Published in final edited form as:

Science. 2019 September 20; 365(6459): 1313–1319. doi:10.1126/science.aax1280.

Structural basis for client recognition and activity of Hsp40 chaperones

Yajun Jiang¹, Paolo Rossi¹, Charalampos G. Kalodimos^{1,*}

¹Department of Structural Biology, St. Jude Children's Research Hospital, Memphis, United States.

Abstract

Hsp70 and Hsp40 chaperones work synergistically in a wide range of biological processes including protein synthesis, membrane translocation, and folding. We used NMR spectroscopy to determine the solution structure and dynamic features of an Hsp40 in complex with an unfolded client protein. Atomic structures of the various binding sites in the client complexed to the binding domains of the Hsp40 reveal the recognition pattern. Hsp40 engages the client in a highly dynamic fashion using a multivalent binding mechanism that alters the folding properties of the client. Different Hsp40 family members have different numbers of client-binding sites with distinct sequence selectivity, providing additional mechanisms for activity regulation and function modification. Hsp70 binding to Hsp40 displaces the unfolded client. The activity of Hsp40 is altered in its complex with Hsp70 further regulating client binding and release.

One Sentence Summary:

The structure of an Hsp40 in complex with an unfolded protein reveals how the chaperone exerts its activity in concert with Hsp70.

Molecular chaperones perform essential tasks in the cell including preventing aggregation of non-native proteins and assisting with their folding (1, 2). They form a large family of unrelated proteins that come in a variety of architectures and sizes (3) resulting in distinct activities that can be further regulated by forming complexes with other chaperones. The dynamic nature of the interaction of chaperones with client proteins has presented a major challenge for the structure determination of their complexes (4–10) that has hindered mechanistic insight.

*Correspondence to: babis.kalodimos@StJude.org.

Author contributions: Y.J. and C.G.K. conceived the study and designed the experiments. Y.J. prepared all protein samples, determined the structures of the chaperone complexes by NMR, and performed the activity assays. P.R. assisted with the structure determination. Y.J. and C.G.K. wrote the manuscript.

Competing interests: The authors declare no competing interests.

Data and materials availability: Atomic coordinates and NMR chemical shifts have been deposited in the Protein data Bank (PDB) and Biological Magnetic Resonance Data Bank (BMRB), respectively, under the following accession codes: 6PSI and 30638 for full-length ttHsp40–PhoA, 6PPT and 30627 for ttHsp40^{CBD1}–PhoA^a, 6PQ2 and 30628 for ttHsp40^{CBD1}–PhoA^b, 6PRQ and 30637 for ttHsp40^{CBD1}–PhoA^c, 6PQE and 30629 for ttHsp40^{CBD2}–PhoA^d, 6PQM and 30632 for ttHsp40^{CBD2}–PhoA^e, 6PRI and 30634 for ttHsp40^{CBD2}–PhoA^f, 6PR5 and 30635 for ttHsp40^{CBD2}–PhoA^g, and 6PRP and 30636 for ttHsp40^{CBD2}–ttHsp70^{C-tail}. All other data are available in the main text or the supplementary materials. Requests for materials should be addressed to the corresponding author.

The Hsp70-Hsp40-NEF (nucleotide exchange factor) is a ubiquitous tripartite chaperone machinery that is involved in critical aspects of proteome quality processes (2, 11–13). Hsp70 uses ATP binding and hydrolysis, coupled to NEF-induced ADP release, to regulate its binding to and release of client proteins. Hsp40 (also referred to as DnaJ or J-domain proteins) serve various roles in the function of the machinery (11, 14). First, their J domain is essential for the stimulation of the ATPase activity of Hsp70. Second, they recognize and bind to non-native proteins, which then present to the Hsp70. Third, they are thought to diversify the function of Hsp70 by determining client specificity and possibly the fate of the client (e.g. refolding vs degradation) (14, 15). Most species carry many more Hsp40 than Hsp70 genes. For example, the human genome encodes more than 50 Hsp40s but only 11 Hsp70s (11). Members of the Hsp40 family can synergize with each other (16) and even function independently of Hsp70 (17) to suppress protein aggregation. Hsp40s have established roles in various pathologies including cancer and neurodegenerative diseases (18, 19). Despite their importance, it remains poorly understood how Hsp40s recognize and engage non-native proteins, how they modify the folding properties of the bound clients, and how they work together with Hsp70 to give rise to a functional chaperone machinery.

There are three major classes of Hsp40s: A, B, and C (14). Classes A and B are the most ubiquitous, being present in all three kingdoms and in most compartments of eukaryotic cells (fig. S1A). We have used NMR spectroscopy to characterize Hsp40 chaperones from various species including the *Thermus thermophilus* type B Hsp40 (*ttHsp40*), the *E. coli* type B Hsp40 (CbpA), the yeast type A (Ydj1) and type B Hsp40s (Sis1), and the human type B Hsp40 (DNAJB1). Type A and B Hsp40s encompass a J domain at the N terminus, a glycine/phenylalanine (G/F)-rich region, two tandem β -barrel domains, and a helix that mediates dimerization (fig. S1A). In addition, type A Hsp40s include a zinc-finger domain and a C-tail (fig. S1A). We selected *ttHsp40* for detailed structural studies because it yielded high quality NMR spectra (fig. S1B,C) in its complex with client proteins. We used NMR data to refine and complete the structure in solution of a previously reported (20) crystal structure of free *ttHsp40* (see Materials and Methods).

To probe the interaction of *ttHsp40* with unfolded proteins, we used two physiological substrates, alkaline phosphatase (PhoA) (471 amino acids) and maltose-binding protein (MBP) (396 amino acids) (21, 22), as client proteins. These proteins have been successfully used as model unfolded client proteins in complex with molecular chaperones to yield suitable samples for high-resolution NMR structural analyses (4, 5, 23). To determine the client-binding sites in *ttHsp40*, we titrated unfolded PhoA (fig. S2A) and MBP to isotopically labeled *ttHsp40*. The *ttHsp40* residues that interact with the client proteins form two distinct surfaces, one in each β -barrel domain (Fig. 1A). On this basis we refer to the β -barrel domains as client-binding domain 1 (CBD1) and 2 (CBD2). Both binding sites consist primarily of hydrophobic residues and are decorated by a number of polar residues (Fig. 1A and fig. S2B-E). Each of the binding sites exposes $\sim 650 \text{ \AA}^2$ of hydrophobic surface. Thus, *ttHsp40* has in total 4 substrate-binding sites that collectively expose $\sim 2600 \text{ \AA}^2$ of hydrophobic surface that can be used to engage unfolded proteins.

To identify the residues of PhoA and MBP that are recognized and bound by *ttHsp40*, we monitored by NMR the interaction of labeled unfolded PhoA and MBP by unlabeled

ttHsp40 (fig. S3), as we showed before for other chaperones (4, 5). NMR analysis indicated that there are several distinct ttHsp40-recognition sites in PhoA and MBP. In PhoA, seven of these sites (labeled *a* through *g*; Fig. 1B) interact relatively strongly with ttHsp40 whereas another eight sites (labeled *w1* through *w8*) interact weakly with ttHsp40 (Fig. 1B). Collectively, the amino acid composition of the ttHsp40-recognition sites within PhoA and MBP show that ttHsp40 has a preference for aromatic and large aliphatic residues (fig. S4A-C), in agreement with earlier peptide array experiments (24).

To better understand how ttHsp40 recognizes and binds to client proteins, we sought to determine whether ttHsp40 CBD1 and CBD2 have similar or distinct specificities. PhoA and MBP fragments encompassing individual chaperone-binding sites were studied for complex formation with isolated CBD1 or CBD2 (fig. S5). Interestingly, PhoA sites *a*, *b* and *c* bind exclusively to CBD1, PhoA sites *d* and *e* bind exclusively to CBD2, whereas PhoA sites *f* and *g* bind either to CBD1 or CBD2 with a preference towards the latter (fig. S5B). Experiments using MBP as client resulted in a similar observation with sites within MBP showing distinct specificity for CBD1 vs CBD2.

Using NMR spectroscopy, we determined the structure of ttHsp40 in complex with full-length unfolded PhoA (Fig. 2). The architecture of the complex depicted in Fig. 2 represents one of the several similar arrangements of PhoA with respect to ttHsp40 as supported by the experimental data (see below and fig. S6). PhoA sites *a* and *b* bind to CBD1 on one of the ttHsp40 subunits whereas PhoA site *c* binds to the CBD1 of the other subunit. PhoA sites *a* and *b* are adjacent in the primary sequence (Fig. 1B) and are thus positioned close to each other in the structure of the complex (Fig. 2). NMR characterization of a PhoA fragment that encompasses both sites *a* and *b* (PhoA^{*a-b*}) demonstrates that these two sites compete with each other for CBD1 and their subcomplexes are in dynamic equilibrium (fig. S7). The relative affinities of PhoA sites *a* and *b* for CBD1 are similar and the subcomplexes are equally populated (fig. S7). PhoA site *d* is bound to CBD2 on one subunit and PhoA sites *e*, *g*, and *f* on the C-terminal region of PhoA share CBD2 on the other subunit (Fig. 2). As with PhoA sites *a* and *b* (fig. S7), PhoA sites *e*, *g*, and *f* are in dynamic equilibrium, with each one of them competing with the other two for CBD2. The weak binding sites (*w1* through *w8*) interact only transiently with ttHsp40 and binding can be observed only in conjunction with the stronger binding sites (*a* through *g*).

To obtain high-resolution information on the recognition of the unfolded PhoA by ttHsp40, we determined the structure of each one of the seven PhoA sites (*a* through *g*) in complex with the corresponding ttHsp40 interacting domain (CBD1 or CBD2) by NMR (fig. S8). The structures are shown in Figure 2B. All ttHsp40–PhoA interacting surfaces are dominated by intimate non-polar contacts. In several of these sub-complexes, hydrogen bonds and salt bridges are formed at the periphery of the binding sites (Fig. 2B). ttHsp40 shows selectivity for side chains without backbone-backbone interactions to the bound client. This binding mode is in agreement with biochemical data (24) but in contrast to a crystal structure of the yeast Hsp40 chaperone Ydj1 (25) in complex with a peptide, wherein bulky hydrophobic residues are exposed to the solvent (fig. S9A). The client-binding sites in ttHsp40 can only accommodate short, typically 8-residue-long sequences consisting of a central cluster of hydrophobic residues that is flanked by polar residues on either side (fig. S4A,B). Formation

of the various sub-complexes buries a relatively small surface area ranging between ~500 and ~760 Å², of which 70–80% is non-polar highlighting the importance of direct interactions with exposed and contiguous hydrophobic surface in the recognition of unfolded proteins by ttHsp40. When all four substrate-binding sites are occupied, ttHsp40 buries up to 2300 Å² of surface area. This is much less than the surface buried by the trigger factor (TF) (4) (up to ~5500 Å²) or the SecB (5) (up to ~17000 Å²) chaperones.

The four binding sites in ttHsp40 span a distance of ~150 Å (Fig. 1C). The simultaneous engagement of several non-polar sites, which typically cluster within the hydrophobic core of the client protein in its native state, prevents the client protein from forming a collapsed structure (native or molten globule) and stabilizes an extended, non-native state (Fig 2). Both PhoA and MBP tend to retain residual secondary structure in their non-native state (fig. S4D,E); ttHsp40 does not show any selectivity for specific secondary structure (fig. S4D,E). Instead, the structural data suggest that binding of any client site results in melting of its residual secondary structure by ttHsp40. Thus, binding of client proteins to ttHsp40 results in disruption of their secondary and tertiary structure.

Analysis of the NMR spectra of PhoA in the presence of ttHsp40 (fig. S3A) shows that one ttHsp40 molecule can interact with all of the binding sites within PhoA. The stoichiometry of the complex is further corroborated by ITC experiments (fig. S2F). Given that ttHsp40 has fewer binding sites than PhoA, this clearly indicates a dynamic formation of the complex, as is typical of chaperone complexes with proteins in their non-native state (10, 26–28). We first sought to determine how fast the individual PhoA sites are released from the corresponding ttHsp40 binding site by using NMR relaxation dispersion (29). The results show that the individual PhoA sites, for example PhoA site *a* (fig. S10A), interact rapidly with ttHsp40 and they stay bound for less than 10 ms ($k_{\text{off}} \sim 100 \text{ s}^{-1}$). Kinetic measurements of the full length PhoA dissociation from ttHsp40 yielded a k_{off} of $\sim 10 \text{ s}^{-1}$ (fig. S10B). Thus, although simultaneous engagement of multiple sites within full-length PhoA by ttHsp40 results in an increased lifetime, the complex remains very dynamic with an overall residence time of ~100 ms (fig. S10C). During its lifetime, there are several dissociation and association events of the individual client sites with the four substrate binding sites in ttHsp40. The chaperone associates with all of the sites within PhoA, including, albeit very transiently, the weak sites. Such a binding mechanism may allow ttHsp40 to alter the structure, both secondary and local higher-order, of a large number of client regions using only four binding sites.

CBD1 and CBD2 adopt a very similar β -barrel-like structure (root-mean-square-deviation (r.m.s.d.) is 1.1 Å) and they share a sequence similarity of ~50% (fig. S11). Surprisingly, CBD1 and CBD2 use distinct surfaces to interact with non-native proteins: CBD1 uses a narrow hydrophobic groove positioned between β -strands 1 and 2 whereas CBD2 uses a wide hydrophobic groove located at the opposite surface of the domain positioned between β -strands 3 and 4 (fig. S11A,B). The domains exhibit subtle differences in amino acid selectivity: CBD2 recognizes hydrophobic sequences that feature at least one aromatic residue, whereas CBD1 prefers hydrophobic sequences that contain no aromatic residues (fig. S11C-E). Negatively charged residues preceding the aromatic residue in the client sequence increase its binding to CBD2 (fig. S11F,G).

Structural analysis of the ttHsp40 client-binding sites, together with sequence alignment data of other Hsp40s indicated that the number of client-binding sites may vary within this chaperone family (fig. S12A-C). To test this hypothesis, we characterized by NMR the binding properties of Hsp40s from various species (fig. S12D-G). The data showed that similarly to ttHsp40, the *E. coli* type B Hsp40 CbpA and the type A Hsp40 Ydj1 from yeast use both CBD1 and CBD2 to bind unfolded proteins. In contrast, type B Hsp40 Sis1 from yeast and the type B human Hsp40 DNAJB1 use only CBD1 (Fig. 1C and fig. S12D-G). The zinc finger domain in Ydj1, a domain that is present only in type A Hsp40s, also binds to clients (Fig. 1C and fig. S12H), corroborating previous biochemical experiments (30). Thus, Ydj1, and in principle other type A Hsp40s, may feature as many as six client-binding sites whereas type B Hsp40s can have as few as two (Fig. 1C). Interestingly, the client-binding site in Ydj1 CBD2 is occluded by an intra-subunit interaction with the C-tail (Fig. 1D and fig. S12I) (31). NMR titration experiments showed that client binding may result in negligible or partial release of this autoinhibitory mechanism depending on the length of the client protein and the availability of the other client-binding sites (fig. S12J). Many type A Hsp40s feature a hydrophobic sequence in their C-tail suggesting that this auto-inhibitory mechanism may be conserved (fig. S11K).

Type A and B Hsp40s, such as Ydj1 and Sis1, typically have distinct functional roles in the cell (14, 32). Our NMR binding data showed that Ydj1 and Sis1 have different sequence preferences and highlighted a number of motifs that are recognized by only one of the chaperones (fig. S4F). Given that CBD1 and CBD2 can exhibit distinct amino acid preferences even within the same Hsp40 (fig. S4), the variation in the number and composition of CBDs and the zinc finger among the various Hsp40s may result in distinct client selectivity and could account for the observed specialized functions of different Hsp40s (14, 33, 34).

Critical to the function of the Hsp70/Hsp40 machinery is the interaction between the two chaperones. Very little is known about how the full-length proteins interact with each other. We have characterized the interaction between full length ttHsp40 and ttHsp70 by NMR (fig. S13A,B). The data show that, in addition to the interaction mediated by the J domain and the nucleotide-binding domain (NBD) of Hsp70 (35), there is also a specific interaction between the C-terminal tail of ttHsp70 and CBD2 of ttHsp40 (Fig. 3A and fig. S13C,D). We determined the NMR structure of the complex between the isolated ttHsp70 C-tail and the ttHsp40 CBD2 (Fig. 3A). The structure shows that the crucial residues in ttHsp70 C-tail are D610 and Y611, which form hydrogen bonds with K238 and R246, respectively, of ttHsp40. In addition, the aromatic ring of Y611 is placed within a hydrophobic pocket wherein it forms intimate nonpolar contacts. The ttHsp70 C-tail substitutions D610K and Y611A (ttHsp70^{CM}) decreases binding to ttHsp40 40-fold (fig. S13E). This binding mode of ttHsp70 C-tail to ttHsp40 appears to be conserved in *E. coli* (fig. S13F). Our NMR data show that the C-tail of human Hsp70 (Hsc70) binds to the CBD1 of DNAJB1 (36) (fig. S13G) and the yeast Hsp70 Ssa1 C-tail binds to the CBD1 of Sis1 (37) and of Ydj1 (fig. S13H).

The binding sites of ttHsp70 C-tail and client proteins on ttHsp40 are overlapping, suggesting that the two proteins compete for binding to ttHsp40 (Fig. 3B,C and S13I). NMR

experiments show that in the presence of both ttHsp70 and PhoA, ttHsp40 CBD2 is occupied by the ttHsp70 C-tail and ttHsp40 CBD1 by PhoA, giving rise to a ternary complex (Fig. 3C and fig. S14A). In Sis1, the Hsp70 Ssa1 C-tail and the client compete for binding to CBD1 (Fig. 3D and S14B). Because Sis1 does not use CBD2 to bind to clients, Ssa1 binding displaces the client from Sis1 (Fig. 3D) as demonstrated by the NMR experiments (fig. S14B). The Ssa1 C-tail also binds to CBD1 of Ydj1 but because of the presence of the zinc finger, which binds to client proteins, a stable ternary complex can form in this case (Fig. 3D and S14C). Thus, Hsp40s can form complexes of varying stability with their cognate Hsp70s and clients depending on the number of client-binding sites within the Hsp40. This provides yet an additional level of activity regulation of the tripartite chaperone machinery.

We examined the role of ttHsp40 in the refolding of chemically denatured luciferase by the Hsp70/Hsp40/NEF (38). ttHsp40 is essential for the refolding activity of the machinery (Fig. 4A and fig. S15A,B). Amino acid substitutions in either of the ttHsp40 CBDs that impair client binding (figs. S1D,G and S15C) decreases dramatically the refolding efficiency, thus highlighting the importance of client binding by Hsp40s, in addition to their role in stimulating the ATPase activity of Hsp70s (Fig. 4A and fig. S15A). These findings may explain why Sis1, which has no client binding-sites available in its complex with the yeast Hsp70 Ssa1 (Fig. 3D), has a much lower refolding capacity than Ydj1 (39).

Disruption of the interaction between ttHsp40 and ttHsp70, by means of mutating its C-tail (ttHsp70^{CM}), has a strong adverse effect on refolding (Fig. 4A). Our data corroborate previous reports on the important role of this interaction (40, 41). The Hsp70/Hsp40/NEF machinery also assists in the refolding of denatured MBP, which, in contrast to denatured luciferase, can fold back to its native state even in the absence of the chaperone machinery (Fig. 4B and fig. S15D). Interestingly, the Hsp70/Hsp40/NEF machinery decreases the apparent folding rate of MBP 4-fold, but it increases almost 3-fold the yield of the folded protein (Fig. 4B). The ttHsp70^{CM} variant confers a strong “anti-folding” activity to the tripartite chaperone machinery by suppressing the rate of refolding and the yield of the refolded protein. The very fast association rate of ttHsp40 with client proteins, makes ttHsp40 as efficient as SecB in retaining a protein in its unfolded state despite its much lower affinity ($K_d \sim 10 \mu\text{M}$ vs $\sim 0.2 \mu\text{M}$) for client proteins (Fig. 4C). The mechanistic basis is that in ttHsp40 the binding sites are flat and fully solvent-exposed, in contrast to TF and SecB where the binding sites are located in deep crevices (Fig. 4D-F).

In light of the present data, the Hsp70/Hsp40-driven client refolding cycle (2, 14) could be further refined as depicted in fig. S16. Hsp40 binds to several hydrophobic regions within the client and presents it to Hsp70 in an unfolded state, which is also the conformational state that Hsp70s prefer to bind (42). The Hsp70 C-tail displaces partially or entirely the client from Hsp40 depending on the number of client-binding sites on Hsp40 (Fig. 3C,D). Hsp70 binding to Hsp40 does not elicit any major conformational changes, as indicated by the NMR spectra (fig. S13A), and thus the client is transferred via a simple displacement mechanism. The Hsp40 in the context of the ternary complex has a lower affinity for the client (fig. S13I) and exhibits weaker holdase activity (fig. S15E) because two of the client-binding sites are occupied by the Hsp70 C-tail. In the state with low affinity for the client (i.e. Hsp70 in the ATP state after NEF has displaced ADP (2, 13)), the ternary chaperone

complex is thus expected to readily release the client for it to either fold to its native state or be captured again by Hsp40 and Hsp70 for another chaperone-mediated cycle.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments:

We wish to thank X. Guan for contributions during the early phase of the project, Z. Xia for help with molecular cloning, Y. Xia for help with the NMR experiments, B. Bukau for providing the CbpA, Ydj1, Sis1 and DNAJB1 plasmids, and L. E. Kay for the tHsp70 plasmid. NMR spectra were acquired at the St. Jude Biomolecular NMR Spectroscopy Center. Peptides were synthesized at the St. Jude Macromolecular Synthesis core.

Funding: This work was supported by the U.S. National Institutes of Health grant R35 GM122462 (C.G.K.) and ALSAC.

References and Notes:

1. Bukau B, Weissman J, Horwich A, Molecular chaperones and protein quality control. *Cell* 125, 443–451 (2006). [PubMed: 16678092]
2. Balchin D, Hayer-Hartl M, Hartl FU, In vivo aspects of protein folding and quality control. *Science* 353, aac4354 (2016).
3. Saibil H, Chaperone machines for protein folding, unfolding and disaggregation. *Nat Rev Mol Cell Biol* 14, 630–642 (2013). [PubMed: 24026055]
4. Saio T, Guan X, Rossi P, Economou A, Kalodimos CG, Structural basis for protein antiaggregation activity of the trigger factor chaperone. *Science* 344, 1250494 (2014).
5. Huang C, Rossi P, Saio T, Kalodimos CG, Structural basis for the antifolding activity of a molecular chaperone. *Nature* 537, 202–206 (2016). [PubMed: 27501151]
6. Deville C, Carroni M, Franke KB, Topf M, Bukau B, Mogk A, Saibil HR, Structural pathway of regulated substrate transfer and threading through an Hsp100 disaggregase. *Sci Adv* 3, e1701726 (2017).
7. Verba KA, Wang RY, Arakawa A, Liu Y, Shirouzu M, Yokoyama S, Agard DA, Atomic structure of Hsp90-Cdc37-Cdk4 reveals that Hsp90 traps and stabilizes an unfolded kinase. *Science* 352, 1542–1547 (2016). [PubMed: 27339980]
8. Rosenzweig R, Moradi S, Zarrine-Afsar A, Glover JR, Kay LE, Unraveling the mechanism of protein disaggregation through a ClpB-DnaK interaction. *Science* 339, 1080–1083 (2013). [PubMed: 23393091]
9. Oroz J, Chang BJ, Wysoczanski P, Lee CT, Perez-Lara A, Chakraborty P, Hofele RV, Baker JD, Blair LJ, Biernat J, Urlaub H, Mandelkow E, Dickey CA, Zweckstetter M, Structure and pro-toxic mechanism of the human Hsp90/PPIase/Tau complex. *Nat Commun* 9, 4532 (2018). [PubMed: 30382094]
10. Burmann BM, Wang C, Hiller S, Conformation and dynamics of the periplasmic membrane-protein-chaperone complexes OmpX-Skp and tOmpA-Skp. *Nat Struct Mol Biol* 20, 1265–1272 (2013). [PubMed: 24077225]
11. Kampinga HH, Craig EA, The HSP70 chaperone machinery: J proteins as drivers of functional specificity. *Nat Rev Mol Cell Biol* 11, 579–592 (2010). [PubMed: 20651708]
12. Nillegoda NB, Wentink AS, Bukau B, Protein Disaggregation in Multicellular Organisms. *Trends Biochem Sci* 43, 285–300 (2018). [PubMed: 29501325]
13. Mayer MP, Gierasch LM, Recent advances in the structural and mechanistic aspects of Hsp70 molecular chaperones. *J Biol Chem* 294, 2085–2097 (2019). [PubMed: 30455352]
14. Craig EA, Marszalek J, How Do J-Proteins Get Hsp70 to Do So Many Different Things? *Trends Biochem Sci* 42, 355–368 (2017). [PubMed: 28314505]

15. Behnke J, Mann MJ, Scruggs FL, Feige MJ, Hendershot LM, Members of the Hsp70 Family Recognize Distinct Types of Sequences to Execute ER Quality Control. *Mol Cell* 63, 739–752 (2016). [PubMed: 27546788]
16. Nillegoda NB, Kirstein J, Szlachcic A, Berynskyy M, Stank A, Stengel F, Arnsburg K, Gao X, Scior A, Aebersold R, Guilbride DL, Wade RC, Morimoto RI, Mayer MP, Bukau B, Crucial HSP70 co-chaperone complex unlocks metazoan protein disaggregation. *Nature* 524, 247–251 (2015). [PubMed: 26245380]
17. Hageman J, Rujano MA, van Waarde MA, Kakkar V, Dirks RP, Govorukhina N, Oosterveld-Hut HM, Lubsen NH, Kampinga HH, A DNAJB chaperone subfamily with HDAC-dependent activities suppresses toxic protein aggregation. *Mol Cell* 37, 355–369 (2010). [PubMed: 20159555]
18. Mitra A, Shevde LA, Samant RS, Multi-faceted role of HSP40 in cancer. *Clin Exp Metastasis* 26, 559–567 (2009). [PubMed: 19340594]
19. Zarouchlioti C, Parfitt DA, Li W, Gittings LM, Cheetham ME, DNAJ Proteins in neurodegeneration: essential and protective factors. *Philos Trans R Soc Lond B Biol Sci* 373, (2018).
20. Barends TR, Brosi RW, Steinmetz A, Scherer A, Hartmann E, Eschenbach J, Lorenz T, Seidel R, Shoeman RL, Zimmermann S, Bittl R, Schlichting I, Reinstein J, Combining crystallography and EPR: crystal and solution structures of the multidomain cochaperone DnaJ. *Acta Crystallogr D Biol Crystallogr* 69, 1540–1552 (2013). [PubMed: 23897477]
21. Wild J, Altman E, Yura T, Gross CA, DnaK and DnaJ heat shock proteins participate in protein export in *Escherichia coli*. *Genes Dev* 6, 1165–1172 (1992). [PubMed: 1628824]
22. Huang HC, Sherman MY, Kandror O, Goldberg AL, The molecular chaperone DnaJ is required for the degradation of a soluble abnormal protein in *Escherichia coli*. *J Biol Chem* 276, 3920–3928 (2001). [PubMed: 11062236]
23. Saio T, Kawagoe S, Ishimori K, Kalodimos CG, Oligomerization of a molecular chaperone modulates its activity. *Elife* 7, (2018).
24. Rudiger S, Schneider-Mergener J, Bukau B, Its substrate specificity characterizes the DnaJ co-chaperone as a scanning factor for the DnaK chaperone. *EMBO J* 20, 1042–1050 (2001). [PubMed: 11230128]
25. Li J, Qian X, Sha B, The crystal structure of the yeast Hsp40 Ydj1 complexed with its peptide substrate. *Structure* 11, 1475–1483 (2003). [PubMed: 14656432]
26. Rosenzweig R, Sekhar A, Nagesh J, Kay LE, Promiscuous binding by Hsp70 results in conformational heterogeneity and fuzzy chaperone-substrate ensembles. *Elife* 6, (2017).
27. Weinhaupl K, Lindau C, Hessel A, Wang Y, Schutze C, Jores T, Melchionda L, Schonfisch B, Kalbacher H, Bersch B, Rapaport D, Brennich M, Lindorff-Larsen K, Wiedemann N, Schanda P, Structural Basis of Membrane Protein Chaperoning through the Mitochondrial Intermembrane Space. *Cell* 175, 1365–1379 e1325 (2018). [PubMed: 30445040]
28. Lee JH, Zhang D, Hughes C, Okuno Y, Sekhar A, Cavagnero S, Heterogeneous binding of the SH3 client protein to the DnaK molecular chaperone. *Proc Natl Acad Sci U S A* 112, E4206–4215 (2015).
29. Suh JY, Iwahara J, Clore GM, Intramolecular domain-domain association/dissociation and phosphoryl transfer in the mannitol transporter of *Escherichia coli* are not coupled. *Proc Natl Acad Sci U S A* 104, 3153–3158 (2007). [PubMed: 17360622]
30. Summers DW, Douglas PM, Ren HY, Cyr DM, The type I Hsp40 Ydj1 utilizes a farnesyl moiety and zinc finger-like region to suppress prion toxicity. *J Biol Chem* 284, 3628–3639 (2009). [PubMed: 19056735]
31. Wu Y, Li J, Jin Z, Fu Z, Sha B, The crystal structure of the C-terminal fragment of yeast Hsp40 Ydj1 reveals novel dimerization motif for Hsp40. *J Mol Biol* 346, 1005–1011 (2005). [PubMed: 15701512]
32. Schilke BA, Ciesielski SJ, Ziegelhoffer T, Kamiya E, Tonelli M, Lee W, Cornilescu G, Hines JK, Markley JL, Craig EA, Broadening the functionality of a J-protein/Hsp70 molecular chaperone system. *PLoS Genet* 13, e1007084 (2017).

33. Sahi C, Kominek J, Ziegelhoffer T, Yu HY, Baranowski M, Marszalek J, Craig EA, Sequential duplications of an ancient member of the DnaJ-family expanded the functional chaperone network in the eukaryotic cytosol. *Mol Biol Evol* 30, 985–998 (2013). [PubMed: 23329686]
34. Reidy M, Sharma R, Shastry S, Roberts BL, Albino-Flores I, Wickner S, Masison DC, Hsp40s specify functions of Hsp104 and Hsp90 protein chaperone machines. *PLoS Genet* 10, e1004720 (2014).
35. Kityk R, Kopp J, Mayer MP, Molecular Mechanism of J-Domain-Triggered ATP Hydrolysis by Hsp70 Chaperones. *Mol Cell* 69, 227–237 e224 (2018). [PubMed: 29290615]
36. Suzuki H, Noguchi S, Arakawa H, Tokida T, Hashimoto M, Satow Y, 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. *Biochemistry* 49, 8577–8584 (2010). [PubMed: 20809635]
37. Li J, Wu Y, Qian X, Sha B, Crystal structure of yeast Sis1 peptide-binding fragment and Hsp70 Ssa1 C-terminal complex. *Biochem J* 398, 353–360 (2006). [PubMed: 16737444]
38. Schroder H, Langer T, Hartl FU, Bukau B, DnaK, DnaJ and GrpE form a cellular chaperone machinery capable of repairing heat-induced protein damage. *EMBO J* 12, 4137–4144 (1993). [PubMed: 7900997]
39. Lu Z, Cyr DM, Protein folding activity of Hsp70 is modified differentially by the hsp40 co-chaperones Sis1 and Ydj1. *J Biol Chem* 273, 27824–27830 (1998).
40. Freeman BC, Myers MP, Schumacher R, Morimoto RI, Identification of a regulatory motif in Hsp70 that affects ATPase activity, substrate binding and interaction with HDJ-1. *EMBO J* 14, 2281–2292 (1995). [PubMed: 7774586]
41. Yu HY, Ziegelhoffer T, Craig EA, Functionality of Class A and Class B J-protein co-chaperones with Hsp70. *FEBS Lett* 589, 2825–2830 (2015). [PubMed: 26247431]
42. Sekhar A, Velyvis A, Zoltsman G, Rosenzweig R, Bouvignies G, Kay LE, Conserved conformational selection mechanism of Hsp70 chaperone-substrate interactions. *Elife* 7, (2018).
43. Sprangers R, Velyvis A, Kay LE, Solution NMR of supramolecular complexes: providing new insights into function. *Nat Methods* 4, 697–703 (2007). [PubMed: 17762877]
44. Tzeng SR, Pai MT, Kalodimos CG, NMR studies of large protein systems. *Methods Mol Biol* 831, 133–140 (2012). [PubMed: 22167672]
45. Gelis I, Bonvin AM, Keramisanou D, Koukaki M, Gouridis G, Karamanou S, Economou A, Kalodimos CG, Structural basis for signal-sequence recognition by the translocase motor SecA as determined by NMR. *Cell* 131, 756–769 (2007). [PubMed: 18022369]
46. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A, NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 6, 277–293 (1995). [PubMed: 8520220]
47. Krois AS, Ferreon JC, Martinez-Yamout MA, Dyson HJ, Wright PE, Recognition of the disordered p53 transactivation domain by the transcriptional adapter zinc finger domains of CREB-binding protein. *Proc Natl Acad Sci U S A* 113, E1853–E1862 (2016).
48. Khanra N, Rossi P, Economou A, Kalodimos CG, Recognition and targeting mechanisms by chaperones in flagellum assembly and operation. *Proc Natl Acad Sci U S A* 113, 9798–9803 (2016). [PubMed: 27528687]
49. Shen Y, Bax A, Protein structural information derived from NMR chemical shift with the neural network program TALOS-N. *Methods Mol Biol* 1260, 17–32 (2015). [PubMed: 25502373]
50. Guntert P, Automated NMR structure calculation with CYANA. *Methods Mol Biol* 278, 353–378 (2004). [PubMed: 15318003]
51. Brunger AT, Version 1.2 of the Crystallography and NMR system. *Nat Protoc* 2, 2728–2733 (2007). [PubMed: 18007608]
52. Mittermaier A, Kay LE, New tools provide new insights in NMR studies of protein dynamics. *Science* 312, 224–228 (2006). [PubMed: 16614210]
53. Palmer AG 3rd, NMR characterization of the dynamics of biomacromolecules. *Chem Rev* 104, 3623–3640 (2004). [PubMed: 15303831]
54. Mulder FA, Mittermaier A, Hon B, Dahlquist FW, Kay LE, Studying excited states of proteins by NMR spectroscopy. *Nat Struct Biol* 8, 932–935 (2001). [PubMed: 11685237]

55. Bieri M, Gooley PR, Automated NMR relaxation dispersion data analysis using NESSY. *BMC Bioinformatics* 12, 421 (2011). [PubMed: 22032230]
56. Chun SY, Strobel S, Bassford P Jr., Randall LL, Folding of maltose-binding protein. Evidence for the identity of the rate-determining step in vivo and in vitro. *J Biol Chem* 268, 20855–20862 (1993).
57. Diamant S, Azem A, Weiss C, Goloubinoff P, Increased efficiency of GroE-assisted protein folding by manganese ions. *J Biol Chem* 270, 28387–28391 (1995).
58. Rodriguez F, Arsene-Ploetze F, Rist W, Rudiger S, Schneider-Mergener J, Mayer MP, Bukau B, Molecular basis for regulation of the heat shock transcription factor sigma32 by the DnaK and DnaJ chaperones. *Mol Cell* 32, 347–358 (2008). [PubMed: 18995833]
59. Marsh JA, Singh VK, Jia Z, Forman-Kay JD, Sensitivity of secondary structure propensities to sequence differences between alpha- and gamma-synuclein: implications for fibrillation. *Protein Sci* 15, 2795–2804 (2006). [PubMed: 17088319]

(A) The structure of ttHsp40 is shown in cartoon and the various domains, with their amino acid residue boundaries, are labeled. The client-binding sites are colored blue on a solvent-exposed surface model of ttHsp40. The residues that make up the client-binding sites are shown in ball-and-stick. **(B)** Hydrophobicity plot of PhoA as a function of its primary sequence. A hydrophobicity score (Roseman algorithm, window = 9) higher than zero denotes increased hydrophobicity. The sites identified by NMR to be recognized by ttHsp40 are highlighted in blue (strong) or grey (weak). The strong sites are labeled *a* through *g* and

the weak sites are labelled $w1$ through $w8$. **(C)** Domains involved in client binding in various Hsp40s as determined by NMR. ZF denotes the zinc finger domain in Ydj1. **(D)** The identified auto-inhibitory mechanism in Ydj1 mediated by the interaction of its C-tail and CBD2.

Fig. 2. Structure of the ttHsp40–PhoA complex.

(A) Solution structure of the ttHsp40–PhoA. ttHsp40 is shown as a space-filling model in grey. The PhoA sites recognized by ttHsp40 are shown as space-filling models and colored per the color code in the graphic in panel B. The flexible regions of PhoA are shown as a pink ribbon. (B) Cartoon representation of the complex between ttHsp40 and PhoA as determined by NMR is shown in the middle of the panel. The strong binding sites in PhoA are colored distinctly, whereas the weak binding sites are colored uniformly in light blue. The inter-subunit distance between the two CBD1s is ~60 Å and the distance between CBD1

and CBD2 within the same subunit is ~ 35 Å. The red arrows denote a dynamic equilibrium of the indicated PhoA sites between ttHsp40-bound and dissociated state. Expanded views of the ttHsp40–PhoA complex highlighting the binding details and contacts that mediate recognition of the seven strong PhoA sites (*a* through *g*) by ttHsp40 are shown. ttHsp40 in the expanded views is shown as grey ribbon and residues contacting PhoA are displayed as grey ball-and-stick. As discussed in the main text and in Material and Methods, the structure of the complex shown here displays one of the several similar structures that are in dynamic equilibrium. For example, PhoA^{*b*} may be bound to CBD1 instead of PhoA^{*a*}, and PhoA^{*e*} or PhoA^{*g*} may be bound to CBD2 instead of PhoA^{*f*}. Simultaneous binding of all four ttHsp40 CBDs by PhoA results in the most stable complex (fig. S10).

Fig. 3. Interaction between ttHsp70 and ttHsp40.

(A) Hsp70 (PDB ID 5NRO) and ttHsp40 shown as space-filling models. The red arrow indicates the interaction between the J domain and the NBD of Hsp70 as observed by NMR and crystallography (35). The ttHsp70 C-tail interacts with CBD2 of ttHsp40 and the NMR structure of their complex is shown in the expanded view. D610 and Y611 in ttHsp70 C-tail are the most important residues for mediating its interaction with ttHsp40 CBD2. (B) Overlay of the structures of the ttHsp70 C-tail and PhxA^d, both in complex with ttHsp40 CBD2, shows that their binding to CBD2 is mutually exclusive. (C) Schematic of ttHsp40 showing the competition between the client and Hsp70 C-tail for CBD2. (D) Summary of the present NMR findings on the competition between client and Hsp70 C-tail in various Hsp40s. Sis1 and DNAJB1 use only CBD1 to interact with either the client or Hsp70 and cannot engage simultaneously both partners.

Fig. 4. Chaperone activities and chaperone-client binding architectures.

(A) Folding of denatured luciferase by the ttHsp70/ttHsp40/ttNEF tripartite chaperone machinery measured as percentage of its enzymatic activity. ttHsp40 superscripts 1M, 2M, and 1,2M refer to amino acid substitutions that impair client binding in the CBD1, CBD2 and both CBD1 and CBD2, respectively. Detailed graphs are shown in fig. S15A. Data are mean $\pm$ SD. (B) Folding of urea-denatured MBP (pre form) by the tripartite chaperone machinery. The folding rates were determined as described in Materials and Methods. ttHsp70^{CM} refers to the C-tail double substitution (D610-Y611 to A610-A611) that impairs binding to ttHsp40. Folding was monitored by Trp fluorescence. Data are mean $\pm$ SD. (C) Folding of urea-denatured MBP (pre form) in the presence of ttHsp40, SecB and TF chaperones. For all experiments the MBP:chaperone ratio is 1:2. (D) Cartoon representation of the structures of ttHsp40-PhoA, (E) TF-PhoA (4), and (F) SecB-PhoA (5). Chaperones are colored in blue and PhoA in orange. The structures display the distinct and very different binding modes that the three chaperones use to engage their client proteins. The K_d and rates of chaperone-PhoA complexes are shown.