 mRNA polyadenylation via interaction with symplekin. J Biol Chem 2004;279:10551–10555.
9. Fu S, Rogowsky P, Nover L, Scanlon MJ. The maize heat shock factor-binding protein paralogs EMP2 and HSBP2 interact non-redundantly with specific heat shock factors. Planta 2006;224:42–52.
10. Hu Y, Mivechi NF. HSF-1 interacts with Ral-binding protein 1 in a stress-responsive, multiprotein complex with HSP90 in vivo. J Biol Chem 2003;278:17299–17306.
11. Kim BH, Schoffl F. Interaction between Arabidopsis heat shock transcription factor 1 and 70 kDa heat shock proteins. J Exp Bot 2002;53:371–375.
12. Tai LJ, McFall SM, Huang K, Demeler B, Fox SG, Brubaker K, Radhakrishnan I, Morimoto RI. Structure-function analysis of the heat shock factor-binding protein reveals a protein composed solely of a highly conserved and dynamic coiled-coil trimerization domain. J Biol Chem 2002;277:735–745.
13. Reindl A, Schoffl F. Interaction between the Arabidopsis thaliana heat shock transcription factor HSF1 and the TATA binding protein TBP. FEBS Lett 1998;436:318–322.
14. Yoshima T, Yura T, Yanagi H. The trimerization domain of human heat shock factor 2 is able to interact with nucleoporin p62. Biochem Biophys Res Commun 1997;240:228–233.
15. Zhang HY, Li YC, Lin JT, Xu CS. Cloning and analysis of rat heat shock factor binding protein 1 cDNA. Yi Chuan 2004;26:647–652.
16. Stoscheck C. Quantitation of protein. Methods Enzymol 1990;182:50–69.
17. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol 1997;276:307–326.
18. Hendrickson WA, Ogata CM. Methods enzymology, Vol. 276: Phase determination from multiwavelength anomalous diffraction measurement. New York: Academic Press; 1997.
19. Terwilliger TC. SOLVE and RESOLVE: automated structure solution and density modification. Methods Enzymol 2003;374:22–37.
20. Perrakis A, Morris R, Lamzin VS. Automated protein model building combined with iterative structure refinement. Nat Struct Biol 1999;6:458–463.
21. Jones TA, Zou JY, Cowan SW, Kjeldgaard M. Improved methods for binding protein models in electron density maps and the location of errors in these models. Acta Crystallogr A 1991;47(Pt 2):110–119.
22. Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr D Biol Crystallogr 1998;54(Pt 5):905–921.
23. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. Electrostatics of nanosystems: application to microtubules and the ribosome. Proc Natl Acad Sci USA 2001;98:10037–10041.
24. Schuck P. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. Biophys J 2000;78:1606–1619.
25. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994;22:4673–4680.
26. Gouet P, Courcelle E, Stuart DI, Metoz F. ESPript: analysis of multiple sequence alignments in PostScript. Bioinformatics 1999;15:305–308.
27. Monera OD, Zhou NE, Kay CM, Hodges RS. Comparison of antiparallel and parallel two-stranded alpha-helical coiled-coils. Design, synthesis, and characterization. J Biol Chem 1993;268:19218–19227.
28. O'shea EK, Klemm JD, Kim PS, Alber T. X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. Science 1991;254:539–544.
29. Tripet B, Wagschal K, Lavigne P, Mant CT, Hodges RS. Effects of side-chain characteristics on stability and oligomerization state of a de novo-designed model coiled-coil: 20 amino acid substitutions in position "d". J Mol Biol 2000;300:377–402.
30. Matthews BW. Structural and genetic analysis of the folding and function of T4 lysozyme. FASEB J 1996;10:35–41.
31. Blaber M, Zhang XJ, Lindstrom JD, Pepiot SD, Baase WA, Matthews BW. Determination of alpha-helix propensity within the context of a folded protein. Sites 44 and 131 in bacteriophage T4 lysozyme. J Mol Biol 1994;235:600–624.
32. Brown JH, Cohen C, Parry DA. Heptad breaks in alpha-helical coiled coils: stutters and stammers. Proteins 1996;26:134–145.
33. Supekar VM, Bruckmann C, Ingallinella P, Bianchi E, Pessi A, Carfi A. Structure of a proteolytically resistant core from the severe acute respiratory syndrome coronavirus S2 fusion protein. Proc Natl Acad Sci USA 2004;101:17958–17963.
34. Lupas A, Van Dyke M, Stock J. Predicting coiled coils from protein sequences. Science 1991;252:1162–1164.