BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

Biochemistry **2010,** *49,* 8577–8584 8577
DOI: 10.1021/bi100876n

Peptide-Binding Sites As Revealed by the Crystal Structures of the Human Hsp40 Hdj1 C-Terminal Domain in Complex with the Octapeptide from Human Hsp70[†]

Hironori Suzuki, Shuji Noguchi, Hiroshi Arakawa, Tadaaki Tokida, Mariko Hashimoto, and Yoshinori Satow*

Graduate School of Pharmaceutical Sciences, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

Received May 31, 2010; Revised Manuscript Received September 1, 2010

ABSTRACT: Heat shock protein (Hsp) 40s play essential roles in cellular processes by cooperating with Hsp70 proteins. Hsp40 proteins recognize non-native polypeptides, deliver these peptides to Hsp70 proteins, and stimulate the ATPase activity of Hsp70 proteins to facilitate the correct folding of the polypeptides. We have determined the crystal structures of the C-terminal peptide-binding domain of human Hsp40 Hdj1 (CTD) and of its complex with the C-terminal octapeptide of human Hsp70, $^{634'}$GPTIEEVD$^{641'}$. CTD exists as a twisted, horseshoe-shaped homodimer. The protomer consists of two domains, I and II, with similar topologies. The octapeptides are located in two sites, 1 and 2, of domain I. In site 1, the octapeptide forms an antiparallel β-sheet with CTD. The negatively charged residues of the EEVD motif in the octapeptide form electrostatic interactions with the positively charged Lys residues of CTD. The Ile side chain of the octapeptide fits into the narrow concave formed by the hydrophobic residues of CTD. In site 2, the octapeptide also forms an antiparallel β-sheet with CTD, and the EEVD motif forms electrostatic interactions. The side chains of Pro and Ile of the octapeptide interact with the hydrophobic surface region of CTD site 2, which is broader and shallower than the concave binding region of site 1. This region seems to be capable of binding hydrophobic side chains that are bulkier than the Ile side chain. The roles of these two peptide-binding sites of Hdj1 are discussed.

The molecular chaperone Hsp70[1] and its cochaperone Hsp40 play important roles in cellular processes that include protein refolding, assembly, trafficking, degradation, and suppression of aggregation of non-native polypeptides (*1*−*4*). The Hsp40 proteins are classified into three subtypes based on their domain composition (*5*). Type I Hsp40 proteins, such as *Escherichia coli* DnaJ, yeast Ydj1, and human Hdj2, contain an N-terminal J-domain, a G/F-rich region, two zinc-finger-like motifs, and a conserved C-terminal peptide-binding domain (CTD). Type II Hsp40 proteins, such as yeast Sis1 and human Hdj1, contain the aforementioned domains but the zinc finger-like motifs are replaced with the G/M-rich region (*6*). Type I and type II Hsp40 proteins are not functionally equivalent although both types of Hsp40 proteins directly bind to non-native polypeptides. Type II Hsp40 proteins require Hsp70 proteins to suppress peptide aggregation (*7*), while type I Hsp40 proteins can suppress the aggregation of non-native polypeptides without the involvement of Hsp70 (*8*). Type III Hsp40 proteins consist of the J-domain and a specialized domain such as a clathrin-binding domain found in auxillin, a mammalian neuronal-specific protein (*9*).

During the protein refolding reaction, Hsp40 proteins recognize the hydrophobic side chains of non-native polypeptides and the well-conserved C-terminal EEVD motif of Hsp70 proteins (*10*, *11*) and facilitate the delivery of the non-native polypeptides to Hsp70 proteins (*12*). Hsp70 proteins then refold the non-native polypeptides using the energy derived from ATP hydrolysis that is activated by the J-domains of Hsp40 proteins (*1*). Deletion of the EEVD motif of Hsp70 abolishes protein refolding (*13*), indicating that the recognition of the EEVD motif by Hsp40 is one of the key steps of the protein refolding reaction.

The crystal structures of yeast Hsp40 proteins have been reported: yeast type II Hsp40 of Sis1 complexed with the C-terminal domain of Ssa1 at 3.2 Å resolution (*14*) (Protein Data Bank (PDB) ID 2B26) and yeast type I Hsp40 of Ydj1 complexed with the substrate model peptide at 2.7 Å resolution (*15*) (PDB ID 1NLT). Based on these crystal structures and the results of the mutation analysis (*16*), it was proposed that the recognition sites toward the EEVD motif and toward non-native polypeptides are common. Moreover, during the refolding reaction, Hsp70 proteins replace the non-native polypeptides bound to Hsp40 proteins with the EEVD motif and receive the non-native polypeptides for refolding. It remains unclear; however, why Hsp40 proteins must homodimerize to function as a cochaperone. Recently, the crystal structure of the CTD of the human type II Hsp40 of Hdj1 (Lys 158−Ile 340) was determined at 2.7 Å resolution (*17*) (PDB ID 2QLD). However, it remains unknown how the human Hsp40 recognizes the Hsp70 and non-native polypeptides.

We report here the crystal structures of the CTD of the Hdj1 (Asp 161−Ile 340; $CTD_{161-340}$) at 1.85 Å resolution, its complexes with the octapeptide GPTIEEVD corresponding to C-terminal residues 634′−641′ of human Hsp70 HSPA1A (*18*) at 1.85 Å, and the CTD of the Hdj1 (Ser 151−Ile 340; $CTD_{151-340}$)

[†]This work was supported by Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT) Grants-in-Aid as well as a Protein 3000 grant (Y.S.) and in part by the Global COE Program "Medical System Innovation on Multidisciplinary Integration" from MEXT, Japan (H.S.).

*To whom correspondence should be addressed. Telephone: +81-3-5841-4842. Fax: +81-3-5841-4891. E-mail: fwkozo@mail.ecc.u-tokyo.ac.jp.

[1]Abbreviations: Hsp, heat shock protein; CTD, C-terminal peptide-binding domain; rms, root mean square; PMSF, phenylmethanesulfonyl fluoride; PDB, Protein Data Bank.

Published on Web 09/01/2010

Table 1: Data Collection and Crystallographic Refinement Statistics[a]

	$CTD_{161-340}$	Au derivative	complex I	complex II
construct	Asp 161−Ile 340	Asp 161−Ile 340	Asp 161−Ile 340	Ser 151−Ile 340
data collection				
space group	$P2_1$	$P2_1$	$P2_1$	$C2$
unit cell parameters				
a (Å)	103.3	103.4	104.7	111.1
b (Å)	40.9	41.0	41.1	50.9
c (Å)	62.7	62.7	62.9	90.2
β (deg)	96.6	97.4	97.4	114.8
no. of protomers in asymmetric unit	2	2	2	2
solvent content	0.62	0.62	0.62	0.54
wavelength (Å)	0.9780	0.9780	0.9780	1.5418
resolution range (Å)	102.6−1.85 (1.90−1.85)	100.0−3.50 (3.73−3.50)	103.7−1.85 (1.90−1.85)	81.9−2.51 (2.57−2.51)
no. of total observations	178574	14448	389035	90145
no. of unique observations	40223	5160	43147	14847
completeness (%)	94.3 (64.8)	68.3 (70.2)	99.1 (96.9)	98.2 (87.1)
redundancy	4.4 (2.2)	2.8 (2.7)	9.0 (3.4)	6.1 (4.6)
R_{merge} (I)[b]	0.101 (0.693)	0.046 (0.056)	0.056 (0.482)	0.058 (0.386)
average $I/\sigma(I)$	22.4 (1.3)	31.8 (20.1)	23.1 (1.5)	21.4 (3.3)
refinement				
no. of protein atoms	2616		2831	2919
no. of water molecules	215		231	96
no. of ligand atoms			158	240
R_{work}[c]/R_{free}[d]	0.228/0.247		0.225/0.247	0.220/0.269
rms deviations from ideal values				
bond lengths (Å)	0.016		0.015	0.016
bond angles (deg)	1.53		1.44	1.70
average B-factor (Å^2)	43.6		39.1	35.0
Ramachandran plot[e](%)				
most favored	92.5		92.8	90.8
additionally allowed	7.5		7.2	9.2

[a]Values in parentheses in second to fifth columns are for the highest resolution shells. [b]$R_{merge}(I) = \sum|I - \langle I\rangle|/\sum I$, where I is the observed diffraction intensity. [c]$R_{work} = \sum||F_o| - |F_c||/\sum|F_o|$. F_o and F_c are observed and calculated structure amplitudes, respectively. [d]R_{free} is the R-factor based on 5% of the data, which was excluded from refiment. [e]Determined by PROCHECK (*31*).

complexed with GPTIEEVD at 2.51 Å resolution. These structures provide detailed information that reveal how Hdj1 recognizes the Hsp70 C-terminus. The structures also reveal a new peptide-binding site that might interact with hydrophobic polypeptides. Based on the structures, the role of homodimeric Hsp40 as a cochaperone during the protein refolding reaction is discussed.

EXPERIMENTAL PROCEDURES

Preparation and Crystallization of CTD. The purification and crystallization of $CTD_{161-340}$ and $CTD_{151-340}$ were described elsewhere. Briefly, $CTD_{161-340}$ is expressed using *E. coli* and purified by ammonium sulfate fractionation, cation-exchange column chromatography, and gel-filtration chromatography. $CTD_{151-340}$ was prepared by trypsin digestion of full-length Hdj1 expressed using *E. coli* and purified in identical manner as the purification method of $CTD_{161-340}$. Crystals of $CTD_{161-340}$ were grown from the equivolume mixture of 10 mg/mL $CTD_{161-340}$ in buffer A (100 mM sodium chloride, 20 mM Hepes, pH 7.5) and the reservoir solution containing 13% (mg/mL) PEG 3350 and 100 mM sodium citrate, pH 6.0 at 20 °C. Crystals of $CTD_{161-340}$ complexed with the C-terminal octapeptide of human Hsp70, complex I crystals, were grown from the equivolume mixture of 10 mg/mL $CTD_{161-340}$ in buffer A containing 4 mM octapeptide and the reservoir solution containing 18% (mg/mL) PEG 3350 and 100 mM sodium citrate, pH 6.0 at 4 °C. Crystals of $CTD_{151-340}$ complexed with the octapeptide, complex II crystals, were grown from the equivolume mixture of 20 mg/mL $CTD_{151-340}$ in buffer A containing 4.5 mM octapeptide and the reservoir solution containing 18% (mg/mL) PEG 3350 and 100 mM CHES, pH 8.8 at 20 °C. The molar ratios of the octapeptide to CTD protomer are 8.6:1 in the crystallization of complex I and 4.8:1 in the crystallization of complex II.

X-ray Diffraction Data Collection, Structure Determination, and Crystallographic Refinement. Diffraction data of crystals of $CTD_{161-340}$ and complex I were collected at beamlines 5A and 6A at the Photon Factory (Tsukuba, Japan). Prior to the data collection, the $CTD_{161-340}$ crystals were soaked for a few seconds in the reservoir solution containing 30% glycerol and flash-cooled to 95 K. The complex I crystals were flash-cooled in the same way. For phasing, Au derivative crystals were prepared by soaking $CTD_{161-340}$ crystals in the reservoir solution containing 10 mM $KAu(CN)_4$ for 1 min and flash-cooled as described above. The diffraction data processing was carried out using the HKL2000 package (*19*). Initial phases for the diffraction data of $CTD_{161-340}$ crystal were calculated using the program MLPHARE implemented in the CCP4 program package (*20*) and then improved by solvent flattening, histogram matching, and molecular averaging using the program DM. Model building and adjustment were carried out using the program COOT (*21*). Crystallographic refinement was performed using the programs CNS (*22*) and REFMAC (*23*) implemented in the CCP4 package in the final stages. A starting model of the refinement of the crystal structure of complex I was the refined crystal structure of $CTD_{161-340}$, since they were nearly isomorphous to each other. The crystallographic refinement was performed using REFMAC.

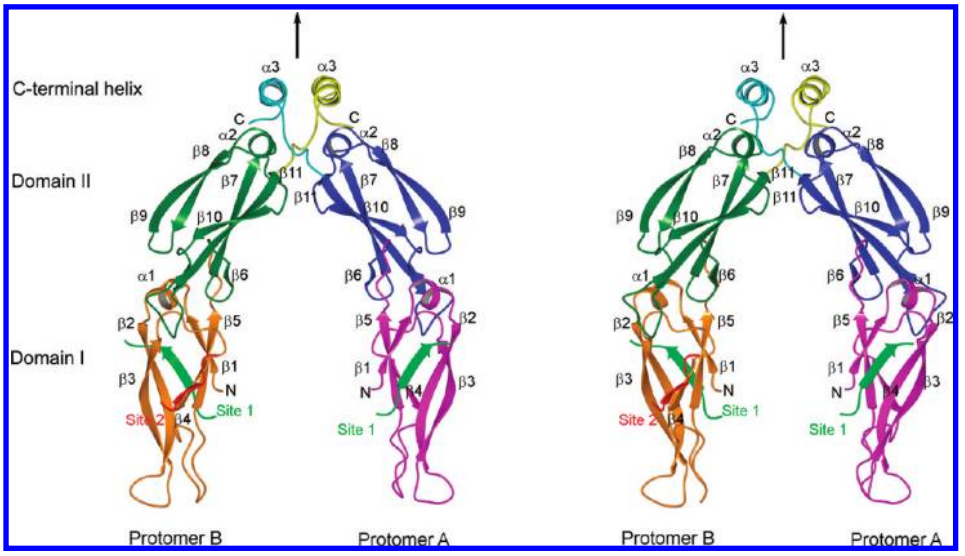

FIGURE 1: Stereo ribbon drawing of the complex I homodimer. Domain I of protomers A and B is drawn in magenta and orange, domain II in blue and green, and C-terminal helix in cyan and yellow, respectively. The octapeptides bound to site 1 and site 2 of protomer B are shown in light green and red, respectively. The octapeptide bound only to site 1 of protomer A is shown in light green. The two protomers are related by a pseudo-2-fold axis indicated with a black arrow on the plane of the drawing. The locations of the N- and C-termini are indicated by N and C, respectively. The α-helices and β-strands are labeled with their identification codes.

Diffraction data of complex II were collected using the Cu rotating-anode X-ray generator and the R-AXIS IV detector (Rigaku, Tokyo, Japan). Prior to the data collection, the crystal of complex II was soaked for a few seconds in the reservoir solution containing 15% PEG 20000 and 10% PEG 400 and was flash-cooled to 120 K. The diffraction data processing was carried out using the HKL2000 package. The crystal structure of complex II was determined with the molecular replacement method using the refined $CTD_{161-340}$ structure as a search model and the program MOLREP implemented in the CCP4 package. The crystallographic refinement was performed as in the refinement of the $CTD_{161-340}$ crystal structure.

The atomic coordinates and diffraction data of the $CTD_{161-340}$, complex I, and complex II crystal structures have been deposited in the Protein Data Bank with PDB IDs of 3AGX, 3AGY, and 3AGZ, respectively.

RESULTS

Overall Structure. The crystal structure of $CTD_{161-340}$ was determined using the single isomorphous replacement method combined with anomalous scattering and was refined at 1.85 Å resolution (Table 1). Two protomers, A and B, exist in the asymmetric unit of the $P2_1$ space group. The N-terminal five residues of each protomer and Ser 186−Ile 203 and Glu 223−Pro 232 regions of protomer B were not visible in the electron density map, and hence they are thought to be disordered in the crystal.

The structure of complex I was refined at 1.85 Å resolution. The Thr 165−Ile 340 region of each protomer and three octapeptides were located in the electron density map. Among the three octapeptides, two were bound to protomer B and one to protomer A. Asp 641′ of the octapeptide bound to protomer A and Gly 634′ and Asp 641′ of the octapeptide bound to site 2, as described below, of protomer B were not included in the model because their electron densities were low.

The crystal structure of $CTD_{151-340}$ complexed with the octapeptide obtained as complex II crystals was determined using the molecular replacement method and refined at 2.51 Å resolution. Two protomers, A and B, exist in the asymmetric unit of the *C*2 space group. Ala 156−Ile 340 of protomer A, Lys 159−Ile 340 of protomer B, and all of the residues of the two octapeptides bound to each protomer were located in the electron density map.

CTD exists as a twisted, horseshoe-shaped homodimer as shown in Figure 1. The CTD protomer consists of 11 β-strands (β1−β11) and 3 α-helices (α1, α2, and α3) and is folded into elongated globular domains I (residues 165−241) and II (250−320) and a C-terminal helix region (321−340). Domain I consists of a three-stranded β-sheet (β1, β4, and β5), a two-turn α-helix (α1), and a two-stranded antiparallel β-sheet (β2 and β3). Domain II consists of a four-stranded β-sheet (β6, β7, β10, and β11), a two-turn α-helix (α2), and a two-stranded antiparallel β-sheet (β8 and β9). The topology of domains I and II is common, except for the short β-strand, β6, of domain II. The C-terminal helix region consists of the C-terminal three-turn α-helix (α3) and the loop region (residues 321−327) preceding the helix. α2 in domain II and the C-terminal helix region constitute the dimerization interface (Figure 2). Hydrophobic side chains of Leu 262 and Leu 266 of α2 and Phe 322, Pro 323, Ile 326, Val 333, Leu 334, Val 337, Leu 338, and Pro 339 from the C-terminal helix region form a rigid hydrophobic core. Protomers A and B are related to each other by a noncrystallographic 2-fold axis in the complex I structure. In the complex II structure, the domains II and the C-terminal helix regions of protomers A and B are related to each other by a noncrystallographic 2-fold axis, but domain I is not.

When the $CTD_{161-340}$ structure is compared with the 2QLD structure (*17*), two differences are found. In the $CTD_{161-340}$ structure, a two-stranded antiparallel β-sheet of domain I comprises two β-strands of 12 amino acid residues (β2, Cys 179−Leu 190, and β3, Ile 197−Val 208), whereas in the 2QLD structure, this region consists of two shorter β-strands of 8 amino acid residues (β2, Cys 179−Ser 186, and β3, Asp 201−Val 208). In the 2QLD structure, the α3 helix at the dimer interface is one turn shorter than that of the $CTD_{161-340}$ structure because Gln 336−Ile 340 are not included in the model. The structure determination using diffraction data of higher resolution and

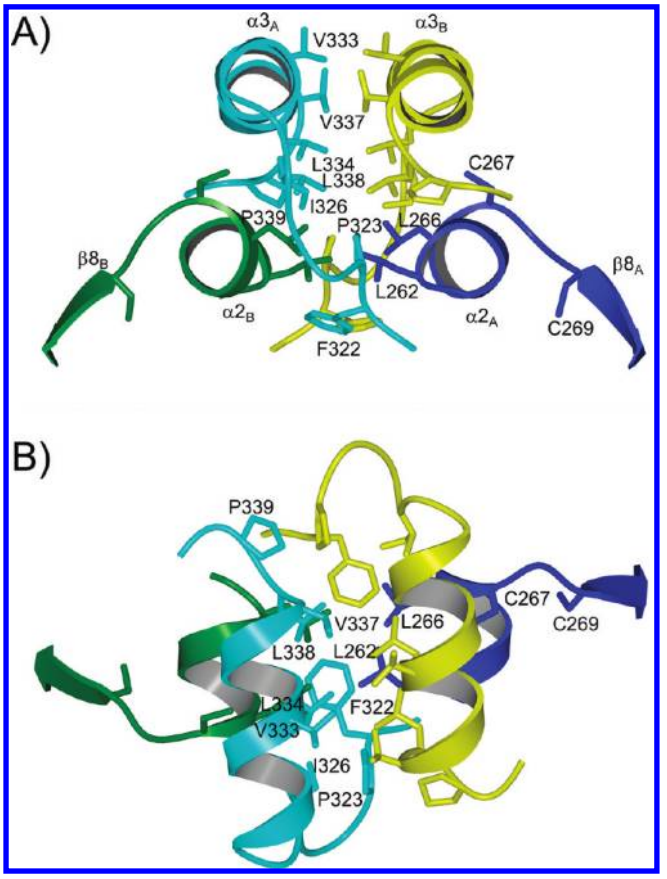

FIGURE 2: The dimeric interface between protomers A and B. (A) The orientations and color representations of the molecules are the same as those in Figure 1. The side chains of the residues at the dimeric interface and the two cysteines are drawn as stick models. (B) A top view of the dimeric interface.

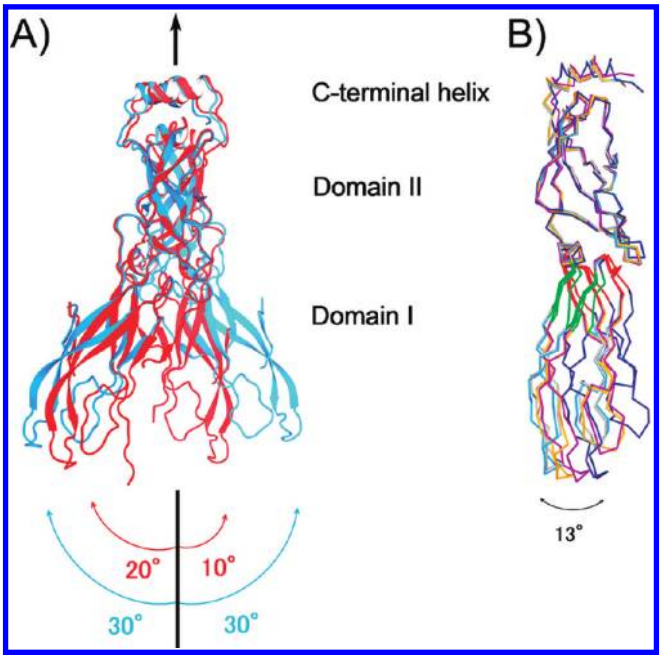

FIGURE 3: Comparison among the CTD structures. (A) The complex I structure (cyan) is superposed with the complex II structure (red) using the main-chain atoms of domain II of protomer A. Protomers B are presented in pale coloring. The view is rotated 90° around the pseudo-2-fold axis with respect to that of Figure 1. The 2-fold axis is shown with a black arrow. The angles formed by the pseudo-2-fold axis and the long axis of both elongated globular domains I of protomers A and B are shown. (B) Superposition of the five protomers of the three CTD structures. The protomer B of the $CTD_{161-340}$ structure is not shown. For clarity, residues Ala 156−Val 164 of protomer A and Lys 159−Val 164 of protomer B in complex II are not shown. Each protomer is superposed using the main-chain atoms of domain II. Protomer A of the $CTD_{161-340}$ structure is shown in gray, protomers A and B of complex I are shown in cyan and orange, respectively, and those of protomers A and B of complex II are shown in violet and blue, respectively. The region of Gly 178−Lys 181 and Ile 206−Lys 209 and that of Leu 168−Ser 171, Leu 239−Lys 242, and Gly 211−Gly 215 are shown in green and red, respectively. The largest angular differences formed by domain I of $CTD_{161-340}$ protomer A and that of complex II protomer B are shown with a black arrow.

higher completeness in this study gave the detailed structural information, and this may explain the structural differences observed between the two structures.

Domain I Orientation. Structures of domain II and the C-terminal helix regions of complexes I and II superimpose well with a root-mean-square (rms) positional difference of the main-chain atoms of 0.37 Å. The orientation of domain I with respect to domain II varies between the two complex structures and between two protomers of the complex II structure. In the complex I structure, the angles formed by the noncrystallographic 2-fold axis and the long axis of the elongated globular domains I of the protomers A and B are approximately 30° as shown in Figure 3A. In the complex II structure, these angles are 20° and 10°, and the cleft between the domains I is narrower than that of the complex I structure. The widths of the cleft represented by the distances between the C_α atoms of Asn 230 of protomers A and B are approximately 45 and 34 Å in complexes I and II, respectively.

When protomers of the complex I and II structures and protomer A of the $CTD_{161-340}$ structure are superposed, using the main-chain atoms of domains II, α1 (Leu 172−Ser 177) in domain I, and the linker regions between domains I and II (Lys 242−Phe 247), a good superposition is observed with rms positional differences of the main-chain atoms in the range of 0.09−0.67 Å. However, two β-sheets in domain I do not superpose well, as shown in Figure 3B. The largest angular difference of 13° is observed between protomer A of $CTD_{161-340}$ and protomer B of complex II. The variations in the orientation of the β-sheets of domain I mainly originate from the differences in the main-chain torsion angles of the regions at the end of the β-sheets near α1: Gly 178−Lys 181, Ile 206−Lys 209, Leu 168−Ser 171, Leu 239−Lys 242, and Gly 211−Gly 215.

Interactions between the CTD and the Octapeptides. The octapeptides were bound to two distinct sites of domain I, sites 1 and 2, of each protomer (Supporting Information Figure S1). These two sites of protomer A are named as sites 1A and 2A and those of protomer B as sites 1B and 2B. Of the sites 1 and sites 2 of complexes I and II, only site 2A of complex I is not occupied by the octapeptide, and instead Pro192 of a neighboring molecule is located there.

The octapeptide bound to site 1B adopts an extended conformation and forms an antiparallel β-sheet with β2 of domain I. The main-chain atoms of Thr 636′, Glu 638′, and Val 640′ form five hydrogen bonds with the main-chain atoms of Lys 182, Lys 184, and Ser 186 of β2 (Figure 4). The negatively charged side chains of Glu 638′ and Glu 639′ form salt bridges with the side chains of Lys 184 and Lys 181, respectively, and the C-terminal carboxyl group of Asp 641′ with Lys 182. Pro 635′ is surrounded by the hydrophobic side chains of Val 164, Ile 185, and Ile 235. The Ile 637′ side chain is located in the hydrophobic concave

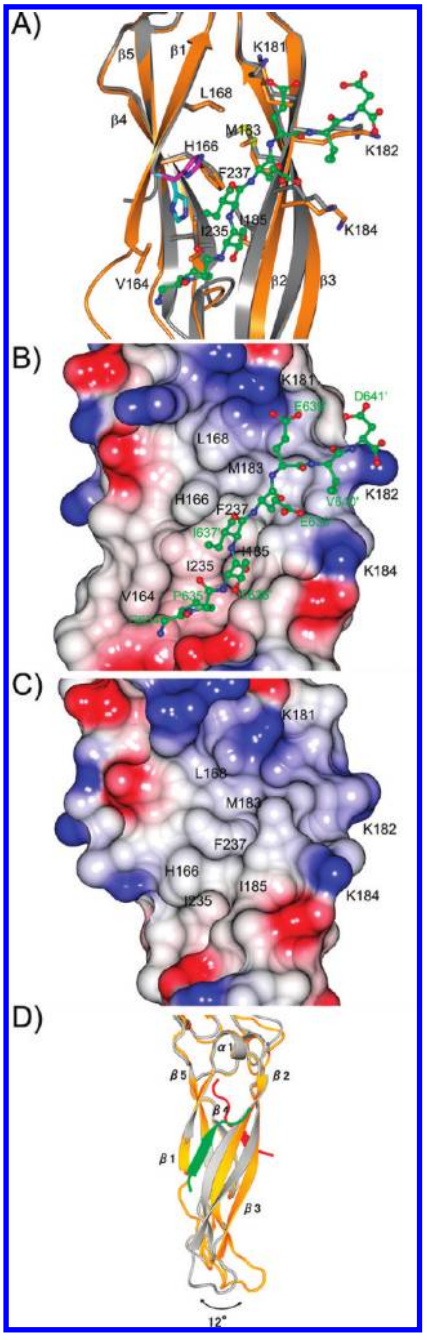

FIGURE 4: Interactions between CTD and the octapeptide bound to site 1. (A) A ribbon drawing of domain I of the $CTD_{161-340}$ protomer A and the complex II protomer B is shown in gray and orange, respectively. The β5 strand of $CTD_{161-340}$ protomer A is superimposed on that of the complex II protomer B. The side chains of the residues interacting with the octapeptide are illustrated as stick models: N in blue, C in gray or in orange, and S in yellow. The bound octapeptide is shown as a ball-and-stick model: O in red, N in blue, and C in green. The carbon atoms of the His 166 side chain of $CTD_{161-340}$ and complex II are shown in cyan and magenta, respectively. (B, C) Electrostatic potential surface of the complex II structure (B) and the $CTD_{161-340}$ structure (C) as viewed in the same orientation as (A). Positive and negative potentials are shown in blue and red, respectively. The bound octapeptide is shown as a ball-and-stick model. (D) The conformational change upon the octapeptide binding to site 1. The domains I of protomer A of the $CTD_{161-340}$ structure (gray) and protomer B of the complex II structure (orange) are superposed using the main-chain atoms of the β5 strand. The octapeptides bound to site 1 and site 2 of protomer A are shown in green and red, respectively. The red arrow shows the angular shift of the β-sheet consisting of β2 and β3 strands of the complex II structure upon octapeptide binding.

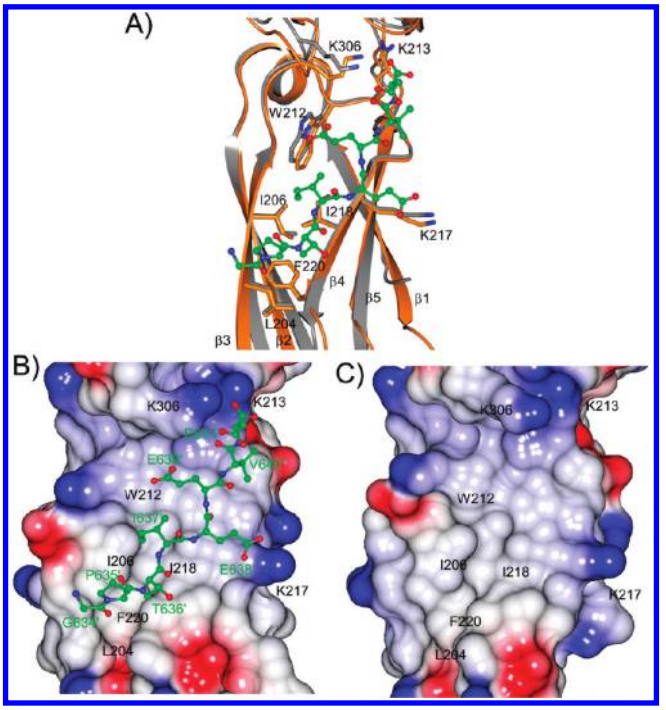

FIGURE 5: Interactions between CTD and the octapeptide bound to site 2. (A) Ribbon drawings of the $CTD_{161-340}$ protomer A and the complex II protomer B are shown in gray and orange, respectively. (B, C) Electrostatic potential surface of the complex II structure (B) and the $CTD_{161-340}$ structure (C) as viewed in the same orientation as (A). The bound octapeptide is shown as a ball-and-stick model. The color representation is the same as that in Figure 4.

surface formed by Val 164, His 166, Met 183, Ile 185, Ile 235, and Phe 237. The His 166 side chain covers the concave surface in the $CTD_{161-340}$ structure. Upon the octapeptide binding, His 166 flips away from the concave surface, and Ile 637′ covers the concave. The C_α and carbonyl C atoms of the N-terminal Gly 634′ are in contact with the Asp 225 side chain, and the Gly 634′ amino group protrudes into the solvent.

The changes in the conformation of domain I are induced by the octapeptide binding to site 1. When β5 of protomer A of the $CTD_{161-340}$ structure is superposed on that of protomer A of the complex II structure, the β-sheet consisting of β1, β5, and β4 strands is also superposed as shown in Figure 4C. In contrast, the β-sheet consisting of β2 and β3 strands of protomer A of the complex II structure shift by approximately 12° as compared with that of protomer A of the $CTD_{161-340}$ structure, and the groove between β1 and β2, where the octapeptide is bound, is wider by about 1.5 Å.

The octapeptide bound to site 2A of the complex II structure is also adopts an extended conformation and forms an antiparallel β-sheet with β4 of domain I (Figure 5). The main-chain atoms of Thr 636′ and Glu 638′ form four hydrogen bonds with the main-chain atoms of Lys 217 and Thr 219 of β4. Of the negatively charged moieties of the EEVD motif, the side chains of Glu 638′ and Asp 641′ form salt bridges with Lys 217 and Lys 213, respectively, and the C-terminal carboxyl group of Asp 641′ with Lys 306. The side chain of Glu 639′ is in contact with the side chain of Trp 212. The side chains of Pro 635′ and Ile 637′ are located on the broad hydrophobic surface, approximately 8 Å in width and 15 Å in length, formed by the side chains of Ile 206, Trp 212, Ile 218, and Phe 220. Two conformers are observed for the Ile 637′ side chain at site 2. In site 2A of the complex II structure, the C_δ atom of Ile 637′ protrudes toward the Ile 218 side chain

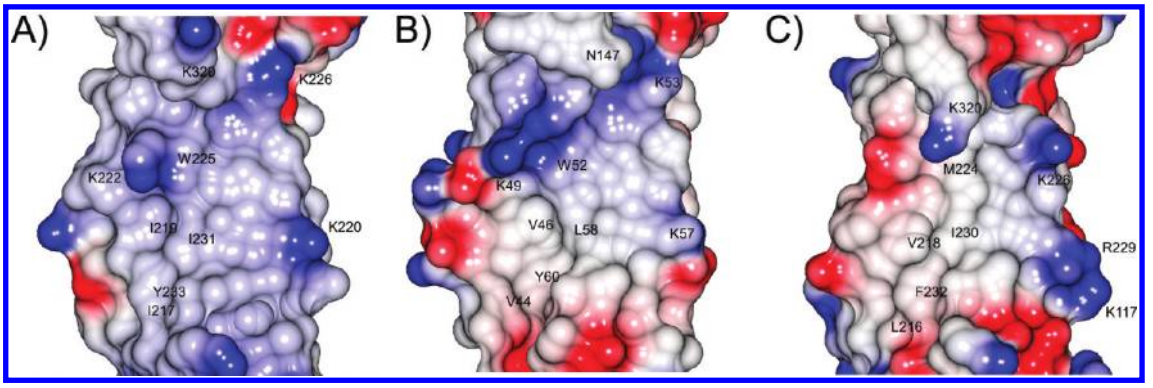

FIGURE 6: Electrostatic potential representation of homologous Hsp40 proteins. Each site 2 is viewed from the same direction as in Figure 5B. (A) Yeast type II Hsp40, Sis1 (PDB ID 1C3G). (B) *C. parvum* type II Hsp40 (2Q2G). (C) Yeast type I Hsp40, Ydj1 (1NLT). Positive and negative potentials are shown in blue and red, respectively. Residues corresponding to those of Hdj1 involved in the interaction with the bound octapeptides are indicated with their labels.

which is located at the bottom of site 2A. On the other hand, the C_δ atoms of Ile 637′ in site 2B of the complex I and II structures protrude into the solvent. These two conformations are in marked contrast to the Ile 637′ side chain in site 1 which is positioned merely in the narrow hydrophobic concave surface. The N-terminal Gly 634′ bound to site 2A is not in direct interaction with CTD. No conformational change is observed around site 2A upon the binding of the octapeptide.

DISCUSSION

In the $CTD_{161-340}$ crystal structure, residues 186−203 and 223−232 of domain I of protomer B were not visible in the electron density map. This is presumably due to disorder in the crystal. In contrast, most of domain I of protomer B in the complex I structure, that is nearly isomorphous to the $CTD_{161-340}$ structure, was visible in the electron density map. These observations suggest that the flexibility of domain I is reduced upon binding of the octapeptide. Structural comparison of complexes I and II reveals the wide variation in the orientation of the two β-sheets of domain I as shown in Figure 3B, suggesting that CTD still retains the ability to rearrange the orientation of domain I even after formation of the complex with the octapeptides.

The octapeptides are bound to the two distinct sites, 1 and 2, of domain I, and the EEVD motif of each octapeptide forms salt bridges with the side chains of Lys residues of domain I. The hydrophobic interactions at sites 1 and 2 are markedly different. At site 1, the Ile 637′ side chain is positioned in the hydrophobic concave surface which emerges due to the side-chain rotation of His 166 and is almost inaccessible to the solvent. As the shape and size of the concave surface are fit for accommodating the Ile side chain, this surface appears to specifically recognize Ile 637′. At site 2, the Ile 637′ side chain, together with Pro 635′, is located on the hydrophobic surface region of domain I, and their side chains are accessible to the solvent. The hydrophobic region of site 2 is broader and shallower than that of site 1. As such, polypeptides with hydrophobic side chains bulkier than isoleucine, such as phenylalanine and tyrosine, could potentially bind in this region. Therefore, site 1 is thought to be the specific recognition site for the Hsp70 C-terminal region, whereas site 2 recognizes non-native polypeptide substrates. In the homologous yeast Hsp40 Sis1, the binding site to the C-terminal region of yeast Hsp70 Ssa1 corresponds to site 1 of Hdj1 (*14*). Molecular chaperones GroEL and DnaK are known to use hydrophobic patches (*24*) and grooves (*25*) to form transient complexes with exposed hydrophobic residues in non-native polypeptides. As in these chaperones, the hydrophobic site 2 of Hdj1 may be involved in the recognition of hydrophobic non-native polypeptides. In the complex I structure, no octapeptide is bound to site 2A. Instead, Pro 192 of a neighboring molecule forms a hydrophobic interaction with Phe 220 at site 2A. This suggests that site 2 has affinity toward hydrophobic residues and supports the idea that site 2 is the recognition site for hydrophobic non-native polypeptides. In previously determined structures of yeast type I and type II Hsp40 proteins, Ydj1 and Sis1, respectively, and *Cryptosporidium parvum* type II Hsp40 (PDB ID 2Q2G), broad hydrophobic regions corresponding to site 2 of Hdj1 are identified, as shown in Figure 6. The broad hydrophobic site 2 may be commonly present among other type I and type II Hsp40 proteins.

Site 1 of Hsp40 has been considered to be the binding site for non-native polypeptides. This is because it has been reported that mutations in site 1 residues of yeast Hsp40 proteins Ydj1 and Sis1 reduce the binding affinity toward non-native polypeptides (*11*, *16*). The crystal structures of Hdj1 reported here clearly demonstrate that site 1 is located just behind the newly found peptide-binding site 2 which is generally present among Hsp40 proteins and that the hydrophobic residues at site 1 form van der Waals contacts with the residues at the bottom of site 2. In the mutational analysis of yeast Hsp40s, Phe 251 of Sis1 and Phe 249 of Ydj1, which correspond to Phe 237 of Hdj1, are mutated to Ser and Ala, respectively. The Phe residues are buried at the bottom of site 1 of Hsp40s and form van der Waals contacts with the site 2 residues of Ile 217, Ile 219, and Tyr 233 of Sis1 and site 2 residues Leu 216, Val 218, and Phe 232 of Ydj1 (Supporting Information Figure S2). These site 2 residues correspond to Leu 204, Ile 206, and Phe 220 of Hdj1. The site 1 residues Ile 203 of Sis1 and Leu 216 of Ydj1, which respectively correspond to Ile 185 and Leu 204 of Hdj1 and are mutated to Thr and Ala in the mutational analysis of yeast Hsp40s, also form van der Waals contacts with the site 2 residues. Mutations of these site 1 residues of yeast Hsp40s, which disrupt the structure of site 1 of yeast Hsp40 proteins, would simultaneously disrupt the structure of site 2. This could explain why mutations at site 1 of yeast Hsp40 proteins result in a reduction in the affinity toward non-native polypeptides. As such, the mutagenesis study of yeast Hsp40 proteins cannot conclude unambiguously that site 1 is the binding site for the non-native polypeptides. It appears that site 2 with the broad hydrophobic surface region is more suitable for recognizing the non-native polypeptides than site 1, which has the narrow hydrophobic concave surface.

Hsp40 proteins function as a homodimer, and the chaperone activity of the monomeric form is dysfunctional (*10*). This indicates that even if one protomer of Hdj1 has both sites 1 and 2, it should not function as a cochaperone and that site 1 of one protomer and site 2 of another protomer cooperate in the protein refolding reaction. The Hdj1 homodimer, which has two site 1−site 2 pairs (site 1A−site 2B and site 1B−site 2A), may have half-of-the-sites binding, as observed in GroEL (*26*). Half-of-the-sites binding at site 1 and site 2 in Hdj1 would be consistent with the observations that the Hsp40 homodimer binds to the monomeric Hsp70 stoichiometrically (*27*) and that one of the sites 2 in the complex I structure is not occupied by the octapeptide. Although octapeptides are bound to both sites 1 in the complex I structure and to all sites 1 and sites 2 in the complex II structure, this observation is probably due to the high molar ratio of the octapeptide to Hdj1.

When sites 1 and 2 represent the binding site for Hsp70 and for the non-native polypeptide, respectively, the role of the Hdj1 homodimer in the chaperone mechanism is thought to be as follows. In the initial stages of the refolding reaction, one protomer of the Hdj1 homodimer binds through site 2 to the hydrophobic polypeptide that is capable of forming an antiparallel β-sheet with β4. Hdj1 recognizes the bound polypeptide as a non-native form to be refolded. The other protomer binds through site 1 to the Hsp70 C-terminal region. Thus, the non-native polypeptide substrate bound to one Hdj1 protomer and Hsp70 bound to the other Hdj1 protomer are brought in close proximity to each other by the Hdj1 homodimer. The flexibility in the orientation of domain I, which is suggested by the comparison of the CTD structure shown in Figure 3, may aid both the binding to the non-native polypeptide substrate that would be diverse in size and shape and the efficient transfer of the substrate to Hsp70. After the formation of the ternary complex of Hdj1, Hsp70, and the non-native polypeptide substrate, Hdj1 stimulates ATP hydrolysis by Hsp70 through the N-terminal J-domain and provides the peptide substrate for Hsp70 to refold.

The assignments of the ligands toward sites 1 and 2 are presumed based on the structural characteristics of the sites where the same octapeptides are bound; therefore, ambiguity in the assignment could not be completely excluded. To establish the roles of these sites, it would be necessary to carefully design Hdj1 point mutants in which site 1 or site 2 is disrupted while the other site remains unchanged, to confirm by the method of X-ray crystallography that only one of the sites is disrupted, and to determine their binding affinities toward Hsp70 and toward non-native polypeptides.

Recently, Hsp40 proteins have been shown to bind to their specific client proteins other than Hsp70 proteins and non-native polypeptides: human Hsp40 DNAJB5 specifically forms a multiprotein complex with the class II histone deacetylase 4 and thioredoxin 1 in heart cells and regulates cardiac hypertrophy (*28*). Two cysteine residues of DNAJB5 are essential for the binding to the class II histone deacetylase 4, and the cysteine residues are conserved among homologous human Hsp40 proteins, Hdj1, DNAJB4 (*29*), and DNAJB5 (*30*). In Hdj1, the conserved cysteines are Cys 267 and Cys 269, and they are located adjacent to the dimerizing C-terminal helices α3, as shown in Figure 2. This suggests that the dimerizing C-terminal helices of the Hsp40 proteins are a part of the binding sites for their specific client proteins. Among the three homologous human Hsp40 proteins, amino acid identities of the C-terminal helices are 38%, whereas those of domains I and II are much higher, 77−82% and 61−67%, respectively. The low sequence identity of the C-terminal helices might reflect the differences in the client protein specificity of human Hsp40 proteins.

ACKNOWLEDGMENT

The cDNA of human Hsp40 Hdj1 was kindly provided by Prof. K. Ohtsuka of Chubu University. The synchrotron radiation experiments were performed under the approval of the Photon Factory Program Advisory Committee (Proposal No. 2008G141).

SUPPORTING INFORMATION AVAILABLE

$F_o - F_c$ Fourier maps of the octapeptides in complex I and II structures; sequential and structural comparison of the peptide-binding sites of Hdj1, Sis1, and Ydj1. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Frydman, J., Nimmesgern, E., Ohtsuka, K., and Hartl, F. U. (1994) Folding of nascent polypeptide chains in a high molecular mass assembly with molecular chaperones. *Nature 370*, 111–117.
2. Hartl, F. U. (1996) Molecular chaperones in cellular protein folding. *Nature 381*, 571–580.
3. Bukau, B., and Horwich, A. L. (1998) The Hsp70 and Hsp60 chaperone machines. *Cell 92*, 351–366.
4. Hartl, F. U., and Hayar-Hartl, M. (2002) Molecular chaperones in the cytosol: From nascent chain to folded protein. *Science 295*, 1852–1858.
5. Cheetham, M. E., and Caplan, A. J. (1998) Structure, function and evolution of DnaJ: Conservation and adaptation of chaperone function. *Cell Stress Chaperone 3*, 28–36.
6. Cry, D. M., Langer, T., and Douglas, M. G. (1994) DnaJ-like proteins: Molecular chaperones and specific regulators of Hsp70. *Trends Biochem. Sci. 19*, 176–181.
7. Minami, Y., Hohfeld, J., Ohtsuka, K., and Hartl, F. U. (1996) Regulation of the heat-shock protein 70 reaction cycle by the mammalian DnaJ homolog Hsp40. *J. Biol. Chem. 271*, 19617–19624.
8. Cry, D. M. (1995) Cooperation of the molecular chaperone Ydj1 with specific Hsp70 homologs to suppress protein aggregation. *FEBS Lett. 359*, 129–132.
9. Holstein, S. E., Ungewickell, H., and Ungewickell, E. (1996) Mechanism of clathrin basket dissociation: Separate functions of protein domains of the DnaJ homologue auxilin. *J. Cell Biol. 135*, 925–937.
10. Sha, B., Lee, S., and Cyr, D. M. (2000) The crystal structure of the peptide-binding fragment from the Hsp40 protein Sis1. *Structure 8*, 799–807.
11. Lee, S., Fan, C. Y., Younger, J. M., Ren, H., and Cry, D. M. (2002) Identification of essential residues in the type II Hsp40 Sis1 that function in polypeptide binding. *J. Biol. Chem. 277*, 21675–21682.
12. Han, W., and Christen, P. (2003) Mechanism of the targeting action of DnaJ in the DnaK molecular chaperone system. *J. Biol. Chem. 278*, 19038–19043.
13. Freeman, B. C., Myers, M. P., Schumacher, R., and Morimoto, R. I. (1995) Identification of a regulatory motif in Hsp70 that affects ATPase activity, substrate binding and interaction with HDJ-1. *EMBO J. 14*, 2281–2292.
14. Li, J., Wu, Y., Qian, X., and Sha, B. (2006) Crystal structure of yeast Sis1 peptide-binding fragment and Hsp70 Ssa1 *C*-terminal complex. *Biochem. J. 398*, 353–360.
15. Li, J., Qian, X., and Sha, B. (2003) The crystal structure of the yeast Hsp40 Ydj1 complexed with its peptide substrate. *Structure 11*, 1475–1483.
16. Li, J., and Sha, B. (2005) Structure-based mutagenesis studies of the peptide substrate binding fragment of type heat-shock protein 40. *Biochem. J. 386*, 453–460.
17. Hu, J., Wu, Y., Li, J., Qian, X., Fu, Z., and Sha, B. (2008) The crystal structure of the putative peptide-binding fragment from the human Hsp40 protein Hdj1. *BMC Struct. Biol. 8*, 3.
18. Hunt, C., and Morimoto, R. I. (1985) Conserved features of eukaryotic *hsp70* genes revealed by comparison with the nucleotide sequence of human hsp70. *Proc. Natl. Acad. Sci. U.S.A. 82*, 6455–6459.
19. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol. 276*, 344–358.

20. Collaborative Computational Project, Number 4 (1994) The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr. D50*, 760–763.
21. Emsley, P., and Cowtan, K. (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D60*, 2126–2132.
22. Brünger, A. T., Adams, P. D., Clore, G. M., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, N., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR System (CNS), a new software suite for macromolecular structure determination. *Acta Crystallogr. D54*, 905–921.
23. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D53*, 240–255.
24. Fenton, W. A., Kashi, Y., Furtak, K., and Horwich, A. L. (1994) Residues in chaperonin GroEL required for polypeptide binding and release. *Nature 371*, 614–619.
25. Zhu, X., Zhao, X., Burkholder, W. F., Gragerov, A., Ogata, C. M., Gottesman, M. E., and Hendrickson, W. A. (1996) Structural analysis of substrate binding by the molecular chaperone DnaK. *Science 272*, 1606–1614.
26. Li, J., and Wang, C. (1999) "Half of the sites" binding of D-glyceraldehyde-3-phosphate dehydrogenase folding intermediate with GroEL. *J. Biol. Chem. 274*, 10790–10794.
27. Qian, X., Hou, W., Zhengang, L., and Sha, B. (2002) Direct interactions between molecular chaperones heat-shock protein (Hsp) 70 and Hsp40: Yeast Hsp70 Ssa1 binds the extreme C-terminal region of yeast Hsp40 Sis1. *Biochem. J. 361*, 27–34.
28. Ago, T., Liu, T., Zhai, P., Chen, W., Li, H., Molkentin, J. D., Vatner, S. F., and Sadoshima, J. (2008) A redox-dependent pathway for regulating class II HDACs and cardiac hypertrophy. *Cell 133*, 978–993.
29. Hoe, K. L., Won, M., Chung, K. S., Jang, Y. J., Lee, S. B., Kim, D. U., Yun, J. H., and Yoo, H. S. (1998) Isolation of a new member of DnaJ-like heat shock protein 40 (Hsp40) from human liver. *Biochim. Biophys. Acta 1383*, 4–8.
30. Chen, M. S., Roti, J. R., and Laszlo, A. (1999) Hsc40, a new member of the hsp40 family, exhibits similar expression profile to that of Hsc70 in mammalian cells. *Gene 238*, 333–341.
31. Morris, A. L., MacArthur, M. W., Hutchinson, E. G., and Thornton, J. M. (1992) Stereochemical quality of protein structure coordinates. *Proteins 12*, 345–364.